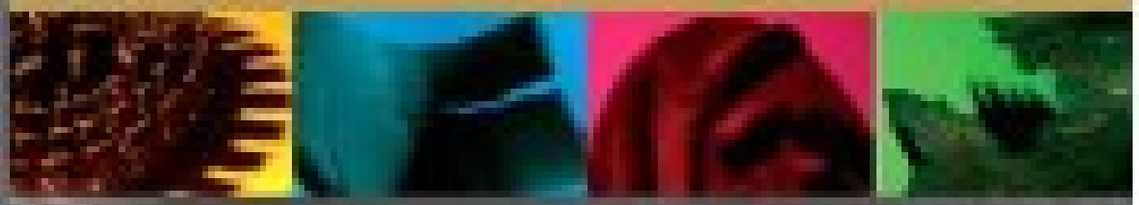
به نام خدا

مرکز دانلود رایگان
مهندسی متالورژی و مواد

www.Iran-mavad.com

COMPREHENSIVE CORROSION



2.01 Principles of Corrosion in Liquids

G. S. Frankel
Fontana Corrosion Center, MSE The Ohio State University, Enarson Hall 154 W 12th Avenue, Columbus, OH 43210, USA
R. A. Cottis
Corrosion and Protection Centre, School of Materials, University of Manchester, P.O. Box 88, Sackville Street, Machester M60 1QD, UK

2.01.1 Thermodynamics

The corrosion reaction is an electrochemical cell. Like other electrochemical cells such as batteries, corrosion requires at least one anodic and one cathodic half reaction, where anodic reactions generate electrons and cathodic reactions consume electrons. In closed electrochemical cells, charge is consumed; there is no net generation or consumption of electrons.

For a generic metal M dissolving in an acidic solution, the dominant half reactions might be

$$\mathrm{M} = \mathrm{M}^{n+} + ne^{-}(\text{anodic half-cell reaction}) \quad [1]$$

$$2\mathrm{H}^{+} + 2\mathrm{e}^{-} = \mathrm{H}_2(\text{cathodic half-cell reaction}) \quad [2]$$

Each half reaction will be in equilibrium (where the rates in the forward and reverse directions are equal) at a reversible potential. For half reactions in which all reactants and products are in their standard state, the reversible potential is given by the standard potential, E°, which can be calculated from the difference in standard free energy of formation of the products and reactants, ΔG°, the number of electrons transferred in the half reaction, n, and Faraday's constant, F, according to

$$E^{\circ} = -\Delta G^{\circ}/nF \quad [3]$$

Reversible and standard potentials of any half reaction versus a suitable reference electrode are given. It is customary to use the standard hydrogen electrode (SHE) shown in eqn [2] as a universal reference when listing the potential of a half-cell. Under standard state conditions, the activity of the protons is unity, $a_{\mathrm{H}^+} = 1$, and the partial pressure of hydrogen gas is one bar, $P_{\mathrm{H}_2} = 1$ bar (1 bar = 0.987 atm). The potential of the SHE is arbitrarily set at zero, $E^{\circ}_{\mathrm{SHE}} \equiv 0$. Correspondingly, this means that the free energy of protons in solution at unit activity is arbitrarily set to 0 as a reference point, which is required for any free energy scale.

In the more general case where species are not present in their standard state concentrations, the reversible potential at equilibrium for a half reaction, E^{rev}, is given by the Nernst equation

$$E^{\mathrm{rev}} = E^{\circ} + \frac{RT}{nF}\ln\left\{\frac{\Pi(a_{\mathrm{ox}})^{j}}{\Pi(a_{\mathrm{red}})^{k}}\right\} \quad [4]$$

where R is the gas constant, T is temperature, and $\Pi(a_{\mathrm{ox}})^{j}$ and $\Pi(a_{\mathrm{red}})^{j}$ are the products of the activities of the oxidized and reduced species respectively, raised to a power equal to their stoichiometric coefficients in the half reaction equations. For dilute solutions, the activity coefficient is approximately equal to one. Therefore

$$a_{\mathrm{A}} = \gamma_{\mathrm{A}} c_{\mathrm{A}} \approx c_{\mathrm{A}} = [\mathrm{A}] \quad [5]$$

where γ_{A} is the activity coefficient of species A, and c_{A} and [A] represent the concentration of species A in units of moles per liter. It should be noted that equating activity to concentration for dilute concentrations is a statement of Henry's law, which utilizes a hypothetical standard state of 1 mol l^{-1}. The Henrian activity is different than that obtained from Raoult's Law, which states that activity equals mole fraction (activity coefficient equals 1) when mole fraction is close to unity. Raoultian activity is used commonly in electrochemistry for gases and solvents. Thus, for the general half reaction

$$a\mathrm{A} + b\mathrm{B} + n\mathrm{e}^{-} = c\mathrm{C} + d\mathrm{D} \quad [6]$$

The Nernst equation is

$$E^{rev} = E^{o} + \left(\frac{RT}{nF}\right) \ln\left(\frac{[A]^{a}[B]^{b}}{[C]^{c}[D]^{d}}\right) \quad [7]$$

And at room temperature

$$E^{rev} = E^{o} + \left(\frac{0.059}{n}\right) \log\left(\frac{[A]^{a}[B]^{b}}{[C]^{c}[D]^{d}}\right)$$

The reversible potentials calculated from Nernst equations are important in corrosion, because half reactions will proceed in the net anodic (oxidizing) direction only at potentials more noble than the reversible potential. Similarly, cathodic (reducing) reactions will occur only at potentials less anodic than the reversible potential. Therefore, an analysis of the reversible potentials for two reactions provides insight into the thermodynamic likelihood for an electrochemical cell to proceed in the direction required for corrosion to occur: the reversible potential for the cathodic reaction must be higher than that for the anodic corrosion reaction.

2.01.2 Kinetics

Since electrochemical half reactions either generate or consume electrons, their rates can be assessed and measured as currents, or current densities when normalized to the area. As mentioned, an electrochemical half reaction will proceed in the net anodic or cathodic direction at potentials above or below the reversible potential respectively. At the reversible potential, the reaction proceeds in both directions at the same rate, which when converted to the units of current density is called the exchange current density, i_0.

Electrochemical kinetics describe the rate of an electrochemical reaction as a function of the potential relative to the reversible potential or a reference electrode. In general, the rate of an electrochemical half reaction increases as the potential difference from the reversible potential increases. At a large enough potential difference from the reversible potential, the Tafel equation describes this relationship when the rate is limited by the transfer of charge across the interface:

$$E - E^{rev} = b \log(i/i_0) \quad [8]$$

where b is the Tafel slope. On a plot of E versus $\log|i|$, called an Evans diagram, this expression is a straight line with the slope of b. In **Figure 1**, the expression is plotted as $\log|i|$ versus E and the slope is $1/b$:

$$\log i = \log i_0 + \frac{E - E^{rev}}{b} \quad [9]$$

As shown in **Figure 1**, the hydrogen evolution reaction, eqn [2], goes in the cathodic direction at potentials below the reversible potential for that reaction, which is given by the Nernst equation. The metal dissolution reaction goes in the anodic direction at potentials above the reversible potential for that reaction.

Many metals will dissolve by uniform corrosion when immersed in an acidic solution. In an acid, the primary cathodic reaction is often the hydrogen evolution reaction, and metal dissolution may be written as eqn [1]. Both reactions will occur on

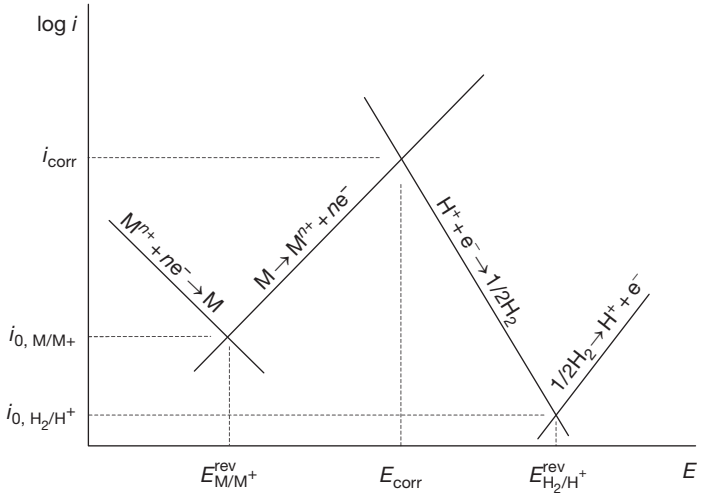


Figure 1 Schematic Evans diagram for corroding active metal in an acid. Reproduced from Frankel, G. S.; Landolt, D. In *Corrosion and Oxide Films*, Stratmann, M., Frankel, G. S., Eds.; 2003; p. 38, with permission from Wiley-VCH Verlag GmbH & Co. KGaA.

the same surface at essentially the same potential. This potential is called a 'mixed potential,' because it is between the reversible potential of the two reactions. It is also called the open circuit potential (OCP) or for a corroding metal, the corrosion potential, E_{corr}.

As shown in **Figure 1**, E_{corr} is the potential at which the anodic and cathodic current densities are equal, which is required for charge neutrality. For the more general case where there might be more than one anodic or cathodic reaction, E_{corr} at any point in time is the potential where the following is obeyed:

$$\sum i_a = \sum i_c \qquad [10]$$

For uniform corrosion to occur, the anodic reaction must not be localized at specific spots on the surface of a sample. Instead, the anodic reaction must occur at multiple sites across the surface and continually move to new sites, thereby evenly removing material across the surface over time. Similarly, the cathodic reaction must not be localized. When these reactions remain at fixed locations, the attack is nonuniform, resulting in a localized form of attack such as pitting, crevice, or intergranular corrosion.

2.01.3 Passivity

Another form of nominally uniform attack involves a passive metal surface rather than the active dissolution described in the previous section. Passivity describes the relatively low rate of corrosion exhibited by metals covered by a thin oxide or hydroxide film, the passive film. Passive films are thermodynamically stable at high potentials for many metals. For example, the Nernst equation can be calculated for a reaction in which a metal oxide is formed from a metal in an aqueous solution:

$$MO_n + 2nH^+ + 2ne^- = M + nH_2O \qquad [11]$$

$$E^{rev} = E^\circ - 0.059pH \qquad [12]$$

where $E^\circ = -(nG^\circ_{H_2O} - G^\circ_{MO_n})/nF$ following eqn [3]. As indicated before, at potentials above this reversible potential, reaction [11] would tend to go in the anodic direction, from right to left or the opposite direction of what is written in eqn [11], resulting in the formation of an oxide film. For many metals, this potential is below the potential for hydrogen evolution associated with proton or water reduction, and therefore, the passivation reaction will proceed spontaneously in aqueous solutions.

Passive films form spontaneously in the majority of practical environments on stainless steels, and alloys of Al, Ni, and Ti. Fe spontaneously passivates at high pH or at high potentials. W, in contrast, spontaneously passivates at low pH. For many metals, including Al, Cu, and Zn, the passive oxide film is stable only in the mid-pH region. These metals are described as being amphoteric in nature.

Stainless steels are iron-based alloys with good corrosion resistance owing to the formation of a Cr-rich passive film. When Fe is alloyed with >13% Cr, such a protective passive film can form, thereby achieving the good corrosion protection associated with Cr while maintaining the mechanical properties of steel. Common grades of stainless steel contain 16–20% Cr, as well as other elements such as Ni and Mo.

Spontaneously formed passive films typically have a thickness of 2–4 nm, although they tend to thicken as the potential is increased. The protectiveness of passive films usually increases with potential, and the application of a high potential often improves the corrosion resistance by thickening the passive film. This is the basis for anodization, which is commonly performed on Al to thicken the aluminum oxide layer. At room temperature, the thickness increases by about 1.2 nm V^{-1}. Iron and nickel oxides are semiconducting so that they support the oxygen evolution reaction at potentials above the reversible potential for that reaction, and therefore cannot be anodized to high potentials. Cr-rich passive films that form on Cr and Cr-containing alloys such as stainless steel dissolve at high potentials forming the soluble chromate ion (a process known as transpassive corrosion), and thus also cannot be anodized.

It should be noted that almost all metals and alloys that can be protected by a spontaneously formed passive film are susceptible to accelerated localized corrosion such as pitting or crevice corrosion as the result of localized breakdown of the otherwise protective passive film. Such breakdown requires the presence of an aggressive species, such as chloride ions, and is accelerated for many passive metals by the presence of high chloride concentration, high temperature, and high potentials.

2.01.4 Uniform Corrosion in Practice

Metals and alloys can corrode uniformly in one of three states:

- Active
  In this case, the metal is in direct contact with the environment, the rate of corrosion being controlled by the kinetics of the anodic and cathodic

reactions, with a contribution (usually small except in pure waters) of solution resistance due to the separation between the instantaneous anodic and cathodic sites on the metal surface. It is normally only practicable to use alloys in this state if the rate of hydrogen evolution is low, which effectively limits alloy systems that can be used in this state in practice to those based on metals with an equilibrium potential that is not too far below that for hydrogen at the relevant pH (such as copper, nickel, and iron), or on metals that have a very low exchange current density for hydrogen evolution (such as lead and zinc). In aerated environments, it is common for the rate of corrosion to be controlled by the transport of oxygen to the metal surface, which leads to a typical uniform corrosion rate in air-saturated solutions such as seawater of 0.1–1.0 mm $year^{-1}$, depending on the flow conditions. In stagnant liquids, soils and similar situations where the flow is restricted, when the surface is covered by a corrosion product film or when the solution has a low conductivity, such as many fresh waters, the corrosion rate will be lower.

- Immune
  In this case the metal is the thermodynamically stable state and corrosion cannot occur at a significant rate (in an environment that is completely free of ions of the metal, there will be a very low corrosion rate as the metal oxidizes to a sufficient extent to achieve the metal ion concentration that is in equilibrium with the metal). In the presence of complexants or species that produce stable salts, the metal ion concentration may be maintained below the value that is in equilibrium with the metal, allowing corrosion to occur. Thus, in cyanide leaching of gold ores, gold ions are complexed by cyanide, allowing continuing corrosion of particles of metallic gold, with oxygen reduction as the cathodic reaction. In this condition, the metal is effectively corroding in the active state, although the kinetics will tend to be rather more complex.
- Passive
  In the case of passive metals, the corrosion rate is controlled by the properties of the passive film. This is usually an oxide or hydroxide, although passivation by salt films is also possible. The kinetics of passive corrosion have been studied in considerable detail in recent years, largely because of the requirements of nuclear waste disposal, for which some approaches depend on the corrosion performance of passive alloys over many thousands of years (see **Chapter 2.39, Predictive Modeling of Corrosion** for a detailed discussion of this work). Other applications where the rate of passive corrosion is important are in the primary circuit of water-cooled nuclear reactors, where the release of radioactive transmutation products from the metals of construction of the reactor leads to significant safety hazards, and in biomedical applications where the release of metal ions (such as nickel and chromium) can lead to allergic or toxic effects on neighboring tissue. Otherwise, the rate of passive corrosion is usually not significant, and the various forms of localized corrosion tend to present much more serious problems for passive alloys.

2.01.5 Uniform Corrosion of Iron

Unprotected iron will corrode and form rust in many environments. Lorenz and Heusler have provided a good summary of their work and others' on the mechanism of iron corrosion.[1] The reactions commonly used to describe this process in an aerated environment are

$$Fe \rightarrow Fe^{2+} + 2e^- \quad [13]$$

$$O_2 + 2H_2O + 4e^- \rightarrow 4OH^- \quad [14]$$

$$Fe^{2+} + 2OH^- \rightarrow Fe(OH)_2 \quad [15]$$

$$4Fe(OH)_2 + O_2 \rightarrow 2Fe_2O_3 + 4H_2O \quad [16]$$

where the final iron oxide product, red rust, is usually a hydrated form of Fe_2O_3 (usually assigned the formula FeO(OH)). However, it seems unlikely that solid $Fe(OH)_2$ is formed, as this would be expected to lead to a form of passivation, with an underlayer of $Fe(OH)_2$ covered by Fe_2O_3, and it is more likely that the oxidation of ferrous species to ferric occurs in solution, leading to precipitation of rust in solution away from the surface of the metal, resulting in the nonprotective corrosion product film that everyone is familiar with. In a deaerated aqueous environment (in the absence of oxygen), the cathodic reaction is hydrogen evolution via proton (eqn [2]) or water reduction and the ferric product will not form. The rate of reaction in this case is low in neutral or alkaline environments.

The mechanism of ferrous ion formation has been shown to be more complicated than eqn [13] as it involves multiple single electron transfer steps and depends on the pH of the environment:

$$Fe + H_2O \rightleftharpoons FeOH + H^+ + e^- \quad [17]$$

$$FeOH \rightarrow FeOH^+ + e^- \quad [18]$$

$$FeOH^+ + H^+ \rightarrow Fe^{2+} + H_2O \quad [19]$$

The second charge transfer step (reaction [18]) is the rate-determining step, and in consequence, the rate of the overall reaction depends on the concentration of FeOH on the iron surface, and hence (by virtue of the equilibrium of reaction [17]) on the pH. Thus, in deaerated environments, the anodic and cathodic reaction kinetics both depend on pH; somewhat counter-intuitively, the rate of the anodic reaction at a constant potential decreases as the solution becomes more acidic (i.e., as the pH decreases), but this is countered in practice by the increase in the rate of the cathodic hydrogen evolution reaction due to the increased hydrogen ion concentration, which leads to an increase in the corrosion potential.

At high pH, iron spontaneously passivates as described earlier, forming a thin oxide layer. In neutral or acidic environments, passivity on Fe is only achieved by the application of a sufficiently high potential. If the potential is scanned, the polarization curve will exhibit an active–passive transition. In the active region, the current can be extremely high, approaching $1\,A\,cm^{-2}$ (corresponding to a corrosion rate in the region of $1\,mm\,h^{-1}$). In acidic solutions, a limiting current associated with mass-transport limited dissolution in the presence of a salt film on the surface can be observed over a range of potential. However, after the active–passive transition, the current density might drop by 4 orders of magnitude as a result of the formation of a thin passive film. On iron, the passive film has been found to be a bilayer, with an inner layer of Fe_3O_4 and an outer layer of γ-Fe_2O_3.

Alloying of iron with carbon and other elements such as Mn and Si to make steel generally has a small influence on the corrosion rate, although sulfide inclusions (typically MnS) act as local sources of sulfide and other species that accelerate corrosion locally. However, alloying elements such as Cu, Cr, P, and Ni added in small quantities to create low-alloy steels can result in considerable improvement in corrosion resistance. One class of low-alloy steels specifically designed for improved atmospheric corrosion resistance is referred to as weathering steels.[2] The rust layer on weathering steels is more compact and protective than that formed on plain carbon steel, although not as thin and protective as the passive film formed on stainless steels having Cr content >13%. The protectiveness of this compact thick film formed on weathering steels is improved by periodic drying, and hence, their primary use is in atmospheric exposure rather than immersed or buried conditions. However, under many atmospheric conditions, the oxide layer on weathering steels provides sufficient protection even without painting, which is a huge advantage as the maintenance costs associated with repainting degraded structures can be eliminated.

2.01.6 Uniform Corrosion of Copper

Copper is a noble metal in that its reversible potential is higher than that for the hydrogen evolution reaction. Therefore, Cu is thermodynamically immune to corrosion in deaerated noncomplexing environments. This is the reason why, at the time of this writing, many countries are considering long-term disposal of high-level radioactive waste in Cu canisters that would be located in the anoxic (oxygen-free) saturated zone (below the water table) of a geologic formation. However, complexation of copper ions by, for instance, chloride or ammonia, or the presence of species, such as sulfide, that form copper salts that are more stable than the oxide, can decrease the reversible potential for Cu dissolution to the extent that it is below that of hydrogen evolution, rendering the Cu susceptible to corrosion even in anoxic environments.[3] Furthermore, Cu will corrode in the presence of oxygen and other oxidizing agents.

Cu is widely used in water distribution piping, both for hot and cold water. Cu water piping will occasionally fail by pitting corrosion. This form of attack is not well understood[4] (there appear to be several different mechanisms of pitting, at least some of which are microbially induced), and usually confined to certain districts or groups of houses. The concentrations of sulfate, chloride, hydroxide, carbonate, and bicarbonate ions can influence Cu pitting susceptibility.[5]

In aerated solutions, the surface of Cu is typically covered with a film of CuO, which provides marginal protection. In atmospheric conditions, Cu eventually develops a protective patina; in most atmospheres (where there is significant SO_2 present as a result of pollution) this has the structure, composition, and characteristic green color of the mineral bronchantite, $Cu_4(SO_4)(OH)_6$.[6,7] The patina can also contain carbonate, or by the seashore, chloride.[8]

The corrosion rate of Cu in the atmosphere has been found to depend on the conditions existing during the initial period of the exposure. When the initial exposure was in the winter, the rate of corrosion over a period of years was greater than when the initial exposure was in the summer, indicating that the nature of the initial surface oxide film plays a critical role in the subsequent development of that film.

Cu is often alloyed for improved mechanical properties. Brasses and Al bronzes are susceptible to dealloying attack whereby the less noble alloying element is dissolved preferentially, leaving a porous structure with little mechanical strength. Dezincification is the most common form of dealloying, caused by the preferential dissolution of Zn from brass. A small addition of As to the alloy, only about 0.04%, will reduce the susceptibility to dezincification; additions of Sn, Sb, or P are also effective.[9]

2.01.7 Uniform Corrosion of Zinc

Zn is an active metal with good corrosion resistance in the atmosphere and in natural waters owing to the formation of a protective oxide or hydroxide surface film. This combination of low corrosion potential and low self corrosion rate makes it an ideal sacrificial anode for steel. Zn and Zn alloys are widely used as corrosion protective coatings on steel, often in combination with paint overcoats. Most steel automotive body panels nowadays are galvanized prior to coating.

The pH of the environment has a large effect on the corrosion rate of Zn as Zn is amphoteric, corroding rapidly in low- and high-pH environments, but slowly in the pH range from about 6 to 12. As a result, the lifetime of galvanized layers is decreased in polluted locations where the mist and dew are acidic.

References

1. Lorenz, W. J.; Heusler, K. E. In *Corrosion Mechanisms*; Mansfeld, F., Ed.; Marcel Dekker: New York, 1987.
2. Murata, T. In *Uhlig's Corrosion Handbook*, 2nd edn.; Revie, R. W., Ed.; John Wiley: New York, 2000.
3. Pourbaix, M. In *Lectures on Electrochemical Corrosion*; Plenum: New York, 1973.
4. Rushing, J. C.; Edwards, M. *Corr. Sci.* **2004**, *46*, 3069–3088.
5. Cong, H.; Michels, H. T.; Scully, J. R. *J. Electrochem. Soc.* **2009**, *156*, C16–C27.
6. Vernon, W. H. J.; Whitby, L. *J. Inst. Met.* **1929**, *42*, 181–202.
7. Vernon, W. H. J. *Whitby, J. Inst. Met.* **1930**, *44*, 389–408.
8. Vernon, W. H. J. *J. Inst. Met.* **1933**, *52*, 93–100.
9. Bengough, G. D.; May, R. *J. Chem. Soc.* **1924**, *32*, 81–269.

2.02 Passivity and Localized Corrosion

G. T. Burstein
Professor of Materials Chemistry and Corrosion, Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK

This article is a revision of the Third Edition article 1.5 by G.T. Burstein, volume 1, pp 1:118–1:150, © 2010 Elsevier B.V.

Abbreviations

AC Alternating current
DC Direct current
SHE Standard Hydrogen Electrode

Symbols

a Reaction coordinate
A Constant
B Constant
E Electrochemical potential
E° Standard equilibrium potential
F The Faraday
i_b Backward current density
i_f Forward current density
i_{pass} Passive current density
k Rate constant
M Atomic weight of metal
R Universal gas constant
t Time
T Temperature
x Metal loss
z Number of electrons released by oxidation
ΔG^* Change in Gibbs free energy
η Overpotential
ρ Density

2.02.1 Introduction

This article was first published in the 3rd edition; it is reprinted here in its entirety with only minor typographic corrections and a few editorial comments incorporated as footnotes.

Passivity is a state of low corrosion rate brought about under a high anodic driving force, or potential, by the presence of an interfacial solid film, usually an oxide. The phenomenon of passivity has been defined in a number of ways,[1,2] most of which are similar in meaning. Although from the more literal sense of the word, passivity could include metals immune from corrosion (such as gold in water); in the parlance of corrosion science, this is specifically excluded. Passive metals are thermodynamically unstable[3]: they possess a kinetic stability, which is engendered by a solid interfacial film, without which corrosion would occur. The oxide film, formed anodically through a mechanism very similar to that of the corrosion process, stifles the rate of further oxidation (corrosion and further passivation) by forming a barrier between the metal surface and its environment. So, many metallic structures and components depend on the state of passivity for their stability. From an engineering point of view, the phenomenon of passivity is a remarkable one due to its requirement of the extremely thin film to procure passivity.

The typical passivating oxide film on many metals is only 1–10 nm in thickness and is produced by the oxidation of the surface to a depth measured in monolayers of atoms. Without the passivating oxide film, many metallic structures would corrode at a very fast rate, phenomenological evidence for this being provided by the very fast rates at which localized corrosion (such as pitting corrosion or stress-corrosion cracking) can propagate when passivity is disrupted, and regeneration of the passivating oxide film cannot take place. This thin oxide film is the stabilizing barrier which separates the potentially active metal from the environment, for structures smaller than a hypodermic needle and larger than an aircraft. Not surprisingly, passivity and passivation have been examined extensively in the past for many decades using a large range of techniques; despite the great deal of information and knowledge that has thereby been generated, there are many questions that remain to be answered before a complete understanding of the subject is achieved.* For example, since the advent of stainless steel, it has been realized that this important class of material owes its stainlessness to the surface oxide film, and even now, despite the large amount of research, using a wide range of *in situ* and *ex situ* techniques, the exact nature of this oxide film remains to be resolved. Passivity of iron has long been recognized[6–8]; the modern theories relating to the passivating oxide film are primarily due to Evans.[9–12]

Passivity of a metal lies in contrast to its activity, in which the metal corrodes freely under an anodic driving force. The passive state is well illustrated by reference to a classical polarization curve prepared potentiostatically or potentiodynamically (**Figure 1**). As the potential is raised (in the anodic or positive sense) above the equilibrium potential between the metal and its dissolved ions, the driving force toward oxidation and the rate of dissolution increase; in the classical case, it represents an exponential rise of current with potential according to the Tafel equation.[13] When the potential is high enough, a dramatic reduction in the dissolution rate occurs, and the rate of dissolution remains low with further increase in potential. The latter state is the state of passivity and the minimum potential at which the low oxidation

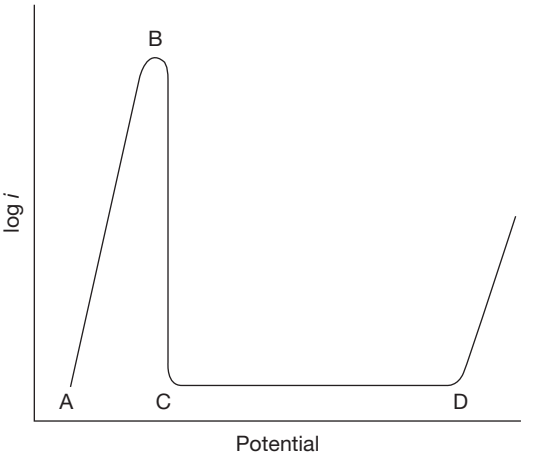


Figure 1 Schematic anodic polarization curve for a metal. Region AB describes active dissolution of the metal. BC is the active–passive transition, with passivation commencing at B. Passivation is complete only at potentials higher than C. The metal is passive over the range CD.

rate exists is the passivation potential. It is a feature of the passive state that the oxidation rate is independent of potential, or nearly so.[14] This feature does not define the state of passivity uniquely, as the diffusion-controlled dissolution reactions are also independent of potential, and passivity is not due to dissolution under diffusion control. Some surface films too, while displaying all the qualitative properties of passivity with respect to the polarization curve, allow a high corrosion current to pass, and cannot be regarded as passivating. An extreme example of this is the phenomenon of electropolishing, in which the rate of dissolution is controlled by the migration of ions through a viscous salt film layer, the rate is independent of potential, and the current density is high.[14–17] It is a primary requirement of passivity that the oxidation rate is small and it depends entirely on the application. The state of passivity is never perfect, and a passive metal always corrodes at a finite rate, albeit that this rate may be very low. The surface film forms a barrier between the metal and its environment, which retards further oxidation. Raising the potential in the passive state, while increasing the driving force toward oxidation, serves to thicken the surface oxide film, thereby increasing the barrier for further oxidation. It is this increase in the barrier film thickness with increase in potential that generates the potential independence of the oxidation rate as long as the metal remains passive. This phenomenon is used in high-voltage anodizing of some metals, in which highly resistive oxide films can be grown to a thickness of several hundred nanometers at high voltages.

*Editor's note: This article embodies the state of the subject in about 1992. The subject has changed relatively little since then and most of the content is consistent with more recent work. For more recent reviews of passivity, the reader may consult 'argentumsolutions'[4] and Staehle *et al.*[5]

The above description of passivity refers to passivity stimulated by an externally applied potential. Passivity of real structures is, however, generally achieved in the absence of externally applied power (although it can be stimulated by application of an anodic voltage, a practice followed for anodic protection).[18] In the absence of an applied potential, the sink for the electrons generated by the anodic oxidation of the metal is a cathodic reaction which occurs on the passive metal surface itself. The cathodic reaction, thereby, provides the driving force. In aqueous systems, this is generally the reduction of dissolved oxygen or the reduction of water to hydrogen (or both simultaneously, depending on the potential involved) occurring on the passive oxide surface, as:

$$O_2 + H_2O + 4e^- \leftrightarrows 4OH^- \quad [1]$$

or

$$2H^+ + 2e^- \leftrightarrows H_2 \quad [2]$$

These same cathodic reactions also fuel the corrosion of metals if corrosion (rather than passivity) is the dominant process. It is to be noted that both cathodic reactions involve an increase in pH, while the anodic site tends to develop a decreased pH. These changes in pH over the anodic and cathodic sites can be important as they can lead to destabilization of the passive state. The state of passivity involving a cathodic reaction on the passive surface is sometimes referred to as 'chemical passivity,' as opposed to 'anodic passivity' or 'electrochemical passivity' (stimulated by an applied potential). The two phenomena are, however, very similar: both are electrochemical in nature. The only mechanistic difference lies in the fact that under chemical passivity, the transferred electrons must pass through the passivating oxide film from the metal atoms to the cathodic reactant, since the anodic reaction (metal oxidation) occurs at the metal–film interface, and the cathodic reaction occurs at the film–electrolyte interface. On the contrary, under externally applied anodic stimulation, the electrons are passed around an external circuit to a separate cathode. Electrons pass rather readily through most passivating oxide films, even those of high resistivity, because they are so thin. Other cathodic reactants may also behave as passivating agents. Their role varies depending on their nature. Thus, for example, the chromates added to the aqueous environment raise the potential of the metal by accelerating the cathodic reaction. If the potential is thereby raised into the passive regime, the metal is passivated essentially by the added chromate, which then functions as a corrosion inhibitor. The reduced chromate also aids passivity by depositing Cr(III) into the passivating film. The earlier theory that only oxygen is required as an adsorbed layer for passivation is, of course, not tenable, since passivity can be achieved readily with alternative oxidizing agents, or with an applied potential in the absence of dissolved oxygen (or any other oxidizing agent): passivity is due to a surface film, usually an oxide, as a distinct solid phase and usually of thickness greater than one monolayer. The oxide ions required to form the film originate from water molecules, and the role of dissolved oxygen is merely that of an oxidizing agent.

Other corrosion inhibitors also enhance passivity without electrochemical reduction, by depositing insoluble oxidation products into the passivating film. For example, benzoate ions cause deposition of ferric benzoate into the oxide, but do not provide any cathodic reaction.

The phenomenon of chemical passivity can be considered by examining the polarization curves for the separate anodic and cathodic reactions, shown in **Figures 2** and **3**. The two forms of passivating agent, one that yields an accelerated cathodic reaction and the other providing insoluble reaction product, are both illustrated. Considering the fact that the anodic and cathodic reactions must occur at an equal rate, the point of intersection of these two kinetic lines in **Figures 2** and **3** describes the condition of the metal.

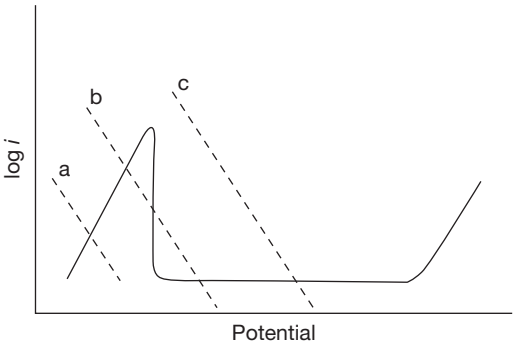


Figure 2 Schematic anodic polarization curve for a passivable metal (solid line), shown together with three alternative cathodic reactions (broken line). Open-circuit corrosion potentials are determined by the intersection between the anodic and cathodic reaction rates. Cathode a intersects the anodic curve in the active region and the metal corrodes. Cathode b intersects at three possible points for which the metal may actively corrode or passivate, but passivity could be unstable. Only cathode c provides stable passivity. The lines a, b and c, respectively, could represent different cathodic reactions of increasing oxidizing power, or they could represent the same oxidizing agent at increasing concentration.

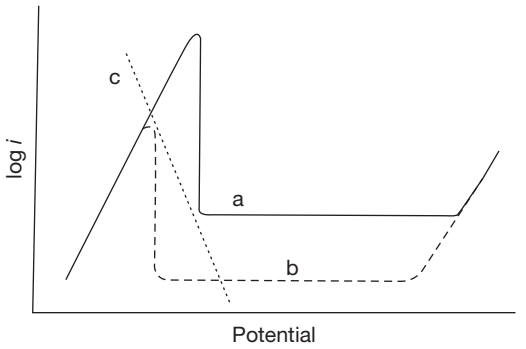


Figure 3 Schematic anodic polarization curves for a passivable metal showing the effect of a passivating agent that has no specific cathodic action, but forms a sparingly soluble salt with the metal cation: (a) without the passivating agent and (b) with the passivating agent. The passive current density, the active–passive transition and the critical current density are all lowered in (b). The effect of the cathodic reaction c, is to render the metal active in case a, and passive in case (b).

To achieve chemical passivity, the potential must be sufficiently high to overcome the active loop (otherwise the oxidizing agent accelerates corrosion) and therefore, there must be sufficient cathodic reactant of sufficiently high driving force.

The process of formation of a passivating oxide film is anodic in nature; the driving force for its formation is raised by increasing the potential anodically across the metal–environment interface. Thus, for example, the driving force for formation of a passivating oxide film on iron by the reaction,

$$2Fe + 3H_2O \leftrightarrows Fe_2O_3 + 6H^+ + 6e^- \qquad [3]$$

is raised as the potential is increased above the equilibrium potential between the metal and its oxide. Raising the potential, however, also raises the driving force toward oxidation of the metal, and thus, increases the propensity toward accelerated corrosion. The two processes of accelerated corrosion and passivation are, therefore, generally in competition, and the balance between the two can be quite delicate. Examples are abound, both in the laboratory and in the field, of passivity failure, leading frequently to rapid and sustained localized corrosion. Mechanisms describing the origins of stress-corrosion cracking and corrosion-fatigue inherently involve rupture of the passivating oxide film, and consequently, the rapid reactions that may ensue on the bare metal surface. The basis on which such forms of localized corrosion occur are well understood. In its passive state, the metal is thermodynamically very unstable. The conditions of the metal, the solution and the potential dictate its state of passivity. If passivity is lost by whatever mechanism (e.g., by oxide film rupture or through reaction of local heterogeneities), the metal reacts anodically at high potential and therefore at a high rate, with the cathodic area occurring mainly on the surrounding undamaged passive surface. One consequence of such anodic reaction is the transport of ions to carry the current, and the generation of acidity through hydrolysis of the reaction products.[19–23] (Note that the reaction above describing ferric oxide formation on iron generates H^+.) If these processes generate a local solution of sufficient aggressiveness, then localized corrosion will ensue: the rate will be fast because of the high potential. If the local solution developed is not sufficiently aggressive or the reaction rate is not high enough (because the potential is too low), the site repassivates (i.e., it regenerates the passivating oxide film). These phenomena are dealt with in more detail below.

Passivity normally exists within a well defined potential range, below which the metal may activate and corrode, and above which it may transpassivate and corrode. The potential range is characteristic of the metal–environment system, and predominantly, of the available cathodic reaction. Because all anodic processes occurring through the passivating film, including the film growth, are controlled by the ionic conductance of the film,[15,24] a low ionic conductivity to metal cations as well as oxide ions is beneficial to passivity. Oxides of high ionic conductivity would grow to be thick, and thick oxide films are more likely to be mechanically unstable, because they are more crystalline, and more defective. The passivating film thickness is also controlled by the rate of dissolution of the oxide into the electrolyte. The electron conductivity of the oxide is of less consequence. Metals such as aluminum and tantalum, whose oxides are of low electronic conductivity, can be as passive as iron and copper, whose oxides are highly electronically conductive. Although passivity in an open circuit requires electron passage through the oxide, this is rarely a limiting step. Consideration of the variables associated with each system is therefore critical in determining the state of passivity.

2.02.2 Determination of the Passive Corrosion Rate

If a metal were perfectly passive and the solubility of its ions were zero, the passive corrosion rate would be

zero but it would take a long time to reach that state. This is never the case, and passivity always implies an ultimate steady state corrosion rate, and it is generally important to know that rate. This is determined as the passive current density, i_{pass}, through polarization experiments. The rate of penetration in the passive state is then $dx/dt = Mi_{pass}/zF\rho$, where M is the atomic weight of the metal, z is the electron number, F the Faraday constant, and ρ the density of the metal. Polarization curves as shown in **Figure 1** are commonly measured by potential sweep measurements (at constant sweep rate, dE/dt); such curves also yield a passive range of potentials. The value achieved, however, depends on the imposed potential sweep rate because at any potential the steady state value of i_{pass} takes some time to be reached. The time taken to reach the steady state, often overlooked or ignored in laboratory studies of passivity, can be measured in minutes, hours, or even days.[†] Fast sweep rates may, therefore, yield an erroneously high value. The passive current density is, thus, more accurately obtained at constant potential. A further feature, which can yield an erroneously low value of i_{pass}, is the presence of a cathodic reactant, such as oxygen. To overcome this, laboratory measurements are made in deaerated solution; full deaeration is, however, not possible. The lower the true passive current density, the greater is the inaccuracy in the measured value.

Measurement of i_{pass} at constant potential does not resolve this inaccuracy. The error can be recognized by measurement using a rotating disc or cylinder electrode. Variation in the rotation rate then alters the rate of oxygen transport, and an apparent reduction in i_{pass} is observed as the rotation rate increases. The constituent of passive corrosion rate forming a viably passive structure is to be questioned, and is dependent on the application. One of the most demanding states of passivity is that required by metallic surgical implant materials, where even small passive corrosion rates of metals (undamaging to the metal), which show toxicity in ionic form and whose ions are not excreted at the appropriate rate, may accumulate to significantly toxic levels. It is, nevertheless, quite common for the passive corrosion rate to be low, that a well engineered structure corroding in its passive state is visibly unchanged after very many years service.

[†]Editor's note: This is probably an understatement; work on Alloy 22, a highly corrosion resistant passive alloy, suggests that the passive corrosion rate falls over months or years.

2.02.3 The Polarization Curve

It is worthwhile to consider the general form of the anodic polarization curve, obtained when the potential of a metal is scanned in deaerated solution. **Figure 4** shows a series of schematic curves which describe the onset of passivity (as shown in **Figure 1**) and the behavior of passive state. Such curves can only be determined potentiostatically or potentiodynamically; galvanostatic measurement of potential cannot provide the polarization curve of a passivating metal because the potential is a three-valued function of the current density. At low anodic potential (region AB) in solutions where the metal ion has appreciable solubility, the metal dissolves actively with the current rising exponentially with increase in potential (see **Figure 1**). Passivation commences at B and is completed at C; region BC comprises the active–passive transition. The passive current then continues to flow with further increase in potential, approximately independent of potential, until point D, F, H, or J depending upon the

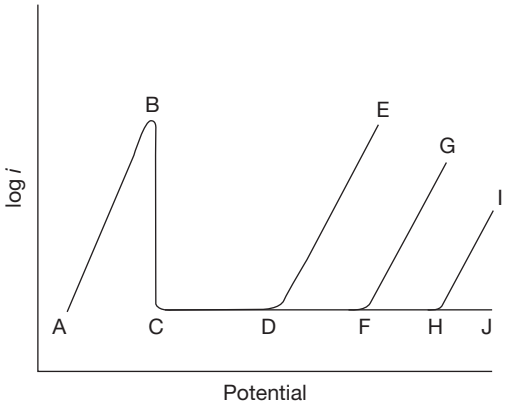


Figure 4 Schematic anodic polarization curve of a metal showing possible behavior of the passive region. AB is the active region and BC the active–passive transition. The metal i is passive at potentials more positive than C. If the solution contains aggressive anions such as chloride, passivity may breakdown at D (the pitting potential), and the current rises with further increase in potential (DE) as corrosion pits propagate and nucleate. The region DE is not necessarily linear. If pitting agents are absent, but the passivating film be transpassively dissolved, then transpassive corrosion commences at F, increasing in rate with increasing potential (FG). If the passivating film is chemically and electrochemically stable, and is conductive to electrons, oxygen evolution commences at H, and increases in rate to I. The metal remains passive. If the film is stable and insulating to electrons, oxide film growth continues with further increase in potential (HJ) and the metal remains passive.

properties of the passivating film with respect to the electrolyte solution.

In solutions with a sufficient concentration of aggressive ions, such as chloride, premature failure of passivity occurs with many metals (such as aluminum, iron, nickel, and their alloys) at point D, by the nucleation of pitting corrosion. Failure of passivity is localized, and the current rises with further increase in potential (region DE), as the pits grow and more pits form. Note that the current density is usually obtained by dividing the measured current by the surface area of the specimen being tested. When localized corrosion, such as pitting occurs, the actively dissolving area is found to be very small (else, corrosion cannot be termed localized). The current density measured as such does not, thus, represent the local corrosion rate, or the penetration rate, which is in fact much faster. The rise in mean current density in region DE of **Figure 4** is not necessarily an exponential function of the potential. The region accommodates an increasing number of pits with increase in potential, as well as their growing surface area, with time. This region also shows a significant noise in the measured current.

In the absence of aggressive anions, the passivating oxide film may begin to be oxidized at point F to produce a soluble anion, and passivity again fails, this time by oxidative dissolution. The current rises with further increase in potential (region FG) as the dissolution rate rises. This occurs, for example, with chromium and manganese by oxidation to CrO_4^{2-} and MnO_4^{2-} or MnO_4^{-}, respectively. If the oxide film is stable, and is also a good electron conductor, such as with iron and nickel in neutral or moderately alkaline solutions, then point H marks the potential at which the solution begins to be oxidized. In an aqueous solution, this is the oxidation of water to oxygen, for which the standard electrode potential is 1.228 V (SHE) at 25 °C. Subsequently, the rise of current with increase in potential in region HI describes the increasing rate of oxygen evolution. Passivity here is not disturbed (unless the oxide film is damaged mechanically by the oxygen bubbles). This process can only occur if the oxide film is an electron conductor, since the oxidation of water requires the withdrawal of electrons from the water molecules through the oxide to the metal: oxidation of water occurs at the film–electrolyte interface.

Finally, if the film is stable in solution, and is a poor electron conductor, as those formed on aluminum and tantalum in halide-free environments, oxygen evolution is not possible. Raising the potential beyond point H in **Figure 4** causes the film to grow further, with the current continuing to be approximately independent of potential (region HJ). It is under these conditions that high-voltage anodizing can be adopted to produce relatively thick oxide films, as described below.

For metal–environment couples, where the passivating oxide film does not break down, does not oxidatively dissolve, and does not conduct electrons readily, the only mechanism for relaxing the increasing electric field as the potential is increased is the thickening of the oxide film. The process occurs by ion conduction through the film. In fact, if the potential of aluminum immersed in a suitable anodizing electrolyte is raised suddenly to a high value, transient evolution of oxygen bubbles is observed. The oxide is, of course, a dielectric, and the sudden potential jump imposes dielectric breakdown with conduction of electrons because of the increased electric field. The effect is rapidly quenched as the barrier oxide film grows simultaneously, relaxing the electric field.

In many cases, passivating and highly electrically insulating oxide films can be grown to thicknesses of several hundred nanometers, and even micrometers; such films can be observed readily in the scanning electron microscope.[24–29] The applied anodic voltages can be in the region of several tens, and even hundreds of volts, and can be applied as a DC current or voltage, or as an AC voltage, depending on the required properties of the oxide. The most widely anodized metal industrially is aluminum. Although AC anodizing of aluminum imposes a cathodic half-cycle for each anodic half-cycle, oxide film reduction does not occur, contrary to what might be expected. This is because the oxide film on aluminum is irreducible. In an appropriate electrolyte, the passivating oxide film can be induced to be overlaid with a much thicker microporous oxide layer. Electrolytes that induce formation of a porous overlayer tend to be more aggressive, for example, aluminum in sulfuric or phosphoric acid. Inevitably, the porous layer is underlaid with a nonporous barrier layer (otherwise the high applied voltage would induce vigorous oxygen evolution in those areas lacking the electron-insulating oxide). The barrier layer, which is the passivating component, is thus named because it carries most of the electric field and controls the film growth current. Its thickness increases linearly with the applied voltage, consistent with the behavior of other passivating oxide films. This duplex oxide film formed on aluminum finds wide technological

applications. Anodizing of titanium to appropriate oxide film thicknesses produces vigorous interference colors, a phenomenon exploited in the ingenious art of Pedeferri.[30]

2.02.4 Mechanism and Kinetics of Passivation

The basic mechanism of passivation is easy to understand. When the metal atoms of a fresh metal surface are oxidized (under a suitable driving force), two alternative processes occur. They may either enter the solution phase as solvated metal ions, passing across the electrical double layer, or may remain on the surface to form a new solid phase, the passivating film. The former case is active corrosion, with metal ions passing freely into solution via adsorbed intermediates. In many real corrosion cases, the metal ions, despite dissolving, are in fact not very soluble, or are not transported away from the vicinity of the surface very quickly, and may consequently deposit on the surface as an oxide or insoluble salt. Often these cases are not passivating because the insoluble corrosion product is deposited from the solution phase as a poorly adherent, loose, crystalline substance. Formation of rust on steel is one such example; blue and green deposits on corroded copper plumbing are another. Oxide films which are passivating are most commonly formed when they are generated directly on the metal surface without the constituent metal ions passing first into solution.[31] While the precise structures of passivating films remain contentious,[32,33] they are known to be coherent and continuous, necessarily so since any lack of continuity would provide paths for active dissolution leading to localized corrosion.

Formation of the first layer (a monolayer) of passivating oxide film on a denuded metal surface occurs easily by the loss of protons from the adsorbed intermediate oxidation products, such intermediates being common to both dissolution and passivation processes.[34–42] Thus, for example, the first oxidative step in the anodic oxidation of nickel is the formation of the unstable adsorbed intermediate NiOH by

$$\mathrm{Ni} + \mathrm{H_2O} = \mathrm{NiOH_{ads}} + \mathrm{H^+} + \mathrm{e^-} \qquad [4]$$

Similar initial reactions occur on many metals such as iron and cobalt. This intermediate can now react further in one of the two ways. Oxidation and protonation of the intermediate to Ni(II) leads to dissolved nickel ions (active corrosion) which are unable to passivate the metal:

$$\mathrm{NiOH_{ads}} + \mathrm{H^+} = \mathrm{Ni^{2+}_{aq}} + \mathrm{H_2O} + \mathrm{e^-} \qquad [5]$$

On the other hand, oxidation and deprotonation of the same intermediate leads directly to passivation by:

$$\mathrm{NiOH_{ads}} = \mathrm{NiO_f} + \mathrm{H^+} + \mathrm{e^-} \qquad [6]$$

where the subscript 'ads' represents the adsorbed intermediate, 'aq' represents the aquated metal ion in solution, and 'f' represents the oxide film. These reactions are able to proceed because the solvent, water, is effectively an infinite source, and sink, for H^+. The intermediate $NiOH_{ads}$ is common to both the dissolution and passivation reactions. It can be regarded as an hydroxide of Ni(I) or an adsorbed OH radical; the distinction probably is not important to the case of passivation. Because of subsequent reactions, it is better considered as the compound of univalent nickel. This intermediate does not exist as a bulk species. Its existence as an intermediate is undoubtedly due to the fact that it is adsorbed to the metal surface on one side, and carries the aqueous double layer on the other. The preferential occurrence of reactions [5] and [6] depends on the solution composition, its pH, the potential, and the temperature. It is apparent that the nickel atoms involved in going from Ni to NiO via steps [4] and [6] never leave the metal surface. Formation of NiO (or the hydrated form, $Ni(OH)_2$) by precipitation from the dissolved nickel ions formed in reaction [5] does not lead directly to passivity. For iron, formation of oxides from dissolved Fe^{2+} via precipitation from solution leads to nonprotective rust. (Brown rusts involve further oxidation of Fe^{2+} by dissolved oxygen). Such precipitated oxides are not usually passivating. The passivating oxide formed by reaction [6] can only occur when the equilibrium potential between the Ni and NiO is reached or exceeded. Higher oxides of nickel may also impart passivity once the appropriately higher equilibrium potential is exceeded.[31] Clearly, such oxides form preferentially in solutions where the oxide is effectively insoluble: for many metals these are solutions of neutral and mildly alkaline pH. Nevertheless, in acidic solutions where the oxide is thermodynamically unstable relative to the dissolved metal cation, passivation can still occur (and does so on many metals, including nickel), as exhibited by chromium. Cr(III) oxide is thermodynamically unstable in acidic systems,[3] but the metal is still very passive in sulfuric acid (at the appropriate potentials); the implication is that the dissolution of

the passivating oxide into acidic solutions, which must inevitably occur, is very slow, and the rate of corrosion of chromium in its passive state is undoubtedly controlled by this dissolution rate. (It is chiefly chromium (III) which gives rise to passivation of alloys containing sufficient quantities of Cr, such as stainless steels and nickel–chromium alloys.)

There are, in fact, many possible steps that could be included in the basic mechanisms described above, for example, involving adsorption processes for H_2O and OH^-, and interactions among the adsorbed species. The most widely studied is the reaction of iron, but similar steps are also encountered in the anodic oxidation of many metals. These mechanisms include the same basic steps as described above.[34–42]

The first monolayer is generally formed before any significant film thickening occurs, because of the very strong dependence of the film growth rate on the electric field across the reacting interface. A mechanism of the type shown in eqns [4] and [6] does not necessarily involve an oxide nucleation step. Rather, it is a random oxidation of the exposed surface metal atoms until the first monolayer of oxide is complete.[41–44] Such a model shows that passive films may, in fact, be amorphous, with no defined crystal structure. Alternatively for some metals, an initial nucleation step may occur with two-dimensional spreading of the oxide over the surface to produce expanding circular 'pancakes' of oxide which link when they impinge on one another. This mechanism, which involves different kinetics, has not been found in repassivation experiments involving mechanically stripped transition metal surfaces.[41–44] Growth of the first monolayer of the passivating oxide causes significant reduction in the rate of oxidation of the metal, but in general, the process does not stop there.

Once the metal surface has become covered with its first monolayer of the oxide film, the metal is then separated from its environment by that oxide, and film thickening ensues: metal cations or oxide anions (or both) are transported through the existing oxide film, thereby relaxing the electric field which drives the mobile ions, until a steady state is reached. Ion migration (or vacancy migration) under the electric field controls the film growth rate.[14,24,45–48] As the film thickens at constant potential, the relaxing electric field across it reduces the rate of field-assisted ion migration, and the rate of film growth declines. The steady state is reached when the rate of film thickening equals the rate of film dissolution into the environment, resulting in a net constant rate dissolution of metal. Different metals display different passive corrosion rates, and it is the numerical value of this rate which dictates whether the metal can be regarded as sufficiently passive for specific engineering purposes. Highly passive metals usually carry very thin oxides, and the electric field sustained by the overall metal–oxide-film–electrolyte interphase region is very high. For example, at a potential of 0.2 V and an oxide film thickness of 2 nm the mean electric field is ca. 108 V m^{-1}. The precise distribution of this electric field, namely the fraction lying across the film and that lying across the two phase interfaces is not known, although several models exist.[47–51] Similarly, the identity of the transporting ions is often not known. Although bulk oxides can often be classified as cation-vacant, anion-vacant or containing excess ions interstitially, thus dictating the mode of ionic charge transport, passive films are so thin that properties of bulk oxides may not necessarily be applicable. The effect is strikingly realized when it is considered that many passivating oxide films are often no thicker than one or two unit cells of the equivalent bulk oxide, and their transport properties may consequently differ from those of the bulk material. Only in the model of Macdonald *et al.*[50,51] is such a distinction drawn for the corrosion and passivation characteristics, where the model predicts that the transport of metal cations across the oxide from the metal interface to the electrolyte interface causes dissolution, and the transport of oxide ions in the reverse direction causes film growth.

Derivation of the rate of oxide film growth proceeds as follows. When the metal carrying its oxide film is at its equilibrium potential, no overpotential exists across the oxide, and the rate at which ions move through the oxide from the metal to the electrolyte equals the rate at which they move back again. Expressed as current densities, the forward rate is i_f and the backward rate i_b; these rates are equal at equilibrium. Ions moving through the film do so via vacancies (anionic or cationic, leading to anionic or cationic migration), or interstitially. They, thereby, encounter an energy barrier which they must surmount in order to move: the width of this energy barrier is the distance between anionic or cationic sites in the oxide lattice if the mechanism is by vacancy migration, or the distance between interstitial sites if the mobile ions move interstitially. This distance is $2a$. The height of the energy barrier is the activation energy, ΔG^*. Thus, at equilibrium

$$i_f = i_b = zFk \exp\left[-\frac{\Delta G^*}{RT}\right] \qquad [7]$$

where k is the rate constant, z is the charge number, F is the Faraday constant, R is the gas constant, and T the temperature. Equation [7] is an expression of the Arrhenius equation.

We now apply an anodic overpotential η, for film growth, and for simplicity, we regard this overpotential as being distributed entirely across the oxide film. If the oxide film has a thickness x, then the electric field across the film is $d\eta/dx$, which equals η/x if the overpotential lies linearly across the oxide film. The potential drop across the distance $2a$ (a single energy barrier) is then $2a\eta/x$. The activation energy in the forward direction (i.e., with the electric field) is thereby reduced by an amount $a\eta zF/x$, and that in the backward direction (i.e., against the electric field) increased by an amount $a\eta zF/x$. Thus, under the applied overpotential η

$$\begin{aligned} i_f &= zFk \exp\left[-\frac{(\Delta G^* - a\eta zF/x)}{RT}\right] \\ &= zFk \exp\left[-\frac{\Delta G^*}{RT}\right]\exp\left[\frac{a\eta zF}{xRT}\right] \end{aligned} \quad [8]$$

Similarly, in the backward direction

$$\begin{aligned} i_b &= zFk \exp\left[-\frac{(\Delta G^* + a\eta zF/x)}{RT}\right] \\ &= zFk \exp\left[-\frac{\Delta G^*}{RT}\right]\exp\left[-\frac{a\eta zF}{xRT}\right] \end{aligned} \quad [9]$$

The following parameters are constants (at constant temperature) and are simplified to

$$A = zFk \exp\left[-\frac{\Delta G^*}{RT}\right] \quad [10]$$

and

$$B = \frac{azF}{RT} \quad [11]$$

The net current density, i, in the forward direction (i.e., producing film growth) is

$$i = i_f - i_b \quad [12]$$

and

$$i = A \exp\left[\frac{B\eta}{x}\right] - A \exp\left[-\frac{B\eta}{x}\right] = 2A \sinh\left[\frac{B\eta}{x}\right] \quad [13]$$

This is the general expression for film growth under an electric field. The same basic relationship can be derived if the forward and reverse rate constants, k, are regarded as different, and the forward and reverse activation energies, ΔG^* are correspondingly different: these parameters are equilibrium parameters, and are both incorporated into the constant A. The parameters A and B are constants for a particular oxide: A has units of current density ($A\,m^{-2}$) and B has units of reciprocal electric field ($m\,V^{-1}$). Equation [14] has two limiting approximations.

(a) When the electric field, η/x, is large (high η and/or small film thickness, x), $2 \sinh y = \exp y$, and

$$i = A \exp\left[\frac{B\eta}{x}\right] \quad [14]$$

This is commonly known as the high field equation. It similar to the Tafel equation for activation controlled electrochemical reactions with the additional feature that the potential driving the reaction lies across an oxide film of variable thickness, x, rather than across the double layer (of constant thickness). The kinetics of oxide film growth according to this equation approximates to (but are not identical to) inverse logarithmic film growth.[46,52,53] The passivity of most passive metals is associated with a high electric field since most passivating films are very thin. A high electric field implies either a thin passivating film (x is small) or a high overpotential (η is large), or both.

(b) When the electric field is small (i.e., η is small and/or x is large), $\sinh y = y$, and

$$i = \frac{2AB\eta}{x} \quad [15]$$

Equation [15] is ohmic ($i \propto \eta$ for constant film thickness): the term $x/2AB$ can be regarded as the film resistance. The equation is identical to parabolic film growth, for which the film thickens with the square root of time at constant potential.

Because passivating films are generally very thin, passivation is normally a high field process and eqn [14] can replace eqn [13]. The same relationships can be applied to high voltage anodizing of metals such as aluminum and titanium. However, under these conditions, the oxide sometimes assumes a duplex structure, with an inner barrier layer and an outer porous layer: for these, the film thickness x, relates to that of the barrier film only, since it is largely this part which carries the applied voltage. The relationships make no assumption about the identity of the mobile ion, and are equally applicable whether the anion or the cation is mobile within the film lattice causing film growth. Because the anodic current in the passive state is approximately independent of the potential in the steady state, the above kinetic equations

imply that the thickness of the passivating oxide film increases linearly with increase in the overpotential, and therefore with the applied potential. This has been confirmed by ellipsometric analysis of the oxide film.

As the film growth rate depends strongly on the electric field across it (eqn [14]), separation of the anodic and cathodic sites for metals in open circuit is of little consequence, provided film growth is an exclusive reaction. Thus, if one site is anodic, and the adjacent site cathodic, film thickening on the anodic site causes the two sites to swap roles so that the film on the former cathodic site also thickens correspondingly. Thus, the anodic and cathodic sites of the stable passive metal 'dance' over the surface. However, if a permanent separation of sites occur, where the anodic site has restricted access to the cathodic component in the electrolyte (as in a crevice), then breakdown of passivity and associated corrosion can follow. The kinetics outlined above, first observed empirically by Güntherschulze and Betz,[54] were modeled by Verwey [45] with the rate-controlling energy barrier being that between two adjacent cation sites within the oxide film. The same basic form can be derived if the rate-controlling energy barrier is that between a metal atom on the metal surface and an adjacent cation site in the film.[14,46] The rate is then limited by ion injection into the film rather than by ion migration through the film. It should also be noted that the above kinetic relationships can be modified to include the other processes occurring during anodic oxide film growth, such as space charges within the film and charging of the phase interfaces.[47,48,55]

The point defect model of Macdonald *et al.*[51] draws a distinction between cation and anion mobility, with mobile cations leading to metal dissolution and mobile anions (oxide ions) leading to passivity: this model has an inherent attraction of incorporating a mechanism for metal dissolution. The procedure leads to direct logarithmic film growth (at constant potential), as does the place exchange mechanism of Sato and Cohen.[56] Apart from the work of Macdonald,[51] the passive state dissolution rate has been studied far less than the oxide film growth rate, and it is not really known why the dissolution rate of the passivating oxide should be so slow. The process of oxide dissolution has been reviewed[57] and is complex. The complexity of the dissolution has been exemplified by Nii,[58] who showed that the rate of dissolution of nickel oxide particles (with no metal substrate) in aqueous solution is strongly dependent on trace quantities (for concentrations down to $\sim 10^{-6}$ M) of ionic impurities in the solution. The rate of dissolution of oxides is influenced by the polarization of the oxides, a notion clearly important to the passivity of metals.

The kinetics outlined above refer strictly to single crystals, with no metallurgical defects, such as grain boundaries. Polycrystalline metals and alloys contain many such defects, and their passivation is expected to be different, especially when the defects may contain high concentrations of impurity atoms, such as sulfur segregated in the grain boundaries in steels. The kinetics of passive oxide film growth, as measured electrochemically, may well be dominated by the passivation of the grain faces because of their overwhelmingly large surface area. However, the long-term stability of the passive state may become dominated by the grain boundaries (as in the case where intergranular corrosion occurs) or by inclusions, precipitates or other microscopic defects. This feature is very important in consideration of measured passive corrosion rates.

Determination of the kinetics of passivation can be carried out potentiostatically, provided any previously formed oxide film, such as that formed by prior exposure to air is first removed. Such removal can be achieved by the reduction at low potential for fairly noble metals, but cannot be readily achieved for the relatively base metals such as chromium, aluminum, or titanium. Some metals, such as iron, become charged with hydrogen during cathodic reduction of their oxide films, a process predicted by the thermodynamics of the two processes. Many alloys may undergo changes in surface composition and structure during applied electrochemical reduction, and subsequent reoxidation kinetics may not be representative of the true alloy surface. Under these conditions, the electrochemical kinetics of oxide film growth can only be determined if the preformed oxide film is stripped away mechanically, by cutting with an inert scribe, such as diamond, sapphire, or boron nitride.[59–66] Such techniques then allow measurement of the kinetics of passivation through potentiostatic current transients or through galvanostatic potential transients. Passivation in open circuit is also amenable to such examination.[66]

2.02.5 Thermodynamics of Passivity

The thermodynamics of passivity, primarily due to Pourbaix,[3] are usefully considered by examining the

equilibria between metals and their oxides in contact with water. Before doing so, however, it is worth bearing a number a points in mind. First, the passivating oxide film is in many cases so thin that its thermodynamics may not be identical to that of bulk oxides. This is already implicit in the discussion of the adsorbed NiOH intermediate in nickel oxidation. The matter is of particular importance when it is realized that the two faces of the oxide film, one in contact with the metal and the other in contact with the environment, are so very different. By nature, the passive film must therefore be regarded as being extremely anisotropic in the sense that its dimensions are very much larger in the two directions parallel to the metal surface than in the third dimension, which is perpendicular to the metal surface. Second, the electric field across any existing passivating oxide is very large in the direction perpendicular to the metal surface and insignificant in the parallel direction, adding to its anisotropy. Finally, in the case of engineering alloys such as stainless steels, it is frequently difficult to ascribe the passivating oxide film to a known stoichiometry or structure, and the use of thermodynamic data of known bulk oxides is open to question. This problem is clearly less pronounced for pure metals. Nevertheless, consideration of the thermodynamics of standard bulk metal oxides can still provide surprisingly good agreement with observation, and can often be used effectively and predictively.

As with all determinations of thermodynamic stability, we commence by defining all possible stable phases, and their standard chemical potentials. For most metals there are many such phases, including oxides, hydroxides, and dissolved ions. For brevity, only the minimum number of phases is considered. The simplest system is a metal M, which can oxidize to form a stable dissolved product, M^{2+} (corrosion), or to form a stable oxide MO (passivation). In aqueous environments, three equilibria can thereby be defined:

$$M^{2+} + 2e^- = M \qquad [16]$$

$$MO + 2H^+ + 2e^- = M + H_2O \qquad [17]$$

$$M^{2+} + H_2O = MO + 2H^+ \qquad [18]$$

The equilibrium potentials (E) for reactions [16] and [17] are then given, respectively, by the Nernst equation:

$$E = E^\circ - \frac{RT}{2F}\ln\frac{[M]}{[M^{2+}]} \qquad [19]$$

$$E = E^\circ - \frac{RT}{2F}\ln\frac{[M][H_2O]}{[MO][H^+]^2} \qquad [20]$$

where the brackets represent the activities of the components and E° is the standard equilibrium potential (or standard electrode potential) for the reaction. The value of E° is readily calculated from the standard chemical potentials of the relevant components. By defining the standard states of the metal, the oxide, and the solvent as the pure components (unit mole fraction), the only solution variables are the M^{2+} and H^+ activities, provided the metal, the oxide and the water are pure. Thus, in decadic logs, eqns [19] and [20], respectively, are

$$E = E^\circ + \frac{2.303RT}{2F}\log[M^{2+}] \qquad [21]$$

and

$$E = E^\circ + \frac{2.303RT}{2F}\mathrm{pH} \qquad [22]$$

For the nonelectrochemical reaction [18], the equilibrium constant (K) for the reaction is written as

$$\mathrm{pH} = -\frac{1}{2}\log K - \frac{1}{2}\log[M^{2+}] \qquad [23]$$

The thermodynamic equilibria are then represented for fixed dissolved metal ion activity as a plot of E° against the pH of the solution, as shown in **Figure 5**, noting that eqn [21] is independent of pH and eqn [23] is independent of potential. The Figure shows that oxide stability is achieved at potentials greater than E° for reaction [22] and in solutions of pH greater than that described by eqn [23]. If the potential is greater than E° for reaction [21] and the pH more acidic than that given by eqn [23] the metal corrodes to give M^{2+}. If the potential is low, the metal is immune to oxidation. The diagram thus describes the thermodynamically stable phases of the metal in terms of regions of passivity, corrosion and immunity as a function of E and pH. The position of the lines representing eqns [21] and [23] depends on the value of the selected M^{2+} activity. It is common (but not necessary) to adopt a concentration of 10^{-6} M for this purpose, but the issue is critical in determining the relative thermodynamic stabilities of the dissolved component and the oxide. The fact that only equilibria with water are considered in the example specifies that any dissolved components present in solution have no effect on the reactions of the metal and its components. Reactive ions such as Cl^- dissolved in the solution modify the equilibria by their own free energy considerations. Although more complicated,

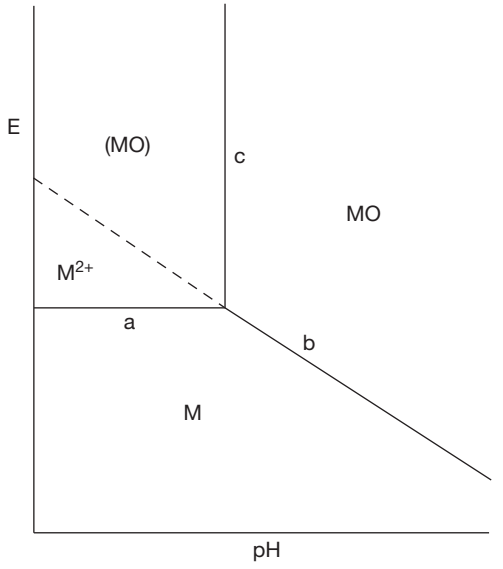


Figure 5 Schematic potential–pH diagram for a metal M in equilibrium with water in the absence of complexing species. Line a represents eqns [16] and [21]. Line b represents eqns [17] and [22]. Line c represents eqns [18] and [23]. The stable phases are marked in bold. The metastable phase is in parentheses. The broken line is an extrapolation of eqn [22] and indicates possible metastable passivity.

such systems can nevertheless be thermodynamically calculated if the formation constants of the associated complexes and compounds can be defined.

The line representing eqn [22] is shown broken for pH more acidic than eqn [23]. The oxide is not thermodynamically stable in this region of pH, and for true thermodynamic equilibrium should not be shown at all. Many metals, however, still show a passive region in this state, provided the potential is higher than that given by eqn [22]. The origin of this is, of course, kinetic. If an oxide film does form, and if it is very slow to dissolve (it must dissolve since the oxide is unstable relative to the dissolved ion) then passivation ensues. The state has sometimes been referred to as 'metastable passivity'.[‡31] While it might reasonably be rationalized that passivity is brought about by a very low solubility of the dissolved metal ion, thereby effectively preventing metal dissolution through formation of a sparingly soluble oxide layer, this is not a necessary condition for passivity. Certainly, a thermodynamically stable oxide layer is more likely to generate passivity. However, the existence of the metastable passive state implies that an oxide film may (and in many cases does) still form in solutions in which the oxides are highly soluble. This occurs for example, on nickel, aluminum and stainless steel, although the passive corrosion rate in some systems can be quite high. What is required for passivity is the rapid formation of the oxide film and its slow dissolution, or at least the slow dissolution of metal ions through the film. The potential must, of course, be high enough for oxide formation to be thermodynamically possible. With these criteria, it is easily understood that a low passive current density requires a low conductivity of ions (but not necessarily of electrons) within the oxide.

It is not possible to predict thermodynamically whether or not a metastable passivating film forms in acidic solutions. Thus nickel, cobalt, and aluminum exhibit metastable passivity in acidic solution (although not necessarily practical); but cadmium does not. Metals which display metastable passivity, particularly in acidic solutions, often show polarization curves whose potential ranges can be correlated with the thermodynamic potentials of the potential–pH (Pourbaix) diagram (**Figure 5**). Thus, for example, nickel in sulfuric acid shows an active loop between the potentials represented by lines a and b in the Pourbaix diagram; as the potential is increased further, passivity commences when the potential exceeds approximately that described by eqn [22], observed as an active–passive transition. Because metastable passivity is a kinetic phenomenon, however, a significant overpotential may be required to produce it. Where more than one oxide may exist, it is not necessarily associated with the lowest metal oxidation state that provides the best state of passivity. On thermodynamic grounds alone, one predicts the classic form of the anodic polarization curve, in which the passive range of potential is preceded (at lower potential) by an active loop as shown in **Figures 1–4**. From the point of view of corrosion engineering for acidic systems, metastable passivity is of the utmost importance. Most structural metals designed for acidic environments are used in this state of metastable passivity. These are often chromium-containing alloys (stainless steels and nickel-based alloys), since chromium shows extremely good metastable passivity in acidic environments.

It is worth emphasizing that the position of lines representing equilibria with the dissolved species, M^{2+}, depend critically on the solubility of the ion, which is a function of pH. For example, iron in moderately alkaline solution is expected to be very passive,[3] and so it is in borate solutions (in the absence of aggressive ions). However, the anodic polarization curve still shows a small active loop at low potential.[68]

[‡]Editor's note: This phenomenon has been analyzed in more detail by Macdonald.[67]

The form of Figure 5 is common among many metals in solutions of acidic to neutral pH of non-complex anions. Some metals such as aluminum and zinc, whose oxides are amphoteric, lose their passivity in alkaline solutions, a feature reflected in the potential–pH diagram. This is likely to arise from the rapid rate at which the oxide is attacked by the solution, rather than from direct attack on the metal, although at low potential, active dissolution is predicted thermodynamically.[3] The reader is referred to the classical work of Pourbaix[3] for a full treatment of potential–pH diagrams of pure metals in equilibrium with water.[§]

2.02.6 Chemical and Electrochemical Passivity

Passivity is provided and maintained by the appropriate cathodic reaction such as reaction [1] or [2]: it is a requirement of this reaction that the potential is held within the passive range. This may be achieved by the cathodic reactant in the solution, such as oxygen, or it may be applied from an external power supply. In the absence of other oxidizing agents, only reaction [1] or [2] can provide and maintain passivity. For more noble metals, reduction of water is not possible, and passivity can be maintained only by reaction [1]. These metals are, of course, immune from corrosion (and passivation) in the absence of dissolved oxygen. Base metals such as aluminum, when bared in solution (these metals inevitably carry a prior air-formed oxide film), are rendered passive by reduction of water because the initial potential is so low.[59,60,66] Passivity can then be maintained by reduction of oxygen because the potential rises significantly during passivation. Passivating agents in solution may be divided broadly into two groups. Those providing the cathodic reactant are oxidizing agents and raise the open-circuit potential of the metal into the passive regime, but only in sufficient concentration. The process is shown schematically in Figure 2. In insufficient concentration, such oxidizing agents raise the potential further into the active region, thereby accelerating corrosion. Such behavior is well typified by the long-known reaction between iron and nitric acid. In concentrated nitric acid, iron is passive, but if the acid is diluted, very powerful corrosion occurs with vigorous evolution of nitrogen oxides. This process was reported earlier by Keir[7] and Faraday.[8] Many oxidizing agents fall into this category, such as chromates, molybdates, oxygen, and hydrogen peroxide acting on steels. Some oxidizing agents become aggressive in reduced form, and cannot be regarded as passivating agents; these include for example, the chlorate anion which is cathodically reduced to chloride on iron. The chloride anion can initiate pitting corrosion, particularly at high potential, and the high potential is itself brought about by the same cathodic reaction.

Other passivating agents, such as phosphates[70–73] and benzoates[74] (for steels) do not fall into this category and they have no oxidizing action. Their role in enhancing passivity is to aid formation of the passive film under the cathode present in solution. In a polarization curve, this would be observed as a decrease in the passive current density, often coupled with a decrease in the potential of the active–passive transition (see Figure 3). These chemical passivating agents act by forming highly insoluble compounds with the dissolving metal (such as ferric phosphates and benzoates). The salts are precipitated over areas of the metal surface which may become active. Presumably, a passivating oxide is then formed beneath the deposited metal salt, since the metal is no longer capable of rapid dissolution, although this has not been firmly established. The formation of azelaic acid from linseed oil[75,76] allows precipitation of sparingly soluble ferric azelate as a passivating agent in linseed oil-based paints. Again, the metal presumably passivates by oxide film formation beneath the deposited azelate. For these passivating agents to be effective, an oxidizing agent, such as oxygen must be present.

Chromate ions fall into both the categories, namely as both an oxidizing agent and a source of insoluble film.[77,78] Their oxidizing power raises the potential for a number of metal systems into the passive region. The reduced form is Cr(III), either as Cr_2O_3 or as a spinel with the metal oxide of the substrate, which is deposited onto the metal surface; this is itself a passivating agent and acts as such for many metals. These prime processes render chromates as effective passivating agents, and are responsible for the use of chromates as inhibitive pigments in metal priming paints (often as zinc chromate).[‖]

[§]Editor's note: Resources are now available, both as commercial programs and on the world-wide web[69] that allow calculation of *E*–pH diagrams for many metals, taking account of a wide range of dissolved and solid species.

[‖]Editor's note: Unfortunately chromate's power as a corrosion inhibitor is matched by its toxicity, and it is now banned for most applications.

Other forms of similar chemical passivating agent act where no oxide is formed. These must be mentioned here, although their action is not truly passivating. Although the dissolution–precipitation mechanism for oxide formation on metals like iron does not generally lead to passivation, some salt films formed by precipitation of dissolved ions from solution may give rise to significant passivity. Thus, silver in acidic chloride solutions forms a very stable insoluble layer of AgCl, which reduces the corrosion rate of the metal substantially. The chloride is formed, only partially, by precipitation from solution. As such the AgCl layer is thick relative to the normally formed oxide film on metals such as aluminum or stainless steel, and results in visible tarnishing of the metal. Although the corrosion rate is thereby rendered low, the metal is not strictly passive, and tarnishing of silver by silver chloride is not only unsightly – it is slowly damaging. The behavior of lead in sulfuric acid is similar because of the low solubility of lead sulfate.

Passivity can also be enhanced by incorporating components that catalyze the cathodic reaction, thereby raising the potential. By alloying titanium with small quantities of platinum or palladium, or by implantation of these elements into the surface, the cathodic hydrogen evolution reaction is thus accelerated, and titanium becomes anodically protected in appropriate environments.[79] Interestingly, this simple form of anodic protection does not involve an external power supply. Platinum and palladium function as good hydrogen catalysts – they raise the exchange current density for hydrogen evolution. A similar mechanism has been proposed for the action of lead-based pigments in priming paints.[75]

2.02.7 Passivation in Nonaqueous Liquids

Far less is known of the corrosion and passivation of metals in nonaqueous solvents and solutions,[80] and the topic can be mentioned only briefly here. The prime requirement for passivity of metals in water is the formation of an oxide film by anodic reaction with oxygen obtained from water. This is not always available in organic solvents. Thus, polarization of nickel in methanol containing sulfuric acid[81] shows no passive region in solutions containing less than 2 wt% water, but shows an increasingly large passive potential range, and lower passive current density, as the water content is increased above 2 wt%. Passive oxide formation arises from reaction with the water component. Reaction of the metal with the methanol component to produce the oxide is clearly not possible. Some metals can react with the methanol component of methanolic solutions, but such reaction produces the nonpassivating methoxide rather than the oxide, and results in corrosion. The effect is seen clearly by the reaction of titanium with alkaline methanol solutions.[65] Generally, the effect of chloride in inducing pitting corrosion and stress corrosion cracking in metals such as titanium is increased substantially as the water content of the methanol solution is reduced, demonstrating the destabilizing effect of the reduced water activity on the stability of the passive state. Both passivation and corrosion can be difficult to predict in nonaqueous liquids in the absence of laboratory tests.

2.02.8 Passivity of Alloys

While the same basic mechanism for passivity of pure metals also applies to alloys, the processes involved in the passivation of alloys have an added complexity. In many cases, only one component of the alloy has the property of being passive in a particular environment. Alloys such as stainless steels, which contain highly passive components (chromium in this case), owe their corrosion resistance to the surface enrichment of the passivating component.[82] Thus, stainless steels resist corrosion in many acidic systems (where iron or carbon steel would be poorly passive or not passive at all) by a passivating oxide film containing Cr predominantly as Cr(III). Surface analytical techniques such as Auger electron and X-ray photoelectron spectroscopies reveal substantial enrichment of chromium in the passivating oxide film on these alloys.[32,33,83–87] There are only two ways by which this enrichment can arise. Either the other component (iron in the case of binary iron/chromium stainless steels) is left behind at the metal surface when the chromium oxidizes, or it is dissolved during passivation. In aqueous corrosion, particularly at ambient temperatures, it is undoubtedly dissolution of the iron component that causes chromium enrichment of passive stainless steel surfaces.[64] This contrasts gaseous oxidation of the same class of metals at high temperature, where chromium-rich oxide films are also formed, but iron cannot dissolve in the environment, and diffusion in the metal phase is much more rapid. Classically, a minimum amount of chromium is required in the alloy before chromium enrichment in the oxide film can be achieved: for stainless steels,

this is determined to be around 12–13 wt%. Steels containing less than this amount of chromium are not regarded as stainless. The basis of this is as follows. In environments, corrosive towards iron but passivating towards chromium (such as sulfuric acid), iron dissolves selectively from the surface of the alloy, and so the surface concentration of chromium increases to a level where the basic chromium-rich oxide film can cover the surface completely: the metal is then passive. This film contains significant quantities of Cr(III), but iron ions are also found in its structure, generally decreasing towards the film–electrolyte interface. Steels containing greater than about 12–13% chromium show a dramatic improvement in their passivating properties, in comparison to those of lower chromium content.[88] Analyses of such oxide films are performed using ex situ surface analytical techniques, such as Auger electron spectroscopy, X-ray photoelectron spectroscopy, and secondary ion mass spectrometry coupled with some form of depth profiling. Most such analyses performed to date have been carried out on surfaces exposed to the solution for relatively less time (seconds or minutes), despite the fact that steady-state passivity of stainless steels takes much longer to achieve. It is not known whether, after a sufficiently long period of exposure, the oxide film on stainless steel is composed entirely of chromium oxides.

The ease with which stainless steels can passivate increases with the level of chromium within the alloy, and hence, materials with higher chromium content are more passive (i.e., conduct a lower passive current density) and passivate more readily (i.e., the critical current density is lower and the active–passive transition is lower in potential). They are also passive in more aggressive solutions: the pitting potential is higher.

The precise mechanism by which surface enrichment of chromium through dissolution of iron in stainless steels occurs is still under some debate.[64,89–91] For alloys which are single-phase solid solutions, the phenomenon is readily understood if the passivating element is the major component of the alloy. The corrodible element then simply dissolves from the metal surface enriching the surface with the passivating element which then completes its passivating oxide. However, far less than 50% chromium is required to achieve passivity. Examination of the repassivation of stainless steels shows evidence of dissolution of the corrodible components from the film as it grows.[64] Excess iron ions in the film are dissolved and replaced by chromium ions from the metal, progressively enriching the film with chromium until passivity is reached. However, the process may not be so simple. Recent ideas of percolation through binary alloys involve dissolution of individual and clusters of iron atoms from the alloy surface.[89–91] The residual chromium atoms are able to diffuse over the surface, progressively uncovering iron atoms and iron atom clusters. These are dissolved until there is sufficient chromium on the surface to link up into a continuous passivating network. The threshold concentration of Cr in the alloy (~12%) is then determined as the percolation threshold below which interlinked networks of iron atom clusters exist through the alloy, and it is no longer stainless. Above this threshold composition, continuous percolation clusters do not exist and the alloy is passive and stainless. The attraction of this model is that its concepts can be extended to the nucleation of pitting corrosion of stainless steels[92] (see below). The distinction between the two modes of iron dissolution and associated chromium enrichment is, however, a fine one.

Stainless steels show another interesting feature. At high potentials, passivation by an oxide rich in Cr(III) is lost since the Cr(III) component of the oxide is transpassively dissolved in chromate ions (line FG in **Figure 4**). In solutions in which iron oxides are stable at high potential (mildly alkaline) and the metal does not undergo pitting corrosion, the resulting iron oxide can then provide some passivity: the roles of the two components of the alloy are thus reversed. If the nonchromium components of the steel are also not passive at high potentials, the metal corrodes transpassively. Very powerful oxidizing agents can bring about transpassive corrosion of stainless steel. Under normal conditions of use in dilute aqueous solutions, however, these high potentials would not be achieved, and chromium provides the prime passivity of these materials.

For stainless steels to remain passive, the level of chromium defined by the bulk composition must be maintained throughout the entire metal matrix. If local regions exist within the alloy microstructure where the chromium level is depleted, these become susceptible to localized corrosion, and increasingly so as the chromium content is reduced. Ultimately, if the chromium content is reduced locally below ~12%, this region no longer carries the composition of a stainless steel, and corrodes as a nonstainless steel alloy. However, any reduction in the local chromium concentration of the alloy matrix is deleterious to passivity, since these alloys become progressively less passive as the chromium concentration is decreased. Such conditions are found under sensitizing heat treatments, commonly in or around welds (the heat-affected zone).

In these regions, at the appropriate temperature, chromium carbides precipitate in the grain boundaries depleting the matrix around the carbides of chromium.[88] The grain boundaries may then activate in service under conditions where the grain matrix is passive.

In multiple-phase alloys, the situation is different. The passivity of each phase is of importance, but there can also be a significant synergy between the behavior of adjacent phases. The presence of copper-rich precipitates in aluminum alloys containing the element accelerates the cathodic reaction locally because the exchange current density for the cathodic reaction on the relatively noble precipitate is higher than it is on the surrounding aluminum matrix. In passivating electrolytes and in the absence of pitting agents such as chloride, passivity is thus enhanced because of the increased potential. However, in the presence of chloride, the potential may be raised into the pitting regime, and passivity then becomes degraded. Aluminum alloys containing more noble alloying elements such as copper or iron are generally less resistant to pitting corrosion than pure aluminum.

2.02.9 Compositions of Passivating Oxide Films

The composition, structure and even thickness of passivating oxide films on metals are in fact extremely difficult to determine, mainly because these oxide films are so thin, and are 'held together' by the compositions of the interfaces (with the metal and with the electrolyte) as well as by the high electric field they carry. Many methods of determination are *ex situ* techniques, requiring removal of the specimen from the electric field as well as from the electrolyte. These methods include Auger electron spectroscopy (AES), X-ray photoelectron spectroscopy (XPS) and secondary ion mass spectrometry (SIMS), as well as X-ray diffraction (XRD).[32,33] A range of in situ methods, such as ellipsometry, frequency response analysis[#] and photocurrent spectroscopy are also used to provide information on the passive film.[32,33] These techniques, however, give more indirect information; they do not present compositional or structural information directly, but probe other properties of the passive interface, from which more detailed understanding must be inferred. The new techniques of scanning tunneling and atomic force microscopies, which can be used in situ, may yield some information in the future.

The thickness of the oxide film is most readily determined by cathodic reduction, provided the film is cathodically reducible and the reduction is 100% current efficient. The cathodic charge involved then determines the film thickness for films of known composition. The method is suitable only for more noble metals, and for most structural metals it is not relevant. Film thicknesses determined by ellipsometry have been used successfully, but assumptions of the refractive index must be made. The passivating film on iron is determined to be Fe_3O_4, γ-Fe_2O_3 or FeOOH, depending on the conditions under which it is formed.[3,32,33] The thickness of the film formed in borate buffer solutions is up to ~4–5 nm, the value depending on the potential at which it is formed. The thickness is a linear function of potential. There is now significant evidence that the Fe(III) film which passivates iron contains some Fe(IV),[93] a phenomenon which may account for the sometimes necessarily high overpotential required to passivate the metal. Surface analyses of passive films on iron, as well as other metals such as aluminum and stainless steel, show that the outer surfaces are usually hydrated, equivalent to hydroxide ions forming part of the outer structure in place of oxide ions.[94–96] This is, of course, only to be expected from film growth in aqueous solution. The inner regions of the films tend to be less hydrated or not hydrated at all. The passivating oxide on aluminum is reckoned to be amorphous, but crystallizes with increasing thickness during anodizing.

The incorporation of anions from the electrolyte, such as borate and carbonate, into the oxide has also been shown to occur on iron and cobalt,[97] such anions being restricted to the outer layers of the film. Attempts to determine chloride incorporation into passive iron surfaces from chloride-containing electrolytes have met with contradictory results, and it is still not known whether or not chloride incorporation is a precursor step to pitting corrosion.[32,33,98–101] Neither is this mechanistically important, but vexed question have been answered unambiguously for other metals.

The oxide film composition on alloys is inevitably more complex; some discussion of chromium-containing alloys is given above. Generally, and not surprisingly, the components of the alloy are to be found in the passivating oxide film,[101] but the role that these play in determining passivity and its breakdown is far

[#]Editor's note: now generally known as electrochemical impedance spectroscopy (EIS).

more open to question. Stainless steels show the presence of iron, chromium, and nickel in the oxide (and molybdenum for alloys containing the element), with elemental ratios that depend on pH and potential. By various ex situ techniques, chromium is found in (III), (IV), and (VI) oxidation states in varying proportions.[32,33,84–87,101] The reasons why changes in the oxide film composition should improve passivity and resistance to breakdown remain speculative.

2.02.10 Breakdown of Passivity and Pitting

Probably less surprising than the phenomenon and characteristics of passivity itself, is the breakdown of passivity, and the very high rate of localized corrosion that can ensue. The passivating oxide film is extremely thin, and its existence and stability depend significantly on the composition of the environment and the very high electric field it carries. It is in essence fragile, chemically, electrochemically, and mechanically, but the fragility depends on the metal and the environment as a system. Mechanical breakdown of the oxide film arises when the material is under tensile stress and can give rise to stress-corrosion cracking or corrosion fatigue. Mechanical breakdown can also arise from an erosive or cavitative action from the environment. These processes generate a fresh metal surface, the reaction of which can be twofold. In suitably quiescent environments, the bare metal surface repassivates, and the electrochemical reaction with the environment causes no significant damage, although a continuing erosive action does lead to loss of material through wear. During corrosion fatigue too, the regrowth of the passivating film over the surface can help wedge open an incipient crack during the compressive half-cycle, allowing mechanical crack growth.**

Chemical breakdown of passivity occurs naturally enough if the potential is lowered into the active regime (region AB in **Figure 4**). This can occur if the cathodic reactant supporting passivity, such as oxygen dissolved in the electrolyte, is depleted, as occurs locally in crevices or occluded regimes of the structure. Chemical breakdown also arises at high potential in the presence of aggressive ions, usually anions, which interact with the metal enhancing dissolution (region DE in **Figure 4**). The most common aggressive anion is chloride, aggressive to very many passive metals. In such cases, absence of the aggressive ion would impose passivity at the same potential: the process is, thus, essentially different from crevice corrosion.

Localized corrosion can occur if the passivating element of an alloy is locally depleted from the matrix. This may occur in the sensitized regions of stainless steels for example, as outlined above. Other alloy components too, may be deleterious towards the passivity of alloys. For example, the presence of significant quantities of sulfur as an impurity in ferrous- and nickel-based alloys is generally degrading towards passivity.[23,33,102,103] Sulfur is incorporated into the oxide film or accumulates at the metal–oxide interface; the element retards the formation of the oxide film and catalyses dissolution of iron and nickel, often remaining adsorbed on the dissolving metal surface. In extreme cases, a sulfide film forms on the metal surface, which is nonprotective. Appropriate heat treatment of such alloys allows segregation of the sulfur to the grain boundaries or precipitation of sulfides, both leading to very high local concentrations of sulfur and premature failure of passivity at these sites.

Where the environment is potentially corrosive, however (i.e., contains aggressive ions), even though the overall surface remains passive, corrosion can advance into the bare metal surface. The question is, how can this happen? In stress corrosion cracking attack occurs mainly around the active–passive transition (BC in **Figure 4**), where oxide film growth still occurs, but the state of passivity is potentially unstable, because of the proximity in potential of the active region (AB in **Figure 4**). Dissolution occurs simultaneously with repassivation (regrowth of the oxide film). A small ohmic resistance in the electrolyte between the external cathodic area and the anodic incipient crack tip may be sufficient to lower the potential of the crack tip into the active regime. Alternatively, a small increase in acidity developed at the crack tip as a consequence of anodic reaction may enlarge the active loop a little by raising the potential of the active–passive transition locally. Some anodic reactions at the crack tip may cause local brittleness, such as dealloying or the diffusion of cathodically reduced hydrogen in atomic form into the metal ahead of the crack tip. These processes allow repeated rupture at the crack tip, even if it

**Editor's note: This statement is a little misleading. Wedging open of a crack by corrosion product increases the minimum stress intensity factor without affecting the maximum stress intensity factor, thereby reducing ΔK, and hence fatigue crack growth rate. This process is known as oxide-induced crack closure. However, wedging by corrosion product can lead to tensile stresses at the crack tip that enhance SCC crack growth rate.

has already repassivated, because of the embrittled local metal matrix. Stress corrosion cracking also occurs near the passive–active transition at higher potential when such a transition is available (near point D in **Figure 4**). In this case, chloride undoubtedly enhances the ability of an incipient crack to propagate by enhancing dissolution at the crack tip, even though the pitting potential (the potential of point D) is not necessarily exceeded.

There are many forms of pitting corrosion (the formation of holes by corrosive action). The various forms have been reviewed by Szklarska-Smialowska.[104] Sometimes activation of a steel brought about by lowering the potential from the passive potential regime into the active regime produces etch pits, a form of uneven corrosion associated with different crystal planes within the metal. This process is less sensitive to the type of anion present, and more associated with the pH changes that occur during activation in unbuffered solutions. The pitting process considered here is a process involving the formation of small holes in an otherwise passive surface as a consequence of the aggressive action of a component of the electrolyte. It is a form of localized corrosion which involves loss of passivity locally, but does not require a stress (although it may be enhanced by a stress). It usually occurs at high potential, where one might expect passivity to dominate. In pitting corrosion, passivity is lost or ruptured through the chemical or electrochemical action of a specific component of the environment. The most commonly encountered pitting agent in aqueous electrolytes is chloride, although some other anions may also induce pitting. Notably, these are the other halides. Sulfur oxyanions such as thiosulphate,[23,105] and indeed sulfide ions, are conducive towards pitting corrosion in ferrous materials and nickel-based materials. It is of some consequence that most anions capable of causing pitting in particular metals also induce stress corrosion cracking of the same metals, although the converse is not necessarily true. Many metals undergo pitting corrosion in chloride solutions of sufficient concentration; these include steels and stainless steels, nickel, copper, and aluminum, and their alloys, and many more.[††]

[††]Editor's note: It is a requirement for pitting by the mechanism discussed here that the bulk of the component is passive in order to provide a large cathode area to drive the local anode. Consequently, carbon steels do not suffer from this form of pitting in neutral chloride solutions, since steel is not passive in such solutions. However, steels may suffer from transient localized corrosion in neutral chloride solutions, such as seawater, due to the presence of mill scale and/or the local effects of sulfur derived from sulfide inclusions.

A few metals are resistant to chloride-induced pitting, including chromium, titanium, and tantalum, but they are not necessarily immune to pitting, only that they are resistant. Some other anions may also induce pitting but outside the normally encountered range of corrosion potentials, and are therefore not regarded as specific pitting agents. For example, nitrate ions, which passivates towards aluminum under normal potentials, induce pitting in dilute nitric acid solutions at high potentials. Such potentials are only accessible through application of an external voltage source.

Precisely how these anions nucleate pitting in a passive metal surface is still largely unsolved, although a range of models exists. Pits in chloride solutions are often rare events.[23,106] In stainless steels, pits can be very widely spaced apart, and most of the surface is passive. Pit propagation rates, are however, very fast: a propagation rate equivalent to a current density of 10–100 kA m^{-2} is common.[23,106] In chloride solutions, pitting is characterized by a minimum potential, called the pitting potential[104] (point D in **Figure 4**). Only above the pitting potential is pitting observed to occur. The metal is in essence passive below this potential, and this is a commonly used design criterion. However, detailed examination of the passive region (i.e., above the passivation potential and below the pitting potential, the region CD in **Figure 4**) of steels and stainless steels shows that the passive current is more noisy in chloride solutions than it is in the absence of chloride.[23,106–108] The origin of this noise is the nucleation and propagation of 'metastable' pits (also sometimes called unstable pits). Metastable pits are those which grow for a short period and then die through repassivation.[23,106–111] These are observed both below and above the pitting potential. Although these metastable pits do not generally cause significant damage to the metal, apart from a degree of microscopic surface roughening, they are evidence that passivity is not fully stable in the presence of chloride ions, even below the pitting potential. For these materials at least, the pitting potential (sometimes called the breakdown potential, erroneously, since breakdown of passivity occurs at lower potentials for ferrous materials) is not a pit nucleation potential: it can only be regarded as the minimum potential for stable pitting. Pits which are stable propagate almost indefinitely, either until total penetration of the metal, or until they are large enough to be washed out by the bulk electrolyte. Such noise in the current can be observed throughout the passive range of potential if the experiment is carried out carefully enough.[106]

The pitting potential (point D in Figure 4) is sometimes treated as a type of equilibrium potential. There are empirical reasons for this. The pitting potential is often observed to decrease linearly with $\log(Cl^-)$ (the chloride activity or concentration), giving an apparently Nernstian form.[104] However, observation of metastable pitting at potentials well below the pitting potential probably precludes this notion. In addition, if the potential sweep direction is reversed at point E, the polarization curve shows strong hysteresis. The negatively moving sweep lies at higher currents, and the potential at which the pitting metal repassivates is much lower than point D. (The point where the metal repassivates in the negative sweep is sometimes called the repassivation potential.) It must also be noted that the pitting potential itself depends on the prior surface treatment of the metal, such as the surface roughness.[112] These facts are inconsistent with the pitting potential describing some simple electrochemical equilibrium. The pitting potential is nevertheless an extremely useful engineering parameter.

Pitting corrosion requires an incubation period or induction time to nucleate.[113,114] Even after initial nucleation, further events are sporadic in time. Several possible mechanisms of nucleation have been put forward. Complexing of metal cations on the surface of the film has been proposed to lead to local dissolution of the oxide,[113] the site then receding to the metal surface. Alternatively, chloride migration through the oxide film may lead to its accumulation as the metal chloride at the metal–film interface. This, if it occurs, could well lead to mechanical failure of the film by bursting, since metal chlorides are of greater molar volume than the corresponding oxides. The notion is appealing since it provides for a high local chloride concentration at the metal–film interface once the film is ruptured, providing an immediate pit propagation path. Film rupture could instead be induced by condensation of vacancies at the metal–film interface.[51] The idea of mechanical film rupture carries the added appeal that pitting events on iron microelectrodes observed potentiostatically show an initiating very fast current spike.[107,115] The event is microscopically violent. Recent observations with stainless steel microelectrodes show that pits here are also initiated with a violent current jump.[116,117] The magnitudes of the events are very small, generally up to some hundreds of picoamperes, indicative not of a low reaction rate, but of the microscopic size of the nucleation event. Interestingly, many of these nucleation events show no propagation stage at all, not even in a metastable pit, but die immediately after nucleation. The implication is that while pit growth may be a rare event over the metal surface, the nucleation event may in fact be rather common. The same nucleation mechanism may not be applicable to all metals which pit in chloride solution, although current fluctuations are observed from aluminum in chloride solutions below the pitting potential. It has also been suggested that passive films may always undergo continuous breakdown/repair events on a microscopic scale, even in solutions which are fully passivating. This interesting idea takes the nucleation event away from the properties of the aggressive chloride anion. Chloride would then serve merely to propagate pits from these nuclei, and prevent their repassivation.

Steels and stainless steels show preferential nucleation of pits at inclusions, most notably sulfide inclusions.[23,118] Other sulfur-rich regions in ferrous and nickel-based alloys may also lead to premature failure. It has been shown that accumulation of sulfur on the surface of these materials retards passivity and enhances dissolution of the metal. These effects occur in any solution in which the metal shows an active region, and are also preferential pitting sites in the presence of chloride. A recent notion[92] for pit nucleation in stainless steels suggests that iron-rich clusters found randomly distributed in the metal structure are active when they become exposed to the electrolyte, since they do not carry sufficient chromium to allow their passivation. These are the pit nucleation sites. They become progressively exposed to the solution through the normal passive dissolution rate. This model cannot be generally applicable to pitting corrosion since it cannot account for pit nucleation in single component metals, such as iron and aluminum. In fact, any site on the metal surface capable of depassivation for any reason, is a potential pit nucleation site. For it to develop into a pit simply requires an anodic reaction capable of procuring a more aggressive anolyte, as outlined below.

The fact that amorphous metals show a higher resistance to pitting corrosion than their polycrystalline counterparts can offer more than one explanation. Such amorphous metals made from iron, chromium, and nickel simulate the stainless steel and show significantly high pitting potentials.[118–121] These alloys are usually made with significant quantities of nonmetallic element additions such as boron in order to retard crystallization during quenching from the melt. Such elements, which are found in the oxide film,[122] may aid

passivity and inhibit pit nucleation by aiding amorphization of the passive film, or by their own specific chemical action. More probably, however, it is the absence of metallurgical defects such as grain boundaries and inclusions which inhibit pit nucleation.

Once nucleated, anodic reaction at the incipient pit then requires a corresponding cathodic reaction on the surrounding passive surface if the metal is pitting in an open circuit, or at the counter electrode if pitting is carried out potentiostatically. Anodic reaction generates cations by dissolution, and H^+ by hydrolysis of the dissolved cations. Acidity is also generated if the metal repassivates. Both the metal cations and H^+ require neutralization, accomplished by ingress of anions. If chloride is the only available anion, its local concentration increases; the locally higher chloride concentration enhances local metal dissolution, and prevents repassivation.[123] The enhanced dissolution then draws more chloride into the now propagating pit, enhancing dissolution even further. Growth of the pit, is in essence, fuelled by its own reaction products: it is a feedback mechanism. The process is sometimes termed autocatalytic. Pit growth is further supported by the lower local pH. The pit anolyte becomes a saturated (or near-saturated) solution of the metal chloride, and is highly acidic. Some 60–70% of saturation of the metal chloride is required to prevent repassivation of stainless steel. It follows from this, that the presence of other anions in solution, in addition to chloride, but which are not aggressive (e.g., SO_4^{2-}) would reduce the ability of the metal to undergo pitting.[‡‡] Both ions are then transportable into the pit anolyte and it is more difficult to build up sufficient chloride to establish propagation: this is found in practice.[124,125]

It is apparent that such a mechanism depends strongly on the transport properties of the electrolyte components into and out of the pit. In stainless steels, pit propagation has been shown to be controlled by diffusion of ions between the pit interior and the external electrolyte.[106,125] The behavior of the pit site depends on its degree of occlusion: pitting is easier on rough surfaces because of the greater occlusion of sites. Surfaces of rougher finish usually display a lower pitting potential. More occluded sites can more readily retain the dissolved cations required to draw chloride into the pit anolyte.

In the early stages of pit propagation on stainless steel in chloride solution, when the pit is small, its depth alone is insufficient to act as a diffusion barrier. Diffusion is then restricted by a perforated cover which exists over the pit mouth; the perforation arises from the initiating event, and the cover exists because of undermining of the surrounding passive film as the pit expands. This perforated cover is critical to survival of the pit through the so-called metastable period of growth. Small ruptures in the cover during the metastable growth period enhance the propagation rate without terminating the pit. However, if this cover is totally lost by rupture before a critical stage in the growth of the pit, the anolyte is washed out (the outward diffusion of cations becomes fast) and the pit repassivates (hence, the term metastable). Pit growth does not then achieve stability. The critical parameter here is the product of the pit growth current density and its radius[106,125]: this pit stability product must reach a value of $\sim 3\ \text{mA cm}^{-1}$ for pit growth on stainless steel to stabilize. Once the critical pit stability product is achieved, the pit can propagate without its cover. The pit depth alone is then a sufficient diffusion barrier to ensure the appropriate high chloride concentration within the anolyte, and pit growth achieves stability. Corrosion pits in chloride solution, thus pass through three consecutive steps of growth. The nucleation event is followed by metastable pit growth. Growth is metastable because the pit depth alone is an insufficient barrier to diffusion and may terminate at any point by rupture of the sustaining cover, resulting in repassivation. If the pit grows to a stage where the pit depth alone is sufficient to maintain the diffusion barrier, it enters the third stage and its growth becomes stable. Most pits die at the nucleation stage, and show no propagation. A few survive to become metastable; of those propagating metastably, only a few achieve stability.

It is worth mentioning that although pitting corrosion is generally undesirable, the process does in fact have its uses. Aluminum sheet used for the preparation of printers' lithographic sheet is subject to a preliminary treatment involving a highly controlled electrochemically induced pitting process in electrolytes based on either hydrochloric acid or nitric acid.[126] This procedure, termed electrograining, is designed to roughen the surface on a microscopic scale prior to anodizing. Pitting corrosion may also be induced into the nickel anodes used in nickel electroplating baths containing chloride in order to prevent their passivation, and allow continuous replenishment of the nickel ions in the electrolyte as they are removed.

[‡‡]Editor's note: Arguably, sulfate also acts to limit the drop in pH in the pit, since the second pK_a for sulfate (the pH at which equal concentrations of SO_4^{2-} and HSO_4^- will be present) is ~1.9, above the pH needed to achieve stable active dissolution.

References

1. Wagner, C. *Corros. Sci.* **1965**, *5*, 751.
2. Brusic, V. In *Oxides and Oxide Films*; Diggle, J. W., Ed.; Marcel Dekker: New York, 1972; Vol. 1.
3. Pourbaix, M. *Atlas of Electrochemical Equilibria in Aqueous Solutions*; Pergamon/CEBELCOR: Oxford, 1966.
4. http://www.argentumsolutions.com/viewed on 7 May 2009.
5. Staehle, R. W.; Forty, A. J.; van Rooyen, D. Eds. *Fundamental Aspects of Stress Corrosion Cracking*; NACE: Houston, TX, 1969.
6. Tomashov, N. D.; Chernova, G. P. *Passivity and Protection of Metals* Plenum Press: New York, 1967; Vol. 9.
7. Keir, J. *Philos. Trans. Roy. Soc. Lond.* **1790**, *80*, 359.
8. Faraday, M. *Experimental Researches in Electricity;* University of London: London, 1844; Vol. 2.
9. Evans, U. R. *Ind. Engng. Chem.* **1925**, *17*, 363.
10. Evans, U. R. *The Corrosion and Oxidation of Metals*; Arnold: London, 1960; Chapter 7.
11. Evans, U. R. *The Corrosion and Oxidation of Metals*; Arnold: London, 1968; First Supplementary Volume, Chapter 7.
12. Evans, U. R. *The Corrosion and Oxidation of Metals*; Arnold: London, 1976; Second Supplementary Volume, Chapter 7.
13. Tafel, J. *Zeit. Physik. Chem.* **1905**, *50A*, 641.
14. Hoar, T. P. In *Modern Aspects of Electrochemistry*; Bockris, J. O'M., Ed.; Butterworths: London, 1959; Vol. 2, p 262.
15. Hoar, T. P.; Mowatt, J. A. S. *Nature* **1950**, *165*, 64.
16. Hoar, T. P.; Mears, D. C.; Rothwell, G. P. *Corros. Sci.* **1965**, *5*, 279.
17. Hoar, T. P.; Mears, D. C. *Proc. Roy. Soc. Lond.* **1966**, *A294*, 486.
18. Edeleanu, C.; Gibson, J. G. *Chem. Ind.* **1961**, *301*(10), 301–308.
19. Turnbull, A. Ed. *Corrosion Chemistry in Pits, Crevices and Cracks*; HMSO: London, 1987.
20. Kolotyrkin, J. M. *Corrosion* **1963**, *19*, 261t.
21. Wranglen, G. *Corros. Sci.* **1974**, *14*, 331.
22. Staehle, R. W.; Brown, B. F.; Kruger, J.; Agrawal, A. Eds. *Localized Corrosion*; NACE: Houston, TX, 1974.
23. Isaacs, H. S.; Bertocci, U.; Kruger, J.; Smialowska, S. Eds. *Advances in Localized Corrosion*; NACE: Houston, TX, 1990.
24. Young, L. *Anodic Oxide Films*; Academic Press: London, 1961.
25. Hoar, T. P.; Mott, N. F. *J. Phys. Chem. Solids* **1959**, *9*, 97.
26. Tajima, S. In *Advances in Corrosion Science and Engineering*; Fontana, M. G., Staehle, R. W., Eds.; Plenum: New York, 1970; Vol. 1.
27. Wood, G. C. In *Oxides and Oxide Films*; Diggle, J. W., Ed.; Marcel Dekker: New York, 1972; Vol. 2.
28. Furneaux, R. C.; Thompson, G. E.; Wood, G. C. *Corros. Sci.* **1978**, *18*, 853.
29. Shimuzu, K.; Kobayashi, K.; Thompson, G. E.; Wood, G. C. *Philos. Mag. A.* **1992**, *66*, 643.
30. Pedeferri, P. *Imagination on Titanium*; Cooperativa Libraria Universitaria del Politecnico: Milan, 1987.
31. de Gromoboy, T. S.; Shreir, L. L. *Electrochim. Acta* **1966**, *11*, 895.
32. Froment, M. Ed. *Passivity of Metals and Semiconductors*; Elsevier: Amsterdam, 1983.
33. MacDougall, B. R.; Alwitt, R. S.; Ramanarayanan, T. A. Eds. *Oxide Films on Metals and Alloys, The Electrochemical Society Proceedings,* 99–22, Pennington: New Jersey, 1992.
34. Kabanov, B.; Burstein, R.; Frumkin, A. *Disc. Farad. Soc.* **1947**, *1*, 259.
35. Heusler, K. E. *Zeit. Elektrochem.* **1958**, *62*, 582.
36. Bockris, J. O'M.; Drazic, D.; Despic, A. *Electrochim. Acta* **1961**, *4*, 325.
37. Weissmantel, C.; Schwabe, K.; Hecht, G. *Werkst Korros.* **1961**, *12*, 353.
38. Sato, N.; Okamoto, G. *J. Electrochem. Soc.* **1964**, *111*, 197.
39. Bockris, J. O'M.; Reddy, A. K. N. *Modern Electrochemistry;* Macdonald: London, 1970; Vol. 2.
40. Bessone, J.; Karakaya, L.; Lorbeer, P.; Lorenz, W. J. *Electrochim. Acta* **1977**, *22*, 1147.
41. Burstein, G. T.; Ashley, G. W. *Corrosion* **1983**, *39*, 241.
42. Misra, R. D. K.; Burstein, G. T. *Corros. Sci.* **1984**, *24*, 305.
43. Burstein, G. T.; Newman, R. C. *Electrochim. Acta* **1980**, *25*, 1009.
44. Burstein, G. T.; Newman, R. C. *J. Electrochem. Soc.* **1981**, *128*, 2270.
45. Verwey, E. J. W. *Physica.* **1935**, *2*, 1059.
46. Cabrera, N.; Mott, N. F. *Rep. Prog. Phys. 1948/49*, *12*, 163.
47. Dignam, M. J. In *Oxides and Oxide Films*; Diggle, J. W., Ed.; Marcel Dekker: New York, 1972; Vol. 1, p 91.
48. Fromhold, A. T. In *Oxides and Oxide Films*; Diggle, J. W., Vijh, A. K., Eds.; Marcel Dekker: New York, 1976; Vol. 3, pp 1.
49. Kirchheim, R. *Electrochim. Acta* **1987**, *32*, 1619.
50. Davenport, A. J.; Brustein, G. T. *J. Electrochem. Soc.* **1990**, *137*, 1496.
51. Chao, C. Y.; Lin, L. F.; Macdonald, D. D. *J. Electrochem. Soc.* **1981**, *128*, 1187, 1194.
52. Ghez, R. *J. Chem. Phys.* **1973**, *56*, 1838.
53. Burstein, G. T.; Davenport, A. J. *J. Electrochem. Soc.* **1989**, *136*, 936.
54. Güntherschultze, A.; Betz, H. *Zeit. Phys.* **1934**, *91*, 70, 92, 367..
55. Dewald, J. F. *J. Phys. Chem Solids* **1957**, *2*, 55.
56. Sato, N.; Cohen, M. *J. Electrochem. Soc.* **1964**, *111*(512), 624.
57. Diggle, J. W. In *Oxides and Oxide Films*; Diggle, J. W., Ed.; Marcel Dekker: New York, 1973; Vol. 2, p 280.
58. Nii, K. *Corros. Sci.* **1970**, *10*, 571.
59. Hagyard, T.; Williams, J. R. *Trans. Farad. Soc.* **1961**, *57*, 2288.
60. Hagyard, T.; Earl, W. B. *J. Electrochem. Soc.* **1967**, *114*, 694.
61. Beck, T. R. *Corrosion* **1974**, *30*, 408.
62. Burstein, G. T.; Ashley, G. W. A. *Corrosion* **1984**, *40*, 110.
63. Burstein, G. T.; Marshall, P. I. *Corros. Sci.* **1983**, *23*, 125.
64. Burstein, G. T.; Marshall, P. I. *Corros. Sci.* **1984**, *24*, 449.
65. Burstein, G. T.; Whillock, G. O. H. *J. Electrochem. Soc.* **1989**, *136*(1313), 1320.
66. Burstein, G. T.; Cinderey, R. J. *Corros. Sci.* **1991**, *32*, 1195.
67. Scully, J. C. Ed. *The Theory of Stress Corrosion Cracking in Alloys*; NATO: Brussels, 1971.
68. Sato, N.; Kudo, K. *Electrochim. Acta* **1971**, *16*, 477.
69. *Stress Corrosion Cracking and Hydrogen Embrittlement of Iron Base Alloys*; NACE: Houston, TX, 1975.
70. Mayne, J. E. O.; Pryor, M. J. *J. Chem. Soc.* **1949**, 1831–1835.
71. Cohen, M. *J. Phys. Chem.* **1952**, *56*, 415.
72. Mellors, G. W.; Cohen, M.; Beck, A. F. *J. Electrochem. Soc.* **1952**, *105*, 332.

73. Mayne, J. E. O.; Menter, J. W. *J. Chem. Soc.* **1954**, 103–107.
74. Brasher, D. M.; Mercer, A. D. *Br. Corros. J.* **1968**, *3*, 120.
75. Appleby, A. J.; Mayne, J. E. O. *J. Oil Col. Chem. Assoc.* **1967**, *50*, 897.
76. Mayne, J. E. O.; Page, C. L. *Br. Corros. J.* **1972**, *7*(111), 115.
77. Hoar, T. P.; Evans, U. R. *J. Chem. Soc.* **1932**, 2476–2481.
78. Szklarska-Smialowska, Z.; Staehle, R. W. *J. Electrochem. Soc.* **1974**, *121*, 1146.
79. Stern, M.; Wissenberg, H. *J. Electrochem. Soc.* **1959**, *106*(755), 759.
80. Heitz, E. In *Advances in Corrosion Science and Technology*; Fontana, M. G., Staehle, R. W., Eds.; Plenum Press: New York, 1974; Vol. 4, pp 149.
81. Mazza, F. *Werkst. Korros.* **1969**, *20*, 199.
82. Vernon, W. H. J.; Wormwell, F.; Nurse, T. J. *J. Iron Steel Inst.* **1944**, *150*, 81.
83. Cahoon, J. R.; Bandy, R. *Corrosion* **1982**, *38*, 299.
84. Olefjord, I. *Mater. Sci. Eng.* **1980**, *42*, 161.
85. Olefjord, I.; Elfstrom, B. O. *Corrosion* **1982**, *38*, 46.
86. Castle, J. E.; Clayton, C. R. *Corros. Sci.* **1977**, *17*, 7.
87. Asami, K.; Hashimoto, K.; Shimodaira, S. *Corros. Sci.* **1976**, *16*, 387.
88. Uhlig, H. H. *Corrosion and Corrosion Control*, 2nd edn.; Wiley: New York, 1971; Chapter 5.
89. Sieradzki, K.; Newman, R. C. *J. Electrochem Soc.* **1986**, *133*, 1979.
90. Song, Q.; Newman, R. C.; Cottis, R. A.; Sieradzki, K. *J. Electrochem. Soc.* **1990**, *137*, 435.
91. Song, Q.; Newman, R. C.; Cottis, R. A.; Sieradzki, K. *Corros. Sci.* **1990**, *31*, 621.
92. Williams, D. E.; Newman, R. C.; Song, Q.; Kelly, R. G. *Nature* **1991**, *350*, 216.
93. Chin, Y.-T.; Cahan, B. D. *J. Electrochem. Soc.* **1992**, *139*, 2432.
94. Okamoto, G. *Corros. Sci.* **1973**, *13*, 471.
95. McBee, C. L.; Kruger, J. *Electrochim. Acta* **1977**, *17*, 1337.
96. Bessone, J. B.; Salinas, D. R.; Mayer, C. E.; Ebert, M.; Lorenz, W. J. *Electrochim. Acta* **1992**, *37*, 2283.
97. Burstein, G. T.; Davies, D. H. *Corros. Sci.* **1980**, *20*, 989; 1143.
98. Goetz, R.; MacDougall, B. J.; Graham, M. J. *Electrochim. Acta* **1986**, *31*, 1299.
99. Mitrovic-Stepanovic, V.; MacDougall, B. J.; Graham, M. J. *Corros. Sci.* **1987**, *27*, 239.
100. Landolt, D.; Mischler, S.; Vogel, A.; Mathieu, H. J. *Corros. Sci.* **1990**, *31*, 431.
101. Mischler, S.; Vogel, A.; Mathieu, H. J.; Landolt, D. *Corros. Sci.* **1991**, *32*, 925.
102. Marcus, P. *C. R. Acad. Sci. Paris., Sér II* **1987**, *305*, 675.
103. Marcus, P.; Moscatelli, M. *Mém. Etud. Sci. Rev. Métall.* **1988**, *85*, 561.
104. Szklarska-Smialowska, Z. *Pitting Corrosion of Metals*; NACE: Houston, TX, 1986.
105. Newman, R. C. *Corrosion* **1985**, *41*, 450.
106. Pistorius, P. C.; Burstein, G. T. *Philos. Trans. Roy. Soc. Lond. A* **1992**, *341*, 531.
107. Bertocci, U.; Koike, M.; Leigh, S.; Qiu, F.; Yang, G. *J. Electrochem. Soc.* **1986**, *133*, 1782.
108. Isaacs, H. S. *Corros. Sci.* **1989**, *29*, 313.
109. Forchhammer, P.; Engell, H. J. *Werkst. Korros.* **1969**, *20*, 1.
110. Williams, D. E.; Westcott, C.; Fleischmann, M. *J. Electrochem. Soc.* **1985**, *132*(1796), 1804.
111. Stockert, L.; Böhni, H. *Mater. Sci. Forum* **1989**, *44*(45), 313.
112. Coates, G. E. *Mater. Perfom.* **1990**, *29*(8), 61.
113. Hoar, T. P.; Jacob, W. R. *Nature* **1967**, *216*, 1299.
114. Heusler, K. E.; Fischer, L. *Werkst. Korros.* **1976**, *27*, 551.
115. Bertocci, U.; Yang-Xiang, Y. *J. Electrochem. Soc.* **1984**, *131*, 1011.
116. Riley, A. M.; Wells, D. B.; Williams, D. E. *Corros. Sci.* **1991**, *32*, 1307.
117. Burstein, G. T.; Mattin, S. P. *Philos. Mag. Lett.* **1992**, *66*, 127.
118. Masumoto, T.; Hashimoto, K. *Ann. Rev. Mater. Sci.* **1978**, *8*, 215.
119. Hashimoto, K. In *Amorphous Metallic Alloys*; Luborsky, F. E., Ed.; Butterworths: London, 1983; pp 471.
120. Diegle, R. B.; Sorensen, N. R.; Tsuru, T.; Latanison, R. M. In *Treatise on Materials Science and Technology*; Scully, J. C., Ed.; Academic Press: London, 1983; Vol. 23, pp 59.
121. Wislawska, M.; Janik-Czakor, M. *Br. Corros. J.* **1985**, *20*, 36.
122. Burstein, G. T. *Corrosion* **1981**, *37*, 549.
123. Galvele, J. R. *J. Electrochem. Soc.* **1976**, *123*, 464.
124. Leckie, H. P.; Uhlig, H. H. *J. Electrochem. Soc.* **1966**, *115*, 1262.
125. Pistorius, P. C.; Burstein, G. T. *Corros. Sci.* **1992**, *33*, 1885.
126. Laevers, P.; Terryn, H.; Vereecken, J. *Trans. Inst. Met. Finish.* **1992**, *70*, 105.

2.03 Crevice Corrosion

N. Corlett
Lloyd's Register EMEA, Denburn House, 25 Union Terrace, Aberdeen, AB10 1NN, UK
L. E. Eiselstein
Exponent, Inc., Menlo Park, CA, USA
N. Budiansky
Exponent, Inc., Natick, MA, USA

Abbreviations

BWR Boiling-water reactor
CCS Critical crevice solution
CCT Critical crevice temperature
CPT Critical pitting temperature
DMC Dissimilar metal crevice
ELI Extra low interstitial
IGA Intergranular attack
MCA Multiple crevice assembly
MIC Microbiologically influenced corrosion
PBS Phosphate-buffered saline
PREN Pitting resistance equivalence number
PWR Pressurized-water reactor
SCC Stress corrosion cracking
TPLO Tibial plateau leveling osteotomy

Symbols

E_b Breakdown potential
E_{corr} Corrosion potential
E_{crit} Potential above which localized corrosion may occur
E_p Protection or repassivation potential
E_r Rest potential
X_{crit} Distance into crevice to achieve the primary pitting potential
ΔE $E_{crit} - E_{corr}$
$\Delta\varphi$ IR Drop into crevice

2.03.1 Introduction

Crevice corrosion is a form of localized attack that occurs within occluded regions or crevices of metallic components. The attack is caused by a change in the environment of the crevice relative to the bulk solution. Usually, the term crevice corrosion is limited to describing the attack in normally passive metals and alloys (e.g., stainless steels); however, in the broader sense, the corrosion of nonpassive metals and alloys (e.g., carbon steel) by differential aeration may also be regarded as a form of crevice corrosion. Crevices occur in many engineered structures; consequently, crevice corrosion is a common form of corrosion in most, if not all, industries. Crevices can be either man-made by design, such as in washers, pipe flanges, or overlapping fuselage skins on aircraft, or they can occur naturally as a result of biofouling, deposits, and debris. Many similarities exist between crevice corrosion and other forms of localized attack, such as pitting, intergranular attack, environmentally-assisted cracking, and dealloying.[1,2] In all these instances, the electrochemical and chemical conditions in the shielded surfaces within the crevice, pit, or crack, become altered when compared with the conditions on the boldly exposed surfaces. Crevice corrosion is usually described in two phases: initiation and propagation. The initiation phase may be further divided into three stages: deoxygenation, alteration of the crevice solution chemistry, and depassivation. Many variables affect crevice corrosion, both in terms of the time-to-initiation and the rate and extent of propagation. For a given material, perhaps the most important factors are the geometry of the crevice and the bulk properties of the environment. In this chapter, the emphasis is on the crevice corrosion of passive metals and alloys in aerated chloride environments since this is one of the most common types of crevice corrosion. The mechanisms of crevice corrosion are discussed, together with the effects of the many variables that affect it. Modeling and testing for resistance to crevice corrosion are discussed, as are practical design aspects and methods to minimize attack. Finally, some industry-specific examples of crevice corrosion are provided.

2.03.2 Theories of Crevice Corrosion

2.03.2.1 Overview

Crevice corrosion was originally described by Mears and Evans[3] and Uhlig,[4] as resulting from the development of a differential aeration cell. Mears and Evans found that accelerated corrosion occurred when a passive metal surface was in intimate contact with that of an inert material. They proposed that oxygen consumed by a cathodic reaction is replenished more slowly in the occluded region between the metal and the inert object, and that the metal surface within the crevice is deoxygenated and anodic compared to the boldly exposed surfaces outside the crevice. This difference in oxygen concentration in the solution inside the crevice relative to the bulk solution outside the crevice leads to accelerated corrosion within the crevice. Following these early observations and theories, substantial research has focused on developing a more detailed explanation of crevice corrosion.

Three general classes of theories have been used to describe the phenomenon of crevice-corrosion:

- the development of a critical crevice solution (CCS);

- the development of a critical ohmic potential difference (IR drop);
- the stabilization of metastable pits by the crevice geometry.

An ideal crevice may be described as one having a uniform gap, typically between 0.1 and 100 μm, along its entire length or depth, and it is formed between the substrate and another surface that may be metallic or nonmetallic, as shown in **Figure 1**.

Ideal crevice geometries rarely exist even in well controlled laboratory experiments. As the crevice gap gets narrower, crevice corrosion is more likely to occur due to restricted mass transport. In addition, as the gap narrows, the surface finishes on the two materials forming the crevice become increasingly important. Also, the effect of surface finish makes it difficult to accurately characterize the crevice gap as is illustrated in **Figure 2**.

2.03.2.2 Critical Crevice Solution

Work on crevice corrosion modeling in the late 1960s and the early 1970s focused on the growth stage of crevice corrosion.[5,6] The work of Crolet and Defranoux[7] in 1973 represents the first attempt to model the initiation stage of crevice corrosion. Crolet and Defranoux modeled the time it takes to develop a critical chemical environment (low pH and high halide concentration) within a crevice as a function of crevice geometry, alloy composition, and bulk solution chemistry. The development of a CCS chemistry has been used successfully by several researchers to rank alloy performance.[6,8–10]

In 1978, Oldfield and Sutton[5,6] further developed the Crolet and Defranoux model to the point that it could be used to predict crevice corrosion susceptibility. Their model considered the effect of the crevice geometry (i.e., depth and width) on the mass transport of chemical species into and out of the crevice, and then, the manner in which this process led to the formation of a critical crevice solution. Their model consisted of four stages: (1) deoxidization, (2) an increase in salinity and acidity, (3) depassivation, and (4) propagation. In their model, crevice corrosion initiation occurs when a creviced metal or alloy is exposed to an aerated, neutral pH, chloride environment. Initially, corrosion occurs over the entire surface, both within the occluded region and over the boldly exposed surfaces. In order for crevice corrosion to occur, the exposed metal must be passive in the bulk environment. Initially, metal dissolution occurs at the passive current density, which is typically of the order of 1 $\mu A\,cm^{-2}$ or less, and both anodic and cathodic reactions occur on all surfaces. As the solution containing dissolved oxygen is consumed at cathodic sites on the noncreviced surfaces, it is replenished from the bulk solution. In occluded regions, however, where mass transport is limited, oxygen becomes depleted and the metal surfaces within the crevice become net anodic, while the boldly exposed surfaces become net cathodic. This difference in the dissolved oxygen concentration sets up an oxygen concentration cell between the passive boldly exposed surfaces and the crevice mouth (cathode), and the occluded region of the crevice (anode). The depletion of dissolved oxygen in the solution contained within the crevice polarizes the metal surfaces in the crevice to lower, more negative potentials.

Although differential aeration is sufficient to induce crevice corrosion in certain metals and alloys, it only provides the initial driving force for crevice

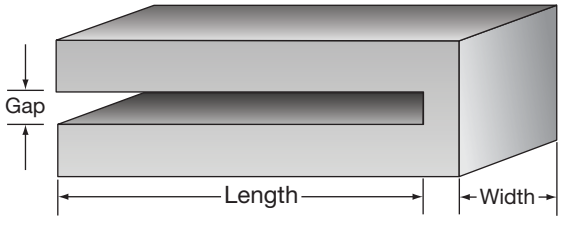


Figure 1 Schematic of an ideal crevice geometry.

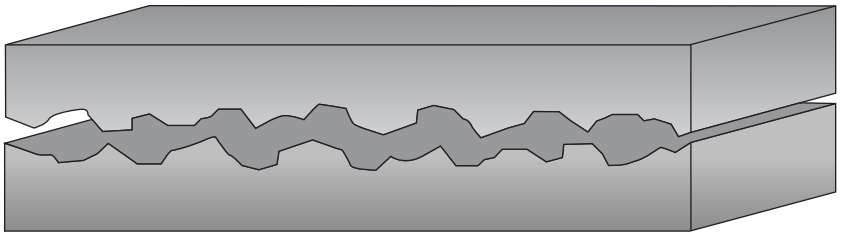


Figure 2 Schematic of hypothetical crevice geometry.

corrosion in stainless alloys. In chloride containing solutions, the iron and chromium cations, which are oxidized as they diffuse through the passive film, dissolve, and accumulate in the crevice. The limited mass transport of metal ions out of the crevice causes metal-ion accumulation, and to preserve electroneutrality, anions (e.g., chloride ions) diffuse into the crevice through electromigration, leading to an increased metal-chloride concentration. The pH within the crevice drops as a result of the metal-ion hydrolysis.

When the concentration of chlorides is sufficiently high and the pH is sufficiently low, the passive film is no longer able to maintain stability, and active corrosion results. This CCS theory relies on the accumulation of aggressive species within the crevice, and subsequent depassivation followed by active dissolution, as indicated by polarization curve a in **Figure 3**. Meanwhile, the boldly exposed surfaces remain in the passive state, as indicated by polarization curve c shown in **Figure 3**. The CCS theory predicts that the most severe crevice corrosion damage should occur in the deepest regions of the crevice since these would be the first areas to become deoxygenated because of the oxygen mass transport limitations. In some systems, however, crevice corrosion is observed at intermediate distances from the crevice mouth. Another limitation of the CCS model, as noted in the Oldfield and Sutton paper (1973), is that no consideration is given to the effect of the IR drop during the initiation stage.

2.03.2.3 IR Drop

The ohmic potential difference or the so-called IR drop theory for describing crevice corrosion was developed by Pickering.[11,12] The IR drop refers to the variation in potential between the crevice interior and the boldly exposed surfaces surrounding the crevice mouth. If the IR drop is sufficiently large, the potential in the crevice can drop below the value required to maintain a stable passive film, thereby initiating crevice corrosion through depassivation. The IR drop increases as the aspect ratio (i.e., length to depth ratio) or tightness of the crevice increases. The IR drop is also affected by other factors such as solution chemistry and conductivity, as well as the presence of other physical objects, such as hydrogen bubbles, solid corrosion products, and salt films that may further block the crevice.[12]

When the metal surfaces within the crevice become net anodic, the local corrosion potential typically becomes more active (i.e., more negative). With the relatively large boldly exposed surfaces supporting the cathodic reactions at more noble (i.e., more positive) potentials, it might be expected that the small surface area within the crevice region would be polarized to higher potentials where it would no longer be active. Due to large IR drops, however, the potential of the crevice region can be hundreds of millivolts lower than that of the boldly exposed surfaces.[11,12]

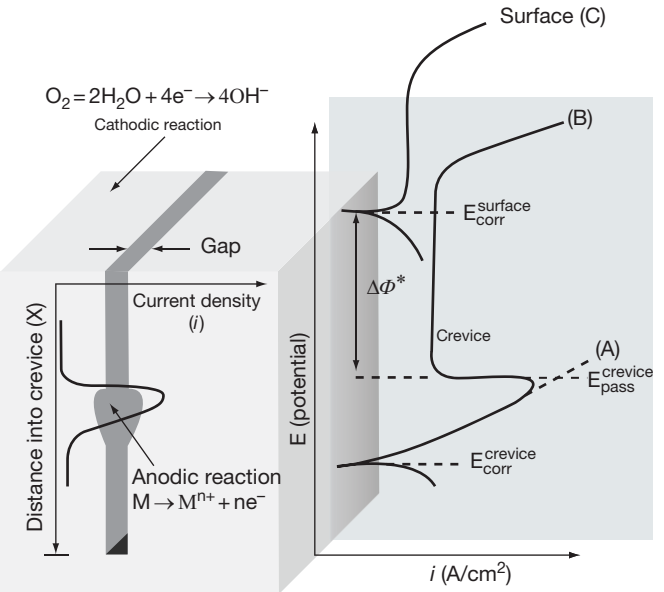


Figure 3 Schematic of crevice corrosion. Net anodic reactions occur within the crevice environment while the boldly exposed surface is net cathodic. Three possible polarization behaviors are shown: (A) active dissolution within the crevice region, (B) active–passive behavior within the crevice region, and (C) passive behavior on the boldly exposed surface.

In metals and alloys that exhibit an active-to-passive transition in bulk environments, the boldly exposed surfaces are maintained in the passive region in aerated solutions, as indicated by polarization curve B in Figure 3. The large IR drop, however, can cause the local potential of the metal surfaces within the crevice to be less than the primary passivation potential, resulting in accelerated metal dissolution. The distance into the crevice where the interfacial potential reaches the primary pitting potential, is referred to as $X_{crit,}$ and the critical IR drop to achieve this potential is called IR* or $\Delta\phi^*$.[11,12] As the distance into the crevice increases, the resistance increases and the current decreases; consequently, the degree of crevice corrosion also decreases. The IR drop theory predicts the intermediate attack that is typically observed in many situations in which the metal or alloy undergoes an active-to-passive transition in the bulk environment, but it does not model the solution concentration process occurring in the crevice.

If the CCS and IR drop theories are combined, the behavior of metals and alloys that do not exhibit an active-to-passive transition in bulk solutions, but do show such behavior in lower pH solutions, can be explained. For instance, by applying both the CCS and the IR drop theories, the crevice corrosion behavior of Inconel 625 can be explained.[13,14]

Indeed, the majority of materials that are susceptible to crevice corrosion are passive in the bulk environment (e.g., stainless steels in aerated chloride environments) and do not normally exhibit an active-to-passive transition. Stewart and Kelly investigated the behavior of 300 series stainless steels taking into consideration the interaction of both the CCS model and the IR drop theory for initiation and propagation.[1,15,16] Their model considered conditions whereby an alloy was passive in the bulk environment, as shown in Figure 3, polarization curve c, but where crevice solution concentration and acidification must occur to allow an active-to-passive transition within the crevice, as shown in Figure 3, polarization curve b, which, in turn, results in crevice corrosion in the active region. With this approach, the Stewart and Kelly model correctly predicts the intermediate crevice-corrosion damage typically observed on many metals and alloys.

2.03.2.4 Stabilization of Metastable Pits

Crevice corrosion occurs at potentials below the pitting potential in most environments. Similarly, metastable pitting occurs on stainless steel alloys at potentials less than the pitting potential.[2,17] Stockert and Boehni[18] have suggested that crevice corrosion is initiated on stainless steels as a result of the occluded environment preventing the repassivation of metastable pits. Typically, repassivation of metastable pits occurs as a result of the pit's inability to maintain a critical solution concentration necessary for autocatalytic pit growth. The geometry of the occluded region prevents mass transport of metal-ion hydrolysis products generated from the metastable pitting, out of the crevice, which leads to an alteration of the crevice chemistry in a manner similar to the CCS model first proposed by Crolet and Defranoux[7] and Oldfield and Sutton.[5,6] Oldfield and Sutton observed that random metastable pits formed within the crevice in areas where mass transport was most limited. As exposure time increases, the density of metastable pitting events within the crevice increases; this is followed by pit coalescence as the density of the pits increases further. This coalescence of metastable pits is observed on electrode surfaces prior to crevice corrosion propagation. Other authors support Stockert and Boehni's contention.[19,20] For example, Laycock *et al.* hypothesized that crevice corrosion could be initiated by a random metastable pitting event occurring at a critical location within a crevice, where the metastable pits would then stabilize.

2.03.2.5 Similarity to Pitting

Crevice corrosion and pitting corrosion are similar in many respects. The ranking of metal alloys for crevice corrosion and pitting corrosion are almost identical, and, therefore, it follows that the effect of alloying elements and the environmental effects on crevice corrosion and pitting are also similar.[2] One of the differences, however, between the two mechanisms is the initiation site for corrosion. In pitting corrosion, the initiation sites are, typically, heterogeneities in the metal or alloy surface (e.g., MnS inclusions in stainless steel or second phase particles in aluminum alloys) that are more susceptible than the metal or alloy matrix (e.g., MnS inclusions in stainless steel or second phase particles in aluminum alloys). In contrast, the initiation sites for crevice corrosion are generally defined by occluded regions as discussed previously, although pitting at surface heterogeneities, such as inclusions, has also been observed within crevices. While pitting generally occurs in halide containing solutions, crevice corrosion can also occur in nonhalide solutions, for example, nickel in sulfuric acid environments.[2,17] Also, the initiation mechanism is different for pitting

and crevice corrosion. Pitting is thought to be a random or stochastic process with respect to time and location, with metastable pits forming below the critical pitting potential, whereas crevice corrosion is generally considered a more deterministic process.[2]

Both pitting and crevice corrosion exhibit critical threshold potentials and temperatures. Typically, the critical crevice potential is lower than the critical pitting potential as the solution chemistry can change more rapidly in a crevice due to metal-ion hydrolysis and electromigration than in a pit on a boldly exposed surface. As a consequence, crevice corrosion usually initiates more readily than pitting. The critical crevice temperature (CCT) is usually lower than the critical pitting temperature (CPT) for many alloys for similar reasons.[2]

2.03.2.6 Variables Affecting Crevice Corrosion

2.03.2.6.1 Crevice geometry

The crevice gap plays a major role in the development of the conditions required for crevice corrosion. Both the gap and the depth of the crevice control the mass transport (solution concentration, acidification, and deoxygenation) and potential gradients that develop within the crevice. In crevices with large gaps the mass transport of oxygen is relatively unaffected compared to bulk solutions, and crevice corrosion is unlikely. Narrow crevice gaps, however, create steep concentration gradients, which are more apt to initiate crevice corrosion closer to the crevice mouth. In addition, in narrow crevices the surface finishes on each side of the crevice become increasingly important. When the crevice surfaces are very close together the minimum crevice gap is dictated by the surface roughness. Large crevice gaps, however, are unaffected by surface roughness.

Various scaling factors are used to describe the effect of the crevice gap and depth on crevice corrosion; as the aspect ratio increases, mass transport gets more difficult and crevice corrosion becomes more likely. Additionally, studies have shown that large aspect ratios lead to crevice corrosion initiation closer to the mouth opening due to IR drop mechanisms.[21] To account for these effects, scaling laws are typically given in terms of the aspect ratio L/G or L^2/G where L is the depth at which crevice corrosion is first observed, and G is the crevice gap. Such scaling allows corrosion engineers and scientists to extrapolate crevice corrosion test results to other proposed design conditions.

2.03.2.6.2 Solution chemistry

For a given material, susceptibility to crevice corrosion and the rate at which it occurs are dependent on the environment to which it is exposed. Temperature plays an important role in a material's susceptibility to crevice corrosion. Below a critical temperature, a material is regarded as resistant to crevice corrosion, and corrosion rates are very low. Conversely, above this critical temperature, the corrosion rate increases sharply with increasing temperature. Additionally, one of the main natural crevice corrosion formers is biofilm growth. In specific temperature ranges, biofilms accelerate crevice corrosion by facilitating the cathodic reactions. Biofilms are less active above and below this temperature range leading to a reduced effect on the cathodic reactions, and therefore, lower crevice corrosion rates.

Solution chemistry plays an important role in the susceptibility of metals and alloys to crevice corrosion. Chloride-containing solutions are the most widely investigated and characterized environments in which crevice corrosion occurs on otherwise normally corrosion resistant materials.[22,23] Chloride ions facilitate crevice corrosion initiation and propagation. A critical concentration of chloride ions has been determined for many materials; below this critical concentration, crevice corrosion does not occur. Above this concentration, however, crevice corrosion is strongly dependent on concentration. As the concentration of chloride ions increases, the potential at which crevice corrosion initiates, decreases.[24] One explanation for this is that chloride-ion concentration plays a significant role in the mechanisms of pitting and metastable pitting. As the chloride ion concentration increases, the potential at which metastable pitting occurs decreases, as does the pitting potential. Metastable pits have been found on surfaces within crevices at the onset of crevice corrosion, and pit coalescence, at later stages.

The solution flow rate has been shown to affect the initiation and propagation phases of crevice corrosion differently. The initiation stages are inhibited by solution flow by enhancing the mass transport of oxygen to creviced regions.[22,23] Additionally, solution flow can help reduce the occurrence of crevice corrosion resulting from natural crevice formers by preventing sediment accumulation or attachment (e.g., biofilms or microorganisms).[23] Once crevice corrosion initiates, however, solution flow can increase the rate of crevice corrosion by increasing the rate-limiting cathodic reaction of oxygen reduction on the boldly exposed

surfaces outside the crevice. As the cathodic reaction increases, so the anodic dissolution reaction in the crevice must also increase.

2.03.2.6.3 Alloying

The alloy additions that impart resistance to pitting corrosion are for the most part the same as those that make alloys more resistant to crevice corrosion. One of the most widely investigated groups of alloys are the stainless steels. The additions of chromium, molybdenum, nitrogen, and nickel have a significant influence on the crevice corrosion resistance of many alloys. The crevice corrosion rate of austenitic stainless steels has been shown to decrease with increasing chromium additions in seawater.[23] The addition of chromium enhances the passivity of the oxide film, making it more resistant to breakdown. As with resistance to pitting, crevice corrosion resistance is enhanced by the addition of molybdenum. The addition of nickel has been shown to be beneficial to the crevice corrosion resistance of 2.5% Mo austenitic stainless steels, while it has shown very little effect for 6.5% Mo stainless steels. Molybdenum and nitrogen additions have been shown to have a synergistic effect. In alloys that contain molybdenum, nitrogen addition results in significant decreases in the crevice corrosion rate. Similar results are observed for ferritic stainless steels as for austenitic stainless steels.[23] Manganese sulfide inclusions have also been shown to influence the crevice corrosion resistance of austenitic stainless steels; they have been found to destabilize passive films, leading to an increased susceptibility to crevice corrosion.[2,25–30]

The pitting resistance of different stainless steel alloys is generally approximated from the alloy composition by means of the pitting resistance equivalence number (PREN) equation.[23] A typical PREN formula for stainless steels[31] is $PREN = Cr + 3.3Mo + 30N$, while, according to Roberge,[32] the most widely used formula for austenitic and duplex stainless steels is $PREN = Cr + 3.3\ (Mo + 0.5W) + xN$, where $x = 16$ for duplex stainless steel, and $x = 30$ for austenitic alloys. This shows that increasing chromium, molybdenum, and nitrogen will increase the pitting and crevice corrosion resistance of a stainless steel. According to Grubb and DeBold,[33] the numerical value of the PREN is approximately equal to the critical crevice temperature (CCT) as measured in degrees Celsius in natural seawater or in ferric chloride solution. This can be seen by comparing the PREN formula with the formula for the CCT in degrees Celsius for stainless steels, which is given by $CCT = 2.5Cr + 7.6Mo + 31.9N - 41.0$.[34] PREN and CCT data are presented for a number of alloys in 30 °C filtered seawater for 30 days in Table 1, which shows the limited equivalence between PREN and CCT.

High performance nickel alloys show improvements in crevice corrosion resistance similar to stainless steels in terms of alloying elements. Alloy C-276, which contains 8% molybdenum, has a relatively higher crevice corrosion resistance, when compared with alloys with lower molybdenum contents.

Table 1 Crevice corrosion rankings in 30 °C filtered seawater for 30 days, with PREN[a] and CCT[b]

Alloy	*UNS designation*	*Crevice corrosion resistance ranking*	*PREN*	*CCT (°C)*
Hastalloy C-276	N10276	1	66.7	
Inconel 625	N06625	1	50.4	
AL 29–4	S44700	1	42.8	
AL 29–4–2	S44735	1	42.7	
Monit	–	1	37.8	
Sea-Cure	S44660	2	35.2	
Ferralium 255	S32550	3	36.8	22.5
Hastelloy G-3	N06985	4	45.9	
Haynes 20 Mod	–	5	38.1	
26–1S	S44626	6	28.3	
20Mo-6	N08026	7	42.4	
E-Brite	S44627	8	29.2	
AL-4X	–	9	34.7	
AL-6X	N08366	10	41.5	
254SM0	S31254	11	46.1	
Hastelloy G	N06007	12	41.3	
904L	N08904	13	36.0	0
AISI 216	S21600	14	38.8	
254 FER	–	15	40.8	
254 SLX	N8904	16	36.6	
Rex 734	S31675	17	42.8	
Type 317LM	S31725	18	34.8	
Nitronic 50	S20910	19	36.5	
JS700	N08700	20	35.2	
Type 316	S31600	21	25.4	−2.5
20Cb-3	N08020	22	26.7	0

[a]Bond, A. P.; Dundas, A. J. *Mater. Perform.* **1984,** *23*(7), 39.
[b]Roberge, P. R. *Handbook of Corrosion Engineering*, McGraw-Hill: New York, 1999; p 734.

Similarly, increases in chromium content lower the rate of crevice corrosion.[13]

Titanium and its alloys are, typically, very resistant to crevice corrosion; for instance, pure titanium (Ti-Grade 2) will not generally undergo crevice corrosion below 70 °C, regardless of the solution pH or chloride concentration. When the pH exceeds 10, pure titanium will not undergo crevice corrosion at any temperature.[35] Although titanium and its alloys have high resistance to crevice corrosion, small additions of alloying elements lead to a marked reduction in the rates of crevice corrosion.[31] The addition of noble alloying elements, such as platinum and palladium, as well as copper, significantly decreases the crevice corrosion rate of titanium alloys by providing additional cathodic reaction sites within the crevice. Increasing the number and area of cathodic reaction sites within the crevice allows for the reduction of hydrogen ions, preventing the large scale separation of anodic and cathodic reaction sites.[31]

2.03.3 Evaluation of Crevice Corrosion

Numerous tests have been developed to investigate and evaluate the crevice corrosion susceptibilities of metals and alloys. The reviews by Kearns and Kain describe many of the test methods.[9,36–38] In general, testing has been focused primarily on normally passive alloys, and in particular, it has been directed towards the identification and development of more corrosion resistant alloys for marine applications and certain process industries, such as the pulp and paper industry.

Of the many tests that have been developed, several broad classifications can be made based on the type of test, for example, field or *in situ* tests, as opposed to those conducted in the laboratory, and real-time tests versus accelerated tests. Another differentiation can be made between electrochemical and nonelectrochemical (e.g., simple immersion) tests, as well as those tests designed specifically to investigate initiation rather than propagation. Further classifications can be made according to the specific environment, the type of crevice former used, and the geometry of the specimen (e.g., flat sheet, cylindrical, etc.)

As is often the case with testing versus real-life experience, the occurrence or absence of crevice corrosion in a given test is no assurance that it will or will not occur under normal service conditions, either as one moves from the laboratory to the field, or as field process conditions change with time.

2.03.3.1 Variables Affecting Crevice Corrosion Testing

In designing or selecting an appropriate test for evaluating crevice corrosion, a number of variables need to be taken into consideration:

- The metallurgical condition of the metal or alloy, including its composition and surface finish.
- The material used for the crevice former, including its surface finish. The nature of the crevice former can affect the crevice geometry, through relaxation of the crevice gap over time. If a metallic crevice former is used, then dissimilar metal corrosion may be an issue to consider.
- Crevice geometry, including the crevice gap, length or depth, and tightness. The geometry of the crevice can affect the initiation and propagation phases differently, and partly determine the crevice-to-boldly-exposed surface area ratio.
- The bulk environment, including solution chemistry, temperature, and flow rate, all of which affect crevice corrosion and can be utilized to accelerate testing.

2.03.3.2 Field or *In Situ* Tests

ASTM International standard G 4-01, Standard Guide for Conducting Corrosion Tests in Field Applications, provides general guidance for such tests, particularly with respect to surface preparation, specimen design, test-rack assembly, and post-test evaluation.[39] The standard mentions crevice corrosion and makes reference to ASTM standard G 78-01. The use of test racks to investigate *in situ* crevice corrosion provides little information regarding initiation and propagation rates.[9,36] However, this approach is effective for screening and ranking different alloys.[40]

2.03.3.3 Laboratory Tests

2.03.3.3.1 ASTM G 48-03, standard test methods for pitting and crevice corrosion resistance of stainless steels and related alloys by use of ferric chloride solution

ASTM standard G 48-03 lists several procedures to determine the pitting and crevice corrosion resistance of stainless steels and other nickel-based corrosion-resistant alloys, when exposed to oxidizing chloride environments.[41] This test describes four procedures, identified as methods A, B, C, and D. Method

A is a ferric chloride pitting test, and Method C is a test developed for determining the critical pitting temperature; therefore, these two methods will not be discussed.

Method B is a ferric chloride test that was developed by Streicher for determining both the pitting and crevice corrosion resistance of iron- and nickel-based alloys.[42] In this test, the crevices are formed by two plastic blocks held against the specimen faces. The maximum depth or mass loss is measured after immersion for a specified time. The recommended temperatures and duration are 23 or 50 °C, and 72 h, respectively.

Method D is used to rank alloys with respect to the minimum critical crevice temperature to cause initiation of crevice corrosion in 72 h, and is similar to the MTI-2 test developed by Treseder.[43] The critical crevice temperature is defined by ASTM G 48 as the minimum temperature to produce crevice attack to a depth of at least 0.025 mm on the surface of the specimen beneath the crevice former in 72 h.

These accelerated test methods (A–D) were designed to minimize corrosion initiation times compared with most natural environments. The rationale for using ferric chloride solutions is that similar solutions are expected to develop within a pit or crevice site.[44] (While this may have been the original rationale, it is not certain why ferric chloride works (rather than any other chloride solution). Rather ferric ion reduction provides a powerful cathodic reaction, leading to a very positive potential that ensures that crevice corrosion will occur if the potential is above the CCT.) The relative performance of alloys in ferric chloride solution tests has been correlated to performance in certain real environments, such as natural seawater at ambient temperature[45] and strongly oxidizing, low pH, chloride-containing environments.[10] Several exceptions to these correlations, however, have been reported.[46,47] This test method notes that surface preparation can significantly influence results. For instance, grinding or pickling of stainless steel surfaces can destroy the passive layer; however, a 24-h air-passivation after grinding or pickling will minimize these differences.

2.03.3.3.2 ASTM G 78-01, standard guide for crevice corrosion testing of iron-base and nickel-base stainless alloys in seawater and other chloride-containing aqueous environments

ASTM G 78-01 was developed to measure the crevice corrosion resistance of iron- and nickel-based alloys to chloride containing solutions for a specified time period.[48] This standard recommends a 30-day exposure. The sample surface is occluded with two multiple-gap plastic washers, or so-called multiple crevice assemblies (MCA), that are firmly pressed against both surfaces. This test was developed in the 1970s and is often cited in the literature. Tests are performed on triplicate specimens to allow for statistical analysis, with the crevice formers tightened using a calibrated torque wrench. Crevice initiation and propagation are addressed by the number of sites initiated, and the maximum depth of attack, respectively. Round-robin testing by both NACE T-5H-6 and ASTM G 1.09 task groups report results illustrating the variability of these types of test. The use of multiple-crevice assemblies can be useful for the ranking of alloys, as well as for investigating the effect of other variables; however, multiple-crevice assemblies should not be used for predicting alloy performance in other conditions.

2.03.3.3.3 ASTM F 746-87 (Reapproved 1999), standard test method for pitting or crevice corrosion of metallic surgical implant materials

ASTM F 746 can be used to evaluate the localized corrosion resistance of alloys for implantable medical devices and surgical instruments.[49] It is a modified version of a test developed by Syrett,[50] and is used as a screening test to rank alloys.

A cylindrical sample of the alloy to be tested is fitted with an inert tapered collar and then immersed in a phosphate-buffered saline (PBS) solution at 37 °C. After stabilization over 1 h, the sample is polarized to a potential much more noble than its rest potential, typically 0.8 V (SCE), in order to stimulate localized corrosion – either pitting or crevice corrosion. The onset of localized corrosion is evident as a rapid increase in the current. If localized corrosion cannot be stimulated, then, after 15 min, the test is terminated. If localized corrosion is stimulated, however, the potential is immediately decreased to the first of several preselected potentials (usually the rest potential plus 0.05 V), and the current is monitored to see whether the sample repassivates, or continues to corrode. If the sample does not repassivate after 15 min, the test is terminated. If the sample does repassivate, the stimulation step is repeated, followed by the next repassivation step. The critical potential for localized corrosion is defined as the most noble preselected potential at which repassivation occurs. Post-test examination of the sample establishes

whether the localized corrosion resulted from pitting of the exposed surface, or from crevice corrosion in the tapered collar.

2.03.3.3.4 ASTM F 2129-06, standard test method for conducting cyclic potentiodynamic polarization measurements to determine the corrosion susceptibility of small implant devices

ASTM F 2129 was developed to evaluate the localized corrosion resistance of medical implants in their finished condition.[51] The method can, however, also be used for screening alloys and evaluating component parts of implants. Implants (or alloy components) are immersed in PBS at 37 °C for up to 1 h to stabilize at the so-called rest potential (E_r) or open circuit potential, after which the potential of the sample is increased slowly (0.167 or 1 mV s^{-1}) in a positive (noble) direction. The breakdown potential (E_b) at which localized corrosion occurs is evident by a rapid increase in the current and, typically for stainless steels or Nitinol in PBS, can be observed by the appearance of a cloud of corrosion product forming around the breakdown location. If no breakdown occurs before a preset vertex potential is reached, usually 0.8 or 1 V (SCE), then the test can be terminated. If breakdown does occur, and the test is designed to be destructive, then the potential sweep is reversed and monitored for repassivation at the so-called protection or repassivation potential (E_p). There remains, however, some controversy among researchers over how to interpret the data obtained from this test. One approach is to consider the gap between the breakdown and rest potentials (i.e., E_b–E_r). This gap provides a measure of the pitting corrosion resistance, whereby, the greater this gap, the greater the resistance to pitting corrosion. Although not universally accepted by researchers, this is a commonly accepted parameter for the ranking of different alloys or material processing.[52–56] Similarly, the parameter $E_p - E_r$ is sometimes used as a measure of crevice corrosion resistance. This last parameter, however, is very dependent upon the amount of pitting or crevice corrosion that occurs before the potential scan is reversed.[57] An alternative approach for finished devices is to consider the parameter $E_b - E_r$ as a measure of localized corrosion resistance, together with a posttest microscopic examination to determine whether breakdown occurred because of pitting or crevice corrosion.

ASTM G 61-86, Standard Test Method for Conducting Cyclic Potentiodynamic Polarization Measurements for Localized Corrosion Susceptibility of Iron-, Nickel-, or Cobalt-Based Alloys, is similar to ASTM F 2129 with the exception that the test is performed at 25°C in 3.5% sodium chloride solution.[58]

2.03.3.3.5 Remote crevice assemblies

The use of remote crevice assemblies is an electrochemical approach that allows for the evaluation of both the initiation and propagation phases of crevice corrosion.[8,59–61] This approach entails the physical separation of the anode (crevice) and the cathode (boldly exposed surface) which are electrically connected via a zero-resistance ammeter; this, therefore, permits the measurement of the current flow between the anode and the cathode. Plotting the current versus the time provides information regarding initiation times, as well as propagation rates (based on the total charge passed), and the reproducibility of results is reportedly good.[9]

2.03.4 Practical Design Aspects

As described above, crevice corrosion occurs in both fabricated and natural crevice regions. In both cases, occluded regions formed by the close contact between a metal surface and a nonmetal crevice former or a second metal surface can lead to accelerated corrosion attack in these regions. In order to prevent crevice corrosion from occurring, occluded regions and sharp inside corners should be avoided, wherever possible.[62,63] When materials are joined by rivets, continuous welding, or soldering, sealants should be used to prevent water uptake into occluded regions. When occluded regions cannot be avoided, decreasing the aspect ratio between the crevice depth and crevice gap can lessen the severity of crevice corrosion damage by reducing the steepness of concentration and potential gradients within the crevice. Metal-to-nonmetal crevices (e.g., gaskets) should be avoided because the crevice gap is typically much less than metal-to-metal crevices; this is because the nonmetal crevice former can conform to the metal surface resulting in tighter crevice geometries.[62,63] Metal-to-nonmetal crevices are typically designed to prevent water uptake; however, there is a fine line between being tight enough to prevent water uptake, and tight enough to facilitate capillary action, leading to severe crevice corrosion. Crevice corrosion can also be mitigated by careful material selection or processing. Materials that have a higher crevice corrosion resistance can be used as an overlay

in areas susceptible to crevice corrosion, to minimize damage. Materials with higher critical crevice temperatures or potentials, or generally higher PREN materials, tend to be more resistant to crevice corrosion. The ranking of crevice corrosion resistance for various stainless steels and nickel based alloys in seawater, along with PREN[64] and CCT[32] values, is given in Table 1.

Materials can also be processed to help avoid crevice corrosion. Metal surfaces outside crevice regions can be painted to prevent cathodic reactions taking place, effectively reducing the area ratio between the anode and cathode regions, thus reducing the rate of crevice corrosion attack.

Additionally, processes can be designed to help prevent crevice corrosion. Natural crevices, such as precipitates and deposits, can be avoided by adding filtration systems to industrial process lines early in the process, to prevent deposits in stagnant regions of the process line.

Solution phase inhibitors or supporting electrolytes can also be added to processing lines to either inhibit corrosion directly, or change solution conductivity to minimize ohmic potential effects, or disrupt ion migration into occluded regions.

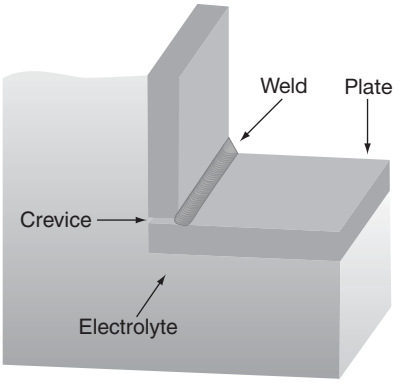


Figure 4 Example of crevice in welded structure.

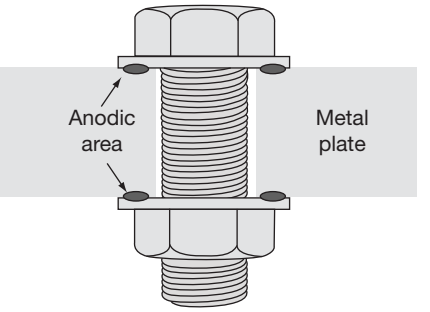


Figure 5 Example of crevice in welded structure.

2.03.5 Industry Specific Examples of Crevice Corrosion

Crevice corrosion is a common form of corrosion that affects a wide range of materials used in almost all industries. Metal-to-metal and metal-to-nonmetal crevices can lead to crevice corrosion in a wide variety of typical designs used for consumer and industrial products. Metal-to-metal crevices typically occur between strands in wire rope, around incomplete fusion welds in pipes and plates (see Figure 4), under o-rings and gaskets, in tube-to-tube sheet connections, in threaded connectors, and in a wide variety of other industrially-used geometries (see Figure 5).[22] Nonmetal-to-metal crevices can occur as a result of o-ring seals, gasket-to-flange joints, tape-coatings, barnacles on metallic surfaces, valve steam packing, dirt, sand, or mud deposits. This section will review how crevice corrosion affects material selection and mechanical design in a few exemplar industries.

2.03.5.1 Medical Devices and Implants

Implantable medical devices are often used in situations where crevices are likely or unavoidable; for instance, where the device is held against tissue or bone, where biofilms develop on the surface, or where two component parts of the device come together, such as screws in bone plates, or the crevice between a tapered stem and ball in a total hip replacement. For this reason, medical devices are required to possess long-term resistance to crevice corrosion during their expected service lifetimes. Common biomedical alloys used for long-term implants are stainless steels, cobalt–chromium–molybdenum alloys (including MP35N and Elgiloy), and titanium alloys. These alloys get their corrosion resistance from the presence of a thin passive film that is generally resistant to uniform corrosion, as well as pitting and crevice corrosion in the warm saline environment of the human body. The pH in the body is maintained at a relatively constant value of 7.4 by phosphate and carbonate buffering, under normal conditions. After surgery, however, the pH at the implant site may range from 5.5 to 9.0, as a result of immunological responses or infections.[65] In the case of a hematoma, the pH can remain at 5.5 for several days.[66] When implant crevice corrosion occurs, the pH may drop to 1 or even lower.[67]

The localized-corrosion resistance of long-term implants is important not only with respect to the maintenance of the mechanical and structural integrity of the implant, but also with respect to the rate at which leachable substances, such as nickel and chromium ions, are released into the body. The release of leachable substances, albeit at very low rates, can elicit an immunological response. This response can vary from person to person and can determine the difference between success and failure of the implant. An example is the corrosion and wear around orthopedic and dental implants that can cause metallosis of the surrounding tissue, and loosening of the implant at the fixation site.[68–71] Although commercial metallic biomaterials are generally corrosion-resistant, there has always been a concern that trace metal ions may be released into the human body because of pitting or crevice corrosion.

Titanium alloys, such as Ti Gr2 and Ti–6Al–4V, are generally ranked as having the greatest resistance to corrosion in saline solutions and are considered resistant to pitting up to the maximum normally-investigated potentials used for testing biomedical implant materials (i.e., 0.8 V SCE).[51] Titanium and its alloys typically show either negative hysteresis or no hysteresis in potentiodynamic polarization tests (i.e., current density falling off as rapidly or more rapidly on the reverse scan than on the forward scan), indicating that passive film growth is enhanced with increasing potential. Consequently, materials exhibiting no hysteresis or negative hysteresis are considered immune to pitting and crevice corrosion.[51,65] Nevertheless, titanium alloys can, under certain *in vivo* conditions such as fretting, show signs of pitting and crevice corrosion as discussed below.[72]

The corrosion resistance of Nitinol (or NiTi), which has an approximately equal nickel and titanium content, is not considered as great as that of titanium or other titanium alloys, such as Ti Gr2 or Ti–6Al–4V. As it exhibits superelastic and shape memory behavior, however, Nitinol is used extensively as an alloy for implantable medical devices, particularly for devices that are deployed endovascularly. There have been some reports that Nitinol will experience *in vivo* breakdown of the passive film, and subsequent localized corrosion. These reports of pitting or crevice corrosion observed on explants,[73–76] however, appear to have been a result of the explants being exposed to decontamination solutions, such as bleach, rather than as a result of actual *in vivo* corrosion.[77–79] Jacobs *et al.* examined Nitinol explants that had been implanted for over 10 years, and that had not been decontaminated with bleach, and found no significant pitting or crevice corrosion.[80] Guidoin *et al.* examined Nitinol stents (with a radiopaque platinum coil wrapped around the Nitinol wire to make the device more visible during X-ray radiography) which they disinfected with bleach, and found little corrosion, except at the crevice between the coil and the Nitinol wire.[73]

The ASTM F 75 specification cobalt–chromium–molybdenum alloys, such as MP35N and Elgiloy, do not exhibit a hysteresis loop in their cyclic polarization curves and are generally thought to be resistant to pitting and crevice corrosion in physiological environments. These alloys exhibit electrochemical behavior in the passive range that differs from that of other biomedical alloys. In particular, their cyclic potentiodynamic polarization curves typically exhibit an increase in current at potentials of about 0.4 V SCE, and perhaps another at about 0.7 V SCE, but they do not show evidence of localized corrosion.[81] The first increase is thought to be associated with a solid-state oxidation reaction involving the conversion of Co(II) to Co(III). The second increase appears to be associated with transpassive dissolution that results in the release of chromium and nickel ions. The transpassive dissolution observed on cobalt–chromium–molybdenum alloys is a general dissolution type of corrosion and does not generally result in crevice or pitting corrosion. These alloys have exhibited pitting and crevice corrosion under certain *in vivo* conditions, however, and these will be discussed below.

Although 316L and ASTM F 138 (ISO 5832-1) specification stainless steels are alloys used for medical implants, they are susceptible to crevice corrosion in human body fluids.[23] Such crevice corrosion can release metallic ions into the tissues surrounding the implants, causing local irritations or systemic effects. The potential for crevice corrosion of 316L stainless steel *in vivo* is not surprising, since 316L has a CCT below body temperature (see **Table 1**).

According to Bai and Gilbert,[65] researchers found that 91% of 316 stainless steel multicomponent explants had undergone visible corrosion, and that crevice corrosion was the primary type of attack.[82] Bai and Gilbert[65] also report the findings of a retrieval study[83] in which it was found that 89% of the plates and 88% of the screws of 250 stainless steel internal fixation devices showed evidence of pitting and crevice corrosion. The environment within the crevices formed between plates and screws of orthopedic devices can be altered from that of the normal *in vivo* environment, such that values ranging from

pH 5 to 7, less than 8–110 mmHg for pO_2, and less than 10–300 mmHg for pCO_2 can be formed.[84]

Some of this corrosion is a reflection of the quality of the stainless steel used in those pioneering days; it had higher carbon and inclusion content than the medical-grade materials used nowadays. Since that time, there have been significant changes in the composition specified for medical grade 316L (ASTM F 138), which now has a specified minimum PREN of 26.[85–87] The first improvement came with the development of AISI 316L, which is still widely used for plates, screws, and nails in orthopedic surgery. As there were continuing reports of *in vivo* stainless steel corrosion, a new type of stainless steel was introduced in the mid-1980s. This alloy has higher chromium and manganese than 316L, and 0.5% nitrogen is added for increased corrosion resistance.[87]

In 1972, Cohen and Wulff described a cobalt-alloy plate and screws that had been removed from a patient; the persistent pain that the patient experienced appeared to be associated with corrosion.[88] The plates and screws were found to have failed because of crevice corrosion between the plate and bone. The plate composition was similar to Stellite 25 (Co–20Cr–15W–10Ni) and the screw composition was similar to Stellite 21 (Co–27Cr–5.5Mo). Cohen and Wulff noted that the corrosion was associated with the plate and not the screws. Reports of corrosion problems and adverse tissue response with Stellite alloys, however, have been extremely rare.[89]

Brown *et al.* discussed crevice corrosion in stainless steel bone plates and screws[69] and modular hip implants.[68] Crevice corrosion tests with bone plates and screws exposed for up to 1 year showed that crevice corrosion does not occur in isotonic saline solutions (0.9% NaCl) for periods of up to 1 year, but can occur in hypertonic saline solutions (7.2% NaCl). Brown *et al.* also observed corrosion of bone plates and screws that had been fixed to the tibia of sheep, but concluded that this corrosion, as well as the corrosion reported on clinically retrieved implants, was initiated by fretting corrosion, which predisposes the contact area to crevice corrosion attack.

Multiple-component systems used in orthopedic surgery provide flexibility in choosing the optimal implant, but introduce the possibility of crevice corrosion that could limit the implant's longevity, either because of unfavorable tissue response to corrosion products or device failure.[68] Based on their high *in vivo* corrosion resistance, both titanium and cobalt–chromium alloys are often used for total joint replacements.[90] Because these joints are articulating, the wear resistance of the alloys is also important. Compared with cobalt–chromium alloys, titanium alloys have relatively poor wear properties, and therefore, the cobalt-base alloys are more commonly used for the articulating components of total joint replacements.[90] Since the 1980s, the modular joint, in which the femoral-head component is fabricated from a cobalt-base alloy and the stem from extra-low interstitial (ELI) Ti–6Al–4V alloy, has gained popularity, particularly for the artificial hip. The galvanic couple created by this arrangement was not considered a problem, as both alloys were passive. By the early 1990s, however, there had been reports of corrosion observed on these implants and Brown *et al.*[68–69] attributed these instances to a fretting-assisted crevice-corrosion mechanism.

The work by Brown *et al.*, in 1995 on 79 retrieved tapered modular hip implants indicated similar behavior, that is, the effect of fretting on crevice corrosion initiation. This work suggested that design changes that decreased the propensity for fretting also reduced the chance of crevice corrosion.[68] In addition, Gilbert *et al.* examined 148 total modular hip implants with mixed (Ti–Al–4V)/Co–Cr and Co–Cr/Co–Cr stem and head combinations,[72] and reported that both combinations showed about the same amount of corrosion in the creviced region between the tapered stem and head. They concluded that galvanic effects were not significant, but that crevice corrosion could occur in cobalt-chromium alloys in fretting and wear situations. Cobalt–chromium–molybdenum alloys, therefore, appear to be resistant to pitting and crevice corrosion in the unloaded condition, but are susceptible to deterioration in the presence of mechanical loads. This form of crevice corrosion has been described by Gilbert and Jacobs as mechanically-assisted crevice corrosion.[67]

As reported by Thomas *et al.*, cemented titanium stems in hip arthroplasty have experienced proximal cement–stem debonding, which can result in a crevice, and early failure.[91] This was well publicized with the 3M Capital hip implant failures. Thomas *et al.* discussed their examination of a series of 12 cemented titanium Furlong Straight Stems that required revision at a mean implant time of 78 months, as a result of thigh pain. They reported that the stems were severely corroded in a pattern that was typical of crevice corrosion. Their conclusion was that the combination of a titanium stem and cement appears to facilitate fretting-induced crevice corrosion for an otherwise crevice corrosion resistant material.

Crevice corrosion has also been found on tibial plateau leveling osteotomy (TPLO) plates and associated bone screws that had been implanted in dogs for ~3 years.[92] It was reported that surface irregularities and porosity of TPLO plates from the casting process provided crevice corrosion initiation sites in the occluded region between the plates and screws.

Crevice corrosion has also been seen on spinal surgery hardware. Tezer *et al.* noted that crevice and fretting corrosion can occur at the junctions of rod-screw, rod-hook, transverse-connector rods, and other connector rods in modular spinal implants.[93] They recommended using alloys with better corrosion resistance to inhibit crevice and fretting corrosion. Modular-spine implants are used to aid fusion, but fretting and corrosion can occur between modular components. Kirkpatrick *et al.* examined the corrosion on 47 explanted spinal implants that were manufactured by a variety of companies.[94] They observed that the stainless steel implants exhibited a range of corrosion, from minor to severe damage, that was consistent with previously observed mechanically-assisted crevice corrosion. Titanium-alloy implants, however, showed no significant corrosion. The single cobalt-alloy explant showed no evidence of corrosion.

2.03.5.1.1 Dental crevice corrosion

With dental implants there is the possibility of microbiologically-influenced corrosion (MIC) affecting crevice corrosion in the creviced region between the dental implant and the gingival epithelium.[71] According to Megremis and Carey, seepage of oral secretions into the crevices created between the restoration and the tooth can lead to microorganism invasion, generation of acidic conditions, the operation of differential aeration cells, and crevice corrosion.[95] The creation of crevice conditions in amalgams can increase acidity to well below a pH of 4. To enhance the crevice corrosion resistance of dental surgical materials, the German Federal Health Department recommended minimum amounts of chromium and molybdenum of 20% and 4%, respectively.[95] In 2006, Wylie *et al.* performed electrochemical testing in artificial saliva at different pHs, on two nickel-based dental casting alloys, before and after porcelain-firing heat treatments.[96] Crevice corrosion was observed on alloys exposed to pH 2.5 saliva. They noted, not surprisingly, that the presence of crevices and lower chromium-content nickel-base alloys could lead to crevice corrosion in dental casting alloys, but this could be avoided by an increase in the chromium content of the alloy.

2.03.5.1.2 Nickel-free alloys

Many medical implants contain nickel, which has led to concerns regarding nickel toxicology and its effect on the implants' biological performance.[70,97–98] As a result, there have been several recent efforts to develop nickel-free biomedical alloys.[97] Along these lines, Reclaru *et al.* analyzed the crevice corrosion behavior and the cation release of five nominally nickel-free austenitic steels in artificial sweat and bone plasma.[98] Generally, the amount of cation release was substantially higher in the artificial sweat solution than in the bone plasma under creviced conditions.[99]

2.03.5.2 Chemical and Power Plants

2.03.5.2.1 Boiler and steam generators

It has been estimated that the second most common type of tube failure in fire-tube boilers is concentration-cell corrosion, frequently referred to as crevice corrosion, due to the build-up of iron oxide sediment or scale.[100] On waterside fire-tubes, corrosion occurs beneath deposits that foul the surface and become anodic relative to the periphery of the deposits. According to Esmacher *et al.*,[100] crevice corrosion also occurs at tube-to-tubesheet roll joints, particularly if high thermal stresses cause flexing at the joint, and crack protective oxides or allow water ingress into the crevice.

2.03.5.2.2 Heat exchangers and condensers

Heat exchangers and condensers are widely used in the power generation and chemical processing industry. Almost all heat exchangers have crevices at the tube-sheet and the tube support-plate locations. For example, in the nuclear power industry, crevice corrosion has resulted in costly tube denting (see **Section 2.03.5.2.3**), as well as intergranular attack (IGA), and stress corrosion cracking (SCC) of Inconel 600 (Alloy 600) in tube-sheet and tube support-plate crevices.[101]

Dissimilar metal crevices (DMC) are common features of heat exchangers and condensers. Tube-to-tube sheet joints frequently contain crevices and dissimilar tube and tubesheet materials. Work by Kearns[102] and Maurer[103] on DMC in seawater have shown that the more corrosion resistant materials in a DMC can be made to corrode as a result of the corrosion products from the less resistant alloys.

2.03.5.2.3 Pressurized-water reactors (PWR) and boiling-water reactors (BWR)

Crevices in oxidizing environments, or boiling at heat-transfer surfaces, can result in a variety of crevice corrosion problems for both BWR and PWR

operations. According to Ford *et al.*, in most BWR designs, the coolant purity approaches that of theoretically pure water (i.e., a room-temperature conductivity of 0.055 $\mu S\,cm^{-1}$), with dissolved chloride and sulfate ion concentrations below 5 ppb, in order to help prevent the concentration of impurities in crevices.[104]

One of the main differences between the PWR and BWR designs is that a PWR has a steam generator that isolates the primary water (i.e., water heated by the reactor core under pressure) from the secondary water that is boiled to make the steam that runs the turbines. A serious crevice corrosion problem appeared in PWRs during the mid-1970s.[105] The phenomenon became known as denting, and it occurred in Inconel 600 tubing because of corrosion in the crevices between the tubes and the carbon steel support plates or tube sheets. In the crevices between the Inconel 600 steam generator tubes and the steel tube-sheets, tube support-plates, and sludge deposits, where water circulation is restricted by geometry and deposits, boiling can concentrate the very low levels of impurities in the secondary water to very high concentrations. This is known as hideout. The concentration of impurities in crevices, under heat transfer conditions, can be 20 000 times that present in the bulk secondary water.[105,106] Denting can restrict primary water flow and result in SCC and IGA, which may result in primary water leaking into the secondary side. Tube damage from hideout, such as wastage, pitting, and denting has been largely eliminated by appropriate management of secondary water chemistry, which includes reducing condenser impurity in-leakage, reducing air in-leakage, using make-up water of higher purity, using boric acid to inhibit acidic chloride attack, and using condensate polishers to purify the feedwater. Newer steam generator designs have used Inconel 690 (Alloy 690) or Incoloy 800 (Alloy 800) that are more corrosion resistant to the highly concentrated environments that can develop in the tube-sheet and tube support-plate crevices, and under sludge piles.[105]

2.03.5.3 Aerospace

Most commercial aircraft utilize aluminum alloys for their airframe structures, which are susceptible to pitting, crevice corrosion, exfoliation, and environment-assisted cracking under normal operating conditions.[107–109] Crevice corrosion on aircraft can occur at unsealed joints or at joints where the fillet seal has broken, on fuselage structures held in intimate contact with wet insulation blankets, and on aluminum and stainless steel tubing under rubber clamps.[107] As described by Inman *et al.*, crevice corrosion and subsequent fatigue cracking of airframe structures can occur and lead to loss of life. They list a number of examples whereby crevice corrosion along riveted structures, which are difficult to inspect, led to fatigue cracking and subsequent failure of sections of the fuselage.[108] For instance, in 1981, the failure of a large portion of the fuselage of a commercial aircraft in Taiwan was attributed to undetected fuselage-skin corrosion that led to fatigue cracking. Another incident in Hawaii in 1988 revealed that fatigue cracking can occur in corroded fuselage structures.[108] Due to its nature, crevice corrosion in aircraft can be difficult to detect until it exceeds the limits allowed by design. As described by Inman *et al.*, moisture is drawn into crevices during normal service, and pollutant gases can dissolve in the water to form corrosive solutions which attack the protective oxide layer. Aircraft operating in marine environments have to contend with additional chloride ion ingress. Also, routine maintenance and cleaning procedures, such as washing and cleaning, can contribute to crevice corrosion.

According to Sankaran *et al.*, crevice corrosion can lead to severe intergranular and exfoliation corrosion, which, in turn, can lead to increased stress and crack growth rates due to the volumetric increase of corrosion products within creviced regions.[110] Faying-surface sealants, such as polysulfide sealants, have dramatically reduced the crevice corrosion of fuselage lap and similar aircraft joints.[107]

2.03.5.4 Nuclear Waste Isolation

High level nuclear waste, such as that resulting from spent nuclear fuel from commercial nuclear reactors used for power generation, needs to be isolated from the environment for 1000–10 000 years or longer.[111,112] Corrosion of the containment materials is considered to be the principal mode of degradation that will limit the lifetime of the waste package.[113] As crevice corrosion is thought to be a part of the limiting lifetime of these containment systems, much of the research on crevice corrosion in the last ten years has been devoted to a better understanding of crevice corrosion mechanisms in this highly engineered containment system.[112] With the proposed repository designs, crevices could be formed between the waste package and the emplacement pallet, by the deposition of mineral precipitates, corrosion products, dust, or by contact with rocks, and the drip shield or ground support materials.[113]

A common approach taken by many nations is to bury the waste underground in corrosion resistant containers. Early work at Yucca Mountain in Nevada, a potential repository site in the United States, initially focused on austenitic stainless steels as the container material[112]; however, this material selection has evolved over the years as the design requirements have increased. Originally, 304 stainless steel was specified as the container material, but as the corrosion resistance requirements increased, first Alloy 825, then Alloy 625, and finally Alloy 22, were selected. Alloy 22, a Ni–22Cr–13Mo–4Fe–3W alloy, was selected in order to have the highest resistance to localized corrosion phenomena, such as stress corrosion cracking, pitting, and crevice corrosion.[112–115] Since 1998, the waste-package design for the high-level waste repository at Yucca Mountain has specified a double-walled cylindrical container with the outer container made from Alloy 22, to resist corrosion, and an inner container made of Type 316 nuclear-grade stainless steel to provide strength and radiation shielding.[112,113] The container will be covered by a Ti Gr7 drip shield to guard against rock fall and prevent contact with water seepage during the first few hundred years.[112] Alloy 22 is extremely resistant to localized corrosion as a result of the formation of a 5–6-nm thick, chromium-rich oxide film. The general corrosion rate of Alloy 22 in acidic to alkaline solutions is expected to be well below 0.1 mm year^{-1} (0.004 mil year^{-1}), and therefore, the life of the containers for Yucca Mountain should not be limited by passive dissolution.[112]

Although the United States is considering burial in unsaturated formations (i.e., above the water table), and will therefore have to consider oxidizing redox potentials, it is unlikely that pitting corrosion will be life limiting, as it is unlikely that in this environment (hot concentrated groundwater), a sufficiently high electrochemical potential would be developed to initiate pits on the exposed surface of Alloy 22.[112]

As described by Rebak and McCright, E_{crit} is considered the potential above which localized corrosion could occur. The margin of safety against localized corrosion, therefore, will be the value of ΔE, where ΔE is the difference between E_{crit} and the uniform or passive corrosion rate, E_{corr} (i.e., $\Delta E = E_{crit} - E_{corr}$).[112] As the value of ΔE increases, so does the margin of safety for localized corrosion. E_{corr} and E_{crit} both depend on the surface condition of the alloy, the environment, and the ambient temperature. Critical temperatures for pitting and crevice corrosion have been measured, and among the highest are those for nickel alloys.

Reassuringly, He and Dunn's work on crevice corrosion propagation indicate that although crevice corrosion of Alloy 22 is possible under some environmental conditions which may develop on the repository waste containers, the maximum penetration depth of localized attack is significantly less than the outer-container wall thickness, even under the aggressive environmental conditions used in their laboratory tests.[113]

2.03.5.5 Underground Structures

Most buried or submerged steel structures, such as pipelines and underground storage tanks, are painted or coated and cathodically protected to protect them from corrosion. If the coating becomes damaged or blistered, crevices can form underneath the protective coating and lead to crevice corrosion.[116–118] To compound the problem, cathodic protection may not be able to penetrate into the crevices formed by disbonded coatings. This so-called shielding is believed to be more severe for certain types of coating (e.g., polyolefin tape) and low conductivity groundwater.[119] Li *et al.*[118] have shown that the steel surface in freshly formed crevices under cathodic protection in dilute chloride environments may not immediately achieve the required protection potential of less than −0.850 V ($Cu/CuSO_4$). They have shown, however, that the crevice will become more alkaline as the cathodic reaction consumes the oxygen and forms hydroxide. The hydroxide ions then attract chloride ions into the crevice to maintain charge balance, which, in turn, increases the conductivity within the crevice, allowing protection potentials to be achieved.

It is important to realize that the foregoing description of corrosion within crevices as a result of damaged coatings should not be taken to imply either that carbon steels are susceptible to conventional crevice corrosion in neutral groundwaters or similar near-neutral solutions, or that crevice corrosion can occur on a steel structure that is cathodically protected. Under these conditions, corrosion does not occur by a conventional crevice corrosion mechanism, as the external potential is more negative than the corrosion potential, and therefore, the anodic reaction under the coating cannot be balanced by the external cathodic reaction. Instead, the corrosion derives from the fact that polymeric coatings are much more permeable to oxygen than to ions, and hence oxygen can penetrate the coating while the cathodic protection current cannot. As a result, the steel under the coating corrodes by a localized or general corrosion

mechanism, controlled by the rate of oxygen permeation through the coating. Similarly, a metal–metal crevice in a carbon steel structure will only suffer from crevice corrosion by the conventional mechanism when the external surface is passive, for example, in alkaline solutions. In neutral or acid solutions, such as acidic groundwaters and seawater, where the steel undergoes active corrosion, some iron dissolution will initially occur in the crevice, but this will quickly raise the ferrous ion concentration to such a level that either the Fe/Fe^{2+} reaction reaches equilibrium or the precipitation of ferrous oxides or hydroxides blocks the crevice and stifles the corrosion. When the external surface is subject to cathodic protection this simply reinforces this tendency, and the steel in the crevice will normally passivate.

Fessler *et al.*[117] noted that cathodic protection potentials should not be decreased to compensate for crevices, as potentials that are too low will form hydrogen bubbles in the crevice, which, in turn, will block the protection current. They suggested that the pipe-to-soil potentials should be maintained roughly between -0.850 and $-1.10\,V$ ($Cu/CuSO_4$) to avoid this problem.

References

1. Kelly, R. G. In *Encyclopedia of Electrochemistry: Corrosion and Oxide Films*; Stratmann, M., Frankel, G. S., Eds.; Wiley-VCH: Weinheim, 2003; Vol. 4, Chapter 4.1, pp 275–307.
2. Szklarska-Smialowska, Z. *Pitting and Crevice Corrosion*; NACE International: Houston, TX, 2005.
3. Mears, R. B.; Evans, U. R. *Trans. Faraday Soc.* **1934**, *30*, 417–423.
4. Uhlig, H. H. *Trans. of AIME* **1940**, *140*, 387–432.
5. Oldfield, J. W.; Sutton, W. H. *Br. Corros. J.* **1978**, *13*(1), 13–22.
6. Oldfield, J. W.; Sutton, W. H. *Br. Corros. J.* **1978**, *13*(3), 104–111.
7. Crolet, J. L.; Defranoux, J. M. *Corros. Sci.* **1973**, *13*(7), 575–585.
8. Kain, R. M. *Mater. Perform.* **1984**, *23*(2), 24–30.
9. Kain, R. M. In *ASM Handbook; Corrosion* ASM International: Metals Park, OH, 1987; Vol. 13, p 405.
10. Oldfield, J. W. *Int. Mater. Rev.* **1987**, *32*(3), 153–170.
11. Pickering, H. W. In *Corrosion and Corrosion Protection*; Frankenthal, R. P., Mansfeld, F., Eds.; The Electrochemical Society: Pennington, 1981; p 85.
12. Pickering, H. W. *Corros. Sci.* **1989**, *29*(2/3), 325–341.
13. Lillard, R. S.; Jurinski, M. P.; Scully, J. R. *Corrosion* **1994**, *50*(4), 251–265.
14. Shaw, B. A.; Moran, P. J.; Gartland, P. O. *Corros. Sci.* **1991**, *32*(7), 707–719.
15. Kelly, R. G. In *Critical Factors and Challenges in Crevice Corrosion*; CORROSION/01 Research Topical Symposium, 2001; Frankel, G. S., Scully, J. R., Eds.; NACE International: Houston, TX, 2001; pp 128–156.
16. Kelly, R. G.; Stewart, J. In *Passivity of Metals and Semiconductors*; Ives, M. B., MacDougall, B. R., Bardwell, J. A., Eds.; The Electrochemical Society: Pennington, 1999; Vol. VIII, pp 546–554.
17. Szklarska-Smialowska, Z.; Mankowski, J. *Corros. Sci.* **1978**, *18*, 953–960.
18. Stockert, L.; Boehni, H. *Mater. Sci. Forum* **1989**, 44/45.
19. Laycock, N. J.; Stewart, J.; Newman, R. C. *Corros. Sci* **1997**, *39*(10–11), 1791–1809.
20. Kehler, B. A.; Ilevbare, G. O.; Scully, J. R. In *Crevice Corrosion Behavior of Ni-Cr-Mo Alloys: Comparison of Alloys 625 and 22*; CORROSION/01 Research Topical Symposium, 2001; Frankel, G. S., Scully, J. R., Eds.; NACE International: Houston, RX, 2001; pp 30–64.
21. Abdulsalam, M. I.; Pickering, H. W. In *The Roles of Aspect Ratio and Oxidizing Potential in Crevice Corrosion*; Critical Factors in Localized Corrosion III, 1998; Kelly, R. G., Frankel, G. S., Natishan, P. M., Newman, R. C., Eds.; The Electrochemical Society: Pennington, 1998; pp 395–408.
22. Sedriks, A. J. In *Overview of Crevice Corrosion Issues for Chromium-Containing Alloys*, CORROSION/96 Research Topical Symposia, 1996; Ives, M. B., Oldfield, J. W., Eds.; NACE International: Houston, TX, 1996; pp 279–310.
23. Sedriks, A. J. *Corrosion of Stainless Steels*, 2nd ed.; Wiley: New York, 1996.
24. Combrade, P. In *Corrosion Mechanisms in Theory and Practice*; Marcus, P., Ed.; Marcel Dekker: New York, 2002; pp 349–397.
25. Brossia, C. S. In *Effects of MnS Inclusions on Crevice Corrosion of Stainless Steels in Neutral Chloride Solutions*, Critical Factors in Localized Corrosion III; Kelly, R. G., Frankel, G. S., Natishan, P. M., Newman, R. C., Eds.; The Electrochemical Society: Pennington, 1998; pp 326–338.
26. Brossia, C. S.; Kelly, R. G. In *On the Role of Alloy Sulfur in the Initiation of Crevice Corrosion in Stainless Steel*, Critical Factors in Localized Corrosion II; Natishan, P. M., Kelly, R. G., Frankel, G. S., Newman, R. C., Eds.; The Electrochemical Society: Pennington, 1995; pp 201–217.
27. Brossia, C. S.; Kelly, R. G. *Corros. Sci.* **1998**, *40*(11), 1851–1871.
28. Brossia, C. S.; Kelly, R. G. *Corrosion* **1998**, *54*, 145–154.
29. Szklarska-Smialowska, Z. In *Sulfide Inclusions in Steel*; deBarbadillo, J. J., Snape, E., Eds.; ASM International: Metals Park, OH, 1975; p 380.
30. Szklarska-Smialowska, Z.; Mankowski, J. *Corros. Sci.* **1978**, *18*, 953–960.
31. Ijsseling, F. P. *Survey of Literature on Crevice Corrosion (1979–1998)*; European Federation of Corrosion Publications, 2000; p 30.
32. Roberge, P. R. *Handbook of Corrosion Engineering*; McGraw-Hill: New York, 1999; p 734.
33. Grubb, J. F.; DeBold, T.; Fritz, J. D. In *ASM Handbook; Corrosion: Materials* ASM International: Metals Park, OH, 2005; Vol. 13B, p 54.
34. Renner, M.; Heuber, U.; Rockel, M. B.; Wallis, D. *Werkst. Korros.* **1986**, *37*, 183.
35. Noel, J. J.; Shoesmith, D. W.; Ikeda, B. M. In *Crevice Corrosion of Alpha Titanium Alloys*, CORROSION/01 Research Topical Symposium, 2001; Frankel, G. S., Scully, J. R., Eds.; NACE International: Houston, TX, 2001; pp 65–102.
36. Kain, R. M. In *ASM Handbook; Corrosion: Fundamentals, Testing, and Protection* ASM International: Metals Park, OH, 2003; Vol. 13A, p 549.
37. Kain, R. M. In *Evaluating Crevice Corrosion Susceptibility*, CORROSION/96 Research Topical Symposium; Ives, M. B., Oldfield, J. W., Eds.; NACE International: Houston, TX, 1996; p 401.

38. Kearns, J. R. In *Corrosion Tests and Standards: Applications and Interpretation*; Baboian, R., Dean, J., S. W., Hack, H. P., Haynes, G. S., Sprowls, D. O., Eds.; ASM International: Metals Park, OH, 1995; Chapter 19, p 175.
39. *ASTM, G 4–01: Standard Guide for Conducting Corrosion Tests in Field Applications*; ASTM International: West Conshohocken, PA, 2001.
40. Tuthill, A. H. *Mater. Perform.* **1985**, *24*(9), 43.
41. *ASTM, G 48–03: Standard Test Methods for Pitting and Crevice Corrosion Resistance of Stainless Steels and Related Alloys by Use of Ferric Chloride Solution*; ASTM International: West Conshohocken, PA, 2003.
42. Streicher, M. A. *Corrosion* **1974**, *30*(3), 77–91.
43. Treseder, R. S. *Method MTI-2 for Laboratory Testing of Wrought Iron- and Nickel-Base Alloys for Relative Resistance to Crevice Corrosion in a Standard Ferric Chloride Solution*; The Materials Technology Institute of the Chemical Process Industries, 1980; pp 16–27.
44. Turnbull, A. *Corros. Sci.* **1983**, *23*(8), 833–870.
45. Garner, A. *Corrosion* **1981**, *37*(3), 178.
46. Tsujikawa, S. *A New Test for Predicting Pitting Corrosion Resistance of CRAs in Sour Environments*, CORROSION/88; Paper no. 64; NACE International: St. Louis, MO, 1988.
47. Wensley, A.; Reid, C.; Brown, D.; Christi, D. In *Corrosion of Stainless Alloys in a Chlorine Dioxide Bleached Pulp Washer*; Proceedings of Technical Association of the Pulp and Paper Industry Engineering Conference, 1991 TAPPI Press: Atlanta, GA, 1991; pp 499–503.
48. *ASTM, G 78–01: Standard Guide for Crevice Corrosion Testing of Iron-Base and Nickel-Base Stainless Alloys in Seawater and Other Chloride-Containing Aqueous Environments*; ASTM International: West Conshohocken, PA, 2001.
49. *ASTM, F 746–87 (Reapproved 99): Standard Test Method for Pitting or Crevice Corrosion of Metallic Surgical Implant Materials*; ASTM International: West Conshohocken, PA, 1999.
50. Syrett, B. C. *Corrosion* **1977**, *33*(6), 221.
51. *ASTM, F 2129–06: Standard Test Method for Conducting Cyclic Potentiodynamic Polarization Measurements to Determine the Corrosion Susceptibility of Small Implant Devices*; ASTM International: West Conshohocken, PA, 2006.
52. Frankel, G. S. In *ASM Handbook; Corrosion: Fundamentals, Testing, and Protection* ASM International: Metals Park, OH, 2003; Vol. 13A, p 236.
53. Leclerc, M. F. In *Corrosion, Metal/Environment Reactions*; Shreir, L. L., Jarman, R. A., Burstein, G. T., Eds.; Butterworth Heinemann: Oxford, 1994; Vol. 1, p 2:164.
54. Pound, B. G. *J. Biomed. Mater. Res.* **2006**, *77A*, 185–191.
55. Silverman, D. C. In *Tutorial on Cyclic Potentiodynamic Polarization Technique*, CORROSION/98; NACE International: San Diego, CA, 1998; Paper no. 299.
56. Wilde, B. E.; Williams, D. E. *Electrochim. Acta* **1971**, *16*, 1971–1985.
57. Wilde, B. E. *Corrosion* **1972**, *28*(8), 283–291.
58. *ASTM G 61–86 (Reapproved 2003): Standard Test Method for Conducting Cyclic Potentiodynamic Polarization Measurements for Localized Corrosion Susceptibility of Iron-, Nickel-, or Cobalt-Based Alloys*; ASTM International: West Conshohocken, PA, 2003.
59. Kain, R. M.; Klein, P. A. In *Crevice Corrosion Propagation Studies for Alloy N06625: Remote Crevice Assembly Testing in Flowing Natural and Chlorinated Seawater*, NACE Corrosion/90: NACE International: Las Vegas, NV, 1990; Paper no. 158.
60. Lee, T. S. In *Electrochemical Corrosion Testing, ASTM STP 727*; Mansfeld, F., Bertocci, U., Eds.; ASTM International: West Conshohocken, PA, 1981; p 43.
61. Lee, T. S. In *Electrochemical Techniques for Corrosion Engineering*; NACE International: Houston, TX, 1986.
62. Fontana, M. G.; Greene, N. D. *Corrosion Engineering*; McGraw-Hill: New York, 1967.
63. Uhlig, H. H. *Corrosion and Corrosion Control*; John Wiley: New York, 1971.
64. Bond, A. P.; Dundas, A. J. *Mater. Perform.* **1984**, *23*(7), 39.
65. Bai, Z.; Gilbert, J. L. In *ASM Handbook; Corrosion: Environments and Industries*, ASM International: Materials Park, OH, 2006; Vol. 13C, pp 837–852.
66. Laing, P. G. In *Corrosion and Degradation of Implant Materials, ASTM STP 684*; ASTM International: West Conshohocken, PA, 1979; pp 199–211.
67. Gilbert, J. L.; Jacobs, J. J. In *Modularity of Orthopedic Implants, ASTM STP 1301*; ASTM International: West Conshohocken, PA, 1997; pp 45–59.
68. Brown, S. A.; Flemming, C. A.; Kawalec, J. S.; Placko, H. E.; Vassaux, C.; Merritt, K.; Payer, J. H.; Kraay, M. J. *J. Appl. Biomater.* **1995**, *6*(1), 19–26.
69. Brown, S. A.; Simpson, J. P. *J Biomed. Mater. Res.* **1981**, *15*(1), 867–878.
70. Eiselstein, L. E.; Proctor, D. M.; Flowers, T. C. *Mater. Sci. Forum* **2007**, 539–543, 698–703.
71. Hanawa, T. *Mater. Sci. Forum* **2006**, *512*, 243–248.
72. Gilbert, J. L.; Buckley, C. A.; Jacobs, J. J. *J. Biomed. Mater. Res.* **1993**, *27*(12), 1533–1544.
73. Guidoin, R.; *et al. J. Endovasc. Ther.* **2000**, *7*(2), 105–122.
74. Guidoin, R.; Zhang, Z.; Dionne, G.; Douville, Y.; King, M.; Legrand, A.-P.; Doppelt, P. In *Corrosion of the Nitinol Wire of Endovascular Prostheses: Does Nickel Ion Release Impair the Devices Performance?* Materials and Processes for Medical Devices Conference ASM International: St. Paul, MN, 2005; pp 284–289.
75. Heintz, C.; Riepe, G.; Birken, L.; Kaiser, E.; Chakfé, N.; Morlock, M.; Delling, G.; Imig., H. *J. Endovasc. Ther.* **2001**, *8*(3), 248–253.
76. Riepe, G.; Heintz, C.; Kaiser, E.; Chakfé, N.; Morlock, M.; Delling, G.; Imig., H. *Eur. J. Vasc. Endovasc. Surg.* **2002**, *24*(2), 117–122.
77. Lasley, C. C.; Mitchell, M. R. In *The Corrosion of Nitinol by Exposure to Decontamination Solutions*, The International Conference on Shape Memory and Superelastic Technologies, 2003; ASM International: Monterey, CA, 2003; pp 375–384.
78. Walak, S. *Analysis of Nitinol Stent-Grafts After Long-Term In-Vivo Exposure*, Materials and Processes for Medical Devices Conference; ASM International: St. Paul, MN, 2005; pp 290–294.
79. Personal communication. Warner, C. P. to Eiselstein, L. E. Menlo Park, CA, 2007.
80. Jacobs, T. S.; Won, J.; Gravereaux, E. C.; Faries, P. L.; Morrissey, N.; Teodorescu, V. J.; Hollier, L. H.; Marin, M. L. *J. Vasc. Surg.* **2003**, *37*(1), 16–26.
81. Pound, B. G. In *Electrochemical Behavior of Cobalt–Chromium Alloys*; Materials and Processes for Medical Devices Conference, 2007 ASM International: Desert Springs, CA, 2007.
82. Colangelo, V. J.; Greene, N. D. *J. Biomed. Mater. Res.* **1969**, *3*(2), 247–265.
83. Cook, S. D.; Thomas, K. A.; Harding, A. F.; Collins, C. L.; Haddad, R. J., Jr.; Milicic, M.; Fischer, W. L. *Biomaterials* **1987**, *8*(3), 177–184.

84. Cahoon, J. R.; Hill, L. D. *J. Biomed. Mater. Res.* **1978**, *12* (6), 805–821.
85. Zardiackas, L. D.; Roach, M.; Williamson, S.; Bogan, J. A. In *Stainless Steels for Medical and Surgical Applications, ASTM STP 1438*; Winters, G. L., Nutt, M. J., Eds.; ASTM International: West Conshohocken, PA, 2003; pp 194–207.
86. Windler, M.; Steger, R. In *Stainless Steels for Medical and Surgical Applications, ASTM STP 1438*; Winters, G. L., Nutt, M. J., Eds.; ASTM International: West Conshohocken, PA, 2003; pp 39–49.
87. Windler, M.; Steger, R.; Winters, G. L. In *Stainless Steels for Medical and Surgical Applications, ASTM STP 1438*; Winters, G. L., Nutt, M. J., Eds.; ASTM International: West Conshohocken, PA, 2003; pp 72–80.
88. Cohen, J.; Wulff, J. *J. Bone Joint Surg. Am.* **1972**, *54*(3), 617–628.
89. John, K. R. In *ASM Handbook; Corrosion: Environments and Industries* ASM International: Materials Park, OH, 2006; Vol. 13C, pp 820–825.
90. Bundy, K. J.; Zardiackas, L. D. In *ASM Handbook; Corrosion: Environments and Industries* ASM International: Materials Park, OH, 2006; Vol. 13C, pp 853–890.
91. Thomas, S. R.; Shukla, D.; Latham, P. D. *J. Bone Joint Surg. Br.* **2004**, *86*(7), 974–978.
92. Charles, A. E.; Ness, M. G. *Vet. Surg.* **2006**, *35*(5), 438–444.
93. Tezer, M.; Kuzgun, U.; Hamzaoglu, A.; Ozturk, C.; Kabukcuoglu, F.; Sirvanci, M. *Arch. Orthop. Trauma Surg.* **2005**, *125*(6), 417–421.
94. Kirkpatrick, J. S.; Venugopalan, R.; Beck, P.; Lemons, J. *J. Spinal Disord. Tech.* **2005**, *18*(3), 247–251.
95. Megremis, S.; Carey, C. M. In *ASM Handbook; Corrosion: Environments and Industries* ASM International: Materials Park, OH, 2006; Vol. 13C, pp 891–921.
96. Wylie, C. M.; Shelton, R. M.; Fleming, G. J.; Davenport, A. J. *Dent. Mater.* **2007**, *23*(6), 714–723.
97. Ossa, C. P.; Rogero, S. O.; Tschiptschin, A. P. *J. Mater. Sci. Mater. Med.* **2006**, *17*(11), 1095–1100.
98. Reclaru, L.; Ziegenhagen, R.; Eschler, P. Y.; Blatter, A.; Lemaitre, J. *Acta Biomater.* **2006**, *2*(4), 433–444.
99. Schutz, R.; Thomas, D. In *Metals Handbook; Corrosion*, 9th ed.; ASM International: Metals Park, OH, 1987; Vol. 13, p 669.
100. Esmacher, M. J.; Gheysen, G.; Rossi, A.; Farraro, J. F.; Sehl, P. In *Firetube Boiler Failures: Review of Case Histories*, CORROSION/05; NACE International: Houston, TX, 2005; Paper no. 433.
101. Cohen, P. Ed. *The ASME Water Handbook on Water Technology For Thermal Power Systems*; AIME: New York, 1989.
102. Kearns, J. R.; Johnson, M. J.; Grubb, J. F. In *Accelerated Corrosion in Dissimilar Metal Crevices*; CORROSION/86; NACE International: Houston, TX, 1986; Paper no. 228.
103. Maurer, J. R. *Mater. Perform.* **1994**, *33*(4), 51–54.
104. Ford, F. P.; Gordon, B. M.; Horn, R. M. In *ASM Handbook; Corrosion: Environments and Industries* ASM International: Materials Park, OH, 2006; Vol. 13C, pp 341–361.
105. Scott, P. M.; Combrade, P. In *ASM Handbook; Corrosion: Environments and Industries* ASM International: Materials Park, OH, 2006; Vol. 13C, pp 362–385.
106. MacDonald, D. D.; Begley, J. A.; Bockris, J. O'M.; Kruger, J.; Mansfeld, J. B.; Rhodes, P. R.; Staehle, R. W. *Mater. Sci. Eng.* **1981**, *50*(1), 19–42.
107. Adjorlolo, A. In *ASM Handbook; Corrosion: Environments and Industries*, ASM International: Materials Park, OH, 2006; Vol. 13C, pp 598–612.
108. Inman, M. E.; Kelly, R. G.; Willard, S. A.; Piascik, R. S. In *Coordinated Metallographic, Chemical, and Electrochemical Analyses of Fuselage Lap Splice Corrosion*, Proceedings of the FAA-NASA Symposium on the Continued Airworthiness of Aircraft Structures, FAA Center of Excellence in Computational Modeling of Aircraft Structures, Atlanta, Georgia, August 28–30, 1997; DOT/FAA/AR-97/2 Office of Aviation Research: Washington, DC, 1997; pp 129–146.
109. Simon, L. B.; Skennerton, G.; Elster, J. L.; Wikswo, J. P.; Kelly, R. G. In *Aging Aircraft 2000*; NASA: St. Louis, 2000.
110. Sankaran, K. K.; Perez, R.; Smith, H. In *ASM Handbook; Corrosion: Environments and Industries*, ASM International: Materials Park, OH, 2006; Vol. 13C, pp 195–204.
111. Kruger, J.; Rhyne, K. *Nucl. Chem. Waste Manage.* **1982**, *3*, 4.
112. Rebak, R. B.; McCright, R. D. In *ASM Handbook; Corrosion: Environments and Industries*; ASM International: Materials Park, OH: 2006; Vol. 13C, p 421.
113. He, X.; Dunn, D. S. *Corrosion* **2007**, *63*, 2.
114. Carranza, R. M. *J. Met.* **2008**, *60*, 1.
115. Carranza, R. M.; Rodri´guez, M. A.; Rebak, R. B. *Corrosion* **2007**, *63*, 5.
116. Beavers, J. A.; Thompson, N. G. *Mater. Perform.* **1997**, *36*(4), 13–19.
117. Fessler, R. R.; Markworth, A. J.; Parkins, R. N. *Corrosion* **1983**, *39*(1), 20–25.
118. Li, Z.; Gan, F.; Mao, X. *Corros. Sci.* **2002**, *44*(4), 689–701.
119. Sridhar, N.; Dunn, D. S. In *Application of a General Reactive Transport Model to Predict Environment Under Disbonded Coatings*, CORROSION/00; NACE International: Houston, TX, 2000; Paper no. 366.

2.04 Pitting Corrosion

M. G. Alvarez
Comisión Nacional de Energía Atómica (CNEA), CAC Dpto. Materiales, Avda. Gral. Paz 1499, (B1650KNA) San Martín, Buenos Aires, Argentina
J. R. Galvele
Instituto Sabato, Universidad Nacional de San Martín, Av. Gral Paz 1499, (B1650KNA) San Martín, Buenos Aires, Argentina

Glossary

Adsorption The surface retention of solid, liquid, or gas molecules, atoms, or ions by a solid or liquid.

Critical pitting temperature The lowest temperature at which the growth of stable pits is possible.

Crystallographic plane A distinct plane in a crystal lattice.

Depassivation Loss of the state of being passive.

Electropolished A smooth and polished surface produced by making the specimen the anode in an electrolytic cell.

Etch A roughened surface produced by chemical, electrochemical, or mechanical means.

Hydrolysis In aqueous solutions of electrolytes, the reactions of cations with water to produce a

weak base or of anions to produce a weak acid.

Inhibition potential The potential above which pits repassivate in solutions containing inhibitors.

Intermetallic compound A type of phase formed when atoms of two or more metals combine in a fixed ratio to produce a crystal different in structure from the individual metals.

Protection potential or repassivation potential Minimum potential below which already existing pits stop growing.

Vacancy A site in a crystal unoccupied by an atom.

Weak acid An acid that dissociates only partially into hydrogen ions and anions in solution.

Abbreviations

CCT Critical crevice temperature
CPT Critical pitting temperature
ESCA Electron spectroscopy for chemical analysis
NMR Nuclear magnetic resonance
SCE Saturated calomel electrode
SHE Standard hydrogen electrode

Symbols

c.d. Current density
C_i Concentration of component *i*
D_i Diffusion coefficient of component *i* ($cm^2\ s^{-1}$)
E_c^* Corrosion potential of the metal in the solution inside a pit
E_i Inhibition potential
E_p Pitting potential
E_r Repassivation potential
F Faraday constant
i Current density ($A\ cm^{-2}$)
j_{pas} Passive current density
j_{pit} Current density inside a pit
J_i Flux of component *i*
K_a Equilibrium constant for the dissociation of an acid
$^*K_1$ Equilibrium constant for the first step of the hydrolysis reaction
K_w Ionic product of water
R Gas constant
T Temperature (°K)
x Pit depth (cm)
η Polarization
ϕ Electrical potential

2.04.1 Introduction

2.04.1.1 Definition of the Problem

Numerous metals and alloys, such as aluminum, stainless steel, titanium, zirconium, and so on have a very high practical interest because in many corrosive environments, they remain passive. For example, a 1 mm thick sheet of any of these metals can resist the action of the corrosive environment for several thousands of years. Nevertheless, under certain environmental conditions, the passive surface of these metals can show very localized corrosion points, where the 1 mm sheet can develop a hole in less than 1 h. This process of localized metal degradation is known as pitting corrosion (Figure 1).

When the possibility of pitting corrosion is suspected, there are several ways by which the problem can be treated. At the design stage of equipment, one possible approach is the selection of alloys that will resist pitting corrosion in the given working environment. If, because of the properties required from the alloy, or because of economical reasons, this solution is not convenient, the other possibility for dealing with the pitting corrosion problem is control of the corrosive

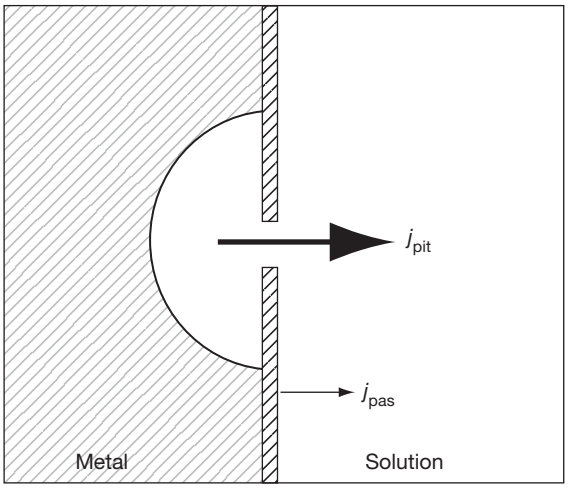


Figure 1 Pitting corrosion. Numerous metals and alloys, such as aluminum, stainless steel, titanium, zirconium, and so on have a very high practical interest because in many corrosive environments they remain passive, with corrosion current densities of the order of $j_{pas} = 10^{-6}\ A\ cm^{-2}$. Nevertheless, under certain environmental conditions, the passive surface of those metals could show very localized corrosion points, where the corrosion rate reaches values close to $j_{pit} = 1\ A\ cm^{-2}$. This process of localized metal degradation is known as pitting corrosion.

environment. This could be done by controlling the concentration of the pitting-inducing anions, by the use of appropriate inhibitors, by controlling the oxygen content of the corrosive environment, and so on. Most of these subjects will be dealt with in this chapter.

2.04.1.2 Anions Producing Pitting in Metals

As shown in **Table 1**, there is a wide number of anions capable of producing pitting of a variety of metals, in aqueous solutions.[1–37]

It is common to find publications in which some specific properties are attributed to chloride anions that would make them particularly aggressive as promoters of pitting corrosion. Nevertheless, this suggestion was dismissed by Hoar[14] as early as in 1947, when discussing the iron pitting mechanism. According to Hoar, the action of the chloride anions should not be related to a strong selective adsorption on the metal, as some authors were suggesting, or to a particularly penetrating power of chlorides into the passive film, as suggested by others. Hoar[14] reported that anions such as ClO_4^- were as aggressive as Cl^-. He suggested that the effect of Cl^-, being a nondepositing ion, had a strong influence on the pH in the electrolyte adjacent to the metal anode. According to Hoar, Cl^- ions replaced OH^- ions in the solution close to iron, which was the metal under discussion, thus hindering the process of precipitation of $Fe(OH)_2$. On the basis of these ideas, Hoar[27] had been able to show even earlier, in 1937, that tin suffered pitting in $NaClO_4$ solutions. Comparing anions of similar aggressive power, the reason for finding chloride anions in most of the practical cases of pitting is the extensive distribution of NaCl in nature. In fact, very special care has to be taken to have chloride free environments.

Table 1 Anions producing pitting in metals

Metal	*Aggressive anion*	*Ref*
Aluminum	Cl^-	1, 2, 3
	Br^-	1, 2
	I^-	1, 2
	ClO_4^-	1, 4
	NO_3^-	1, 5
	SCN^-	6, 7
Cadmium	Cl^-	8, 9
	Br^-	8
	ClO_4^-	8
	$SO_4^=$	8, 9
Iron	Cl^-	10, 11, 12
	Br^-	10, 11
	I^-	10, 13
	ClO_4^-	10, 11, 14, 15
	$SO_4^=$	11, 16, 17
Nickel	Cl^-	13, 18
	Br^-	13, 19
	I^-	13
Stainless steel	Cl^-	20, 21, 22, 23, 24
	Br^-	20, 21, 22, 23, 24
	SCN^-	25
Tin	Cl^-	26
	ClO_4^-	27
	$SO_4^=$	26
	NO_3^-	26
Titanium	Cl^-	28, 29, 30, 31
	Br^-	28, 29, 31
	I^-	29, 32, 33
Zinc	Cl^-	8, 34
	Br^-	8, 35
	I^-	34, 35
	NO_3^-	34, 35
	$SO_4^=$	8, 34, 35
	ClO_4^-	35
Zirconium	Cl^-	29, 32, 33, 36
	Br^-	29, 32, 33
	I^-	29, 32, 33
	ClO_4^-	37

2.04.1.3 Three Types of Pitting

When reviewing the literature on pitting, it was found that the mixture of different processes, with similar corrosion shapes, was the source of serious confusion because the same mechanism was attributed to all of them.[38] The review showed that, at least three different types of pitting were being discussed, without any clear distinction among them. They can be identified as:

1. electrochemical depassivation,
2. chemical depassivation,
3. etch pitting.

Electrochemical depassivation is the most common of the three, and from a practical point of view, it is also the most relevant. It is found in neutral and alkaline solutions. It develops on otherwise passive metals, and it is characterized as being associated with a pitting potential. It is observed only above this potential, remaining passive below it. In this chapter, when talking about pitting corrosion, we will be referring to this type of attack.

In certain acid solutions, iron, nickel, and other metals show polarization curves of the type shown in **Figure 2**. Curves similar to those shown in **Figure 2** were reported by Piron *et al.*[39] for nickel and for Inconel alloy, by Nobe and Tobias[40] for iron,

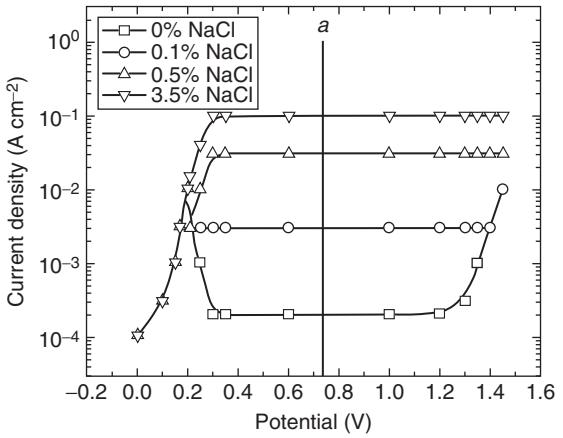


Figure 2 Typical shapes of the polarization curves for iron, nickel, and their alloys, in aqueous solutions of pure 1 M H_2SO_4 and in acid with various contents of NaCl. In pure acid the metal shows an active zone, followed by a passive zone, and finally a current increase because of transpassivity or to oxygen evolution. If the polarization curves are measured in solutions containing various amounts of NaCl, it is observed that the current in the passive zone increases and eventually, at a given NaCl content, passivity fails to appear. If the metal, in pure H_2SO_4, is kept at a potential such as *a*, it becomes passive. If, at this point NaCl is injected into the solution, keeping the potential constant, a current increase will be observed, until the stationary polarization curve for the new solution composition is reached. In the transition from the initial passive state to the new corrosion current, pits are formed on the passive surface, which, in the present chapter, are called chemical depassivation. The potential and the current density values in this figure are arbitrary. The values are in the order of magnitude expected for these metals, but were chosen only for illustration purposes.

by Condit[41] for iron–nickel alloys. Engell and Stolica[42] reported that if iron is passivated in chloride-free sulfuric acid, at a potential somewhere between 0.5 and 1.0 V(SCE), and chloride ions are then injected into the solution, pits are formed on the metal. This type of pitting, resulting from a change in the chemical composition, is classified as chemical depassivation. No pitting potential was found in these experiments, and from **Figure 2**, it can be concluded that it is the transition from one polarization curve to the curve pertaining to the new solution. From a practical point of view, it is a very uncommon case of corrosion. On the other hand, contrary to the earlier opinion of some authors, there is no evidence that would suggest that the same mechanism is acting both in electrochemical depassivation and chemical depassivation.

A third type of pitting is etch pitting. It is crystallographic in nature, and it was developed mainly as a technique for metallography. It was described in detail by Ives,[43] and falls outside the scope of this chapter.

This chapter deals with electrochemical depassivation, with some references to chemical depassivation.

2.04.2 Experimental Aspects of Pitting

2.04.2.1 Pitting Morphologies

In general, when corrosion develops spontaneously, without external currents being applied, pits develop in a very localized way in irregular shapes. Usually, it is observed that once a pit nucleates, no other pits are formed in the surrounding areas. The reason is that while the anodic reaction is concentrated in the pit, the surrounding areas are the location for the cathodic reaction, usually oxygen reduction, which produces local changes in pH, inhibiting the formation of new pits close to the active one.

If, on the other hand, pits are formed by the application of an external current, once they are formed the pits can spread laterally to most of the metal surface. The rate of the spreading of corrosion is controlled by the solubility of the passive film in the acid solution produced inside the pit. For example, pits on cadmium,[9] iron,[12] and zinc[34] were found to spread sidewise very easily, while pits on aluminum and stainless steels remained localized, because of the resistance of the passive film to dissolution. In the particular case of stainless steels, pits could remain closed, and the magnitude of corrosion is made visible only after a slight abrasion of the surface film.[44] Closed pits have been reported on stainless steels by Rosenfeld and Danilov,[45] while Yahalom *et al.*[46] observed the initial passive film covering the surface of closed pits. An example of covered pits in stainless steel could be found in the paper by Laycock *et al.*[47] Another example of covered pits, in this case for high purity iron pitted in a sodium chloride solution with a borate buffer, could be found in the paper by Alvarez and Galvele.[48]

When working under controlled electrochemical conditions, it is frequent to find that hemispherical electropolished pits are formed, as shown in **Figure 3**. These pits were produced on pure aluminum, in a $NaNO_3$ solution, at constant potential. Similar pits were reported for iron in chloride and in sulfate solutions, for stainless steels in chloride solutions, and so on. The rounded polished morphology of pitting is the result of the random dissolution of the metal atoms, inside the pit, particularly when such a

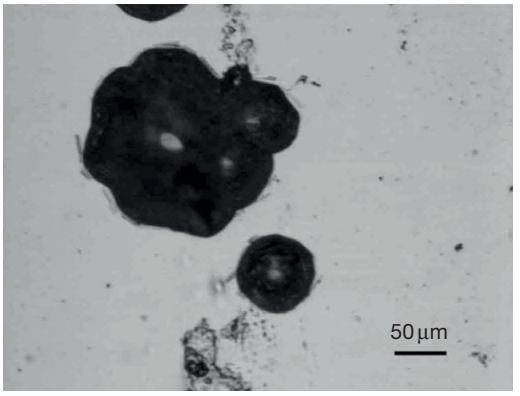


Figure 3 Hemispherical pits on aluminum 99.999% in 1.0 M $NaNO_3$ solution at $E = 1.84\ V_{SHE}$. (Courtesy S. Fernández and R. M. Carranza).

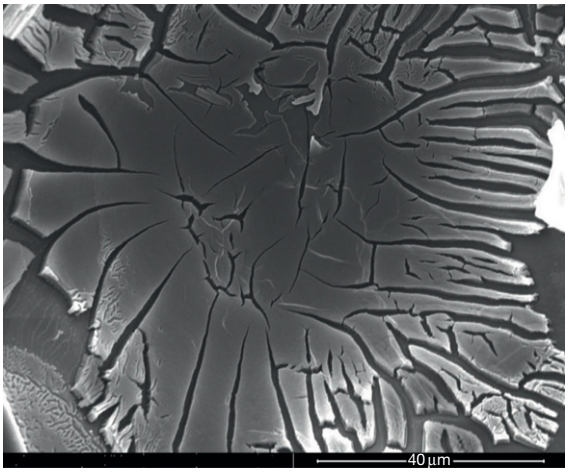


Figure 4 SEM picture of the interior of a pit of aluminum 99.999% in 1.0 M $NaNO_3$ solution at $E = 1.84\ V_{SHE}$. The bottom of the pits covered with corrosion products showing a 'dried mud' structure. (Courtesy S. Fernández and R. M. Carranza).

dissolution takes place through a film of corrosion products, as shown in Figure 4. On the other hand, it is also common to observe that pits could grow developing crystallographic planes. An interesting case is that of pitting of aluminum in chloride solutions. In this system, pits grow developing {100} planes, Figure 5. In the particular case of aluminum, this type of attack can lead to the development of deep tunnels, as shown in Figure 6. The production of these tunnels has a practical use, because it is applied to the production of foils for electrolytic condensers.[49] Crystallographic etching, like that shown for aluminum in chloride solutions, was also found for aluminum in bromide, iodide, and perchlorate solutions.[6] Crystallographic planes inside the pits were also reported for iron in

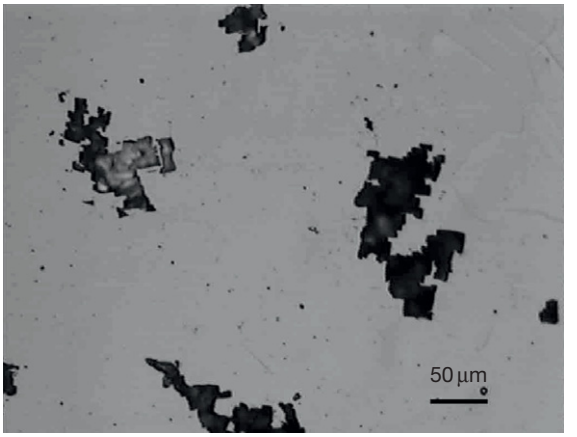


Figure 5 Pitting of aluminum in chloride solutions. Aluminum 99.999% in 1.0 M NaCl solution at $E = -0.420\ V_{SHE}$. (Courtesy S. Fernández and R. M. Carranza).

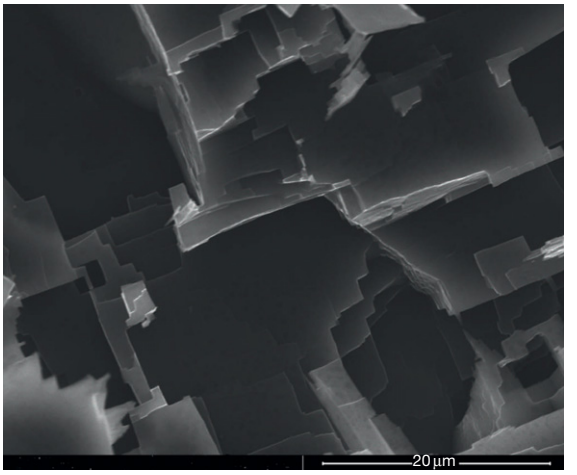


Figure 6 SEM picture showing the {100} planes developed inside pits of aluminum 99.999% in 1.0 M NaCl solution at $E = -0.420\ V_{SHE}$. (Courtesy S. Fernández and R. M. Carranza).

perchlorate solutions,[15] for cadmium in chloride solutions,[9] for zinc in chloride solutions,[34] and so on.

2.04.2.2 Characteristic Electrochemical Parameters for Pitting

When the use of electronic instruments, such as potentiostats, was introduced in corrosion studies, several potential values, characteristic of the pitting process, were found. The most relevant ones were: pitting potential, repassivation potential, and inhibition potential.

2.04.2.2.1 Pitting potential

It was found that electrochemical depassivation, or pitting, did not appear at any arbitrary potential value, but only above a certain potential called pitting potential (E_p). Below this potential, no pitting was observed, no matter how long the metal was exposed to the action of the environment. Above E_p, on the other hand, pitting would start almost immediately, as it was the case with aluminum in chloride solutions, or after a short induction time, as in the case of stainless steels. This pitting potential is a function of the composition of the environment, the temperature, the composition of the alloy, and so on. Figure 7 shows a typical example of an anodic polarization curve, where the pitting potential of aluminum in a NaCl solution is shown.

Curves like that shown in Figure 7 could be drawn by applying a constant potential, and waiting for the current to reach a stationary value. Then, the potential is increased, in steps of ~10 mV, and the new current value is recorded. When the pitting potential is reached the current starts to increase, and pits start to develop on the metal surface. This technique for measuring E_p is usually called the potentiostatic technique. In the early studies of pitting, some authors had doubts about the validity of the pitting potential, suggesting that if the waiting time at constant potential was extended long enough, pitting would be found at potentials below E_p. In order to test this possibility a constant potential technique was applied to measure E_p. In this case, a different sample was used at each potential, the samples were kept for an extended time, the current was measured and after the measurement the samples were observed under the microscope. Figure 8 is an example of this technique. The constant potential measurements confirmed the validity of the pitting potential as a transition potential between passivity and localized corrosion.

Horvath and Uhlig,[50] working with stainless steels and related alloys, defined the pitting potential as the lowest potential at which pitting could not be observed after a 10 h or longer period at constant potential. According to these authors, reproducibility was in the order of ± 5 mV.

Instead of the above-mentioned potentiostatic technique, which tends to be time consuming, another technique frequently used is the potentiokinetic technique, where the electrode potential is changed at a constant scanning rate, while the current is being recorded. One limitation of this technique is that the pitting potential values found can be slightly high, because not enough time is given for the pits to nucleate. Another technique for measuring the pitting potential was introduced by Pessall and Liu.[51] In this technique, the specimen is exposed to the solution, at constant potential, while the metal surface is scratched with a diamond point, or another sharp instrument, to break the passive oxide film. If the potential is lower than the pitting potential, the exposed metal repassivates in a very fast rate. On the other hand, if the potential is above the pitting potential, the repassivation rate is very slow or no repassivation is observed. By measuring the repassivation time, at different potentials, it is possible

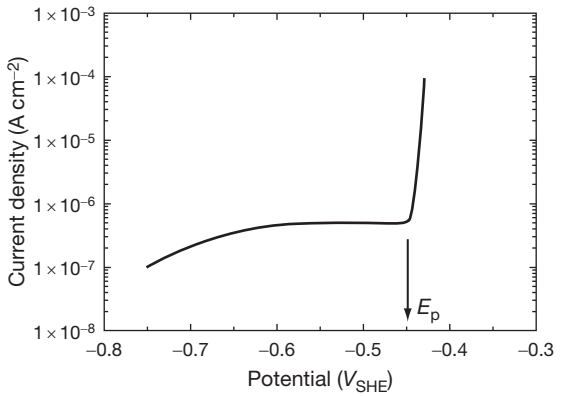


Figure 7 Pitting potential (E_p). Curve based on measurements for aluminum in NaCl solutions. Reprinted from Galvele, J. R.; De Micheli, S. M. de. *Corros. Sci*. **1970**, *10*, 795–807, with permission from Elsevier.[60]

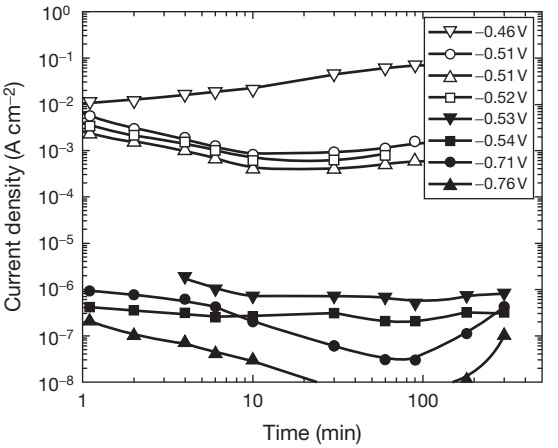


Figure 8 Constant potential technique for measuring E_p. Curves based on measurements for aluminum in NaCl solutions. Reprinted from Galvele, J. R.; De Micheli, S. M. de. *Corros. Sci*. **1970**, *10*, 795–807, with permission from Elsevier.

to measure E_p with dispersion lower than 10 mV. Pessal and Liu developed this technique to solve the problem caused by high dispersion in E_p found for various metals when using potentiokinetic techniques. For the same reason some authors, when measuring pitting potentials, used various techniques in each system, to make sure that the E_p values found were reliable.[9,34,48]

2.04.2.2.2 Repassivation potential

Another characteristic potential of the pitting process was described by Pourbaix[52] and it is known as the repassivation potential. Pourbaix reported that pits, which will start to grow at the pitting potential, or above, will keep growing even when the potential is lowered below E_p, **Figure 9**. The pits will stop growing only when the potential is lower than a certain value called repassivation potential, or protection potential, E_r. As found by Suzuki and Kitamura,[53] the value of the repassivation potential is related to the occlusion of the pit. As described later, Wilde and Williams[54] found, for a series of steel alloys, that the bigger the difference between E_p and E_r, the more susceptible the alloy was to crevice corrosion.

2.04.2.2.3 Inhibition potential

The third characteristic potential was first described by Schwenk,[24] for stainless steels in chloride solutions containing nitrates. It was reported that in these solutions, when increasing the potential above the pitting potential, the current increased. But eventually a potential was reached above which the pits were repassivated. This potential was called inhibition potential, E_i, **Figure 10**, and it has been reported in numerous systems. Kuzub and Novitskii[55] found that E_i, for stainless steels in chloride plus nitrate solutions was a function of the chromium content of the steel. At higher chromium contents E_i was lower, and the inhibition of pitting was easier.

Vetter and Strehblow[11] measured the inhibition potential for iron in perchlorate solutions, for iron and nickel in chloride plus nitrate solutions, and for iron in chloride plus perchlorate solutions. Strehblow and Titze[56] also reported inhibition potentials for iron in bromide plus perchlorate solutions, in iodide plus nitrate solutions, and in iodide plus perchlorate solutions and for nickel in bromide plus nitrate solutions.

While the pitting potential, E_p, has important practical applications, this does not seem to be the case with the pitting inhibition potential. Keitelman and Galvele[57] studied the anodic behavior of high purity iron in a 0.5 M sodium sulfate solution. They measured the pitting potential of iron in pH 9.0 and pH 10.0 solutions. The pitting potential was found to be very close to the corrosion potential measured in a Na_2SO_4, pH 2.7, pit-like solution. On the other hand, the pitting inhibition potential measured in a 0.5 M Na_2SO_4, pH 10.0, solution was very close to the passivation potential found for the metal in the pit-like solution. The practical use of the pitting inhibition potential was limited, because susceptibility to crevice corrosion was observed at potentials above the pitting inhibition potential. Keitelman and Galvele[57] concluded that the pitting inhibition potential could

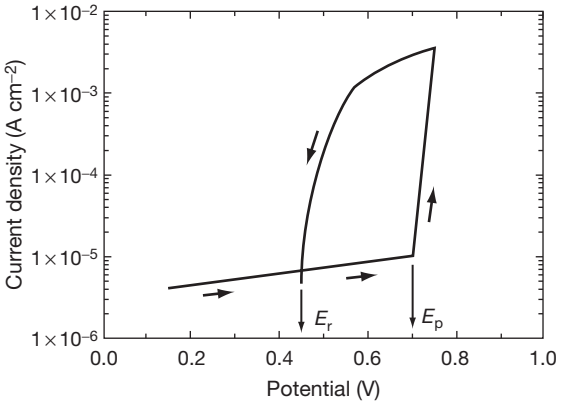


Figure 9 Typical polarization curve showing E_r: repassivation potential; E_p: pitting potential for a metal, like stainless steel, in a chloride solution.[38] Reprinted from Galvele, J. R. *Corros. Sci*. **2005**, *47*, 3053–3067, with permission from Elsevier.

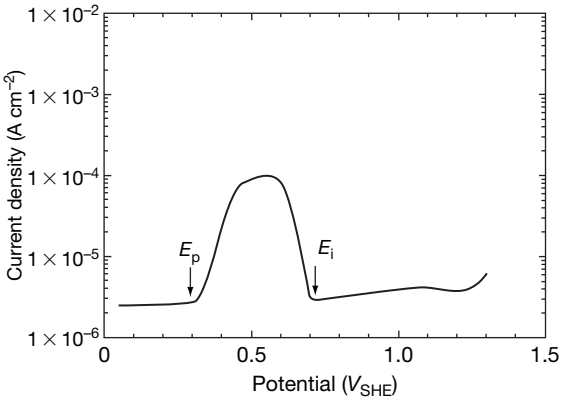


Figure 10 Typical anodic polarization curve showing the presence of a pitting potential, E_p, and an inhibition potential, E_i.[38] Reprinted from Galvele, J. R. *Corros. Sci*. **2005**, *47*, 3053–3067, with permission from Elsevier.

not be used to stop pitting corrosion, because in those areas where a high potential drop could be found, like in a crevice, the pitting inhibition potential was not reached and crevice corrosion developed.

It is important to notice that two different forms of crevice corrosion have just been mentioned. The first one is related to the repassivation potential, E_r, which can be exemplified by the work by Wand and Newman,[58] and the second one, related to the inhibition potential, E_i, which is exemplified by the work of Al-Zaharani and Pickering.[59] These two types of crevice corrosion were pointed out recently by Galvele.[60]

2.04.3 Effect of the Bulk Solution Composition

2.04.3.1 Aggressive Solutions

Table 1 shows the wide variety of ions capable of producing pitting in various metals. What is common to all of them is that none of those ions produce insoluble compounds with the metal under consideration. Another general property is that all the anions involved produce strong acids and they do not interfere with the hydrolysis reaction of the dissolved metal ions with water. The same metal shows different pitting potentials in different environments, and the differences can be substantial, as shown in Figure 11 for aluminum. The case of titanium is of high practical interest, because Beck[28] reported that the pitting potential of titanium in chlorides was very high, being ~10 V higher than that in other halide solutions. In some cases, the differences in pitting potentials could be small and with specific deviations. For example, for zinc[34] it was reported that the pitting potential was very similar in solutions containing chlorides, iodides, or sulfates, but the pitting potential measured in nitrate solutions was almost 200 mV higher than that in the previously mentioned solutions.

The pitting potential of a metal is also a function of the aggressive anion concentration. It is a general rule that the higher the aggressive anion concentration, the lower the pitting potential. Figure 12 shows an example of the relation between pitting potential and NaCl concentration for pure aluminum and some of its alloys with copper. In general, relationships of the type

$$E_p = A - B \log C_x$$

are reported, A and B being constants, E_p the pitting potential, and C_x the concentration in moles of the aggressive anion. A review of the literature[38] shows that the value of B varies from metal to metal, going from ~50 mV up to ~100 mV for pure metals. For Fe–Cr alloys, typical values around 200 mV were reported.

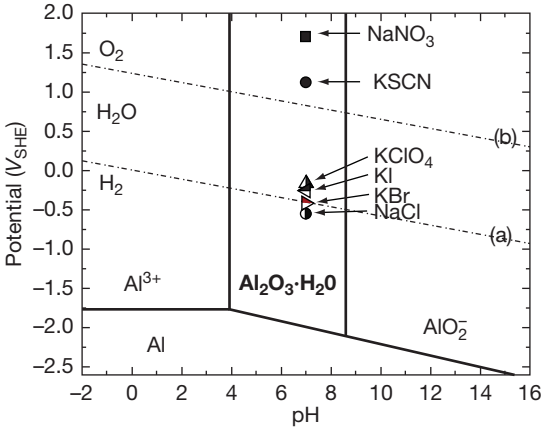


Figure 11 Pitting potentials reported for aluminum in various solutions. For references purposes, the pitting potential values were superimposed on an Al–H_2O potential–pH diagram. (a) Equilibrium values for the reaction of hydrogen evolution. (b) Equilibrium values for the reaction of oxygen evolution. From the pits produced near line (a) it is expected the hydrogen will evolve from the pits. On the other hand, from pits produced on aluminum in nitrate solutions the evolution of nitrogen bubbles was detected.

2.04.3.2 Inhibitors

When the metal is exposed to the action of a mixed electrolyte, different behaviors are observed. If only one of the anions is aggressive and the other anions are inactive from the pitting point of view, the pitting potential will change as if only the aggressive ions

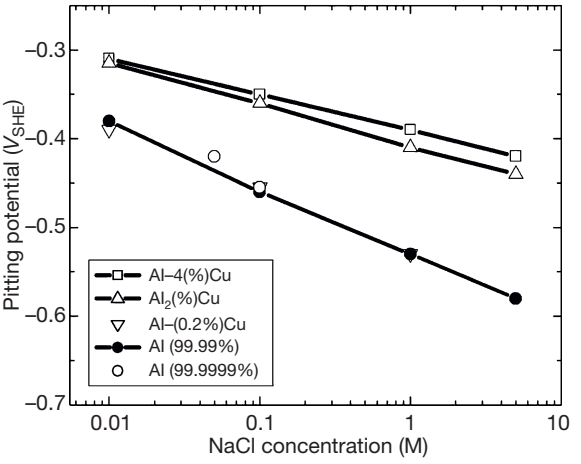


Figure 12 Effect of NaCl concentration on the pitting potential of aluminum and some of its alloys. Reprinted from Galvele, J. R.; De Micheli, S. M. de. *Corros. Sci*. **1970**, *10*, 795–807, with permission from Elsevier.

were present. If the content of the aggressive anion is reduced, the pitting potential will decrease. But there is a minimum concentration of aggressive anion, below which no pitting can be produced. This minimum value changes from metal to metal. Table 2 shows the minimum chloride concentrations found for pitting of various alloys in mixtures of sulfuric acid plus sodium chloride solutions. In this case, only chemical depassivation was considered.

As for electrochemical depassivation, it was observed that a nonaggressive anion could increase the pitting potential value. Leckie and Uhlig[62] as well as Rosenfeld and Danilov[45] studied these effects on austenitic stainless steels. Leckie and Uhlig[62] reported that for chloride ion concentrations higher than 0.1 M, the pitting inhibition efficiency of the studied anions was in the following decreasing order:

$$OH^- > NO_3^- > SO_4^{2-} > ClO_4^-$$

Schwenk[24] reported that while NO_3^- ion inhibited pitting of stainless steels in chloride-containing solutions, it did not inhibit pitting in bromide containing solutions.

Böhni and Uhlig[63] reported for aluminum similar results to those found for stainless steels in chloride solutions. In the case of aluminum, it was assumed that the metal was immune to pitting when the pitting potential was above $0.8\,V_{SHE}$. Böhni and Uhlig[63] found that the inhibition efficiency for aluminum in chloride containing solutions was in the following decreasing order:

$$NO_3^- > CrO_4^{2-} > \text{acetate} > \text{benzoate} > SO_4^{2-}$$

De Micheli and Galvele[6] studied the effect of mixtures of anions on the pitting potential of aluminum. They found that in mixtures of chloride plus perchlorate anions, a smooth transition was observed on the pitting potential values. In mixtures of chloride plus nitrate solutions, a clear inhibition effect of the nitrate ion was observed. Nevertheless, one point of interest is that aluminum, in a 1 M NaCl solution, showed a pitting potential of $-0.64\,V_{SHE}$, and the pitting was inhibited by NO_3^- ions, but on the other hand, if the potential was above $+1.84\,V_{SHE}$, the NO_3^- ion itself produced pitting on aluminum.[1,5] In mixtures of sulfate plus chloride anions, the sulfate ion had little effect on the pitting potential of aluminum.[6] The pitting potential values were the same as those found in chloride solutions without sulfate ions. It was only at very low chloride concentrations, less than 10^{-3} M, that sulfate showed some effect on the pitting potential.

The inhibiting effect of sulfate ions in chloride-containing solutions was reported for the pitting of nickel,[64] zirconium,[32] and titanium.[31]

Numerous authors reported that weak acid salts acted as pitting inhibitors for stainless steel and for aluminum in chloride containing solutions. Similar observations were made for other metals. For example, the pitting potential of zinc, in sodium chloride solutions and in sodium perchlorate solutions, was increased by adding borate ions to the solutions.[35] The pitting potential of iron in sodium chloride solutions was also increased by adding borate salts.[48] In numerous applications chromates were found to be very efficient as pitting inhibitors, but the carcinogenic nature of chromate requires the search for alternative solutions.

Table 2 Minimum chloride concentration necessary for chemical depassivation of various metals and alloys in mixtures of $H_2SO_4 + NaCl$ solutions

Metal	*Minimum Cl^- concentration (N)*	*Reference*
Iron	0.0003	61
Fe–5.6Cr	0.017	61
Fe–20Cr	0.1	61
Fe24.5Cr	1.0	61
Fe–18.6Cr–9.9Ni	0.1	61
Nickel	0.001	18

2.04.3.3 pH

According to Kaesche,[65] for aluminum in NaCl aqueous solutions, the pH of the solution had no effect on E_p, for pH values ranging from pH 2 up to pH 11. For stainless steels, on the other hand, Leckie and Uhligh[62] found that the pitting potential was pH independent up to pH 8, but increased sharply with the pH for higher pH values. Similarly, Alvarez and Galvele[48] found that an increase in pH produced an increase in the pitting potential for iron in NaCl solutions, Figure 13. Inhibiting effects due to alkaline solutions were also reported for nickel,[66] cadmium,[9] zinc,[35] and so on.

2.04.4 Composition of the Solution inside Pits

2.04.4.1 Electrolyte inside Pits

From the earliest studies on pitting it was assumed that the composition of the solution inside the pits

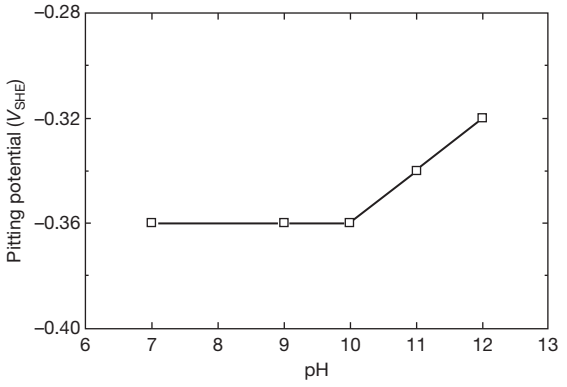


Figure 13 Pitting potential of high purity iron, in 1.0 M NaCl at various pH values. Reprinted from Alvarez, M. G.; Galvele, J. R. *Corros. Sci*. **1984**, *24*, 27–48, with permission from Elsevier.

was different from that outside. A pH decrease in the solution inside aluminum pits was reported by Edeleanu and Evans in 1951.[67] Hagyard and Santhiapillai,[68] prepared very simple artificial pits on an aluminum sample, and reported that, while the pH of the NaCl solution was 6.2, the pH inside the pit was ~3.5. They also found that the chloride concentration inside the pit increased, reaching values about ten times higher than those in the bulk solution. They also found that the concentrated aluminum chloride solution produced inside the pit could flow by gravity outside the pit, activating the metal surface it touched. Wong and Alkire[69] measured the pH of solution inside naturally occurring pits in aluminum, and confirmed that the pH of the solution inside the pits was between 3 and 4. They also found that the solution flowing from the pits, mentioned by Hagyard and Santhiapillai, was not an $AlCl_3$ solution, as earlier believed by various authors, but the dominant species was either $Al(OH)Cl_2$ or $Al(OH)_2Cl$. The distinction between the two species was uncertain because their NMR spectra were similar. According to Wong and Alkire,[69] the efficiency of the anodic reaction of metal dissolution inside the aluminum pits was found to be close to 80%, the side reaction being the evolution of hydrogen gas.

Abundant work on actual pits, as well as crevices, has shown beyond any doubt that the solution inside the pits and crevices has a lower pH and a higher ionic concentration than the bulk solution. For copper in sea water, pH 8, the pH found inside the pits was of the order of 5[70] and for iron, the pH measured inside the pits was ~3.5, while the pH of the bulk solution could have values between 2 and 10.[71,72] For stainless steels exposed to neutral 0.5 M NaCl solutions, the pH inside the pit was found to be lower than 1.0, while the chloride concentration was 6.2 M.[72]

2.04.4.2 Gas Evolution from Pits

The evolution of gas bubbles from pits was reported in the early studies on pitting. Kaesche reported the evolution of hydrogen from pits of aluminum, growing in a NaCl solution.[65] Evolution of hydrogen was also found from aluminum pits produced in perchlorate solutions.[4] Hydrogen evolution was also reported from pits in iron[73] and titanium.[28] Wexler and Galvele[5] reported gas evolution from pits in aluminum in sodium nitrate solutions, but in this case the gas evolving was found to be nitrogen. This observation was confirmed by Bargeron and Benson.[74]

An evaluation made of the gases evolved from the pits in titanium and in aluminum showed that the hydrogen evolving was equivalent to ~10% of the total anodic current in the pit.

2.04.5 Effect of Alloying Elements

It was observed that the pitting potential of a metal could be strongly modified with the use of alloying elements. This is an aspect of high technological interest and abundant work has been done on this subject. It is interesting to point out that there are cases in which the aims of the technological application could be pointing to completely opposite directions. For example, for stainless steels, the main interest was to find alloying elements that could lead to the highest possible pitting potential, and thus obtain a very resistant alloy. On the other hand, with aluminum, one of the technological interests was to find an alloy with the lowest possible pitting potential in sea water. The reason was that aluminum was a very attractive metal to be used as sacrificial anode, for cathodic protection. The limitation was that the pitting potential of aluminum, in sea water, was ~1 V higher than the equilibrium potential for the Al/Al^{3+} reaction. Consequently, much work was oriented to the search of alloying elements that could lower the pitting potential of aluminum.

The extensive work on the effect of alloying elements on increasing the pitting potential of stainless steels has been reviewed by various authors.[75–79] In general, chromium, vanadium, molybdenum,

tungsten, rhenium, nickel, and nitrogen were found to be beneficial alloying elements, as all of them increased the pitting potential of stainless steel. On the other hand, alloying elements such as manganese or sulfur were found to be detrimental. As pointed out by Horvath and Uhlig,[51] care should be taken when making extrapolations. These authors found that while molybdenum was beneficial for stainless steels in chloride-containing solutions, it was detrimental in bromide-containing solutions.

Some elements, to be effective, need the presence of a third element. This is the case of molybdenum, in ferrous alloys. It requires the presence of chromium to inhibit pitting, and the higher the chromium content the more effective molybdenum would be.[80] It was found, with binary Fe–Mo alloys, that the pitting potential of Fe–5Mo was only slightly higher than that of pure iron.[81]

Nitrogen strongly increases the pitting potential of austenitic stainless steels in chloride solutions. Osozawa *et al.*[77] studied stainless steels containing from 0.02% to 0.2% N. They found that during corrosion all the nitrogen was transformed into ammonium ion, preventing the lowering of pH before the sites could grow as ordinary pits.

Another metal about which abundant information on the effect of alloying elements is available is aluminum. For aluminum alloys with ~5% of the alloying metal, it is known that copper increases the pitting potential of aluminum in chloride solutions, while zinc decreases it, and magnesium has almost no effect, **Figure 14**.[82,83]

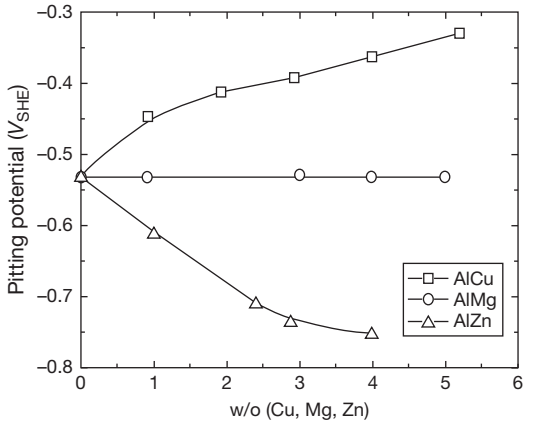


Figure 14 Pitting potential of binary Al–Cu, Al–Mg and Al–Zn alloys in 1 M NaCl solution. Reprinted from Muller, I. L.; Galvele, J. R. *Corros. Sci.* **1977**, *17*, 179–193, with permission from Elsevier.

Reding and Newport[84] studied ~2500 aluminum alloys in sea water. The alloys were made with practically all the metals of the periodic table. The purpose of their work was to develop aluminum anodes for cathodic protection in sea water. To measure the efficiency of the alloys, the authors circulated a constant anodic current density and measured the potential as function of time. As an old technique for measuring pitting potentials, the galvanostatic technique, consisted in applying a constant current to the sample in the corrosive environment and measuring the potential reached by the sample,[85] Reding and Newport's results could be used to have a rough estimate of the effect of those alloying elements on the pitting potential of aluminum in sea water, as shown in **Figure 15**. It is only a rough estimate, as in the galvanostatic technique, various anodic currents were applied to make sure that the same potential was reached. Nevertheless **Figure 15** gives an initial idea of what the effect of those alloying elements will be on the pitting potential of aluminum in chloride solutions.

2.04.6 Pitting Nucleation Sites

The nucleation pit sites have frequently been found to be related to microscopic features on the metal surface. In stainless steels, the pits are usually related to the presence of inclusions. Smialowski *et al.*[86]

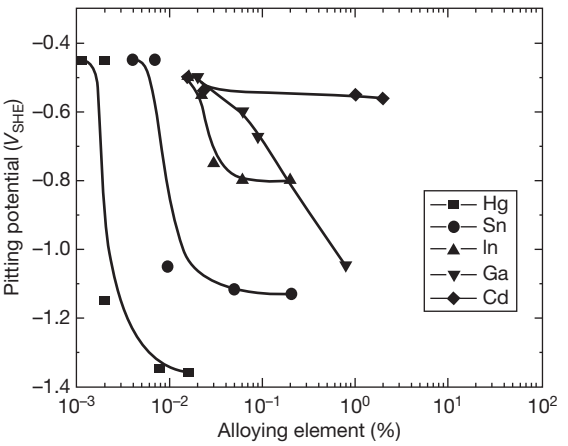


Figure 15 Estimated pitting potential values for binary aluminum alloys in sea water. The estimation is on the basis of the assumption that the measurements made by Reding and Newport[84] lead to electrode potentials close to those obtained in galvanostatic techniques used for pitting potential measurements.

reported that pits are initiated at nonmetallic inclusions, mixed manganese sulfide inclusions being the most effective ones. According to these authors, in the absence of the sulfide inclusions, chromium oxide particles can nucleate pitting.

For low alloy steels, the preferential sites for nucleation of pits were found to be manganese sulfide particles.[87] For nickel, it was reported that pits nucleated either on surface scratches, on grain boundaries or on nickel sulfide inclusions.[88,89]

Suter *et al.*[90] studied the initiation of individual pits on MnS inclusions on 304 stainless steel, with and without mechanical stress. Working in 1 M NaCl solution, the authors found that in those areas where there were no inclusions, no pitting was nucleated at potentials below oxygen evolution. Specimens with shallow MnS inclusions showed metastable pitting but not stable pitting. Specimens with deep MnS inclusions showed active pitting.

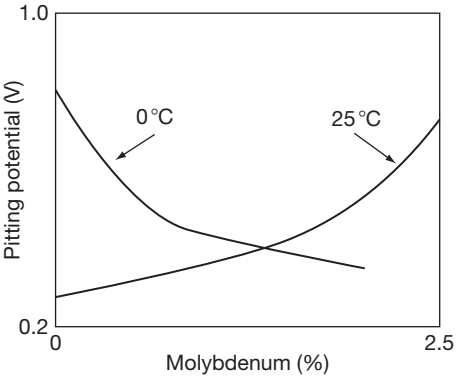


Figure 16 Schematic view of the effect of Mo and temperature on the pitting potential for 15% Cr, 13% Ni stainless steel, in 0.1 N NaCl solution, as described by Horvath and Uhlig. Reproduced from Horvath, J.; Uhlig, H. H. *J. Electrochem. Soc.* **1968**, *115*, 791–795.

2.04.7 Effect of Temperature

The pitting potentials of stainless steels[53] and of aluminum[91] were found to decrease with an increase in temperature. The case of stainless steels is complicated by various factors. Horvath and Uhlig[50] studied the pitting potentials of Ni, of Cr–Ni, Cr–Fe alloys, and of various stainless steels. When studying the behavior of 15% Cr, 13% Ni stainless steel in 0.1 N NaCl, they reported that the pitting potential measured at 0 °C was 0.5 V higher than that measured at 25 °C. They also found that on adding Mo, the pitting potential at 25 °C increased, and at 0 °C, the pitting potential decreased. Eventually, with near 2.4% Mo a reversion in the pitting potential values was reached, as schematically shown in **Figure 16**. The authors reported that for 1.3% Mo the pitting potentials at 0 °C and at 25 °C were the same. Another observation made by these authors was that the results obtained in chloride solutions could not be extrapolated to other halide solutions. In 0.1 N NaBr solution, the pitting potential at 25 °C was higher than that at 0 °C, and also the addition of Mo decreased the pitting potential value at both temperatures.

Another example of complex relation between pitting potential and temperature is described by Beck[28] for the pitting of titanium in chloride, bromide, and iodide solutions. According to Beck, in chloride solutions, the pitting potential at room temperature is close to 10 V, and drops to ~1 V at 200 °C. On the other hand, in bromide solutions, the pitting potential of Ti starts at 1.7 V at 0 °C, drops to 0.9 V at room temperature, and remains constant up to above 100 °C. Finally, in iodide solutions, the pitting potential of Ti starts at ~1.0 V at 0 °C, increases to a couple of V at 50 °C, and from this temperature to at least 100 °C it remains almost constant at 0.45 V.

It is taken into account that in most of the cases, polyvalent cations are involved, which means that a variety of oxy- and hydroxyl-compounds can be produced, each of them with a different thermal stability. Besides, there are anions in solutions such as iodides that decompose by oxidation, while others such as nitrates decompose by reduction, giving a variety of chemical products. Considering also that most of these reactions will have different behaviors at different temperatures, it seems very improbable that a unique explanation could be found for the effect of temperature in all the cases of pitting. Most probably, each case will require a separate study.

Related to temperature and pitting, a parameter called critical pitting temperature (CPT) has been introduced in recent times, as described below, and is extensively used for grading stainless steels.

2.04.8 Mechanism of Stable Pitting Growth

A mechanism that attempts to explain stable pit growth, should also explain:

1. the nature of the pitting potential, and its properties;
2. the nature of the repassivation potential, and its properties;
3. the nature of the inhibition potential; and so on.

Historically, the development of mechanisms for pitting was simultaneous with the mechanistic studies of passivity. Pitting being a process of failure of passivity, it is natural that the mechanisms suggested for pitting were closely related to the authors' ideas on passivity. Consequently, in the older literature, it is found that those authors who sustained that passivity was the result of adsorption of oxygen atoms on the metal surface, favored the idea that pitting was the result of preferential adsorption of the aggressive anions, displacing the oxygen atoms.[20,50,62,63,92] Hoar,[14,27] as early as in 1937, questioned the mechanism based on competitive adsorption, supported by Kabanov and the concept of high penetrating power of Cl^- for the oxide films, supported at that time by Britton and Evans.

The idea that aggressive anions led to pitting by penetrating the passive film was supported by various authors. But, Vetter and Strehblow[93] pointed out that a migration of polyatomic aggressive anions such as $SO_4^=$, ClO_4^-, or SCN^- through a solid layer was not conceivable. **Figure 17** illustrates the point made by these authors.

Hoar[14,27] described a different view for the action of the so-called 'aggressive ions.' The new mechanism introduced by Hoar, allowed him to predict new cases of pitting, stating that the only condition required from the aggressive anion was that it would neither interfere with the process of anodic dissolution of the metal, nor with the hydrolysis reaction of the metal cations with the water molecules. According to Hoar, the Cl^- replaced the OH^- ions consumed by the metal ions and allowed the pH at the anode to fall to values where no solid products would be formed on the metal surface, but in the bulk of the solution.

Hoar's model of pitting became known as the localized acidification model and received reasonable wide acceptance, because it gave a good explanation of the nature of the aggressive anions, the change in composition inside the pits, and so on. Following Hoar's description, it was perhaps more appropriate to call it the OH^- depletion model, but it is just a semantic problem.

The introduction of potentiostats in corrosion studies and the discovery of pitting potentials were causes of concern for the supporters of this model, because it did not predict the existence of a pitting potential. Van Muylder *et al.*[70] were the first to attempt an explanation for the existence of a pitting potential value with an acidification process. They used the E–pH equilibrium diagram for $Cu–H_2O$ to explain pitting of copper in chloride solutions. **Figure 18** could be used to explain their ideas. Let us assume that pH_{SOL} is the pH of the bulk solution. Above the potential E_1 the anodic reaction of the metal begins as

$$Me + zH_2O = Me(OH)_z + zH^+ + ze^-$$

leading to localized acidification.

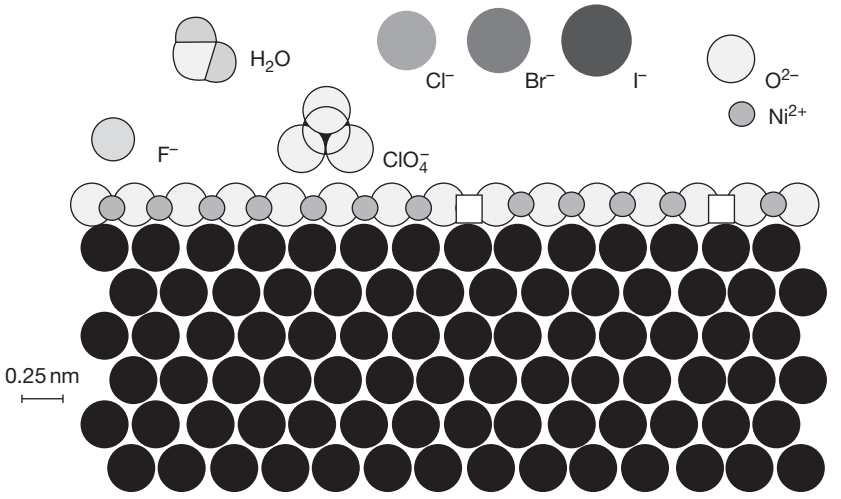


Figure 17 Relative sizes of metal and ions, for Ni covered with a single layer of NiO. Sizes and bond angles from: Linus Pauling, The Nature of the Chemical Bond, Cornell University Press, New York, 1960.

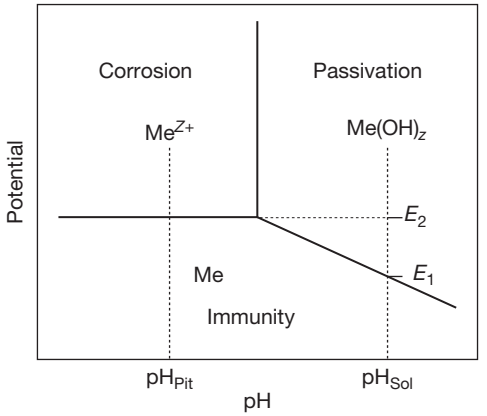


Figure 18 Explanation of the nature of E_p, on the basis of the ideas of Van Muylder *et al.*[70] Pits became stable only at potentials above E_2.

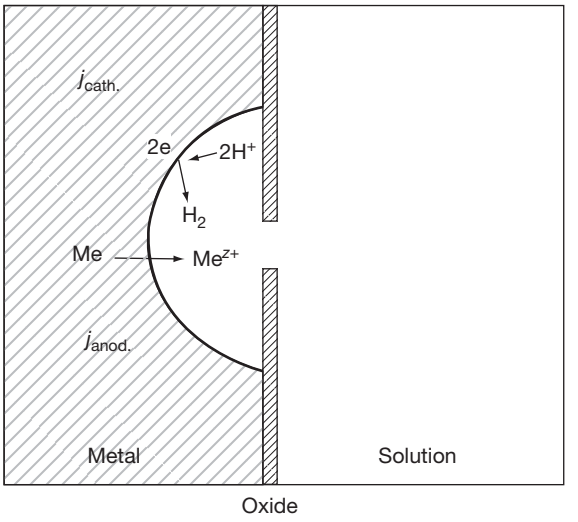


Figure 19 Anodic and cathodic reactions inside a pit.

Nevertheless, if the electrode potential is lower than E_2, the acidification will shift locally the system to the immunity region, see **Figure 18**, and the anodic reaction of metal dissolution will stop. Then, any localized acidification will disperse by diffusion of the protons to the bulk solution. According to Van Muylder *et al.*,[70] it is only when the metal reaches the potential E_2 that the solution at the anodic region could reach the value pH_{PIT}, **Figure 18**, and the reaction of anodic dissolution becomes thermodynamically possible:

$$Me = Me^{z+} + ze \qquad [1]$$

Soluble products will be produced at the anode, and they will not interfere with the corrosion process. This reaction will be followed by the hydrolysis equilibrium:

$$Me^{z+} + zH_2O \leftrightarrow Me(OH)_z + zH^+ \qquad [2]$$

which will perpetuate the localized acidification in the anodic region.

According to Van Muylder *et al.*[70] mechanism, passivity of the metal by insoluble hydroxides or oxides becomes thermodynamically unstable at potentials above E_2. In other words, the pitting potential is a thermodynamically based value and is given by eqn [1]. This mechanism was successfully used to explain the pitting potential of copper in chloride solutions. Nevertheless, it failed when applied for example to aluminum, as can be seen in **Figure 11**. It also failed when applied to iron, zinc, and so on. It was found that it did not work with those metals whose Me/Me^{z+} equilibrium potential was lower than that for hydrogen evolution. The model did not take into account the reactions of proton consumption that could take place inside the pit, as shown in **Figure 19**.

This mechanism was modified by Galvele and coworkers to explain the pitting potentials of such metals as zinc, cadmium, aluminum, or iron and the pitting potentials found in the presence of reducible anions or inhibitors.[5,9,12,34,82,83,94–98] These authors introduced an electrochemical kinetic explanation for the pitting potential.

The key point of the mechanism is that there is a pH drop on the anode surface and that this pH drop is a necessary condition to sustain pitting. The rate at which the protons are produced will be given by eqns [1] plus [2]:

$$Me = Me^{z+} + ze \qquad [1]$$
$$Me^{z+} + zH_2O \leftrightarrow Me(OH)_z + zH^+ \qquad [2]$$

On the other hand, the consumption of protons can take place by various processes:

by direct reduction:

$$2H^+ + 2e = H_2 \qquad [3]$$

by reaction with reducible anions:

$$NO_3^- + 10H^+ + 8e = 3H_2O + NH_4^+ \qquad [4]$$

by reaction with weak acid anions (L^-):

$$H^+ + L^- \leftrightarrow HL \qquad [5]$$

or by diffusion of protons from the anode region to the bulk solution. The rates for eqns [1], [3], and [4] are

potential dependent, while eqn [5] and diffusion are potential independent.

Assuming that the metal is in an alkaline solution, in absence of reducible anions or buffer ions, but in presence of a strong acid salt, which acts only as a supporting electrolyte, the processes inside the pit can be analyzed. Protons will be produced by reactions [1] and [2], and will be consumed by reaction [3]. Reaction [1] is potential dependent, and the higher the potential, the higher the rate of production of protons. On the other hand, with reaction [3], the lower the potential the higher the rate of consumption of protons. If diffusion of protons from the pit to the bulk solution is ignored, there will be a potential at which the rate of production of protons will be equal to the rate of consumption of protons. This is the potential at which reaction [1] and reaction [3] are equal, and it is the corrosion potential of the metal in the acidified solution inside the pit (E_c^*). According to Galvele and co-workers, the corrosion potential of the metal in the pit-like solution is the minimum potential at which pitting can be sustained, so the pitting potential will be

$$E_p \geq E_c^*$$

Actually, to compensate for the protons diffusing from the pit to the bulk solution, E_p should be slightly higher than E_c^*.

If reducible anions are present, the pitting potential will have to increase to values at which the anion reduction becomes negligible. For example, the pitting potential for pure Al in 1 M NaCl solutions is $\sim -0.53\ V_{SHE}$, while in 1 M $NaNO_3$ solutions it is $+1.70\ V_{SHE}$. Wexler and Galvele[5] reported that during pitting of Al in nitrate solutions bubbles of N_2 are evolved from the pits. This indicates that the following reaction is taking place:

$$2NO_3^- + 12H^+ + 10e = N_2 + 6H_2O$$

This reaction produces a heavy consumption of protons, and as shown by the N–H_2O Pourbaix E–pH equilibrium diagrams, a high potential has to be reached for this reaction to become negligible. This is the explanation offered for the high pitting potentials found for Al in nitrate solutions.

One serious criticism to any acidification mechanism for pitting was raised by Vetter and Strehblow.[99] They studied the pitting of iron in a mixture of chloride plus sulfate solutions with phthalate buffer (pH = 5), and observed the formation of small corrosion pits with a diameter of ~1–50 μm. The authors calculated the composition changes taking place in a 2 μm pit. Their calculations showed that the current density inside the pit was $\sim 9\ A\ cm^{-2}$, the pH change was +0.85, the potential change was −0.018 V, and the change in iron concentration was $+0.9\ mol\ l^{-1}$. These authors concluded that no significant changes of pH, or any ion concentration could be expected in very small pits. Consequently, a second mechanism, other than that of localized acidification, should be assumed.

Reviewing Vetter and Strehblow's results, it is observed that, contrary to what experience has shown about the composition inside the pits, their calculations did not show a decrease in the pH, but an increase of it. As mentioned above, for iron the pH measured inside the pits was ~3.5, while the pH of the bulk solution could have values between 2 and 10.[71,72] The reasons were that the authors chose to ignore the hydrolysis of the metal ions and that they were studying a case of chemical depassivation.

Galvele[95,96] made an analysis of the transport processes taking place inside a pit, taking into account the hydrolysis of the metal ions. This ion transport analysis showed that not only significant acidification was possible in very small pits, but also that a quantitative explanation could be found for the effect of pH, inhibitors, aggressive ion concentration, and so on, on the pitting potential.

The transport analysis was made in the following way. Let us assume, for the sake of simplicity, that we have a pit like the one in **Figure 20**. At the bottom of this pit an anodic reaction of the type

$$Me = Me^{n+} + ne^- \qquad [6]$$

will take place. No reaction takes place at the walls of the pit. Thus, a unidirectional pit model is obtained, with considerable simplification in the mathematical treatment that will follow. If pit initiation took place, for example, at the bottom of a crack in the surface oxide film, **Figure 20** would be an exact description of the process. For a hemispherical pit, **Figure 20** would seem to be a gross oversimplification. But, according to Vetter and Strehblow,[93] the mathematical problem of a hemispherical pit could be transformed into an equivalent plane problem by increasing the diffusion path by a factor of the order of 3. As shown below, when calculating the concentration gradients inside pits, only orders of magnitude are considered, so this correction factor can be ignored.

Reaction [6] is assumed to take place in a sodium or potassium salt of an aggressive but noncomplexing

anion. The pH of the bulk solution, which could have any value, is given as a boundary condition. It is also assumed that reaction [6] is followed by a series of hydrolysis equilibrium reactions of the type:

$$\begin{aligned} Me^{n+} + H_2O &= Me(OH)^{(n-1)+} + H^+ \\ Me(OH)^{(n-1)+} + H_2O &= Me(OH)^{(n-2)+} + H^+ \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ Me(OH)^{2+}_{n-2} + H_2O &= Me(OH)^{+}_{n-} + H^+ \\ Me(OH)^{+}_{n-1} + H_2O &= Me(OH)_n + H^+ \end{aligned} \qquad [7]$$

This equilibrium is known to be very quickly reached,[100] so the reactions in eqn [7] will always be in equilibrium. In the first publication[95] only the first step in reactions in eqn [7] was taken into account. In the second publication[96] all the hydrolysis steps from eqn [7] were considered.

To give an idea of the method used in the calculation, one example of the first publication will be described. The same procedure was used in the other cases. In the first publication[95] only the first step of hydrolysis from eqn [7] was considered. This is a simplified description of the processes taking place inside a pit; but it will give the minimum degree of acidification expected inside the pit.

The flow of ions inside the pit is given by the following equation[101]:

$$\mathcal{J}_i = D_i\frac{dC_i}{dx} + D_i\frac{z_iF}{RT}C_i + \frac{d\phi}{dx} \qquad [8]$$

where $\mathcal{J}_i$ is the flow of ions i, D_i the diffusion coefficient of the same ions, C_i their concentration, x the pit depth as shown in **Figure 20**, $(d\phi/dx)$ the gradient of potential ϕ, z_i the valence of ions i, F the Faraday constant, R the gas constant, and T the absolute temperature.

As shown by Vetter and Strehblow,[99] small ionic concentration changes should be expected at pit initiation. So the sodium salt, in the bulk solution, will act as a supporting electrolyte for the species formed in reactions [6] and [7]. This means that the transport of all these species will be due to ion diffusion. The contribution of ion migration can be ignored, which again contributes to a simplification in the mathematical treatment. Consequently, reaction [8] is reduced to:

$$\mathcal{J}_i = D_i\frac{dC_i}{dx} \qquad [9]$$

As it is assumed that the bulk solution could have any pH value, the first step in reaction [7] has to be rewritten to account for the contribution of the OH-ions in alkaline solutions:

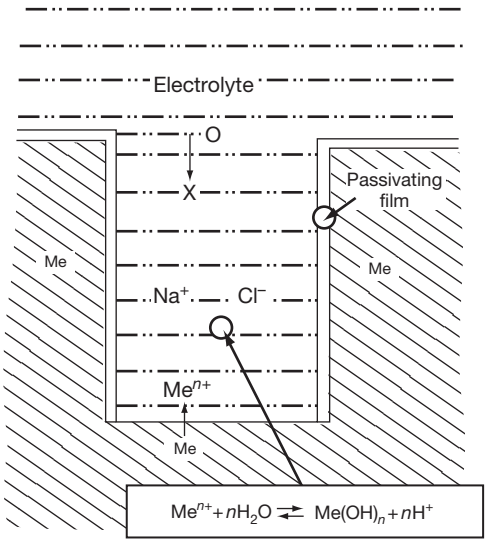


Figure 20 Unidirectional pit model used for calculation of the corrosion species present inside a pit. Reprinted from Galvele, J. R. *Corros. Sci.* **1981**, *21*, 551–579, with permission from Elsevier.

$$2Me^{n+} + H_2O + OH^- = 2Me(OH)^{(n-1)+} + H^+$$

Species S: 1 2 3 4 5 [10]

From eqn [10] we find that inside the pit there are five species (S1–S5), the concentrations of which should be calculated. The detailed mathematical treatment for a system like this can be found in Vetter's book on Electrochemical Kinetics.[101] The five unknown concentrations are calculated by resolution of the five following equations:

the flow of the species containing Me atoms will be given by

$$D_1\frac{dC_1}{dx} + D_4\frac{dC_4}{dx} = \frac{i}{nF} \qquad [11]$$

the flow of the species containing O atoms will be given by

$$D_2\frac{dC_2}{dx} + D_3\frac{dC_3}{dx} + D_4\frac{dC_4}{dx} = 0 \qquad [12]$$

and the flow of the species containing H atoms will be given by

$$2D_2\frac{dC_2}{dx} + D_3\frac{dC_3}{dx} + D_4\frac{dC_4}{dx} + D_5\frac{dC_5}{dx} = 0 \qquad [13]$$

Finally, the two following equilibrium relations must be considered

$$^*K_1 = \frac{C_4C_5}{C_1} \qquad [14]$$

which is the law of mass action applied to the first step of hydrolysis in eqn [7], and

$$K_w = C_3 C_5 \qquad [15]$$

which is the ionic product of water.

The boundary conditions used in the calculations were $C_1 = C_4 = 0$, the values for C_3 and C_5 were given by the pH of the bulk solution. As for the ionic diffusion coefficients, since they were very similar in aqueous solutions, they were taken as $D_1 = D_2 = D_4 = 10^{-5}\,cm^2\,s^{-1}$. The only exceptions were OH^- ions, for which a value of $D_3 = 5.3 \times 10^{-5}\,cm^2\,s^{-1}$ was used, and H^+ ions for which a value of $D_5 = 9.3 \times 10^{-5}\,cm^2\,s^{-1}$ was used. The equilibrium constants for the hydrolysis reactions of eqn [7] were taken from the literature.[102–104]

The concentrations of all the species inside the pit were calculated as a function of the parameter $x.i$, x being the length of the pit with $x = 0$ at the opening of the pit and i the current density of the reaction [6]. **Figure 21** shows the results of the calculations for iron, in a pH 10 solution. The concentration of protons, which remains constant at low $x.i$ values, shows a sharp increase as soon as $x.i$ reaches a value of $10^{-6}\,A\,cm^{-1}$. The next question is to find out what pH is necessary to start pitting. From the Pourbaix diagrams a pH value is found for the transition between corrosion and passivity. This value is shown as a cross identified as (a). As pits initiate with current densities close to $1\,A\,cm^{-2}$, if only $x.i = 10^{-6}\,A\,cm^{-1}$ is required to reach the necessary acidification, pits as small as 0.01 μm could be initiated by this mechanism.

One criticism that could be made to the calculations in **Figure 21** is that the choice of the critical pH is arbitrary. This objection can be disregarded if instead of using the first hydrolysis step, in eqn [7], all the hydrolysis steps are taken into account.[96] **Figure 22** shows an example of these calculations. Again the example given is for iron in a pH 10 solution. For the sake of simplicity, only the iron compounds are shown. For low $x.i$ values the almost exclusive product is solid $Fe(OH)_2$, but above $x.i$ $10^{-6}\,A\,cm^{-1}$ Fe^{2+} becomes the main corrosion product. Another example of the full hydrolysis calculations, in this case for aluminum, can be seen in **Figure 23**.

These calculations were made for the ionic concentration diagrams of the following systems: Zn/Zn^{2+}; Fe/Fe^{2+}; Ni/Ni^{2+}; Al/Al^{3+}; Cr/Cr^{3+}; and Co/Co^{2+}. It has also been shown[95] that these ion concentration diagrams give a quantitative explanation of the influence of solution pH, buffer ions, and inhibitors, on the pitting potential and they also give an explanation of the existence of a pitting repassivation potential.

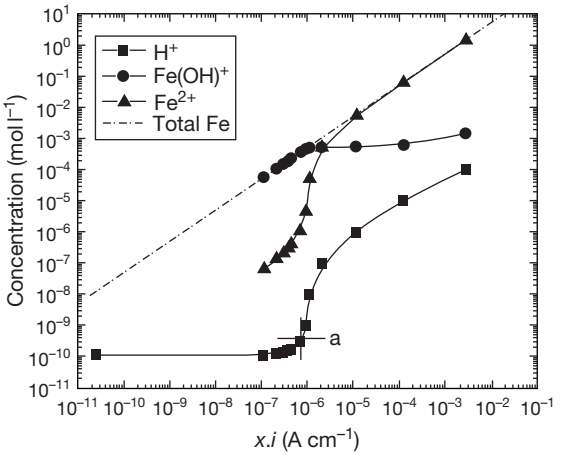


Figure 21 Unidirectional pit model for Fe, assuming that only the first hydrolysis step is relevant. The critical H^+ concentration is shown by the cross (a). Reproduced from Galvele, J. R. *J. Electrochem. Soc*. **1976**, *123*, 464–474, with permission from the Electrochemical Society.

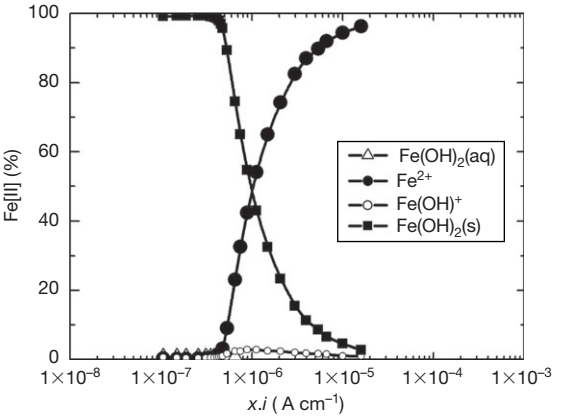


Figure 22 Percentage of metal compounds inside an Fe pit, according to the full hydrolysis pit model. The critical $x.i$ value, for pits to start developing, is taken at the point when solid species concentration is equal to soluble species concentration. Reprinted from Galvele, J. R. *Corros. Sci.* **2005**, *47*, 3053–3067 and Galvele, J. R. *Corros. Sci*. **1981**, *21*, 551–579, with permission from Elsevier.

By changing the external pH value in the boundary conditions of the above mentioned calculations, it is possible to evaluate the effect of pH on the pitting potential. **Figure 24** shows the proton concentrations for unidirectional pits of iron, with external pH values going from 9 to 12. It is observed that by increasing the external pH value, the proton

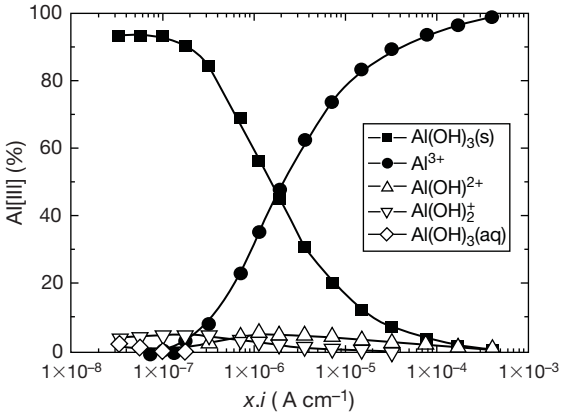


Figure 23 Percentage of metal compounds inside an Al pit, according to the full hydrolysis pit model. Reprinted from Galvele, J. R. *Corros. Sci*. **2005**, *47*, 3053–3067 and Galvele, J. R. *Corros. Sci*. **1981**, *21*, 551–579, with permission from Elsevier.

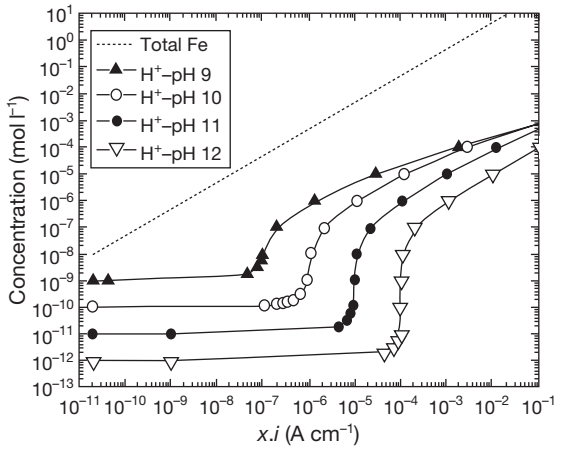


Figure 24 Effect of the pH of the bulk solution on *x.i*. The higher the pH, the more difficult it is to reach the critical *x.i* value for pitting. Reproduced from Galvele, J. R. *J. Electrochem. Soc.***1976**, *123*, 464–474, with permission from the Electrochemical Society.

concentration curve is shifted to higher *x.i* values. For a given *x* value, this means that with higher bulk solution pH values, higher current density values will be required to attain the same acidification. If the current density inside the pit is potential dependent, as it is generally the case, this higher current density requirement will be reflected as higher pitting potential values.

Using the full hydrolysis model the following critical *x.i* values were found:[96]

$$\text{pH} = 0;\ x.i = 5.0 \times 10^{-7}\text{A cm}^{-1}$$
$$\text{pH} = 11;\ x.i = 5.0 \times 10^{-6}\text{A cm}^{-1}$$
$$\text{pH} = 12;\ x.i = 5.0 \times 10^{-5}\text{A cm}^{-1}$$

For pH values above 9 it is observed that an increase in one unit in the pH value leads to an increase of one order of magnitude in the *x.i* value. If the relation between potential and current density inside the pit follows a logarithmic law, the pitting potential should be expected to follow a law of the type

$$E_p = A + B.\text{pH} \qquad [16]$$

where *B* is equal to the Tafel slope for the metal in the acid pit-like solution. Such a relation was found for iron in NaCl solutions (**Figure 13**)[48] and for cadmium in NaCl and in Na_2SO_4 solutions,[9] both at pH values above 10. In all these cases, the *B* value of eqn [16] was very close to the Tafel slope of the metal in the pit-like solution.

In the case of trivalent metals, as it is the case with aluminum, the presence of a further step of hydrolysis slows down the effect of the external pH on the *x.i* value. As shown in **Figure 23**, the transition from passivity to pitting is not as sharp as for divalent metal ions, **Figure 22**. Consequently, for comparison reasons, the critical *x.i* value chosen was the point where the concentrations of solid $Al(OH)_3$ and soluble Al^{3+} were equal, that is, the intersection point of both concentration curves. The results found were[96]:

$$\text{pH} = 5;\ x.i = 1.4 \times 10^{-6}\text{A cm}^{-1}$$
$$\text{pH} = 8;\ x.i = 2.0 \times 10^{-6}\text{A cm}^{-1}$$
$$\text{pH} = 10;\ x.i = 3.1 \times 10^{-6}\text{A cm}^{-1}$$

According to the full hydrolysis calculation,[96] for divalent metal ions an increase of only one unit in the pH of the bulk solution produced an increase as big as one order of magnitude in the *x.i* value. On the other hand, for aluminum an increase of five pH units produced just a two-fold increase in the *x.i* value. The conclusion is that the critical *x.i* value, for aluminum, is practically independent of the external pH. This conclusion is in agreement with the experimental observations made by Kaesche,[65] who found that the pitting potential of aluminum was not affected by the solution pH.

The presence of anions of weak acid salts is also very important for the localized acidification at the bottom of the pit. If weak acid salts are present, the anions of those salts will react with the protons, and hinder the localized acidification process. The equilibrium involved will be

$$HL \leftrightarrow L^- + H^+ \qquad [17]$$

where HL is the weak acid and L is the weak acid anion. The equilibrium constant of this reaction will be

$$K_a = \frac{C_{L-}.C_{H+}}{C_{HL}} \qquad [18]$$

If reaction [17] is introduced as a boundary condition in the ion concentration diagrams, the effect of weak acid salts on the localized acidification can be evaluated. **Figure 25** shows the effect of 10^{-2} M borax on the concentration of the various species in a unidirectional zinc pit. **Figure 26** shows how the concentration of protons is shifted to higher *x.i* values when the concentration of borax is increased. The higher the borax concentration, the higher the *x.i* values necessary to reach the same acidification. An increase of one order of magnitude in the borax concentration leads to an increase of about one order of magnitude in the *x.i* value. If the relationship between the potential and the current density inside the pit follows a logarithmic law, we conclude that the pitting potential of zinc should be related to the borax concentration by an equation of the type

$$E_p = A + B \log C_b \qquad [19]$$

C_b being the concentration of the borax. The value of B should be equal to the Tafel slope of the metal in the pit-like solution. Logarithmic relationships of this type between the pitting potential and the buffer concentration for zinc have been reported by Augustynski *et al.*[35] Relationship [19] was also found to work in the case of pitting of high-purity iron in NaCl solutions in the presence of sodium borate as inhibitor.[84]

The efficiency of a weak acid salt as a pitting inhibitor will depend on the way in which the equilibrium of eqn [17] is shifted. The more the equilibrium is shifted to the left in eqn [17] the more efficiently the salt acts as a pitting inhibitor. This is equivalent to saying that the lower the K_a, value, in eqn [18], the more efficient the inhibitor will be.[95]

The ion concentration diagrams can also explain the existence of a pitting repassivation potential, lower than the pitting potential, **Figure 9**. **Figure 27** is a simplified picture of **Figure 21**. As soon as the system reaches a critical *x.i* value, $x.i_1$ in **Figure 27**, the pit will start to grow. If the potential remains constant, the current density will also be constant, but *x* will increase with time and so will *x.i* ($x.i_2$ in **Figure 27**). If the potential is then lowered, the current density will drop, but the pit will continue to grow while the *x.i* value is higher than $x.i_1$. This means that pits will grow at potentials lower than the initiation potential. Finally the pit will stop growing either because the value *x.i* is lower than the minimum required for the pit growth to occur, or because the electrode potential is lower than the corrosion potential in the solution inside the pit, E_c^*. This mechanism explains the observation made by Suzuki and Kitamura[53]: the deeper the pit the lower the repassivation potential.

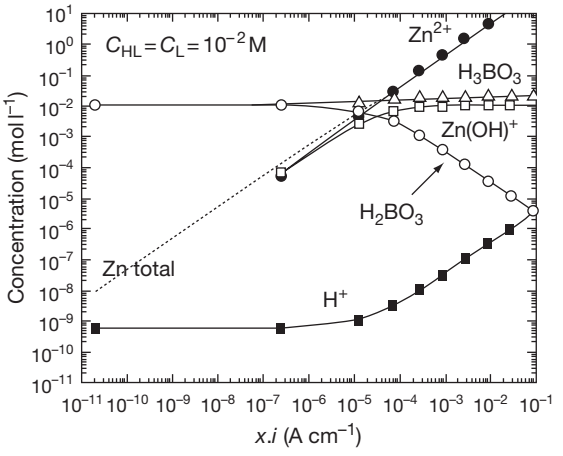


Figure 25 Effect of a buffer on the proton concentration inside a pit of for Zn. The maximum buffer capacity is found when the concentrations of HL and L^- are equal. When the concentration of L^- begins to decrease, the concentration of H^+ increases and the pit becomes stable. Reproduced from Galvele, J. R. *J. Electrochem. Soc.***1976**, *123*, 464–474, with permission from the Electrochemical Society.

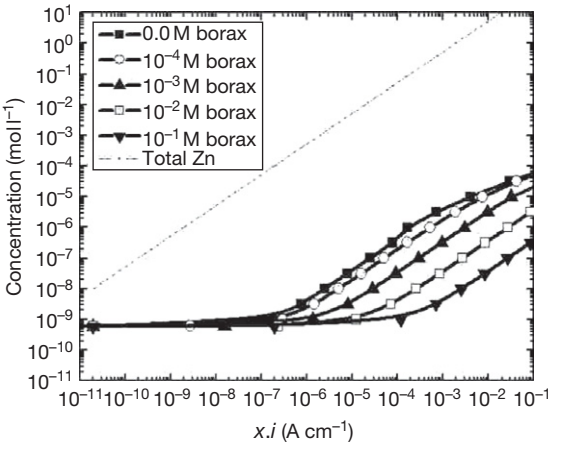


Figure 26 Effect of buffer concentration on the *x.i* for Zn. Reproduced from Galvele, J. R. *J. Electrochem. Soc.***1976**, *123*, 464–474, with permission from the Electrochemical Society.

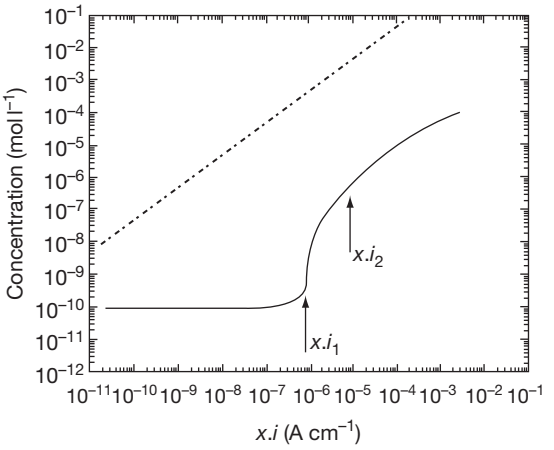


Figure 27 Origin of the repassivation potential, E_r. At $x.i_1$ a pit starts to grow and the potential is E_p. At constant potential i remains constant but x increases, leading the system to $x.i_2$. If the potential is lowered, the growth of the pit will not stop until $x.i_1$ is reached. As the value of x has increased, the value of i must be lower, which means a lower E value. The new E value at which pitting stops is known as E_r.

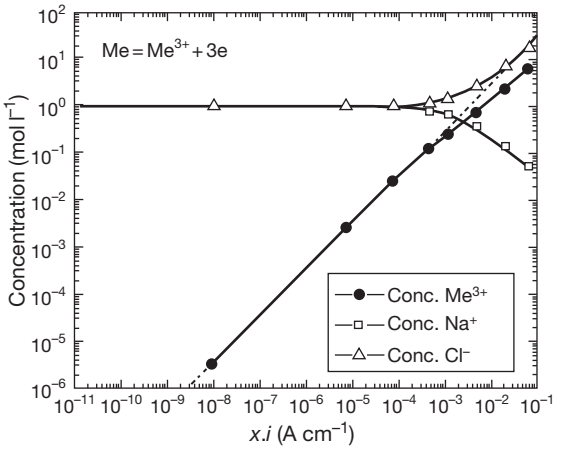


Figure 28 Migration effect inside a pit in a 1 M NaCl solution. The contribution of migration begins when the concentrations of Na^+ and Cl^- start to diverge. The migration also affects the movement of Me^{3+} ions, observed in the deviation of the concentration line at lower $x.i$ values. Reproduced from Galvele, J. R. *J. Electrochem. Soc.***1976**, *123*, 464–474, with permission from the Electrochemical Society.

If the concentration of the aggressive anion salt is changed over a wide range of values, the assumption that this salt is acting as a supporting electrolyte, which leads to the use of eqn [9], is not valid. Eqn [8] is to be used, which indicates the presence of an extra electrical potential inside the pit. **Figure 28** is a simplified diagram where the appearance of a migration phenomenon at high $x.i$ values is shown. For the sake of simplicity, the hydrolysis reactions have been ignored. The electrical potential, due to the migration phenomenon, starts to increase as soon as the concentrations of Na^+ ions and Cl^- ions from the NaCl solution start to diverge. The electrical potential values, for different NaCl concentrations, and $x.i$ values are shown in **Figure 29**.

An important conclusion from the transport process studies[94–97] was the nature of the pitting potential. It was concluded that the pitting potential had no relation with thermodynamic equilibrium, as suggested by Van Muylder *et al.*,[70] but it was given by the following equation:

$$E_p = E_c^* + \eta + \phi \qquad [20]$$

E_p being the pitting potential, E_c^* the corrosion potential in the pit-like solution, η the polarization needed to reach the critical $x.i$ value, and ϕ the electrical potential induced by the migration of the aggressive anions to the bottom of the pit. The validity of eqn [20] was tested in numerous

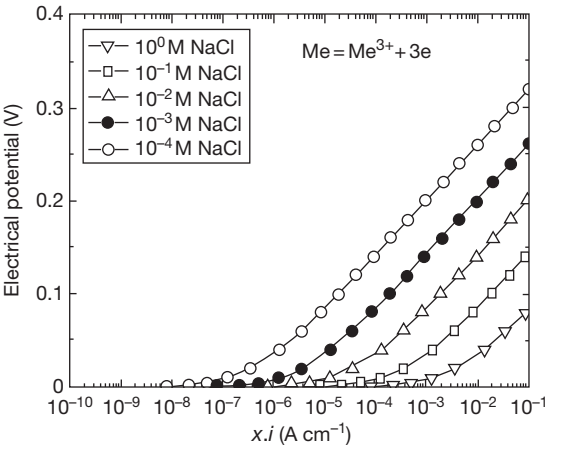


Figure 29 Effect of the NaCl concentration on the potential ϕ inside the pit because of migration. Reproduced from Galvele, J. R. *J. Electrochem. Soc.* **1976**, *123*, 464–474, with permission from the Electrochemical Society.

publications[38,44,82,83,97,98] and some of the results are shown in **Table 3**. Eqn [20] was also confirmed in an independent work by Newman *et al.*[105] Besides, calculations and measurements made on high purity zinc in NaCl solutions with various concentrations of borate buffer, under conditions where the aggressive solution was not acting as a supporting

electrolyte, and eqn [8] had to be used, also confirmed the validity of eqn [20] for the calculation of the pitting potential.

2.04.9 Passive Film Breakdown

In the previous section, the mechanism of propagation of stable pits was discussed. But, a previous step is the passive film breakdown, a problem that was anticipated in **Figure 17**. The breakdown of the passive film and the very initiation of pits are two of the not clearly understood aspects of the pitting process.[106] One source of confusion was the fact that no distinction was made between the three different types of pitting described at the beginning of this chapter. Thus, very often mechanisms were suggested for all types of pitting only on the basis of experiments either on chemical depassivation or on electrochemical depassivation. Another source of confusion was related to the nature of passivity. As pitting is a process of passivity breakdown, any mechanism of pitting had to be related to some mechanism of passivity. For a long time, there were strong discussions between those who maintained that passivity was due to an adsorbed monolayer of oxygen and those who claimed that passivity was due to the formation of a three-dimensional oxide film on the metal surface. Those supporting the idea of an adsorbed monolayer based their pitting mechanisms on the idea of competitive adsorption between the passivating and the aggressive species on the metal surface.[107,108] Those in favor of a passive three-dimensional oxide film tried to find a way to undermine that film. As it is now known that the oxide films present on passive metals are at least several monolayers thick, in this chapter it will be assumed that the metal is initially covered by a three-dimensional passivating oxide film. Theories of pit initiation have been classified into two main groups: film penetration and film breaking.

2.04.9.1 Ion Penetration Mechanisms

The early literature on passive films assumed that the oxide film was acting as a static barrier that isolated the metal from the solution. In order to initiate pitting under such conditions some reaction had to take place to bring the metal in contact with the aggressive ions of the solution. One possible way to initiate pits was for the aggressive ions to migrate into the passive film, thus contaminating it. Anion migration would be stimulated by the high electric field in the film. As a result of changes of the properties of the passive film, high currents could start to circulate through the contaminated zones and pitting would start.[109] Nevertheless, Vetter and Strehblow[93] pointed out that the migration of polyatomic aggressive anions such as $SO_4^{=}$, ClO_4^{-}, or SCN^{-} through a solid layer was not conceivable.

The contamination of the passive films by aggressive anions has been confirmed in several cases. In others, the results obtained by different techniques are conflicting. Basing his studies on impedance measurements, Pryor[110] suggested that migration of chloride ions took place into anodized aluminum films. McBee and Kruger,[111] by ellipsometric studies on pure iron, and on Fe–Cr and Fe–Cr–Mo alloys, found that the injection of chloride ions in the medium produced changes in the properties of the oxide film, probably by penetration of the film. These changes were reversible and the properties of the oxide film were recovered after removing the chloride ions from the medium. Nevertheless, it could not be confirmed that the observations made by McBee and Kruger were due to migration of chloride ions into the film. Marcus *et al.*[112] found chloride present in passive films on nickel using electron spectroscopy for chemical analysis (ESCA) and *in situ* radiotracer measurements. The establishment of a critical concentration of chloride in the film is a prerequisite for the localized breakdown of the film. On the other hand, other authors[113–115] found no incorporation of aggressive ions into the passive film prior to pitting.

The point-defect model for passive film growth developed by Lin *et al.*[116] has been used to explain pit initiation by assuming that the aggressive anion adsorption at the outer layer of the barrier film results in the formation of cationic vacancies. The excess of vacancies at the metal/film interface will lead to the formation of voids and pitting initiation.

2.04.9.2 Film Breakdown Mechanisms

There are two different film breakdown models. In one of them, pitting is considered to start at flaws always present in the oxide film[117,118] while in the other pit initiation results from a continuous state of breakdown and repair of the passive film.[119,120] In these models the role of the aggressive anions is to prevent repassivation at the local breakdown sites.

There are theoretical reasons to expect the existence of flaws in the passive film. Fromhold[121]

Table 3 Comparison between experimental pitting potential values, measured in 1 M NaCl solution and theoretical values calculated with Eqn [20]. Potentials in V_{SHE}

Alloy	E_p *theoretical*	*Ref.*	E_p *experimental*	*Ref.*
Al 99.99%	−0.54	83	−0.53	1
Al–3Cu	−0.43	83	−0.43	82
Al–3 Mg	−0.56	83	−0.53–0.57	83
Al–3Zn	−0.76	83	−0.75	83
Fe–18Cr	0.60	44	0.58–0.62	80
Fe–18Cr–1Mo	0.73	44	0.64–0.70	80
Fe–18Cr–2Mo	0.88	44	0.70–0.82	80
Fe–18Cr–5Mo	1.03	44	1.00–1.04	80

pointed out that the act of forcing an ionic species through a compact passive film to produce additional film growth will seriously affect the compact film itself. He found that the electrochemical potential gradients inside the film give rise to extremely large stress values, high enough to produce the mechanical breakdown of the passive film and to transform the crystalline structure of the film into an amorphous one. On the other hand, Sato[122] found that mechanical stresses resulting from electrostriction and surface tension effects may exceed the breakdown stress of the oxide film. In a later paper, the author predicted[123] that there is a critical potential above which breakthrough pores would be formed in the film. This critical film breakdown potential is lower as the surface tension at the metal–electrolyte interface is smaller, and hence, the anion adsorption is stronger.

2.04.9.3 Passive Film Breakdown and Metastable Pits

If, as suggested by the film breakdown theories, flaws really exist in the passive film, and if for sustained pit growth a critical $x.i$ value has to be reached, it is reasonable to expect that, below the pitting potential, incipient processes of pit initiation should exist. Indications of passive film breakdown and repassivation events were reported by Galvele *et al.*[124,125] when studying the repassivation rate of various metals and alloys. The current decay, when measuring repassivation rates, could be described by a relation of the type

$$i = A.t^{b}$$

where, i is the current density, t is the time and A and b are constants. A value of $b = -1$ is typical of a high field film formation reaction, and it is usually related to the formation of a compact highly protective passive film. Galvele *et al.*[124,125] studied the repassivation rate for zirconium in various halide solutions, as well as for brass in pH 12, 1 M $NaNO_2$ solution. They observed that at potentials well below the pitting potential the value of b was −1. Then there was a region of potentials where an initial value of $b = -0.5$, typical of a diffusion controlled process,was observed, but after a few seconds the current density decreased and returned to values expected for a repassivation rate of $b = -1$. At higher potentials, close to the pitting potential, the value of b increases above $b = -0.5$, reaching eventually positive values, as is typical in pitting processes. The intermediate potential region, in which the initial repassivation rate is not as fast as at lower potentials but after a short time the highly protective oxide film is reformed, was identified by the authors as the region where passivity breakdown takes place. They also concluded that the repassivation rate technique was very convenient for detecting the passivity breakdown. No other references were found in the literature where this line of work was followed.

On the other hand, many investigators have observed that the formation of a stable pit is preceded by electrochemical noise. The noise events took the form of transient fluctuations of potential under galvanostatic control, or of current fluctuations under potentiostatic control. Such transients have been reported for many years in stainless steel,[51,126,127] and Al,[7,118,128] and were originally attributed to passive film breakdown and repassivation events.[51,122,123,129]

Afterwards, it was found that each of these electrochemical noise events originates from the formation, growth, and repassivation of a microscopic pit.[130–134] These short life pits, now referred to as metastable pits, can form at potentials well below the pitting potential (which is associated with onset of stable pit growth) and during the induction time before the initiation of stable pitting at potentials above the pitting potential.

Metastable pit growth is a characteristic feature of the pitting corrosion of stainless steel in chloride solutions,[130–135] although indications of metastable pitting have also been reported for other metals and alloys such as iron–chromium alloys,[136] aluminum,[137] and titanium.[138] Metastable pits grown on stainless steel are covered by a layer that is a remnant of the passive film[130,131]; the rupture of their covers results in dilution of the local pit environment and repassivation of the pits.[45,139]

Metastable pitting transients exhibited by stainless steels in chloride solutions have a charasteristic shape

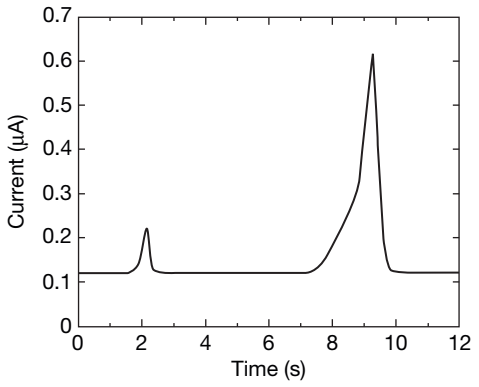


Figure 30 Current transients typical of pre-pits.

(**Figure 30**). The current increases above the background passive current as the pit nucleates and begins to grow. After a short growth period, the metastable pit repassivates and the current quickly decreases to the level of the original passive current. At open circuit, metastable pitting events result in potential transients with a typical shape consisting of a rapid potential drop followed by a slow increase.[140]

There appear to be no discernible differences between metastable pit growth and the early growth of a stable pit.[134] Therefore, studies of metastable pits were performed in order to gain knowledge about the critical factors that determine the stable growth of pits. Two main lines of information may be obtained from such investigations: stochastic data such as a relationship between the nucleation frequencies of metastable and stable pits,[133,141,142] and kinetic data deduced from the analysis of individual current transients.[130,131]

Frankel *et al.*[130] showed that during the metastable pit growth the pit current increases approximately with the square of time. Making the assumption that the pits associated with these transients are hemispherical in shape, they concluded that the current density was approximately constant with time during most of the metastable pit growth. The transients showed lifetimes as long as 15 s, but typically less than 5 s, and their magnitude varied from 4 μA to 20 μA. Other authors found a growth-rate distribution of metastable pits at each potential and that the metastable pit current follows a power law relationship $I \sim t^n$ with n between 0.5 and 1.5.[133,142–144]

According to Frankel *et al.*,[130] metastable pit growth is controlled by the Ohmic resistance of the porous pit cover. Pits repassivated when their cover ruptured and the Ohmic barrier was lost. On the other hand, if a salt film is present when the cover breaks, pit stability is maintained.

Pistorius and Burstein[131,145–147] believe that the rate of growth of individual pits is controlled by diffusion of metal cations from the pit interior, the surface of which is covered with a salt film even before the cover is lost. During metastable growth, the pit depth is too small to account for the diffusion barrier but the perforated cover over the pit mouth provides an additional diffusion barrier which enables the concentrated pit solution to be maintained. The transition to stable pit growth occurs when the pit has grown sufficiently for its own depth to act as a diffusion barrier. The critical condition under which this transition occurs is determined by the magnitude of the product of the pit current density, i, and its depth, a, called the 'pit stability product.' All pits grow initially in the metastable condition, with a pit stability product which increases linearly with time, but below the critical value. Once the critical value of ia is exceeded, pit growth is stable and the pit propagates in the absence of the cover.

Because of the small anodic currents generated during individual metastable events and to avoid these transients being obscured by the passive current or simultaneous transients, experiments must be performed by exposing a restricted surface area of the specimen tested. Most metastable pitting research has been carried out using embedded wires of less than 1 mm diameter (10–500 μm).[130,131,134,135,137,145–148] Events of amplitude down to *ca.* 1pA in amplitude can be detected if the electrode is small enough. Using very low noise systems, some authors[138,149–151] have observed small current transients, which exhibit a shape different from that of typical metastable pits in that the current increased sharply and then decayed relatively slowly to the background. These peaks were attributed to nucleation and subsequent repassivation events.

Williams *et al.* have presented a statistical model relating the probability of forming a large propagating pit to the nucleation, death, and propagation characteristics of metastable pits.[132,139,142] The proposed statistical model states that the probability that a stable pit is formed is equal to the probability of forming a metastable pit times the probability that the metastable pit survives to become a stable pit. Thus, the nucleation frequency for stable pits is expressed in terms of the nucleation frequency for metastable pits and a probability of repassivation.[133,141,142,152] Underlying this approach is the

concept that if metastable pits form then a stable pit is sure-to-happen at some time.

2.04.10 Pitting of Aluminum Alloys

2.04.10.1 Pitting and Intergranular Corrosion

In spite of having a rather low pitting potential, aluminum and certain aluminum alloys are successfully used in marine environments, the reason being that the passive film on aluminum is not a good electron conductor and the electrochemical reduction of oxygen is hindered. This is not the case if copper is present in the alloy. It has been found that the presence of copper in the alloy facilitates the reduction of oxygen, **Figure 31**. Thus copper becomes a deleterious contaminant in these environments, because through the cathodic reaction of oxygen, the presence of copper causes an increase in the electrode potential, which gets above the pitting potential of the aluminum alloy.

Pitting can also be the cause of other types of localized corrosion in aluminum–copper alloys. For example, it has been shown[1] that the susceptibility to intergranular corrosion of aged Al–4%Cu alloys is related to the difference in pitting potentials between a solute-depleted zone, surrounding the intermetallic precipitates along the grain boundaries, and that of the grain bodies. As shown in **Figure 12**, the pitting potentials of the grains in chloride containing solutions were almost 100 mV higher than those of the copper-depleted zone at the grain boundaries. **Figure 12** can be divided into three zones. The first one is below the pitting potential of Al–0.2%Cu, which is in equilibrium with the intermetallic Al_2Cu. In this zone, the alloy will remain passive. The second zone in **Figure 12**, is between the pitting potentials of Al–0.2%Cu and Al–4%Cu. In this area of potentials only the copper depleted zone along the grain boundaries will corrode, giving intergranular corrosion. Finally, in the third zone, in the upper part of the figure, all the phases of the alloy will corrode. Intergranular corrosion would not appear in every corrosive medium, but only in those where the aluminum alloys showed pitting, such as Br^-, I^-, and ClO_4^-, as shown in **Table 4**. The difference in the pitting potentials was a function of the heat treatment of the alloys and for certain heat treatments the difference was so small that the alloys were rendered immune to intergranular corrosion.[82]

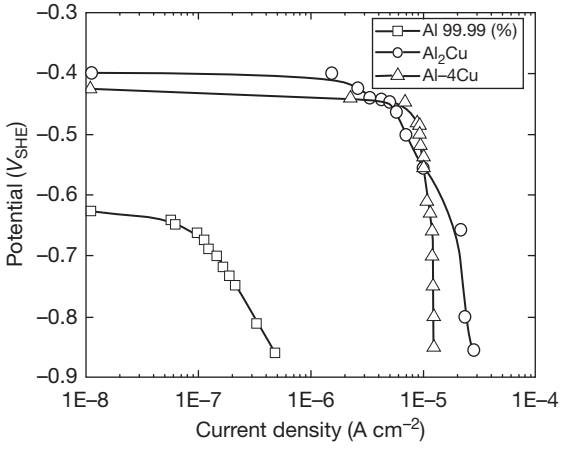


Figure 31 Cathodic polarization curves in air-saturated 1 M NaCl of pure aluminum, solubilized Al–4%Cu, and intermetallic Al_2Cu. Reprinted from Galvele, J. R.; De Micheli, S. M. de. *Corros. Sci.* **1970**, 10, 795–807, with permission from Elsevier.

2.04.11 Pitting of Stainless Steels

2.04.11.1 Pitting and Crevice Corrosion

Pitting and crevice corrosion of stainless steels in chloride solutions are frequently found together, the most frequent observation being that stainless steels will fail most frequently by crevice corrosion, rather than by pitting. The prevalence of crevice corrosion, over pitting, has led to the development of a special cell, known as the Avesta cell, where pitting of stainless steels could be measured without the interference of crevice corrosion.[153]

Susceptibility to crevice corrosion was found to be related with η in eqn [20],[38] as shown in **Table 5**. This observation, together with the explanation of the nature of the repassivation potential, E_r, given in **Figure 27**, helps to explain the empirical relation found by Wilde and Williams between E_r and the susceptibility to crevice corrosion of stainless steels in sea water.[54]

Wilde and Williams found an empirical relation that related the difference E_p–E_r and the susceptibility to crevice corrosion in sea water. The anodic reaction inside the pit allows to explain the data in **Table 5**, as well as the empirical relation found by Wilde and Williams. These authors showed that the larger the difference between the pitting potential and the repassivation potential, E_p–E_r, the more susceptible the alloy was to crevice corrosion. **Figure 32** shows some of their results. It is important to notice that the technique developed by Wilde and Williams

Table 4 Relation between pitting potentials and intergranular corrosion of aged Al–4%Cu alloys in solutions other than NaCl

Electrolyte	*Pitting potentials (V_{SHE})*		*Intergranular corrosion*[a]
	99.99% Al	*Al–4%Cu (solub.)*	*Al–4%Cu (aged)*
1 M KBr	−0.42	−0.28	Yes (−0.40)
1 M KI	−0.26	−0.13	Yes (−0.23)
1 M $NaClO_4$	−0.17	+0.03	Yes (−0.14)
1 M $NaNO_3$	+1.80	+1.85	No (+1.85)
1 M Na_2SO_4	>+3.0	>+3.0	No (+3.0)
1 M H_2SO_4	>+3.0	>+3.0	No (+3.0)
0.1 M $NaIO_3$	>+3.0	>+3.0	No (+3.0)
1 M Ammonium tartrate	>+3.0	>+3.0	No (+3.0)

[a]In parentheses, potentials at which the intergranular corrosion susceptibility was tested.
Reprinted from Galvele, J. R.; De Micheli, S. M. de. *Corros. Sci.* **1970**, *10*, 795–807, with permission from Elsevier.

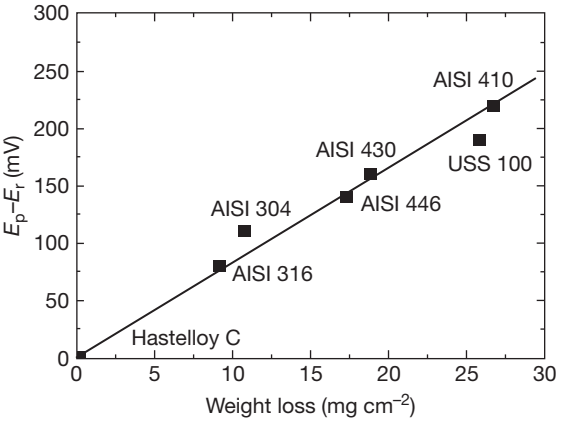


Figure 32 Comparison between the difference E_p–E_r, measured potentiokinetically in a NaCl solution, and crevice corrosion measured after long exposures in sea water. Reprinted from Wilde, B. E.; Williams, E. *Electrochim. Acta* **1971**, *16*, 1971–1985, with permission from Elsevier.

Table 5 Overpotential values, η, from eqn [20], *versus* crevice corrosion susceptibility for various metals and alloys in NaCl solutions

Metal	*η mV*	*Crevice corrosion*	*Reference*
Al–3Zn	10	No	83
Cadmium	~20	No	9
Zinc	~20	No	34
Iron	~30	No	48
Al–3Cu	40	No	82
Al–3 Mg	150	No	83
Aluminum	170	No	82
Nickel	470	Yes	154
Fe–18Cr	670	Yes	44
Fe–18Cr–1Mo	740	Yes	44
Fe–18Cr–2Mo	880	Yes	44
Fe–18Cr–5Mo	1000	Yes	44

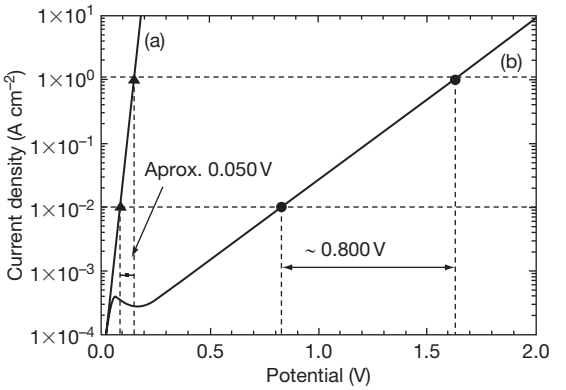


Figure 33 Susceptibility to crevice corrosion in NaCl solutions based on the relation inside the pit, for iron (a)[48] and for stainless steel (b)[44] the potential scale is arbitrary, and was chosen for illustrative purposes. In this example the following assumptions are made: critical value $x.i = 10^{-6}$ A cm^{-1}, pit size $x = 10^{-6}$ cm, crevice size $x = 10^{-4}$ cm. In case (a) the difference in potential required to start pitting corrosion from that to start crevice corrosion will be of only 50 mV. In case (b) crevice corrosion will start at a potential 800 mV lower than that required to start pitting.

assumes that the depth of the pits is the same in all the samples under comparison and that this condition is achieved when the current density reaches 0.2 A cm^{-2}. Nevertheless, this condition will be true only when the number of pits in all the tested samples is similar.

The explanation for Wilde and Williams's correlation could be found in **Figure 33**. We will assume that the current density, when pitting starts, is of the order of 1 A cm^{-2}. We will also assume that curve (a) shows the Tafel curve for iron in a pit-like solution in NaCl and curve (b) is a pseudo-Tafel curve for stainless steel, this last value being on the basis of Galvele *et al.*[44] If the critical *x.i* value is of the order of $x.i = 10^{-6}$ A cm^{-1}, pitting could start on any flaw of 10^{-6} cm. If the pit is allowed to grow up to 10^{-4} cm, to reach again the critical *x.i* value the current density will have to drop to 10^{-2} A cm^{-2}. In the case of curve (a) the difference E_p–E_r will be of only 50 mV. On the other hand, the difference E_p–E_r for curve (b) will be

of 800 mV. If, instead of a pit of 10^{-4} cm, we have a crevice of the same size, in case (a) the difference between pitting and crevice corrosion will be of only 50 mV, a value usually obscured by the corrosion potential oscillations in practical cases, and it could be concluded that metal (a) is not susceptible to crevice corrosion. On the other hand, for metal (b), crevice corrosion will start at a potential 800 mV lower than that required for pitting, and this metal will be more susceptible to crevice corrosion than to pitting.

The relation between pitting and crevice corrosion was supported by Wood *et al.*[118] They studied the evolution of crevice corrosion of AISI type 304 austenitic stainless steel under a 13 mm diameter × 0.3 mm glass microscope cover slide, at constant potential, in an air saturated 5% NaCl solution of pH 8.0. The development of crevice corrosion was followed with an appropriate optical system. They observed that while at 600 mV_{SCE}, pits nucleated on the entire metal surface, at 400 mV_{SCE}, pits nucleated only under the glass cover and spread laterally to take the shape of crevice corrosion. According to these authors, their study confirmed the close link between pitting and crevice corrosion in the case of stainless steels, and it strongly suggested that crevice corrosion is really no more than 'lateral pitting' occurring within an occluded area.

2.04.11.2 Critical Pitting Temperature (CPT)

Pitting corrosion of stainless steel and nickel base alloys is characterized by the existence of a CPT, which is defined as the lowest temperature at which the growth of stable pits is possible. At temperatures lower than CPT no stable pitting can occur and only transpassive corrosion takes place. Brigham and Tozer[78] introduced the concept of the CPT as a criterion for ranking resistance of different austenitic stainless steels to pitting corrosion: the higher the CPT, the more resistant the alloy is to pitting.

CPT can be determined by exposing steel samples to an oxidized chloride solution such as $FeCl_3$ and increasing the temperature in a sequence of test periods until pitting is observed.[156]

The CPT can also be determined electrochemically, either by varying the potential at different constant temperatures or by varying the temperature at different constant potentials.[156–159] At temperatures lower than CPT, passivity breakdown is observed at extremely high anodic potentials, corresponding to transpassive dissolution. Just above the CPT, pitting can occur and breakdown potentials far below those required for transpassivity are measured. The potential independent CPT is then determined as the temperature at which a sudden decrease in breakdown potential from the transpassive range to the pitting range is observed (**Figure 34**). The CPT of many stainless steels falls in the range of 10–90 °C.[78,159]

If the simultaneous occurrence of crevice corrosion is eliminated, as is the case when the Avesta cell[153] is used, the CPT can be defined with an absolute accuracy of 1–2 °C.[158,159] On the other hand, when crevice corrosion occurs, localized corrosion initiates at a temperature lower than the CPT, which is called critical crevice temperature (CCT) and the transition from high to low breakdown potentials happens over a temperature interval in which intermediate values of the breakdown potential are observed.[160]

The CPT value is virtually independent of pH in the range 1–7,[158] and of chloride concentration in the range 0.01–5 M.[47,158] However, the CPT is affected by the sulfate concentration,[161] and changes appreciably with surface roughness.[162]

Metastable pitting could occur at temperatures below the CPT, but stable pitting was only detected at the conventional CPT value.[135,160,163] This suggested that the CPT transition was related to the ability of a metastable pit to become stable. Assuming that an anodic salt film must be maintained in all stable pits, CPT has been interpreted as a competition between salt film and oxide formation within the pit.[47,163] Below CPT, passivation would prevent pits from achieving the current density necessary to maintain a concentrated local chemistry and the salt

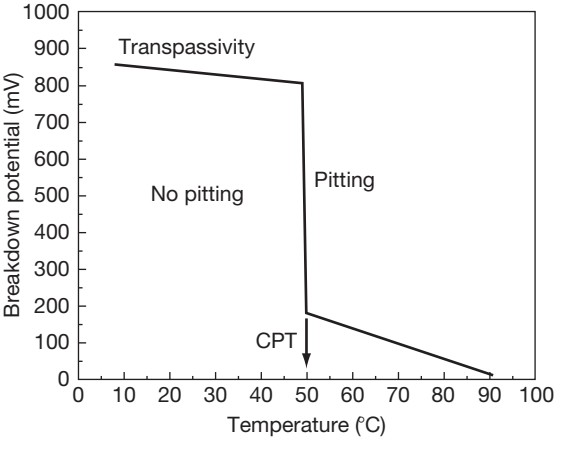


Figure 34 Critical pitting temperature.

layer precipitation is impossible. A slight modification of this model proposed that the CPT is the temperature below which the salt precipitation, rather than stabilizing the anodic process, mediates the oxide passivation in a manner similar to that of the passivation of iron in sulfuric acid.[46] A further refinement of this theory suggests that a stable pit must be able to undercut the surface, thus creating a passivated lacy metal cover that retains the pit contents;[144,164,165] thus the critical event that cannot occur below the CPT is a high rate active film-free dissolution at the undercutting site inside the pit.

References

1. Galvele, J. R.; De Micheli, S. M. de. *Corros. Sci.* **1970**, *10*, 795–807.
2. Dallek, S.; Foley, R. T. *J. Electrochem Soc.* **1976**, *123*, 1775–1779.
3. Metzger, M. *J. Phys. (Colloque C4, supplément 10).* **1975**, *36*, C4–387; C4–393.
4. Freiman, L. I.; Kolotyrkin, Ya. M. *Z Metallov* **1966**, *2*, 488–490.
5. Wexler, S. B.; de Galvele, J. R. *J. Electrochem. Soc.* **1974**, *121*, 1271–1276.
6. De Micheli, S. M.; de Galvele, J. R. *Metalurgia (ABM)* **1971**, *27*, 589–596.
7. Galvele, J. R.; De Micheli, S. M.; de Muller, I. L.; Wexler, S. B.; de Alanis, I. L. In *Localized Corrosion*; U. R. Evans Conference on Localized Corrosion, Williamsburg, Virginia, USA, December 6–10, 1974; Staehle, R. W., Brown, B. F., Kruger, J., Agrawal, A., Eds.; National Association of Corrosion Engineers: Houston, Texas, USA, 1971; pp 580–599.
8. Augustynski, J. *Corros. Sci.* **1973**, *13*, 955–965.
9. Alvarez, M. G.; Galvele, J. R. *J. Electrochem. Soc.* **1980**, *127*, 1235–1241.
10. Kolotyrkin, Ya. M.; Freiman, L. I. *Dokl. Akad. Nauk. SSSR* **1965**, *162*, 376–379.
11. Vetter, K. J.; Strehblow, H. H. *Ber. Bunsen-Ges. Phys. Chem.* **1970**, *75*, 449–455.
12. Semino, C. J.; Galvele, J. R. *Corros. Sci.* **1976**, *16*, 297–306.
13. Strehblow, H. H. Abstract 96. *The Electrochemical Society Extended Abstracts*; Fall Meeting, Las Vegas, Nevada, Oct. 17–22, 1976; p 264.
14. Hoar, T. P. In *Electrode Processes (Discussions of the Faraday Society, No 1, 1947)*; Reprinted by Butterworths: London, UK, 1961; pp 299–300.
15. Freiman, L. I.; Kolotyrkin, Ya. *Corros. Sci.* **1965**, *5*, 199–209.
16. Tousek, J. *Corros. Sci.* **1972**, *12*, 15–28.
17. Kodama, T. *Proceedings Fifth International Congress Metallic Corrosion*; Tokyo: Japan, May 1972; p 223, National Association of Corrosion Engineers, Houston, Texas, USA, 1974.
18. Ammar, I. A.; Darwish, S. *Electrochim. Acta* **1968**, *13*, 781–789.
19. Ammar, I. A.; Darwish, S.; Riad, S. *Electrochim. Acta* **1968**, *13*, 1875–1882.
20. Böhni, H.; Uhlig, H. H. *Corros. Sci.* **1969**, *9*, 353–355.
21. Glazkova, S. A.; Shvarts, G. L.; Freiman, L. I.; Tavadze, F. N. *Z. Metallov* **1974**, *10*, 9–16.
22. Hoar, T. P.; Jacob, W. R. *Nature.* **1967**, *216*, 1299–1301.
23. Kolotyrkin, Ya. M.; Golovina, G. V.; Florianovich, G. M. *Dokl. Akad. Nauk SSSR* **1963**, *148*, 1106–1109.
24. Schwenk, W. *Corrosion* **1964**, *20*, 129t–137t.
25. Ijzermans, A. B *Corros. Sci.* **1970**, *10*, 607–615.
26. Brennert, S. *Tech. Publ. Inst. Tin Res. Dev. Council D.* **1935**, No. 2. Quoted by: Hoar, T. P. In *Electrode Processes* (*Discussions of the Faraday Society, No. 1, 1947*), pp 299–300; Butterworths: London, UK, 1961.
27. Hoar, T. P. *Trans. Faraday Soc.* **1937**, *33*, 1152.
28. Beck, T. R. *J. Electrochem. Soc.* **1973**, *120*, 1310–1324.
29. Abd Rabbo, M. F.; Boden, P. J. In *Localized Corrosion*; U. R. Evans Conference on Localized Corrosion, Williamsburg, Virginia, USA, December 6–10, 1971; Staehle, R. W., Brown, B. F., Kruger, J., Agrawal, A., Eds.; National Association of Corrosion Engineers: Houston, Texas, USA, 1974; pp 653–660.
30. Breiter, M. W. *Electrochim. Acta* **1970**, *15*, 1195–1200.
31. Dugdale, I.; Cotton, J. B. *Corros. Sci.* **1964**, *4*, 397–411.
32. Kolotyrkin, Ya. M.; Gilman, V. A. *Dokl. Akad. Nauk. SSSR* **1961**, *137*, 642–645.
33. Kolotyrkin, Ya. M. In *Proceedings First International Congress Metallic Corrosion*; Butterworths: London, 1962; p 10.
34. Alvarez, M. G.; Galvele, J. R. *Corrosion* **1976**, *32*, 285–294.
35. Augustynski, J.; Dalard, F.; Sohm, J. C. *Corros. Sci.* **1972**, *12*, 713–724.
36. Cragnolino, G.; Galvele, J. R. In *Passivity of Metals*; Proceedings of The Fourth International Symposium on Passivity, Warrenton, Virginia, USA, October 17–21, 1977; Frankenthal, R. P, Kruger, J., Eds.; The Electrochemical Society, Inc.: Princeton, New Jersey, USA, 1978; pp 1053–1057.
37. Gilman, V. A.; Kolotyrkin, Ya. M. *Z. Metallov* **1966**, *2*, 360–361.
38. Galvele, J. R. In *Passivity of Metals*; Proceedings of The Fourth International Symposium on Passivity, Warrenton, Virginia, USA, October 17–21, 1977; Frankenthal, R. P, Kruger, J., Eds.; The Electrochemical Society, Inc.: Princeton, New Jersey, USA, 1978; pp 285–327.
39. Piron, D. L.; Koutsoukos, E. P.; Nobe K. *Corrosion* **1969**, *25*, 151–156.
40. Nobe, K.; Tobias;, R. F. *Corrosion* **1964**, *20*, 263t–266t.
41. Condit, D. O. In *Proceedings Fifth International Congress Metallic Corrosion*; Tokyo, Japan, May 1972; pp 160–163, Published by the National Association of Corrosion Engineers: Houston, Texas, USA, 1974.
42. Engell, H. J.; Stolica, N. D. *Arch. Eisenhüttwes* **1959**, *30*, 239–248.
43. Ives, M. B. In *Localized Corrosion*; U. R. Evans Conference on Localized Corrosion, Williamsburg, Virginia, USA, December 6–10, 1971; Staehle, R. W., Brown, B. F., Kruger, J., Agrawal, A., Eds.; National Association of Corrosion Engineers: Houston, Texas, USA, 1974; pp 78–103.
44. Galvele, J. R.; Lumsden, J. B.; Staehle, R. W. *J. Electrochem. Soc.* **1978**, *125*, 1204–1208.
45. Rosenfeld, I. L.; Danilov, I. S. *Corros. Sci.* **1967**, *7*, 129–142.
46. Yahalom, J.; Ives, L. K.; Kruger, J. *J. Electrochem. Soc.* **1973**, *120*, 384–386.
47. Laycock, N. J.; Moayed, M. H.; Newman, R. C. *J. Electrochem. Soc.* **1998**, *145*, 2622–2628.

48. Alvarez, M. G.; Galvele, J. R. *Corros. Sci.* **1984**, *24*, 27–48.
49. Good, D. *J. Electrochem. Soc.* **1997**, *144*, 1965–1971.
50. Horvath, J.; Uhlig, H. H. *J. Electrochem. Soc* **1968**, *115*, 791–795.
51. Pessall, N.; Liu, C. *Electrochim. Acta* **1971**, *16*, 1987–2003.
52. Pourbaix, M. *Corrosion* **1970**, *26*, 431–438.
53. Suzuki, T.; Kitamura, Y. *Corrosion* **1972**, *28*, 1–6.
54. Wilde, B. E.; Williams, E. *Electrochim. Acta.* **1971**, *16*, 1971–1985.
55. Kuzub, V. S.; Novitskii, V. S. *Z. Metallov.* **1976**, *12*, 177–180.
56. Strehblow, H. H.; Titze, B. *Corros. Sci.* **1977**, *17*, 461–472.
57. Keitelman, A. D.; Galvele, J. R. *Corros. Sci.* **1982**, *22*, 739–751.
58. Wang, S.; Newman, R. C. *Corrosion* **2004**, *60*, 448–454.
59. Al-Zahrani, A. M.; Pickering, H. W. *Electrochim. Acta.* **2005**, *50*, 3420–3435.
60. Galvele, J. R. *Corros. Sci.* **2005**, *47*, 3053–3067.
61. Stolica, N. *Corros. Sci.* **1969**, *9*, 455–470.
62. Leckie, H. P.; Uhlig, H. H. *J. Electrochem. Soc.* **1966**, *113*, 1262–1267.
63. Böhni, H.; Uhlig, H. H. *J. Electrochem. Soc.* **1969**, *116*, 906–910.
64. Szklarska-Smialowska, Z. *Corros. Sci.* **1971**, *11*, 209–221.
65. Kaesche, H. Z. *Phys. Chem.* NF. **1962**, *34*, 87–108.
66. Prazhak, M.; Toushek, Ya.; Spanilyi, V. *Z. Metallov.* **1969**, *5*, 371–375.
67. Edeleanu, C.; Evans, U. R. *Trans. Faraday Soc.* **1951**, *47*, 1121–1135.
68. Hagyard, T.; Santhiapillai, J. R. *J. Appl. Chem.* **1959**, *9*, 323–330.
69. Wong, K. P.; Alkire, R. C. *J. Electrochem. Soc.* **1990**, *137*, 3010–3015.
70. Van Muylder, J.; Pourbaix, M.; Van Laer, P. In *Rapport Technique No 127*; CEBELCOR: Brussels, 1965.
71. Smith, J. A.; Peterson, M. H.; Brown, B. F. *Corrosion* **1970**, *26*, 539–542.
72. Suzuki, T.; Yamabe, M.; Kitamura, Y. *Corrosion.* **1973**, *29*, 18–22.
73. Pickering, H. W.; Frankenthal, R. P. *J. Electrochem. Soc.* **1972**, *119*, 1297–1304.
74. Bargeron, C. B.; Benson, R. *C. J. Electrochem. Soc.* **1980**, *127*, 2528–2530.
75. Kolotyrkin, Ya. M. *Corrosion* **1963**, *19*, 261t–268t.
76. Szklarska-Smialowska, Z. *Corrosion* **1971**, *27*, 223–233.
77. Osozawa, K.; Okato, N.; Fukase, Y.; Yokota, K. *Corros. Eng. (Boshoku Gijutsu),* **1975**, *24*, 1–7.
78. Brigham, R. J.; Tozer, E. W. *Corrosion* **1974**, *30*, 161–166.
79. Szklasrska-Smialowska, Z. In *Pitting and Crevice Corrosion*; NACE In *International: Texas, 2005*.
80. Lizlovs, E. A.; Bond, A. P. *J. Electrochem. Soc.* **1975**, *122*, 719–722.
81. Stout, D. A.; Lumsden, J. B.; Staehle, R. W. *Corrosion* **1979**, *35*, 141–146.
82. Muller, I. L.; Galvele, J. R. *Corros. Sci.* **1977**, *17*, 179–193.
83. Muller, I. L.; Galvele, J. R. *Corros. Sci.* **1977**, *17*, 995–1007.
84. Reding, J. T.; Newport, J. J. *Mater. Prot.* **1966**, *5*, 15–18.
85. Smialowska, S.; Czachor, M. In *Localized Corrosion*; U. R. Evans Conference on Localized Corrosion, Williamsburg, Virginia, USA, December 6–10, 1971; Staehle, R. W., Brown, B. F., Kruger, J., Agrawal, A., Eds.; National Association of Corrosion Engineers: Houston, Texas, USA, 1974; pp 353–362.
86. Smialowski, M.; Szklarska-Smialowska, Z.; Rychcik, M.; Szummer, A. *Corros. Sci.* **1969**, *9*, 123–125.
87. Wranglen, G. *Corros. Sci.* **1974**, *14*, 331–349.
88. Tokuda, T.; Ives, M. B. *Corros. Sci.* **1971**, *11*, 297–306.
89. Schatt, W.; Worch, H. *Corros. Sci.* **1971**, *11*, 623–629.
90. Suter, T.; Webb, E. G.; Böhni, H.; Alkire, R. C. *J. Electrochem. Soc.* **2001**, *148*, B174–B185.
91. Weinstein, S. *Tesis*; Universidad Nacional de Rosario: Argentina, 1975.
92. Kabanov, B.; Burstein, R.; Frumkin, A. In *Electrode Processes (Discussions of the Faraday Society, No 1, 1947)*; Reprinted by Butterworths: London, UK, 1961; pp 259–269.
93. Vetter, K. J.; Strehblow, H. H. In *Localized Corrosion*; U. R. Evans Conference on Localized Corrosion, Williamsburg, Virginia, USA, December 6–10, 1971; Staehle, R. W., Brown, B. F., Kruger, J., Agrawal, A., Eds.; National Association of Corrosion Engineers: Houston, Texas, USA, 1974; pp 240–251.
94. Galvele, J. R. In *Passivity and its Breakdown on Iron and Iron Base Alloys*; Staehle, R. W., Okada, H., Eds.; NACE: Houston, TX, 1976; pp 118–120.
95. Galvele, J. R. *J. Electrochem. Soc.* **1976**, *123*, 464–474.
96. Galvele, J. R. *Corros. Sci.* **1981**, *21*, 551–579.
97. Gravano, S. M.; Galvele, J. R. *Corros. Sci.* **1984**, *24*, 517–534.
98. Keitelman, A. D.; Gravano, S. M.; Galvele, J. R. *Corros. Sci.* **1984**, *24*, 535–545.
99. Vetter, K. J.; Strehblow, H. H. *Ber. Bunsen-Ges. Phys. Chem.* **1970**, *74*, 1024–1035.
100. Hunt, J. P. In *Metal Ions in Aqueous Solutions*; Benjamin: New York, 1963; pp 46–48.
101. Vetter, K. J. In *Electrochemical Kinetics*; Academic Press: New York, 1961; p 171.
102. Sillen, L. G.; Martell, A. E. In *Stability Constants of Metal-Ion Complexes, Special Publication No. 17*; The Chemical Society: London, 1964.
103. Sillen, L. G.; Martell, A. E. In *Stability Constants of Metal-Ion Complexes, Special Publication and Suplement 1, Special Publication 25*; The Chemical Society: London, 1964.
104. Smith, R. M.; Martell, A. E. In *Critical Stability Constants;* Plenum Press: New York, 1976; Vol. 4.
105. Newman, R. C.; Ajawi, M. A. A.; Ezuber, H.; Turgoose, S. *Corros. Sci.* **1988**, *28*, 471–477.
106. Frankel, G. S. *J. Electrochem Soc.* **1998**, *145*, 2186–2198.
107. Uhlig, H. H. *J. Electrochem. Soc.* **1950**, *97*, 215C–220C.
108. Kolotyrkin, Ja. M. *J. Electrochem. Soc.* **1961**, *108*, 209–216.
109. Hoar, T. P.; Mears, D. C.; Rothwell, G. P. *Corros. Sci.* **1965**, *5*, 279–289.
110. Pryor, M. J. In *Localized Corrosion*; U. R. Evans Conference on Localized Corrosion, Williamsburg, Virginia, USA, December 6–10, 1971; Staehle, R. W., Brown, B. F., Kruger, J., Agrawal, A., Eds.; National Association of Corrosion Engineers: Houston, Texas, USA, 1974; pp 2–11.
111. Mc. Bee, C. L.; Kruger, J. In *Passivity and its Breakdown on Iron and Iron Base Alloys*; Staehle, R. W.; Okada, H., Eds.; NACE: Houston, TX, 1976; pp 131–132.
112. Marcus, P.; Herbelin, J. M. *Corros. Sci.* **1993**, *34*, 1123–1145.
113. Szklarska-Smialowska, Z.; Viefhaus, H.; Janik-Czachor, M. *Corros. Sci.* **1976**, *16*, 649–652.

114. Janik-Czachor, M.; Kaszczyszyn, S. *Mater. Corros. (Werkst. Korros.)* **1982**, *33*, 500–504.
115. Bardwell, J. A.; MacDougall, B.; Sproule, G. I. *J. Electrochem. Soc* **1989**, *136*, 1331–1336.
116. Lin, L. F.; Chao, C. Y.; Macdonald, D. D. *J. Electrochem. Soc.* **1981**, *128*, 1194–1198.
117. Hoar, T. P. *Trans. Faraday Soc.* **1949**, *45*, 683–693.
118. Wood, G. C.; Richardson, J. A.; Abd Rabbo, M. F.; Mapa, L. B.; Sutton, W. H. In *Passivity of Metals*; Proceedings of The Fourth International Symposium on Passivity, Warrenton, Virginia, USA, October 17–21, 1977; Frankenthal, R. P., Kruger, J., Eds.; The Electrochemical Society, Inc.: Princeton, New Jersey, USA, 1978; pp 973–988.
119. Videm, K. *The electrochemistry of uniform corrosion and pitting of aluminum, Kjeller Report KR-140*; Institut for Atomenergy: Kjeller, Norway, 1974.
120. MacDougall, B. *J. Electrochem. Soc.* **1979**, *126*, 919–925.
121. Fromhold, A. T. In *Passivity and its Breakdown on Iron and Iron Base Alloys*; Staehle, R. W.; Okada, H., Eds.; NACE: Houston, TX, 1976; pp 161–164.
122. Sato, N. *Electrochim. Acta.* **1971**, *16*, 1683–1692.
123. Sato, N. *J. Electrochem. Soc.* **1982**, *129*, 255–260.
124. Galvele, J. R.; Torresi, M.; Carranza, R. M. *Corros. Sci.* **1990**, *31*, 563–571.
125. Galvele, J. R. Analysis of the Critical Factors in Pit Initiation and Propagation, In Frankel, G. S., Newman, R. C., Eds.; *Proceedings of the Symposium on Critical Factors in Localized Corrosion*, The Electrochemical Society, Inc.: Pennington, NJ, 1992; pp 94–108.
126. Hisamatsu, Y.; Yoshii, T.; Matsumura, Y. In *Localized Corrosion*; U. R. Evans Conference on Localized Corrosion, Williamsburg, Virginia, USA, December 6–10, 1971; Staehle, R. W., Brown, B. F., Kruger, J., Agrawal, A., Eds.; National Association of Corrosion Engineers: Houston, Texas, USA, 1974; pp 427–436.
127. Forchhammer, P.; Engell;, H. J. *Werkst. Korr* **1969**, *20*, 1–12.
128. Wood, G. C.; Sutton, W. H.; Richardson, J. A.; Riley, T. N. K.; Malherbe, A. G. In *Localized Corrosion*; U. R. Evans Conference on Localized Corrosion, Williamsburg, Virginia, USA, December 6–10, 1971; Staehle, R. W., Brown, B. F., Kruger, J., Agrawal, A., Eds.; National Association of Corrosion Engineers: Houston, Texas, USA, 1974; pp 526–546.
129. Okamoto, G.; Tachibana, K.; Nishiyama, S.; Sugita, T. In *Passivity and its Breakdown on Iron and Iron Base Alloys*; Staehle, R. W.; Okada, H., Eds.; NACE: Houston, TX, 1976; pp 106–109.
130. Frankel, G. S.; Stockert, L.; Hunkeler, F.; Boehni, H. *Corrosion* **1987**, *43*, 429–436.
131. Pistorius, P. C.; Burstein, G. T. *Phil. Trans. R. Soc. Lond.* **1992**, *341*, 531–559.
132. Williams, D. E.; Westcott, C.; Fleischmann, M. *J. Electrochem. Soc.* **1985**, *132*, 1804–1811.
133. Williams, D. E.; Stewart, J.; Balkwill, P. H. *Corros. Sci.* **1994**, *36*, 1213–1235.
134. Isaacs, H. S. *Corros. Sci.* **1989**, *29*, 313–323.
135. Garfias-Mesias, L. F.; Sykes, J. M. *Corros. Sci.* **1999**, *41*, 959–987.
136. Bertocci, U.; Koike, M.; Leigh, S.; Qui, F.; Yang, G. *J. Electrochem. Soc.* **1986**, *133*, 1782–1786.
137. Pride, S. T.; Scully, J. R.; Hudson, J. L. *J. Electrochem. Soc.* **1994**, *141*, 3028–3040.
138. Burstein, G. T.; Souto, R. M. *Electrochim. Acta.* **1995**, *40*, 1881–1888.
139. Isaacs, H. S. In *Localized Corrosion*; U. R. Evans Conference on Localized Corrosion, Williamsburg, Virginia, USA, December 6–10, 1971; Staehle, R. W., Brown, B. F., Kruger, J., Agrawal, A., Eds.; National Association of Corrosion Engineers: Houston, Texas, USA, 1974; pp 158–167.
140. Isaacs, H. S.; Ishikawa, Y. *J. Electrochem. Soc.* **1985**, *132*, 1288–1293.
141. Williams, D. E.; Westcott, C.; Fleischmann, M. *J. Electrochem. Soc.* **1985**, *132*, 1796–1804.
142. Balkwill, P. H.; Wescott, C.; Williams, D. E. *Mater. Sci. Forum* **1989**, *44–45*, 299–313.
143. Ezuber, H.; Newman, R. C. Growth Rate Distribution of Metastable Pits. In *Critical Factors in Localized Corrosion*; Frankel, G. S., Newman, R. C., Eds.; The Electrochemical Society Inc.: New Jersey, 1992; pp 120–133.
144. Moayed, M. H.; Newman, R. C. *Corros. Sci.* **2006**, *48*, 1004–1018.
145. Pistorius, P. C.; Burstein, G. T. *Corros. Sci.* **1992**, *33*, 1885–1897.
146. Burstein, G. T.; Pistorius, P. C.; Mattin, S. P. *Corros. Sci.* **1993**, *35*, 57–62.
147. Pistorius, P. C.; Burstein, G. T. *Corros. Sci.* **1994**, *36*, 525–538.
148. Laycock, N. J.; Newman, R. C. *Corros. Sci.* **1997**, *39*, 1771–1790.
149. Burstein, G. T.; Vines, S. P. *J. Electrochem. Soc.* **2001**, *148*, B504–B516.
150. Ilevbare, G. O.; Burstein, G. T. *Corros. Sci.* **2001**, *43*, 485–513.
151. Burstein, G. T.; Liu, C. *Corros. Sci.* **2007**, *50*, 2–7.
152. Stewart, J.; Wiliams, D. E. *Corros. Sci.* **1992**, *33*, 457–463.
153. Qvarfort, R. *Corros. Sci.* **1988**, *28*, 135–140.
154. Alvarez, M. G.; Galvele, J. R. In *IV Reunión Latinoamericana de Electroquímica y Corrosión*; Caracas: Venezuela, 1979.
155. Carranza R. M. Rodríguez M. A. Environmental and Geometrical Conditions to Sustain Crevice Corrosion in Alloy 22, NACE 2007, National Association of Corrosion Engineers, October 31–November 3, Las Vegas, Nevada Paper No 08175.
156. Brigham, R. J.; Tozer, E. W. *Corrosion* **1973**, *29*, 33–36.
157. ASTM-G150–99. *Standard Test Method for Electrochemical Critical Pitting Temperature Testing. of Stainless Steels*; American Society of Testing and Materials, 1999; pp 638–649.
158. Qvarfort, R. *Corros. Sci.* **1989**, *29*, 987–993.
159. Arnvig, P. E.; Bisgard, A. D. *Paper No 437 presented at Corrosion'96*; National Association of Corrosion Engineers: Houston, Texas, USA, 1996.
160. Jakobsen, P. T.; Maahn, E. *Corros. Sci.* **2001**, *43*, 1693–1709.
161. Deng, B.; Jiang, Y.; Liao, J.; Hao, Y.; Zhong, Ch.; Li, J. *Appl. Surf. Sci.* **2007**, *253*, 7369–7375.
162. Moayed, M. H.; Laycock, N. J.; Newman, R. C. *Corros. Sci.* **2003**, *45*, 1203–1216.
163. Salinas-Bravo, V. M.; Newman, R. C *Corros. Sci.* **1994**, *36*, 67–77.
164. Ernst, P.; Newman, R. C *Corros. Sci.* **2002**, *44*, 927–941.
165. Ernst, P.; Newman, R. C *Corros. Sci.* **2002**, *44*, 943–954.

2.05 Dealloying

R. C. Newman
University of Toronto, Department of Chemical Engineering and Applied Chemistry, 200 College Street, Toronto, Ontario M5S 3E5, Canada

Glossary

Critical potential for dealloying Potential above which nanoporosity is formed and the rate of dealloying levels out with time rather than decaying to a very low value.

High-density site percolation threshold A site percolation threshold with the additional requirement that all sites forming a percolation path have at least m like-neighbors, where $m > 2$. Written as $p_c(m)$.

Parting limit Atom fraction of less-noble element in a binary solid–solution alloy below which dealloying does not occur, irrespective of how oxidizing is the potential. Usually lies in the range 0.5–0.6; can sometimes be *ca*. 0.2, for example in α-brass.

Percolation theory Mathematical theory, applied to more-or-less disordered geometrical systems, that deals with the statistics of interconnection of sites or bonds with a specified property (e.g., for a binary metallic solid solution the relevant property is atom identity and the relevant version of the theory is lattice-site percolation.

Site percolation threshold Fractional occupancy of sites, usually on a lattice, by a particular type of entity (e.g., atom of one kind in a binary solid solution) at which long-range connectivity of such sites appears. Written as $p_c(site)$.

Abbreviations

NAB Nickel–aluminum bronze
SCC Stress corrosion cracking

Symbols

$\boldsymbol{a}$ Atom diameter or spacing (m)
$\boldsymbol{b_a}$ Anodic Tafel coefficient (V)
$\boldsymbol{D_s}$ Surface diffusion coefficient ($m^2\,s^{-1}$)
$\boldsymbol{E_0}$ Equilibrium metal–metal ion electrode potential (nonstandard metal ion activity)
$\boldsymbol{E_c}$ Critical potential for dealloying (V)
$\boldsymbol{F}$ Faraday's constant ($C\,mol^{-1}$)
$\boldsymbol{i}$ Anodic current density ($A\,m^{-2}$)
$\boldsymbol{i_0}$ Exchange current density ($A\,m^{-2}$)
$\boldsymbol{j}$ Diffusion flux ($mol\,m^{-2}\,s^{-1}$)
$\boldsymbol{p}$ Atom fraction in a solid solution alloy
$\boldsymbol{p^*}$ The parting limit, expressed as an atom fraction
$\boldsymbol{S}$ Excess interfacial energy per volume of a nanoscale pit in an alloy surface ($J\,m^{-3}$)

η_{th} Thermodynamic shift in equilibrium electrode potential due to curvature (V)
γ Solid–liquid interfacial energy ($J\,m^{-2}$)
Ω Molar volume (m^3)
ξ Blob size or percolation backbone dimension in percolation theory (m)

2.05.1 Introduction and Background

2.05.1.1 Definition of Dealloying

Dealloying is the selective electrolytic dissolution of a less-noble element from a metallic solid solution or intermetallic compound. The product of this dissolution, which usually requires an oxidizing acidic solution, is a nanoporous metal consisting mostly of more noble element(s). **Figure 1** show the morphology usually seen in dealloyed noble-metal alloys: an interconnected ligament/pore structure with nearly zero net curvature. The size scale of the porosity in these pictures is typical for silver–gold alloys dealloyed in perchloric acid; the more common nitric acid immersion usually gives a larger pore size.

2.05.1.2 Use of Dealloying to Make Interesting Nanomaterials

Recently, there has been a resurgence of interest in dealloying,[1–3] driven by potential technological applications and the remarkable properties of the nanoporous products such as near-theoretical strength in compression,[4] a size-scale-dependent elastic modulus,[5] and under a particular alloy composition and dealloying procedure, the formation of a material with an unusual combination of exceptionally low modulus and relatively good mechanical integrity.[6] Many potential applications of dealloyed materials are under investigation, particularly, in the areas of catalysis[7–9] and sensing.[10,11] Metal foils that are dealloyed, on one side, show potential-dependent bending in electrolytes due to surface stress effects and are thought to have potential as microactuators.[12] This literature is quite different from that which is concerned with dealloying as a corrosion problem.

2.05.1.3 History of Dealloying

Dealloying is an ancient technology, and was used by several civilizations to gild the surfaces of objects that were made mainly of copper and/or silver.[13] Such treatments generally involved the application of heat as well as chemical attack, and the resulting surface would have had a good luster due to some consolidation or coarsening of the nanoporous material – an important factor, because ordinary dealloying performed at room temperature gives a brittle, brownish product. There appears to have been a spectrum of techniques, ranging from true dealloying (with a liquid, but not necessarily very aqueous, environment present on the metal throughout) to a dry-oxidation/aqueous oxide-dissolution cycle. A closely related use of dealloying was for a more macroscopic separation or 'parting' of metals – this was also done by a kind of pack cementation, where, for silver–gold alloys, silver was converted into silver chloride by salt, acid, oxidant, and heat. Later – at an uncertain date, some time in the early 1500s – nitric acid first came into use. But, before that, reagents that contained nitric acid had been invented. There are several different recipes in the *Codex Atlanticus* (*ca.* 1508) on the use of saltpeter (potassium nitrate) as the source of the nitrate ions (Leonardo da Vinci may have written these recipes himself, but they are not in his usual mirror writing).[13] Since then, nitric acid has been the preferred reagent for dealloying of gold alloys.[14]

The key discovery about the parting of gold alloys was that the less-noble metal or metals had to be present above a certain percentage; otherwise, there was no dealloying. So, if one has, say, an alloy with 40 at.% silver and 60 at.% gold, one has to melt it with more silver so that the silver content exceeds ~60 at.%. This discovery may be very ancient, and the 'parting limit' was the most striking feature of dealloying. We will return to the parting limit (or rather limits, because there is another one near 20%) later on.

2.05.1.4 Dealloying of Industrial Alloys

Dealloying is also a major corrosion problem in industrial alloys – not so much in itself, as a mechanism of general deterioration, but for its role in stress corrosion cracking (SCC). Dezincification of brass must have been noticed in Roman times; but, the first scientific records date from the early twentieth century, when this was a severe problem affecting α-brass seawater condenser tubes.[15–17] Bengough and May first systematized the addition of arsenic (in amazingly small amounts, less than 0.05 wt%) for dezincification resistance.[18] Another remedial method, more rarely applied, was to reduce the zinc content to below

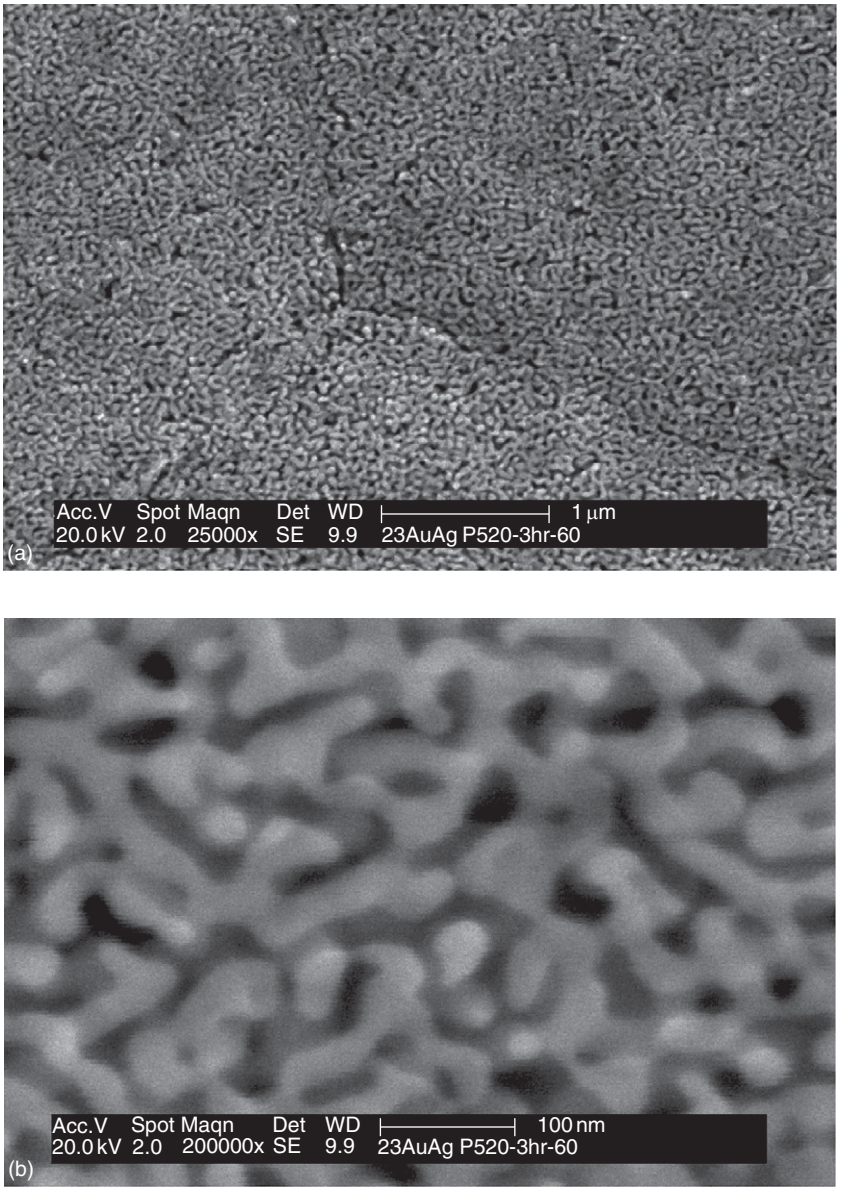


Figure 1 Morphology of Ag–Au alloys dealloyed in perchloric acid showing an interconnected ligament/pore structure with nearly zero net curvature. (a) Reproduced from Senior, N. A.; Newman, R. C. *Nanotechnology* **2006**, *17*, 2311–2316. (b) Reproduced from Senior's PhD thesis (University of Manchester, *ca*. 2005).

20%; but, this adversely affected other properties including general corrosion resistance. β-Brass is more susceptible than α-brass, as one would expect given its higher zinc content. More recently, other copper alloys such as nickel–aluminum bronze (NAB) have shown dealloying,[19] but these are more complex systems as they are usually multiphase. Dealloying has occurred, but rarely, in various other commercial solid-solution alloys that satisfy the basic compositional requirement. Alloys used as thin films for magnetic or other applications could potentially suffer from atmospheric dealloying.

There are analogous phenomena that occur in gaseous environments, or in vacuum where one element is volatile.[20] For the most part, these appear to involve lattice diffusion, which places them apart from normal electrolytic dealloying, where transport is by surface diffusion, as discussed below.

2.05.2 Mechanisms of Dealloying

2.05.2.1 Transport of the More Noble Metal

Historically, three types of mechanisms have been proposed to explain how two alloying constituents, intermixed on an atomic scale, can be separated so effectively by electrolytic action. Perhaps the most obvious is that the dissolution is not really selective at all, but that both elements dissolve, then one replates. But, while there is some evidence that this may occur – very locally – during dezincification of brass,[21] it is clearly not operating in the more general case. The second mechanism, promoted by Pickering and Wagner, involves divacancy injection and room-temperature lattice diffusion of these vacancies.[20,22,23] This appears to be inconsistent with the rapid kinetics of dealloying, but as Pickering pointed out, it would account for the common observation of a significant residue of less-noble metal in the dealloyed material, whereas (in his opinion at the time) it would appear impossible that any less-noble metal could survive at such high anodic overpotentials. Later, Pickering rationalized this by postulating a large IR potential drop within the porous structure, so that the overpotential was actually small at the corrosion front.[24] Apart from graded solid solution compositions, he also observed intermediate phase formation in dealloyed ε-brass by X-ray diffraction.[25] (Remarkably, this exact experiment was first reported in 1933![26]) This strongly impressed Pickering with the necessity of postulating some bulk diffusion mechanism.

Recently, most authors have preferred to appeal to surface diffusion as the operative transport mechanism for the more noble metal (the intermediate-phase problem has not, to our knowledge, been addressed from this standpoint and would make an interesting study). A.J. Forty and his group made transmission electron microscopy (TEM) observations of the early stages of porosity formation in silver–gold alloys immersed in nitric acid, and developed a model for the parting limit based on surface diffusion and filling of incipient pores.[27–33] They were also very interested in the formation of gold oxide during dealloying, something that affects the mobility of the gold. Sieradzki *et al.*[34] carried out atomistic Monte Carlo simulations of dealloying and reproduced several of the main features including parting limits, which were ascribed (not quite accurately in the case of the '55%' threshold) to site percolation thresholds in the lattice. Erlebacher developed a much more sophisticated simulation scheme, focusing mainly on morphology, kinetics, and potential dependence.[35]

There may be a niche for bulk diffusion as a partial explanation of the volume shrinkage (a few percent to almost 30%) that occurs during dealloying of silver–gold alloys, but some kind of plastic contraction is almost certainly involved.[36]

2.05.2.2 Potential Dependence of Dealloying and the Critical Potential

Pickering made comprehensive studies of the electrode potential dependence of dealloying, and the *critical potential* for dealloying (E_c). These experiments did not give a meaningful E_c for α-brass because copper dissolution intervened[25]; but, in binary gold alloys, he observed the now-famous shape of curve shown in **Figure 2**. Up to a certain potential (E_c), the current density is low and almost potential independent; then it increases steeply at E_c.[29] Uniform nanoporosity forms only above E_c, although Pickering often referred to 'pitting' or localized porosity below E_c.[37] As more gold is added to the alloy, E_c increases until, at the parting limit, only the flat, 'passive' region is observed. Up to a certain gold content, this behavior was reproduced, with good quantitative agreement with experiment, by Erlebacher's simulations.[35] Then, a small refinement (prohibition of dissolution of atoms with 10 or 11 neighbors) reproduced the experimentally observed '55%' parting limit almost exactly[38] – **Figure 3**. It was shown that this was close, but not identical, to a high-density percolation threshold for the lattice.[39] The effect of varying the surface diffusion kinetics was surprisingly small – a few percent – **Figure 4**.

Recently, there has been much discussion of the true value and meaning of E_c. In 1993, Sieradzki gave a model for E_c based on two effects: thermodynamic (necessity of dissolving less-noble metal from very small clusters of atoms, thus creating a gold surface with very high curvature) and kinetic (necessity of dissolving these atoms at a certain *rate* to avoid infilling by surface diffusion of gold).[40] More recently, it is proposed that an intrinsic critical potential can be identified, in simulation and experiment, that corresponds only to the thermodynamic (curvature) effect.[35,41–43]

A simplified version of Sieradzki's original model can be written down in a few lines. The random arrangement of atoms in the solid solution can be characterized,[44] in the language of percolation theory, as a 'blob and string' arrangement – 'blobs' are regions where, statistically, one kind of atom (say, Ag in Ag–Au) is densely clustered, while the 'strings' are chains of Ag atoms joining the blobs together.

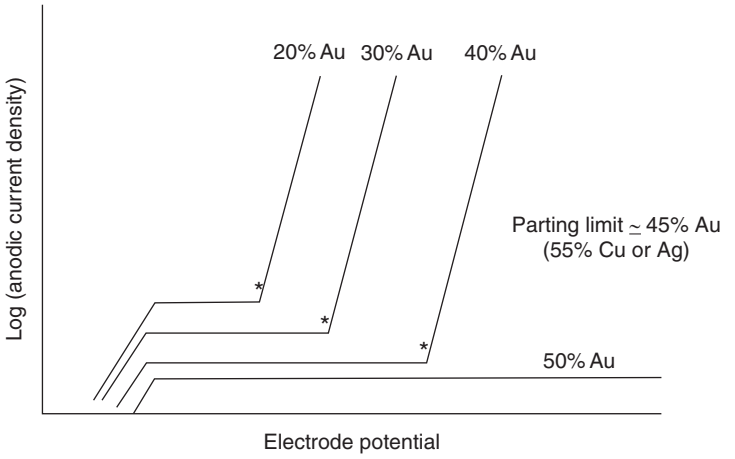


Figure 2 Schematic anodic dissolution behavior of a silver–gold or copper–gold alloy undergoing dealloying in a nonoxidizing aqueous acid, showing critical potentials for nanoporosity generation (*) and the inability to form nanoporosity when the silver or copper content falls below the parting limit.

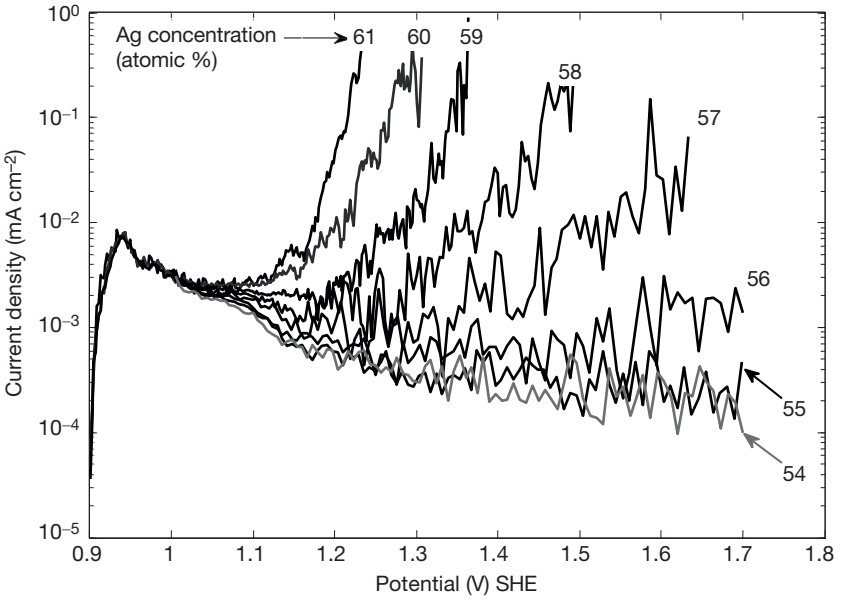


Figure 3 Dynamic current–voltage curves generated by the modified MESOSIM code for the indicated Ag contents in Ag–Au (at.%). Dissolution from 10- and 11-coordinated sites is prohibited. Potential ramping rate is 2 mV s^{-1}. Reproduced form Artymowicz, D. M.; Newman, R. C.; Erlebacher, J. *Philos. Mag.* **2009**, *89*, 1663–1693.

Another name for the 'blob' size is the percolation backbone diameter. We call this quantity ξ. This can be shown to be approximately given by

$$\xi = \frac{1+p}{1-p} \qquad [1]$$

in units of atom diameters (a), where p is the atom fraction of the less-noble element (Ag).

So, the initial surface of the alloy can be considered to be intersected by numerous blobs of Ag that, when dissolved, create tiny pits with highly curved Au surfaces. These surfaces have a solid–liquid interfacial free energy γ. The excess interfacial energy S per unit volume of a hemispherical pit created by the dissolution of a typical 'blob' (diameter ξa, radius $\xi a/2$) is

$$S = \frac{2\gamma}{\xi a/2} = \frac{4\gamma}{\xi a} \qquad [2]$$

(neglecting the slight difference in γ that must exist between the gold-rich pit and the original alloy

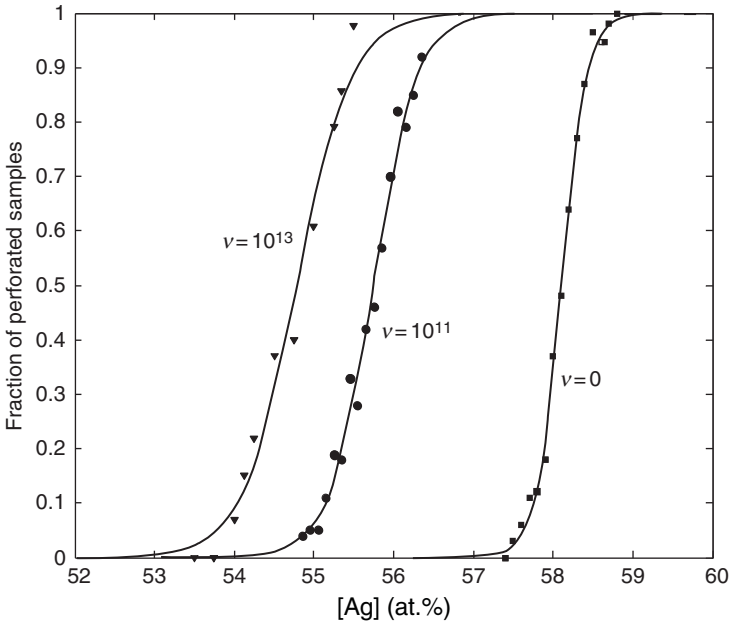


Figure 4 Fraction of samples perforated by dealloying as a function of the 'Ag' content with dissolution of 10- and 11-coordinated atoms prohibited and surface diffusivity set to high, intermediate and negligible values, by changing vs. Samples were 100 atomic layers thick and simulations run for a maximum time of 10^5 s. Perforation was defined as dissolution of the first atom from the bottom layer. Simulations stopped at perforation or when the elapsed time exceeded 10^5s. ν is the value of the pre-exponential constant $\nu_{diffusion}$ in eqn [1]. Reproduced from Artymowicz, D. M.; Newman, R. C.; Erlebacher, J. *Philos. Mag.* **2009**, *89*, 1663–1693.

surface.) So, if we multiply S by the molar volume Ω, we obtain the energy per mole. Dividing by Faraday's constant gives us a shift in the equilibrium electrode potential, η_{th}, for the Ag that was in the pit compared with ordinary bulk Ag:

$$\eta_{th} = \frac{S\Omega}{F} = \frac{4\gamma\Omega}{F\xi a} \quad [3]$$

This shift is positive, so it represents a thermodynamic contribution to the difference between the ordinary Ag equilibrium potential and the critical potential. So, for $p > p^*$, where p^* is the parting limit (*c.* 0.55),

$$\eta_{th} = \frac{4\gamma\Omega(1-p)}{Fa(1+p)} \quad [4]$$

This agrees quite well with observation for Au and Pt alloys, provided that the E_{crit} value for the former is measured extremely slowly, to give a so-called intrinsic critical potential. To have easy dealloying, though, it is necessary for the *rate* of dissolution of the Ag blob to exceed the rate at which the pit is infilled by surface diffusion of Au. This requires an additional overpotential η_k that can be written as

$$\eta_k = b_a \log\frac{i}{i_0} \quad [5]$$

with i being equated to the flux of Au diffusing into the pit at the critical potential (Sieradzki used a linearized kinetic expression rather than this Tafel version). This flux j is given by a Fickian steady-state diffusion flux of the form

$$j = \frac{\alpha D_s}{\Omega\xi a} \quad [6]$$

where α is a geometrical coefficient and D_s the surface diffusivity of Au. So η_k is given by

$$\eta_k = b_a \log\frac{\alpha F D_s(1-p)}{\Omega(1+p)ai_0} \quad [7]$$

and the critical potential E_c is given by

$$E_c = E_0 + \frac{4\gamma\Omega(1-p)}{Fa(1+p)} + b_a \log\frac{\alpha F D_s(1-p)}{\Omega(1+p)ai_0} \quad [8]$$

2.05.2.3 Order and Disorder

Most studies of dealloying have used more or less disordered binary alloys, but comparison of ordered and disordered material have been made for the copper–gold system. The finding is that dealloying is more difficult (requires a higher potential) in the

ordered material.[45] This can be explained by considering the connectivity of less-noble metal atoms in the lattice – the disordered alloy has a cluster structure in which less-noble atoms comprising the dissolution path have many like neighbors, facilitating penetration of the electrolyte.

2.05.2.4 Ternary Element Effects

Ternary element effects in dealloying are both intriguing and practically important. A long time ago, it was suggested that the beneficial effect of arsenic in brass was associated with the blocking of step edges, slowing down the surface diffusion of copper[21,46]; this stops macro-dealloying, but can be detrimental for SCC resistance, as explained below – very fine porosity makes for brittle behavior. Lately, Erlebacher has explored the behavior of Ag–Au–Pt alloys, which show a refined porosity compared with Ag–Au, for possible applications.[47]

2.05.2.5 Why does Brass Dealloy at 20% Zinc?

So why do some alloys show dealloying at the fcc site percolation threshold (20 at.%), while others require 55% of less-noble element? The conditions for dezincification of α-brass are quite specific – corrosion has to proceed for some time (usually in an aerated chloride solution), then dealloying starts under the resulting porous corrosion product – an extremely old observation. Lucey,[48,49] and later Newman *et al.*,[21] gave different variants of the same explanation – there is an enrichment of copper ions (Lucey referred to solid CuCl) under the corrosion product, then the conditions become very close to *equilibrium* for the copper, and the exchange current density somehow enhances the mobility of copper, exposing more zinc for dissolution (but still respecting the site percolation threshold). Newman *et al.*[21] showed that dezincification could be monitored electrochemically at the equilibrium potential of a copper electrode in solutions containing $CuCl_2^-$, which is the dissolution product of Cu in chloride solutions. An example of one of their dezincified samples is shown in Figure 5 – in industrial parlance this would be called 'plug' or localized type dezincification.

2.05.3 Dealloying in Stress Corrosion Cracking

The role of dealloying in SCC is an old subject of discussion. For binary gold alloys, the correlation between the two is transparent,[50] but the actual mechanism of crack growth (and this may be different for intergranular and transgranular SCC) remains controversial. According to Sieradzki and Newman,[50] both forms of SCC require only a thin dealloyed layer, which triggers a brittle substrate microcrack; such crack jumps in fcc metals are impossible according to traditional mechanics, but evidence for the phenomenon has mounted – it is especially easy to demonstrate for the intergranular variant in silver–gold alloys,[51–54] but also for brass in ammonia.[55] Even those who prefer another microscopic mechanism admit a very high crack velocity, for example, 20 mm s^{-1}, in the case of silver–gold, which cannot be rationalized by any continuous-dissolution type of mechanism.[56]

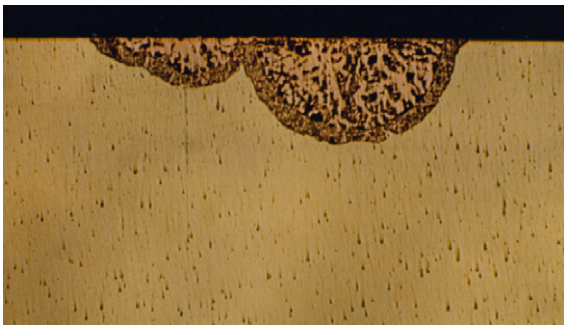


Figure 5 Plug-type dezincification in α-brass. Reproduced from Newman, R. C.; Shahrabi, T.; Sieradzki, K. *Corros. Sci.* **1988**, *28*, 873–886.

The important industrial alloys where dealloying has been blamed for SCC are austenitic and duplex stainless steels (in hot chloride, chloride–H_2S, or caustic solutions) and copper alloys including α-brass (traditionally in aqueous ammonia but also in simple solutions like copper sulfate). These are more challenging to study, because there is always some simultaneous oxidation of all the alloy components, and the dealloyed layers are thin.[57] A key finding was that the compositional dependences of dealloying and transgranular SCC in brass, exposed to a solution with which copper is equilibrated, are identical.[58] Research on stainless steels using caustic solutions, where nickel is the 'noble' metal,[59] has led to another current thread of study in noble-metal alloys – namely the ability to produce a nanoporous dealloyed layer at very low contents of noble metal. If the anodic overpotential is small enough, dealloying can occur in an alloy with only a percent or two of noble metal, and a nanoporous layer can be produced whose composition is not far removed from that of the original alloy.

2.05.4 Summary and Future Prospects

At the time of writing (Spring 2009) we are in the midst of a flood of new literature on dealloying. Over the next few years, practical applications of nanoporous metals (catalysts, sensors, actuators, etc.) will flourish. In turn, long-standing practical corrosion phenomena will be illuminated, especially with regard to the micromechanisms of SCC. We can anticipate that practically every case of SCC in solid-solution alloys will be understood better, and improved preventive measures devised, now that such a wide range of new techniques and concepts is being brought to bear upon the properties and applications of nanoporous metals. Corrosion is just like any other subject – it tends to grow at its interfaces with other disciplines. Also, its standing in the wider scientific community grows with the variety of scientists from remote fields who become interested in corrosion phenomena. Just to give one example, attempts by the author to understand the role of minor alloying elements in dealloying and SCC (arsenic in brass; phosphorus in austenitic stainless steel[46]) were quite primitive at the time, but are now of renewed interest with the discovery of ternary Au–Ag–Pt alloys that show very stable nanoporosity.[47] Also, corrosion scientists can contribute to solving new problems that will come with new technologies. When we start to use nanoporous metals in devices, or alloy nanoparticles as electrocatalysts, a range of new corrosion issues will arise, such as the gradual dissolution of residual less-noble metal that may be essential for the desired function (a Pt–Co alloy is under consideration for hydrogen fuel-cell cathode catalyst formulations, and some candidate alloy electrocatalysts are being made by dealloying).

References

1. Weissmueller, J.; Newman, R. C.; Jin, H. J.; Hodge, A. M.; Kysar, J. W. *MRS Bull.* **2009**, *34*, 577–586.
2. Schofield, E. *Transact. Inst. Metal Finish.* **2005**, *83*, 35–42.
3. Newman, R. C.; Corcoran, S. G.; Erlebacher, J.; Aziz, M. J.; Sieradzki, K. *MRS Bull.* **1999**, *24*, 24–28.
4. Volkert, C. A.; Lilleodden, E. T.; Kramer, D.; Weissmueller, J. *Appl. Phys. Lett.* **2006**, *89*, 061920.
5. Mathur, A.; Erlebacher, J. *Appl. Phys. Lett.* **2007**, *90*, 061910.
6. Senior, N. A.; Newman, R. C. *Nanotechnology* **2006**, *17*, 2311–2316.
7. Xu, C. X.; Su, J. X.; Xu, X. H.; Liu, P. P.; Zhao, H. J.; Tian, F.; Ding, Y. *J. Am. Chem. Soc.* **2007**, *129*, 42–43.
8. Zeis, R.; Mathur, A.; Fritz, G.; Lee, J.; Erlebacher, J. *J. Power Sources* **2007**, *165*, 65–72.
9. Zeis, R.; Lei, T.; Sieradzki, K.; Snyder, J.; Erlebacher, J. *J. Catal.* **2008**, *253*, 132–138.
10. Qian, L. H.; Yan, X. Q.; Fujita, T.; Inoue, A.; Chen, M. W. *Appl. Phys. Lett.* **2007**, *90*, 153120.
11. Yu, F.; Ahl, S.; Caminade, A. M.; Majoral, J. P.; Knoll, W.; Erlebacher, J. *Analyt. Chem.* **2006**, *78*, 7346–7350.
12. Kramer, D.; Viswanath, R. N.; Weissmueller, J. *Nano Lett.* **2004**, *4*, 793–796.
13. Reti, L. *Isis* **1965**, *56*, 307–319.
14. Seath, J.; Beamish, F. E. *Ind. Eng. Chem. Analyt. Ed* **1938**, *10*, 639–641.
15. Bengough, G. D. *J. Inst. Met.* **1922**, *27*, 51–72.
16. Bengough, G. D.; Jones, R. M. *J. Inst. Met.* **1913**, *10*, 13–118.
17. Gibbs, W. E. *Transact. Faraday Soc.* **1916**, *11*, 2–3.
18. Bengough, G. D.; May, R. *J. Inst. Met.* **1924**, *32*, 81–269.
19. Lenard, D. R.; Bayley, C. J.; Noren, B. A. *Corrosion* **2008**, *64*, 764–772.
20. Pickering, H. W.; Kim, Y. S. *Corros. Sci.* **1982**, *22*, 621–635.
21. Newman, R. C.; Shahrabi, T.; Sieradzki, K. *Corros. Sci.* **1988**, *28*, 873–886.
22. Pickering, H. W. *J. Electrochem. Soc.* **1968**, *115*, 143–147.
23. Pickering, H. W.; Wagner, C. *J. Electrochem. Soc.* **1967**, *114*, 698–706.
24. Pickering, H. W. *Mater. Sci. Eng. A* **1995**, *198*, 213–223.
25. Pickering, H. W. *J. Electrochem. Soc.* **1970**, *117*, 8–15.
26. Stillwell, C. W.; Turnipseed, E. S. *Phys. A. J. Gen. Appl. Phys.* **1933**, *4*, 263.
27. Forty, A. J.; Durkin, P. *Philos. Mag. A* **1980**, *42*, 295–318.
28. Forty, A. J. *Nature* **1979**, *282*, 597–598.
29. Forty, A. J. *Philos. Mag. A* **1982**, *46*, 521–530.
30. Forty, A. J. *Gold Bull.* **1981**, *14*(1), 25–35.
31. Forty, A. J.; Rowlands, G. *Philos. Mag. A* **1981**, *43*, 171–188.
32. Durkin, P.; Forty, A. J. *Philos. Mag. A* **1982**, *45*, 95–105.
33. Smith, D. J.; *et al. Inst. Phys. Conf. Series* **1982**, *61*, 483–486.
34. Sieradzki, K.; Corderman, R. R.; Shukla, K.; Newman, R. C. *Philos. Mag. A* **1989**, *59*, 713–746.
35. Erlebacher, J. *J. Electrochem. Soc.* **2004**, *151*, C614–C626.
36. Parida, S.; Kramer, D.; Volkert, C. A.; Rosner, H.; Erlebacher, J.; Weissmueller, J. *Phys. Rev. Lett.* **2006**, *97*, 035504.
37. Fritz, J. D.; Pickering, H. W. *J. Electrochem. Soc.* **1991**, *138*, 3209–3218.
38. Artymowicz, D. M.; Newman, R. C.; Erlebacher, J. *Electrochem. Soc. Transact.* **2006**, *3*(31), 499–506.
39. Artymowicz, D. M.; Newman, R. C.; Erlebacher, J. *Philos. Mag.* **2009**, *89*, 1663–1693.
40. Sieradzki, K. *J. Electrochem. Soc.* **1993**, *140*, 2868–2872.
41. Rugolo, J.; Erlebacher, J.; Sieradzki, K. *Nat. Mater.* **2006**, *5*, 946–949.
42. Dursun, A.; Pugh, D. V.; Corcoran, S. G. *J. Electrochem. Soc.* **2005**, *152*, B65–B72.
43. Dursun, A.; Pugh, D. V.; Corcoran, S. G. *Electrochem. Solid State Lett.* **2003**, *6*, B32–B34.
44. Erlebacher, J.; Aziz, M. J.; Karma, A.; Dimitrov, N.; Sieradzki, K. *Nature* **2001**, *410*, 450–453.
45. Parks, B. W., Jr; Fritz, J. D.; Pickering, H. W. *Scr. Metall.* **1989**, *23*, 951–956.
46. Newman, R. C. *Corros. Sci.* **1992**, *33*, 1653–1657.
47. Snyder, J.; Asanithi, P.; Dalton, A. B.; Erlebacher, J. *Adv. Mater.* **2008**, *20*, 4883–4886.
48. Lucey, V. F. *Br. Corros. J.* **1965**, *1*(2), 53–59.
49. Lucey, V. F. *Br. Corros. J.* **1965**, *1*(1), 9–14.
50. Sieradzki, K.; Newman, R. C. *J. Phys. Chem. Solids* **1987**, *48*, 1101–1113.

51. Kelly, R. G.; Frost, A. J.; Shahrabi, T.; Newman, R. C. *Metall. Trans. A* **1991**, *22A*, 531–541.
52. Saito, M.; Smith, G. S.; Newman, R. C. *Corros. Sci.* **1993**, *35*, 411–416.
53. Friedersdorf, F.; Sieradzki, K. *Corrosion* **1996**, *52*, 331–336.
54. Barnes, A.; Senior, N. A.; Newman, R. C. *Metall. Trans. A* **2009**, *40*, 58–68.
55. Newman, R. C.; Shahrabi, T.; Sieradzki, K. *Scr. Metall.* **1989**, *23*, 71–74.
56. Serebrinsky, S. A.; Galvele, J. R. *Corros. Sci.* **2004**, *46*, 591–612.
57. Newman, R. C.; Corderman, R. R.; Sieradzki, K. *Br. Corros. J.* **1989**, *24*, 143–148.
58. Sieradzki, K.; Kim, J. S.; Cole, A. T.; Newman, R. C. *J. Electrochem. Soc.* **1987**, *134*, 1635–1639.
59. Deakin, J.; Dong, Z. H.; Lynch, B.; Newman, R. C. *Corros. Sci.* **2004**, *46*, 2117–2133.

2.06 Intergranular Corrosion

D. L. Engelberg
Materials Performance Centre, The Mill/A11, School of Materials, University of Manchester, P.O. Box 88, Sackville Street, Manchester M60 1QD, UK

Abbreviations

CSL Coincidence site lattice
DOF Degree of freedom
FCC Face centered cubic
GB Grain boundary
HAZ Heat affected zone
IGA Intergranular attack
IGC Intergranular corrosion
NHE Normal hydrogen electrode
SCE Saturated calomel electrode
TTP Time–temperature–precipitation
UVW Miller indices

Symbols

E Potential (V)
i Current density ($A\,cm^{-2}$)
t Time (s)
O Charge density ($C\,cm^{-2}$)
θ Rotation angle (°)

2.06.1 Nature of Intergranular Corrosion

Intergranular corrosion (IGC) is a form of localized corrosion that proceeds along grain boundaries, synonymously referred to as intergranular attack (IGA). Localized corrosion at grain boundaries is caused by the anodic dissolution of (i) regions depleted of alloying elements, (ii) second phase precipitates, or (iii) regions with segregated alloying or impurity elements. The remaining part of the exposed surface typically functions as the cathode, and large cathodic areas support the anodic dissolution process. The ratio between cathode and anode is generally in excess of unity, and is dependent on factors such as the volume fraction and distribution of electrochemically active phases, the distribution of deleterious alloying and impurity elements, as well as grain size, particularly in nanocrystalline materials. The corrosion rate is dependent on the prevailing corrosion mechanism, and factors such as the diffusion of species to or from the anodic front can govern the dissolution kinetics.

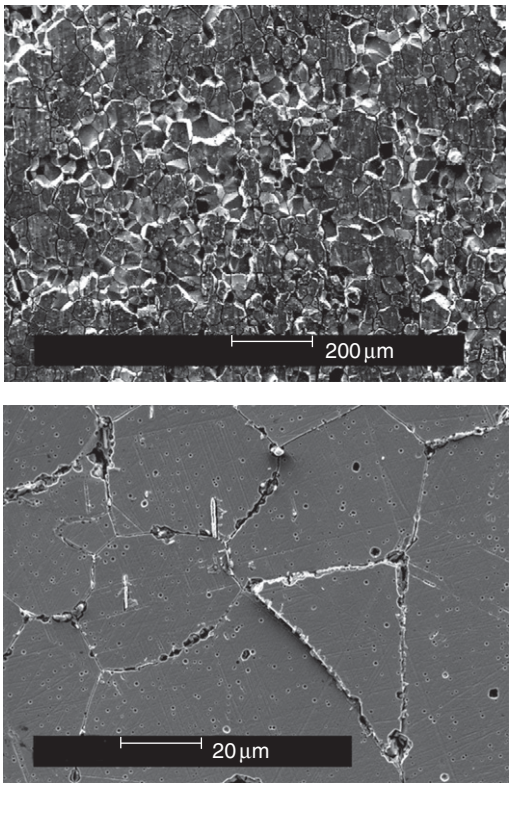


Figure 1 Scanning electron micrographs of sensitized austenitic stainless steels with (a) a severely attacked surface after grain dropping, and (b) a dislodged grain on an attacked surface.

IGC occurs in pure metals as well as engineering alloys, and the term sensitization is frequently used to describe microstructure sensitivities to intergranular attack. In stainless steels, for example, sensitization can be caused through the precipitation of chromium-rich carbides at grain boundaries (**Section 2.06.2.2.1**). A major characteristic of IGC is the development of a relatively homogenous and uniform depth of attack. The dissolution of grain boundaries causes the dislodgement of grains, often referred to as grain dropping. Grain dropping is responsible for most of the weight loss observed after IGC exposure, and corrosion rates can therefore be several orders of magnitude higher than during general corrosion. The nature of intergranular attack with grain dropping events is shown in **Figure 1**. A metallographic section with a uniform depth of attack is shown in **Figure 2**. Nonuniform IGC attack can also occur, and is typically related to microstructure heterogeneities, as well as factors that locally impair the corrosion process. The latter may include, for instance, the coverage of part of an exposed surface with an electrochemically nonactive deposit, or local differences in solution chemistry.

2.06.1.1 Theories of Intergranular Corrosion

Three main theories summarize the occurrence of IGC, comprising (i) the element depletion theory, (ii) the microcell or electrochemical theory, and (iii)

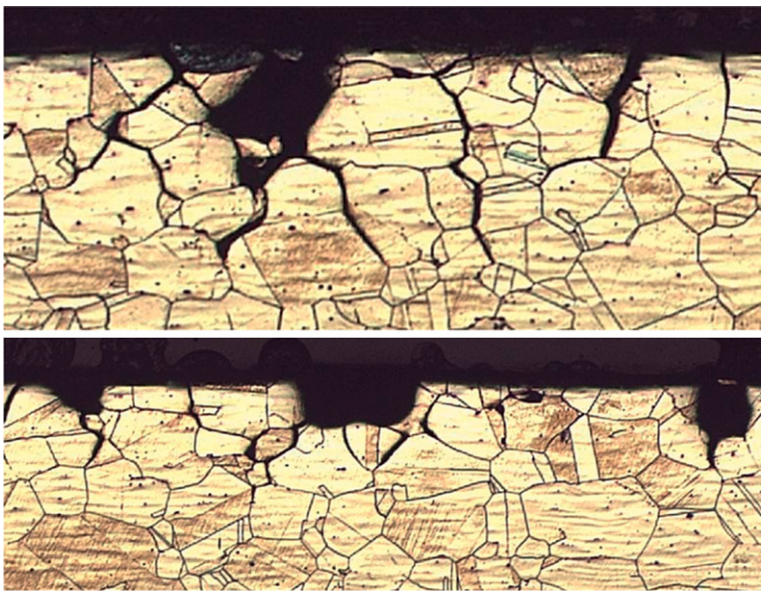


Figure 2 Optical micrographs of two metallographic sections after IGC testing of sensitized austenitic stainless steel. The images highlight the intergranular nature of attack with grain dropping events close to the specimen surface. After IGC testing the sample was etched in 10% oxalic acid to reveal the intergranular nature of attack.

the segregation theory.[1,2] The depletion theory is based on the local impoverishment of an alloy element, such as chromium in the vicinity of chromium-rich carbides. This results in the breakdown of the passive film in the vicinity of these carbides, producing preferred sites for anodic dissolution. The microcell or electrochemical theory is based on the formation of a local electrochemical cell between precipitate and matrix, followed by dissolution of the less noble element or phase. A typical example is the anodic dissolution of second phase precipitates, such as Al_3Mg_2 in the 5XXX series aluminum alloys. The segregation theory is based on the presence of an increased concentration of alloying or impurity elements at or in the vicinity of the grain boundary, such as sulfur or phosphorus in stainless steels. Dissolution of these regions is related to the adsorption of these elements at the grain boundary, ultimately changing the chemical and electrochemical activity.[3]

Another theory of IGC also considers local stresses at boundaries, introduced through the nucleation of second phase precipitates.[1,4] Preferred local attack then takes place at these distorted lattices, close to the second phase precipitate. It is also hypothesized that stresses and strains may lead to a higher degree of atomic disorder of the boundary, resulting in a higher driving force for anodic dissolution. Microscopic stresses or strains at grain boundaries possibly play a secondary role in IGC, for example, they may be responsible for enhanced diffusion pathways.

2.06.1.2 Metallurgical Aspects

The material composition and associated crystallographic structure have a large effect on the IGC behavior including, (i) the type of alloying and impurity element, (ii) the associated concentrations, (iii) their spatial distributions, and (iv) individual electrochemical activities. The type and concentration of alloying or impurity elements generally determines the crystallographic equilibrium structure of a material, and possible phase transformations change element solubilities and associated diffusivities. For example, increasing the concentration of carbon or the concentration of chromium in austenitic stainless steels, respectively, decreases and increases the time to form chromium depleted zones in the vicinity of chromium-rich carbide precipitates.[4] The spatial distribution and connectivity of susceptible microstructure constituents is also important, since IGC cannot be sustained without an inter-connected network of susceptible entities. Large distances between susceptible constituents impede or even prevent the formation of a connected intergranular pathway throughout the microstructure. The electro-chemical activity determines the tendency to dissolve individual elements, even if present only in solid solution or bound to complex intermetallic phases. For example, in the 2XXX series aluminum–copper alloys the presence of copper in solid solution raises the electrochemical potential of the matrix phase, and therefore reduces the electrochemical potential difference to copper-rich precipitates.[5]

The IGC behavior is a strong function of cold work and grain size.[1,4,6] Cold work introduces strain energy into the material, producing changes in the microstructure, such as the occurrence of slip bands, microcell arrays, or phase transformations. These changes are strongly dependent on the crystallographic structure and material composition. For instance, in leaner grade austenitic stainless steels (i.e., Type 304), a considerable volume fraction of deformation induced martensite can develop with the introduction of medium to large amounts of cold work (i.e., $>20\%$). The susceptibility to IGC is then governed by the stability and sensitization propensity of the phases present. The application of small amounts of cold work ($\leq 20\%$) prior to sensitization enhances the diffusion of elements along slip lines, and promotes precipitation at grain boundaries. Subsequent sensitization heat treatments have a strong influence on the precipitation kinetics, and the choice of sensitization temperature can change the location for precipitation, from predominantly at grain boundaries to precipitation at slip bands. The application of higher sensitization temperatures can even induce recrystallization. The effect of grain size on the IGC behavior is dependent on the investigated corrosion system,[4] and differing IGC behaviors with grain size have been observed, for example, with exposures to mildly or strongly oxidizing media.[1] In stainless steels, the sensitization propensity has been reported to increase with decreasing grain size, associated with greater amounts of chromium that can enter the grain boundary region and form chromium-rich precipitates in a given time interval.[6] Contrarily, decreasing sensitization propensities with decreasing grain sizes have also been found, attributed to larger grain boundary surface areas and the formation of a less continuous network of precipitates.[1,4]

2.06.1.3 Grain Boundary Structure and Networks

Grain boundaries are in a higher state of disorder than the crystal lattice, that is, they represent 2-dimensional

lattice imperfections, with increased surface energies compared to the grain interior. This explains why grain boundaries are more active than the adjacent matrix phase, and can therefore easily be revealed with metallographic etching techniques. A simplified geometric description of a grain boundary crystallographic structure uses five macroscopic degrees of freedom (DOF).[7,8] These include three DOF for the relative misorientation between the crystal lattices of two neighboring grains, and the remaining two DOF for referencing the grain boundary plane to the grain orientations. The relative misorientation is described with a rotation axis (UVW) common to both adjoining grains, accounting for two DOF, and a rotation angle (θ) around this axis, defining the third DOF. Grain orientation relationships (misorientations) can be obtained from 2-dimensional assessments (i.e., a plain metallographic section), but information about the grain boundary plane typically involves 3-dimensional material characterization.[7]

The grain boundary network is heterogeneous, and consists of an assembly of different grain boundary crystallographic structures, with inherently different grain boundary free energies. These energy variations have a large influence on material properties, such as, segregation, precipitation, diffusion, and migration behavior. Different grain boundary crystallographic structures therefore have different susceptibilities to IGC, and their resistance can vary as a function of exposure condition. In austenitic stainless steels, for example, the application of an electrolytic nitric acid etch (ASTM E 112),[9] to determine the ASTM grain size, outlines the grain boundary network, without attacking coherent annealing twin boundaries. The coherent twin boundaries have more than one order of magnitude lower surface energy ($\approx 0.02\,J\,m^{-2}$) than general high angle grain boundaries ($\approx 1\,J\,m^{-2}$), and are therefore more resistant to attack. The application of an electrolytic oxalic acid etch,[10] in contrast, outlines the complete grain boundary network, including coherent twins. This behavior highlights the relationship between exposure condition and grain boundary free energy on material susceptibility.

The connectivity of grain boundaries also plays an important part. Crystallographic constraints, in the form of clustering of crystallographically related boundaries, can create large regions with corrosion resistant or susceptible grain boundary clusters. This is often observed in low- to medium stacking fault energy face centered cubic (FCC) materials, such as copper or austenitic stainless steels. Large regions with clusters of low energy boundaries, such as twins, can impede the growth of IGC. Further information about grain boundary structure, surface energy and physical property relationships,[11,12] as well as the influence of grain boundary connectivity,[7] can be found elsewhere in the literature.

2.06.1.4 Grain Boundary Susceptibility

Localized attack at grain boundaries is a function of the 3-dimensional grain boundary structure (5 DOF), and reduced susceptibilities to intergranular corrosion are associated with low energy grain boundaries. The group of low energy grain boundaries comprises low angle grain boundaries, as well as certain high angle grain boundary configurations, such as twin boundaries. Low angle grain boundaries, with misorientations below 15 °, are typically less susceptible to attack than high angle grain boundaries. The resistance of low angle grain boundaries increases with decreasing misorientation, and threshold angles for grain boundary resistance have also been reported.[13] The presence of a significantly increased fraction of low angle grain boundaries after plastic deformation, however, is responsible for enhanced diffusion pathways and increased IGC propensities.

High angle grain boundaries, with misorientations in excess of 15 °, also possess certain low energy configurations which are more resistant to IGC. Figure 3 shows a grain boundary network of a sensitized austenitic stainless steel with a mixture of attacked and resistant high-angle grain boundaries. Low energy grain boundaries, such as twin boundaries,

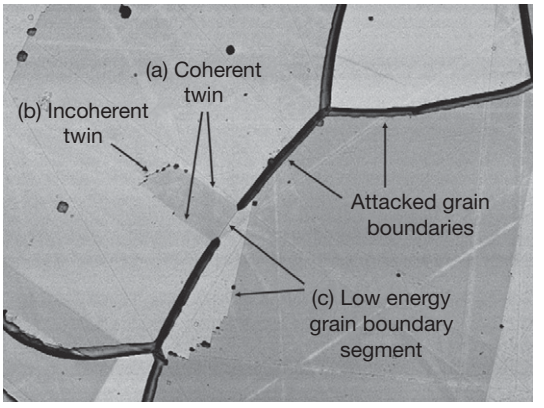


Figure 3 Backscattered scanning electron micrograph of a grain boundary network in a sensitized austenitic stainless steel, with susceptible and immune grain boundaries. (a) Coherent and (b) incoherent twin grain boundaries, as well as some (c) low energy grain boundary segments show full or partial immunity to attack.

generally show a high resistance to IGC. To characterize different grain boundary structures numerous geometric models are available. For microstructure-property investigations, in particular corrosion studies in cubic crystal systems, the coincidence site lattice (CSL) model has frequently been used. This model is based on three macroscopic DOF ((θ)/(UVW)), and does not take the grain boundary plane into consideration. In the CSL model, grain boundaries are defined with the inverse volume density of coinciding lattice atom sites ($\sum$), that is, a high density of coinciding atoms is associated with low $\sum$ boundaries. Grain boundaries up to $\sum$29 are typically considered to be resistant, but large variations in IGC susceptibilities of different grain boundary types have been observed. The misorientation is only an indicator for grain boundary resistance. Increased microstructure resistance is frequently associated with the presence of twin ($\sum$3) grain boundaries, which are usually the most prominent group in the low $\sum$ boundary population.[7]

The most resistant grain boundary is the coherent twin boundary, defined as $\sum$3 in the CSL model. Coherent twins can be expressed as a symmetrical tilt grain boundary with low-index {111} grain boundary planes. These boundaries therefore have a high density of coinciding atom sites, and a low energy grain boundary plane configuration. Other $\sum$3 grain boundaries, such as the incoherent twin, are more susceptible to IGC than coherent $\sum$3 twins. Incoherent twin boundaries are frequently associated with symmetrical tilt boundaries on {112} planes, and have a higher free energy ($\approx 0.5\,\mathrm{J\,m^{-2}}$) than their coherent counterpart ($\approx 0.02\,\mathrm{J\,m^{-2}}$). **Figure 3** highlights the difference in twin boundary susceptibility and also gives examples of other resistant low-energy grain boundary segments. Certain low $\sum$ grain boundaries as well as some general high angle grain boundary configurations can also possess low index habit planes, leading to an increased IGC resistance. Additional factors, such as (i) the deviation from exact reference misorientation, (ii) the deviation from low index grain boundary plane configuration (i.e., curved boundary plane), or (iii) grain boundary faceting, can also influence the IGC resistance.

2.06.2 Intergranular Corrosion of Metals and Alloys

Aluminum alloys, stainless steels, and nickel alloys are the most common material systems associated with intergranular corrosion. This section gives a brief review of IGC in common alloy systems, with strategies to mitigate these problems.

2.06.2.1 Aluminum Alloys

The occurrence of IGC in aluminum alloys is a function of the material composition, and certain alloying elements are responsible for the prevailing corrosion mechanism. The IGC susceptibility of heat treatable aluminum alloys is mainly related to the quenching and ageing treatment after solution annealing and small changes in these routines can cause significant changes in material resistance.[14] Another important factor for the IGC behavior is also the thermomechanical process history. Alloy groups that have been generally associated with IGC include, precipitation hardenable 2XXX series Al–Cu alloys, work hardenable 5XXX series Al–Mg alloys, and precipitation hardenable 7XXX Al–Zn–Mg alloys. The 6XXX precipitation hardenable Al–Mg–Si series has also shown IGC, but to a far smaller extent than the other alloy groups. The above-mentioned aluminum alloys are also associated with exfoliation corrosion, which is a type of intergranular corrosion which proceeds parallel to the specimen surface. Exfoliation corrosion is related to the development of excessive corrosion products at grain boundaries, indicating a layer-wise spallation of plate-like grain segments. This type of attack occurs in highly orientated microstructures of wrought aluminum alloys. IGC and exfoliation corrosion proceed in a similar manner, but exfoliation has also been found in materials which were not susceptible to IGC.

In 2XXX series Al–Cu alloys, including Al–Cu–Mg and Al–Cu–Li alloys, IGC behavior is primarily influenced by the copper concentration.[5,14,15] Copper in solid solution increases the corrosion potential of the matrix phase, and copper segregation or depletion can cause local variations in the electrochemical potential. Copper-depleted regions are less noble and susceptible to anodic dissolution in particular, if adjacent to copper-rich intermetallic phases such as Al_2Cu, Al_2CuMg, or Al_2CuLi. Dissolution takes place in narrow regions close to precipitated phases at the grain boundary. Dealloying of copper-containing intermetallic phases, and copper re-precipitation has also been reported to support local dissolution reactions.[5] To reduce IGC susceptibilities, specifically designed heat and quenching treatments can be employed, which produce a uniform distribution of precipitates.[15] In the 5XXX series Al–Mg alloys with more than 3% Mg, the susceptibility to IGC

is typically based on the dissolution of the intergranularly precipitated β-Al_3Mg_2 phase.[5,15] This phase predominantly precipitates in the temperature range between 50 and 200 °C, due to a rejection of magnesium out of solid solution. Mg-depleted regions have also been found in the vicinity of the β-phase. Electrochemical measurements on bulk β-Al_3Mg_2 in NaCl solution indicated more electro-negative corrosion and passivity breakdown potentials than the aluminum matrix phase. These potential differences are responsible for the anodic dissolution of the β-Al_3Mg_2 phase.[5] Microalloying with zinc produces a stable Al–Mg–Zn phase (τ-phase) and precludes the precipitation of β-Al_3Mg_2 in the 5XXX series alloys. In the 7XXX series Al–Zn–Mg alloys, IGC has frequently been attributed to the presence of Mg-rich grain boundary precipitates, as well as intergranular enrichments of Mg and Zn. Copper bearing 7XXX series Al–Zn–Mg–Cu alloys have been reported to contain copper depleted regions at grain boundaries, with an IGC mechanism similar to the 2XXX series Al–Cu alloys.[15] The application of over-aging heat treatments, and the resultant uniform distribution of precipitates can prevent IGC.[15] In the 6XXX series, Al–Mg–Si alloys IGC has been attributed to the anodic dissolution of the hardening phase Mg_2Si.[15]

2.06.2.2 Stainless Steels

Stainless steels can broadly be divided into five groups, comprising (i) ferritic, (ii) austenitic, (iii) martensitic, (iv) duplex (austenitic–ferritic), and (v) precipitation hardenable materials. Different crystallographic structures have different corrosion behavior, often related to variations in diffusivity and solubility. Chromium depletion is the most common cause for IGC in these alloys, and degradation at grain boundaries is normally referred to as sensitization.[1,4,6] Sensitization is brought about by the presence of chromium-rich phases at grain boundaries, such as carbides, nitrides or intermetallic phases.[16] Exposure of sensitized stainless steel to oxidizing media produces intergranular attack at these chromium depleted regions.[1,4] In highly oxidizing media, such as boiling nitric acid, IGC can also occur in annealed stainless steel. This form of attack is generally related to the segregation and preferred dissolution of impurity elements at grain boundaries.[2]

2.06.2.2.1 Sensitization of austenitic stainless steel

Sensitization is caused by thermal or thermomechanical process treatments, such as annealing, welding, or hot working. Exposure of austenitic stainless steels to temperatures between 450 and 850 °C can induce sensitization, due to the precipitation of chromium-rich carbides.[6,16] **Figure 4(a)** schematically shows sensitized grain boundaries decorated with chromium carbide precipitates and associated chromium depleted zones. Adverse service conditions, in the form of exposure to neutron irradiation in nuclear plant, can also cause sensitization, but without the nucleation of carbides. The amount of depletion typically varies between different boundary types, due to different grain boundary crystallographic structures.

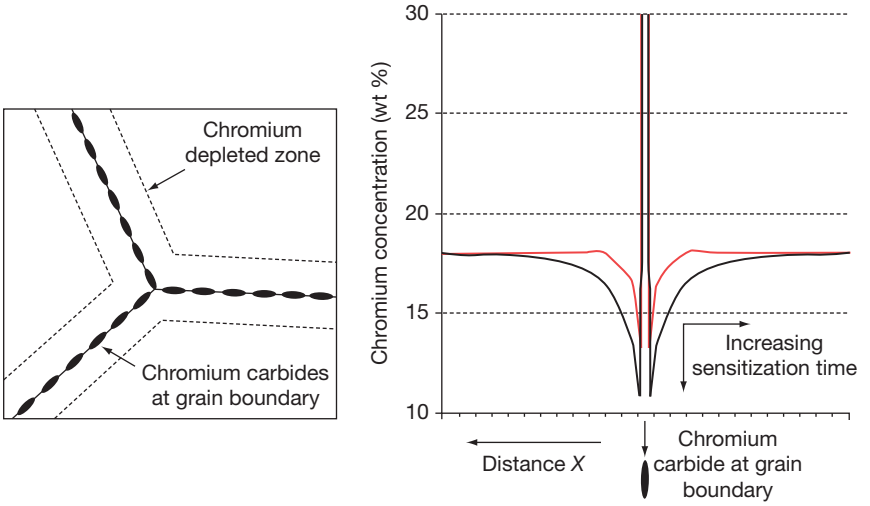


Figure 4 Schematic illustration of a sensitized grain boundary with chromium carbide precipitates and typical chromium depletion profiles. Longer sensitization treatments increase both the width and depth of the chromium depletion profiles, until chromium replenishment takes place.

The depletion profiles in **Figure 4(b)** show a chromium composition gradient in the direction perpendicular to the boundary, with the shape and depth dependent on applied heat treatment parameters. The depletion is caused by the slow diffusivity of chromium, compared to carbon, producing regions with local chromium impoverishment, followed by a breakdown of the passive film at these regions. In **Figure 4**, the high concentration of chromium at the grain boundary is associated with the presence of chromium-rich carbide precipitates. The influence of reduced chromium concentrations on the anodic behavior of these regions, for example, if exposed to oxidizing environments, is shown in **Figure 5**. Lower chromium concentrations show increased current densities in the active as well as passive regions of the anodic polarization curve.

A major factor for the development of chromium depleted regions is the chromium and carbon content, as well as the solubility of carbon in the matrix phase. Nucleation of carbides is caused by a decreasing solubility of carbon in the austenitic matrix phase with decreasing temperature, for instance on cooling through the sensitization temperature regime.[16] The dominant carbide in austenitic stainless steels is $Cr_{23}C_6$, but impurity elements, such as nitrogen and boron, as well as alloying elements usually alter this composition. The presence of high nitrogen contents, for example, may cause sensitization through the nucleation of Cr_2N. The addition of alloying elements, such as molybdenum in higher alloyed stainless steels (i.e., type 316), may also promote the nucleation of intermetallic phases, such as σ-phase $(FeNi)_x(CrMo)_y$[16] or χ-phase ($Fe_{36}Cr_{12}Mo_{10}$).[16]

A typical method to illustrate the sensitization behavior is the time–temperature–precipitation (TTP) diagram, shown in **Figure 6**. TTP diagrams give an indication of the influence of material parameters on the sensitization propensity. **Figure 6(a)** shows the influence of carbon content, and **Figure 6(b)** the influence of grain boundary structure on sensitization. Increasing the carbon content increases the propensity to sensitization at shorter exposure times. The same is observed with different grain boundary structures. Precipitation occurs in the order of decreasing grain boundary energy, starting with a rapid decoration of δ-ferrite–austenite interfaces, followed by precipitation at general high-angle, then incoherent twin and finally coherent twin boundaries. Heat treatments in the lower temperature regime, that is, below 450 °C, usually have no significant effect on sensitization due to markedly reduced diffusivities. In the presence of preexisting carbide nuclei, however, sensitization has been reported below 450 °C after prolonged exposure. This is referred to as low-temperature sensitization, and the occurrence may take years before it becomes significant.[6]

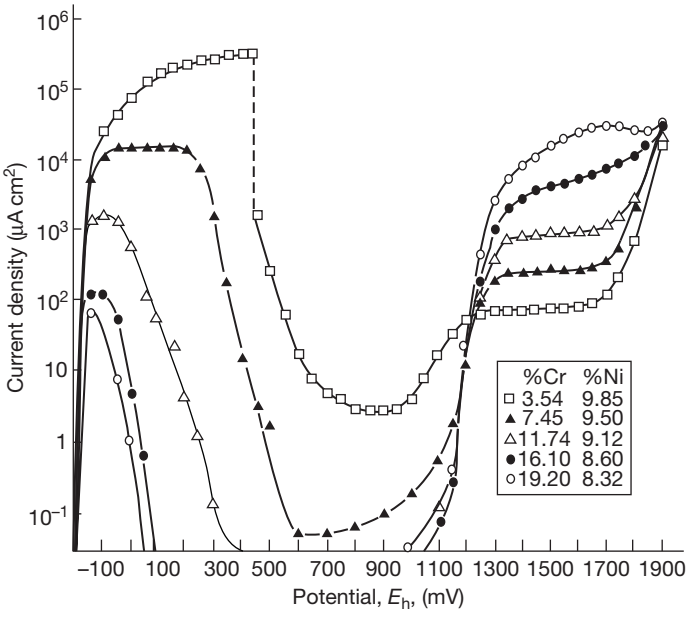


Figure 5 Anodic polarization curves of Fe–9% Ni–*x*% Cr alloys in deaerated 2 N H_2SO_4 solution at 25 °C. Potentials are given versus normal hydrogen electrode (NHE, E_h). Reproduced from Osozawa, K.; Engell, H.-J. *Corros. Sci.* **1966**, *6*, 389–393.

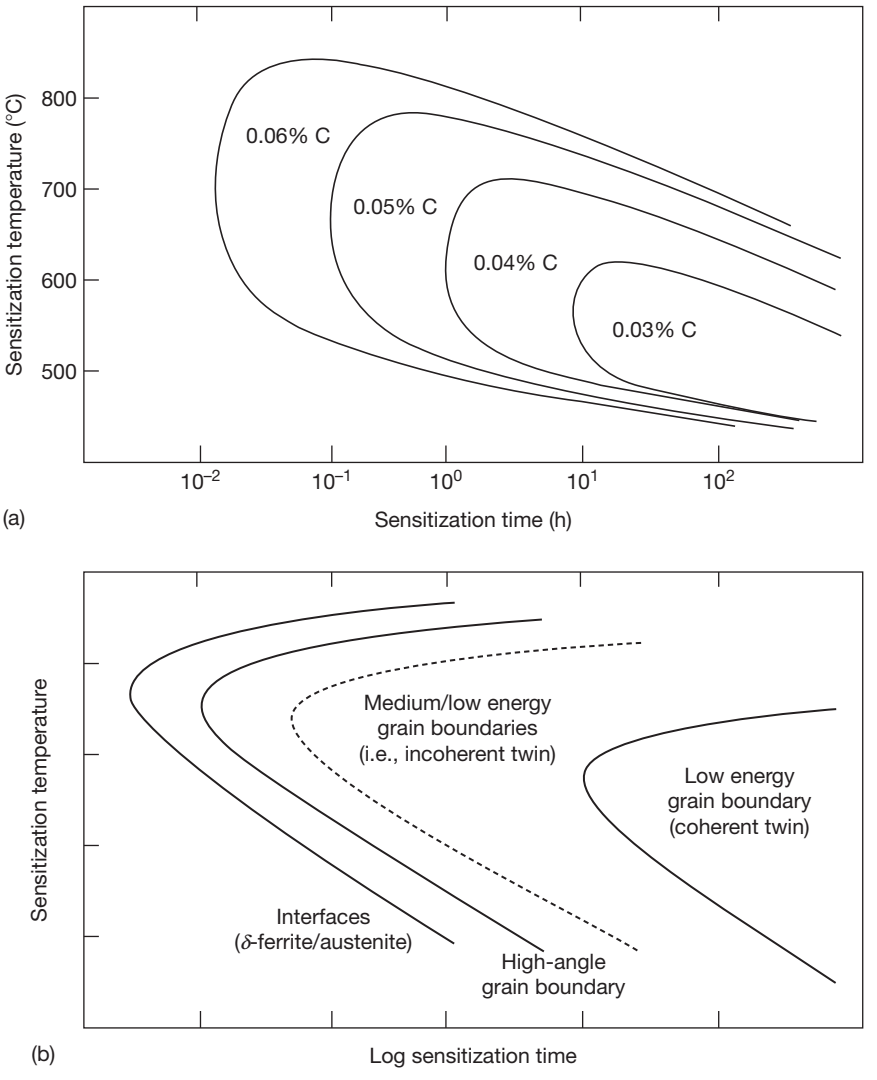


Figure 6 Schematic time–temperature–precipitation (TTP) diagram of austenitic stainless steels with the influence of (a) different carbon contents (based on data from Trillo *et al.*[48]), and (b) different grain boundary structures (based on data from Marshall[16] and Bruemmer and Was[49]).

2.06.2.2.2 Prevention of sensitization

For the prevention of sensitization in austenitic stainless steels, numerous possible mitigation strategies are available. (i) A reduction of the carbon content to concentrations below 0.03% (L-grade) usually prevents sensitization, but prolonged exposures in the sensitization temperature regime may still cause chromium depletion. (ii) The addition of stabilizing elements, such as titanium or niobium (stabilized grades, i.e., Type 321, Type 347), bind carbon into MC-type carbides. This reduces the available carbon in solid solution, and prevents the nucleation of intergranular $M_{23}C_6$ carbides. Stabilizing heat treatments at temperature in the region of 800–900 °C are typically applied to precipitate intragranular MC-type carbides. (iii) The application of a solution annealing treatment dissolves existing carbides, and fast cooling/quenching through the sensitization temperature regime prevents carbide nucleation. Typical solution annealing heat treatments for stainless steels are carried out between 1000 and 1150 °C. Sensitization can occur if the material is afterwards heated into the sensitization temperature regime. (iv) Prolonged exposure periods within the sensitization temperature regime can replenish chromium depleted regions with chromium from the grain interior. This is referred to as desensitization treatment. (v) Metallurgical alterations, such as alloying with higher chromium

contents, or the addition of molybdenum, as well as (vi) changes of manufacture or service conditions, such as the reduction of critical exposure temperatures, are possible measures to prevent sensitization.

2.06.2.2.3 Weld decay and knife line attack

During welding, in particular fusion-based welding processes, material situated in the fusion zone is locally exposed to temperatures close to the melting point. This induces a temperature gradient with distance from the fusion zone, referred to as the heat affected zone (HAZ), and material within the HAZ can be exposed to temperatures that cause sensitization. On exposure to IGC promoting environments, attack usually takes place in a band parallel to the fusion line. This process is known as weld decay. Figure 7 shows a typical appearance and location of weld decay, and gives three typical temperature profiles (1, 2, 3) with distance from the fusion zone. The time (t) at elevated temperature, specimen dimensions, welding process, and associated welding parameters are crucial factors for the occurrence of weld decay.

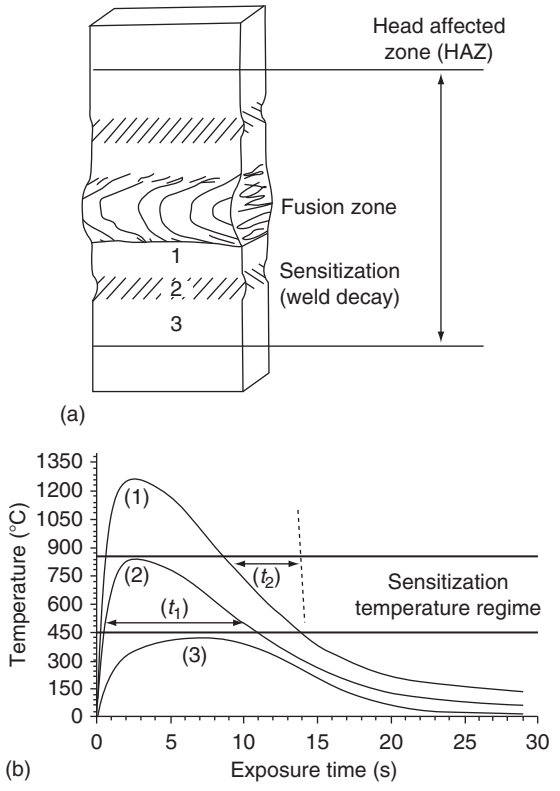


Figure 7 Schematic drawing of the occurrence of weld decay. (a) Weld line with fusion zone, heat affected zone (HAZ) and sensitized region. (b) Typical temperature–time profiles at points 1, 2, and 3 highlight different exposure times within the sensitization temperature regime ($t_1 > t_2$). Reproduced from Procter, R. P. M. Section 1.3 – Effects of Metallurgical/structure on Corrosion. In *Shreir's Corrosion Handbook*, 3rd ed. Metal/Environment Reactions; Shreir, L. L., Jarman, R. A., Burstein, G. T., Eds.; 2000; pp 1–44.

Knife line attack is associated with IGC in weldments of stabilized stainless steels. This phenomenon occurs after exposing stabilized materials, containing intragranular TiC or NbC precipitates, to temperatures that dissolve these stabilizing MC-type carbides. The required temperature range is typically in excess of 1150 °C, and can occur during fusion-based welding processes. On rapid cooling, neither the MC-type carbides, nor the $M_{23}C_6$ are precipitated and the carbon remains in solid solution. MC-type carbides are preferentially nucleated at temperatures between 800 and 900 °C, whereas intergranular $M_{23}C_6$-type carbides have a stronger driving force for nucleation at lower temperatures. Re-heating to temperatures within the sensitization regime, that is, 450–750 °C, then preferentially precipitates intergranular $M_{23}C_6$-type carbides in the HAZ, rendering the microstructure susceptible to IGC. Knife line attack occurs close to the weld bead, since high temperatures are required for the dissolution of the MC-type carbides.

2.06.2.2.4 Ferritic, martensitic, duplex, and precipitation hardening stainless steels

Sensitization in ferritic stainless steels is usually caused by the formation of intergranular chromium-rich carbides ($M_{23}C_6$) and nitrides (M_2N).[17] Similarly to austenitic stainless steels, chromium depletion is the cause for IGC in mildly oxidizing solutions, and the dissolution of intergranular carbides or segregated elements causes IGC in solutions with a high oxidizing power.[17,18] Rapid precipitation of intergranular carbides has been reported to occur in the sensitization temperature regime between 430 and 930 °C,[17] but annealing in excess of 930 °C, followed by rapid cooling, cannot prevent sensitization. This occurs due to significantly faster diffusivities of chromium, carbon, and nitrogen in ferrite compared to austenite.[17,19] For IGC resistance, desensitization treatments between 700 and 800 °C can be used to replenish the chromium depleted regions.[17] The addition of stabilizing elements, in the form of titanium or niobium, is also common practice to prevent IGC in these materials.

The susceptibility of martensitic stainless steels to IGC is a strong function of the microstructure

condition, and is dependent on the preceding thermal history. Peak hardened microstructures after quenching have an increased resistance to IGC. An increased sensitivity to IGC has been reported after temper annealing at temperatures between 450 and 650 °C, due to the precipitation of intergranular carbides.[1,19,20] Exposure of martensitic stainless steels to temperatures in excess of 700 °C equilibrates local differences in the chromium content, leading to desensitization.[1]

Austenitic–ferritic (duplex) stainless steels contain ~50% ferrite and ~50% austenite and usually have a good resistance to IGC. These alloys possess higher chromium contents (above 20%), with carbon concentrations far below 0.03%, and modern duplex stainless steels have insufficient carbon for sensitization. Susceptibility of duplex stainless steels is frequently associated with the presence of a critical fraction of austenite–ferrite interfaces.[21,22] In the presence of sufficient carbon, these interfaces are prone to carbide precipitation after short exposure times, due to rapid diffusion of carbon and chromium in ferrite. Narrow chromium depletion regions form at the austenitic side of the interface, due to the slower diffusion kinetics of chromium in austenite.[21,22] Sensitization is not confined to ferrite–austenite interfaces, and chromium depletion also occurs at austenite–austenite boundaries. Desensitization proceeds faster than in austenitic stainless steel, and duplex stainless steels can undergo rapid healing of chromium depleted regions. Generally, intergranular carbide precipitation occurs at temperatures between 450 and 850 °C,[4] with solution annealing usually carried out at temperatures in excess of 1000 °C. Duplex stainless steels also undergo complex precipitation and decomposition reactions, with the formation of intermetallic phases such as σ-phase and χ-phase.[16,20] These intermetallic phases are rich in chromium and molybdenum, usually nucleate at grain boundaries, and can cause local element depletions, rendering duplex stainless steels susceptible to IGC.

Precipitation hardenable stainless steels, such as the martensitic 17-4 PH grade (17%Cr/4%Ni/4% Cu) derive their strength from the nucleation of intermetallic strengthening phases on ageing. These strengthening phases are typically soluble in austenite, and precipitate on the formation of martensite. Fully austenitic alloys are also available, which are strengthened by intermetallic precipitates, such as Ni_3Ti. Sensitization in these materials typically occurs on heat treating and reversion of martensite into austenite, often associated with the HAZ after welding.[1] Precipitation hardenable stainless steels are susceptible to sensitization similar to their austenitic counterparts, with temperatures up to ≈850 °C.

2.06.2.3 Nickel and Nickel Alloys

Structural nickel alloys can broadly be grouped according to their main compositional constituents, comprising nickel–chromium–iron alloys (Ni–Cr–Fe), nickel–chromium–molybdenum alloys (Ni–Cr–Mo), and nickel–molybdenum alloys (Ni–Mo).[1,3,4] These alloy systems behave similarly to austenitic stainless steels, and the main alloying elements for improved IGC resistance are chromium and molybdenum. Chromium gives resistance in oxidizing media, whereas molybdenum increases the resistance in reducing media. Nickel has a low solubility of carbon, and even concentrations below 0.03% can cause sensitization due to the precipitation of chromium-rich carbides. Chromium nitrides and the presence of certain intermetallic phases can also cause sensitization, as well as segregated elements, such as sulfur or phosphor. For instance, IGC in pure nickel can occur through the presence of sulfur segregation at grain boundaries, with the extent dependent on the grain boundary crystallographic structure and the exposed environmental condition.[23]

Ni–Cr–Fe[1,3,4] alloys are the largest group of nickel-based alloys. The chromium content typically varies from 15% in the basic grades, for example, alloy 600, to far in excess of 30% in the higher-alloyed materials, such as alloy 690. For increasing the corrosion and pitting resistance, 2–3% molybdenum and copper are frequently added. Molybdenum bearing alloys, however, can develop complex carbides and intermetallic phases. Depending on the carbon content, Ni–Cr–Fe alloys are susceptible to sensitization in the temperature range between 450 and 850 °C, forming carbides of type $M_{23}C_6$ and M_7C_3. Solution annealing with fast quenching routines can be used to prevent sensitization.

Ni–Cr–Mo[1,3,4] alloys have a higher resistance to localized corrosion, due to the presence of increased concentrations of molybdenum, typically between 10% and 20%, such as alloy 22. The combination of high chromium and increased molybdenum contents make these alloys highly resistant to a broad spectrum of oxidizing and reducing environments. In contrast, the presence of molybdenum creates more complex carbides and intermetallic phases. Sensitization depends on the carbon content, but can occur between 600 and 1150 °C. The microstructure

nucleates carbides of type M_6C, $M_{23}C_6$, and M_2C, producing regions depleted in chromium and molybdenum. Solution annealing only becomes effective at temperatures in excess of 1220 °C.[4] Additional elements such as copper, tantalum, and tungsten are often incorporated to improve the corrosion resistance.

Ni–Mo[3] alloys typically contain 25–30% molybdenum, and can form ordered intermetallic phases of the types Ni_4Mo and Ni_3Mo. These alloys have an excellent resistance to reducing media, and IGC occurs in regions depleted of molybdenum. Intergranular carbides, such as M_2C, M_6C, $M_{12}C$, can also form. To prevent sensitization, materials with carbon contents below 0.005% are utilized. Knifeline attack can occur if these alloys are heated in excess of 1250 °C.[3]

2.06.2.4 Other Metals and Alloys

Most metals and alloy systems are susceptible to IGC, and a few examples of different alloy systems are briefly summarized. Copper is susceptible to IGC with exposure to certain etching solutions, such as mixtures of hydrochloric and acetic acid, as well as in solutions containing high concentrations of ammonia.[24] Brass is intergranularly attacked by ferric sulfate and cobalt sulfate,[24] and iron is susceptible to IGC in ammonium nitrate solutions.[25] Pure zinc,[26] containing small amounts of lead, tin, aluminum, or other impurity elements, shows a strong sensitivity to IGC when exposed to hot water or steam environments. Zinc alloys with 4% and 12% aluminum also suffer from IGC in concrete, which is mostly attributed to the segregation of impurity elements and the formation of aluminum rich grain boundary phases.[26] Zirconium shows IGC after immersion in high concentrations of hydrochloric acid containing small amounts of oxidizers, such as ferric- or cupric-ions.[27] More examples of localized attack can be found in the individual material sections of this book.

2.06.3 Intergranular Corrosion Testing

Intergranular corrosion testing is carried out either under real service conditions, in simulated service environments, or in accelerated test environments. Accelerated tests are required to rank exposed materials in a similar order as exposure to real service environments. However, the rates of degradation in these tests may not have any relation to the actual service environment. For accelerated IGC testing, numerous standard practices and test methods are available, but IGC testing can also be carried out in nonstandardized environments. Calculated corrosion rates after IGC exposures are based on the same principles as for general corrosion, implying a homogeneous dissolution. Localized attack must be reported separately.

2.06.3.1 Standard Practices and Test Methods

A range of standard practices and test methods for the assessment and ranking of IGC behaviors are available, and have been published under the auspice of different organizations and institutions. These include the International Standard Organization (ISO),[28–33] the National Association of Corrosion Engineers (NACE) and the American Society for Testing Materials (ASTM).[10,34–41] A large number of national standards are also available. In this section, ASTM and ISO standards for assessing the IGC behavior of aluminum alloys, stainless steels, and nickel alloys are summarized. **Table 1** gives the standard practices for IGC of aluminum alloys, **Table 2** for IGC testing of stainless steels and **Table 3** for nickel alloys. Exfoliation tests for aluminum alloys are also included in **Table 1**.

IGC tests in **Table 1** dissolve specific material constituents that are characteristic of individual aluminum alloys. The tested specimens are either assessed with weight loss measurements, in the case of exposure to nitric acid (ASTM G 67),[38] or using metallographic sectioning methods. Both exfoliation tests in **Table 1** are assessed by visual examination of the exposed surfaces. IGC tests in chloride containing solutions (ISO 11846, ISO 15329)[32,33] produce, with anodic polarization, a localized breakdown of the passive film, which results in either intergranular or intragranular pit initiation sites. The location of these initiation events is then used for material discrimination, with intergranular sites identifying IGC susceptible materials. Further analysis may also comprise an assessment of the extent of intergranular attack, in order to categorize the material with susceptibility ratings. Information about the influence of individual alloying elements, changes in electrolyte concentrations, or the influence of exposure parameters can be found in the literature, including the ASSET[42] (ASTM G 66),[41] EXCO[42] (ASTM G 34),[34] nitric acid[43] (ASTM G67),[38] and chloride/peroxide[14] (ASTM G 110/ISO 11846)[33,36] test.

Table 1 ASTM and ISO practices for IGC and exfoliation testing of aluminum alloys

Standard	*Test solution*	*Materials*	*Nature of attack*	*Duration / temperature*
ASTM G 34 (2001)	Sodium chloride, potassium nitrate, nitric acid -EXCO-	Wrought; AA2XXX, AA7XXX	Exfoliation	≤48–96 h/25 °C
ASTM G 66 (1999)	Ammonium chloride, ammonium nitrate, ammonium tartrate, hydrogen peroxide -ASSET-	Wrought, AA5XXX (>2% Mg)	Exfoliation	24 h/65 °C
ASTM G 67 (2004)	Nitric acid -NAMLT-	Wrought, AA5XXX (Al–Mg; Al–Mg–Mn)	GB precipitation (Al_3Mg_2)	24 h/30 °C
ASTM G 110 (2003)	Sodium chloride, hydrogen peroxide	Wrought + cast; AA2XXX, AA6XXX, AA7XXX	GB precipitation	>6 h/30 °C
ISO 11846 (2008)	(Method A) sodium chloride, hydrogen peroxide	Wrought + cast; AA2XXX, AA6XXX, AA7XXX	(Method A–C) GB precipitation	>6 h/30 °C
	(Method B) sodium chloride, hydrochloric acid	Wrought + cast; Solution heat-treatable Aluminum alloys		24 h/RT
	(Method C) anodic polarization in sodium chloride	Aluminum–lithium alloys		Polarization/RT
ISO 15329 (2008)	Sodium chloride	Wrought + cast; Solution heat-treatable AA2XXX, AA6XXX, AA7XXX, AA8XXX	GB precipitation	Polarization/RT

RT = Room temperature, GB = Grain boundary.

IGC tests for austenitic, ferritic, and duplex stainless steels are summarized in **Table 2**. Two different test methods are typically distinguished, on the basis of either material quality testing or sensitization propensity testing. The later assessment involves the exposure of materials to short sensitization heat treatments before IGC testing, with the aim to provide information about the propensity of the material to develop a sensitized grain boundary network. Material quality testing is typically carried out without any additional heat treatment. The IGC tests in **Table 2** are assessed either with weight loss measurements (ASTM A 262 – Method B, C, F and ISO 3651-1),[10,29] or by visual inspection of the exposed specimen surface after bending (ASTM A 262 – Method E and ISO 3651-2).[10,30] Bending tests only provide a qualitative measure of the IGC susceptibility, and these test environments are typically less aggressive. Metallographic sectioning can also be used for further susceptibility discrimination. Tests summarized in **Table 2** are all sensitive to intergranular chromium depletion, but individual tests also show sensitivities to chromium-rich phases, molybdenum-depletion around σ-phase, or other microstructural constituents, such as end-grains.

The oxalic acid test in **Table 2** is an electrolytic screening test, and is used for rapid assessment of cast and wrought austenitic (ASTM A262 – Method A)[10] and ferritic (ASTM A763 – Method W)[39] stainless steels. The recommended current density and exposure time severely over-etch the microstructure, in order to facilitate a simple microstructure susceptibility classification.[17] The test reveals chromium depleted regions in the vicinity of sensitized boundaries or intermetallic phases.[17] Experiments with pure chromium have also indicated an attack of chromium rich phases,[22] making the oxalic acid test suitable for accepting resistant materials, but not for the rejection of susceptible material heats.[17] The oxalic acid test discriminates three different microstructure conditions comprising, (i) an IGC resistant step-like microstructure, (ii) a dual microstructure, containing a mixture of stepped and ditched grain boundaries, and (iii) a ditched structure, associated with deep, grain-surrounding etch trenches. **Figure 8** compares these classifications for a wrought austenitic stainless steel plate. The standard oxalic acid test yields only qualitative information, but quantitative information about the attacked grain boundary network can also be extracted. The extent of susceptible grain boundaries can be determined, for instance, by using the oxalic acid test in conjunction with image analysis methods, or with mean linear intercept procedures. These assessments can provide valuable information to quantify material susceptibilities.

Table 2 ASTM and ISO practices for IGC testing of stainless steels[10,29,30,39]

Standard	*Test solution*	*Materials*	*Nature of attack*	*Duration/ temperature*
ASTM A 262 (2008)	(Method A) 10% Oxalic acid or ammonium persulfate	(Methods A–C, E) Wrought + cast; austenitic SS	(A) Chromium depletion end grains	Polarization/ RT
	(Method B) Ferric sulfate, 50% sulfuric acid -STREICHER-		(B) Chromium depletion σ-phase in Ti + Nb stabilized and Cr–Ni–Mo containing SS	120 h/boiling
	(Method C) 65% Nitric acid -HUEY-		(C) Chromium depletion σ-phase in Cr–Ni–Mo containing SS, end-grains	240 h/boiling
	(Method E) Copper, copper sulfate, 16% sulfuric acid -STRAUSS-	(E) Weldments	(E) Chromium depletion	24 h/boiling
	(Method F) Copper, copper sulfate, 50% sulfuric acid	(F) Cast austenitic SS	(F) Chromium depletion in Cr–Ni–Mo containing castings	120 h/boiling
	[*(Method D) 10% Nitric acid, 3% hydrofluoric acid] -WARREN-	(D) Type 316, 316L, 317, 317L	(D) Chromium depletion	4 h/70 °C
ASTM A 763 (2004)	(Method W) 10% Oxalic acid	(Methods W–Z) Ferritic SS	(Methods W–Z) Chromium depletion	Polarization/ RT
	(Method X) Ferric sulfate, 50% sulfuric acid		(X) Intermetallic phases in Ti/Nb stabilized SS	24–120 h/ boiling
	(Method Y) Copper, copper sulfate, 50% sulfuric acid			96–120 h/ boiling
	(Method Z) Copper, copper sulfate, 16% sulfuric acid	(Z) Weldments		24 h/boiling
ISO 3651-1 (1998)	65% Nitric acid -HUEY-	Wrought + cast; austenitic and duplex SS	Chromium depletion	240 h/boiling
ISO 3651-2 (1998)	(Method A) Copper, copper sulfate, 16% sulfuric acid -STRAUSS-	(Methods A–C) Wrought + cast + weldments; (A) Austenitic, ferritic, duplex SS #	(Methods A–C) Chromium depletion	20 ± 5 h/ boiling
	(Method B) Copper, copper sulfate, 35% sulfuric acid	(B) Austenitic, duplex SS #		20 ± 5 h/ boiling
	(Method C) Ferric sulfate, 40% sulfuric acid	(C) Austenitic, ferritic, duplex SS #		20 ± 5 h/ boiling

*Excluded from ASTM A262, # Material range defined in standards, SS = Stainless steel.

The ferric sulfate–sulfuric acid test is manifest in a broad range of standard procedures (such as ASTM A262 – Method B), and can be used to detect chromium depletion in austenitic, ferritic and duplex stainless steels (**Table 2**), as well as in nickel–chromium and nickel–chromium–molybdenum based alloys (**Table 3**). Generally, an increase in the sulfuric acid concentration from 10% to 60% significantly increases the corrosion rate in this test,[17] whereas the ferric sulfate functions as an inhibitor for stainless steel corrosion.[17] It is therefore recommended to regularly replenish ferric sulfate, to maintain passive conditions and avoid excessive general corrosion.[17]

Nitric acid IGC tests, such as ASTM A 262 – Method C (Huey),[10] are highly oxidizing and attack chromium depleted regions, regions with segregated impurity or alloying elements, as well as exposed end grains. After deformation, attack at slip bands has also been reported.[1] Experiments with pure chromium also indicated that the nitric acid test readily attacks chromium rich phases.[22] The accumulation of hexavalent chromium ions, typically brought about by leaching out

Table 3 ASTM and ISO practices for IGC testing of nickel-based alloys. (GB = Grain boundary/# Material range defined in standards)[28,35]

Standard	*Test solution*	*Materials*	*Nature of attack*	*Duration/ temperature*
ASTM G 28 (2008)	(Method A) Ferric sulfate, 50% sulfuric acid	Wrought + weldments; Ni-based, Cr-bearing alloys #	GB precipitates	24–120 h/ boiling
	(Method B) Sulfuric acid, hydrochloric acid, ferric chloride, cupric chloride	Wrought; Ni-base, Cr-bearing alloys #	GB precipitates	24 h/boiling
ISO 9400 (1996)	Method (A) Ferric sulfate, 50% sulfuric acid	Ni–based, Cr-based alloys #	GB precipitates	24–120 h/ boiling
ISO 9400 (1996)	Method (B) Copper, copper sulfate, sulfuric acid	Wrought + weldments; Ni-based, Cr-bearing alloys #	GB precipitates, excluding χ-, σ-phase, TiC + TiN	$\geq$ 24 h/boiling
	Method (C) Hydrochloric acid	High Mo-containing Ni and Ni–Cr base alloys #	GB Precipitates, Mo depletion	168 h/boiling
	Method (D) Nitric acid	Wrought; Ni-based, Cr-bearing alloys #	GB Precipitates, χ-, σ-phase, TiC + TiN	$\geq$ 48 h/boiling

from dissolved second phases, significantly increases the corrosion rate in this solution.[17,22]

The copper sulfate–16% sulfuric acid IGC test (Strauss) was initially recommended for the assessment of stainless steels with test durations of several hundreds of hours.[17] For accelerated testing, either metallic copper, or an increased concentration of sulfuric acid can be utilized.[17] Synergistic effects of the addition of higher concentrations of sulfuric acid and the contact to copper are utilized for IGC testing of cast stainless steels, such as ASTM A262 – Method F.[10] In sulfuric acid, cupric ions function as an inhibitor.[17]

The IGC resistance of stainless steels in mildly oxidizing solutions is a function of the chromium content, as shown in **Figure 5**.[44] An increase in the chromium content reduces the active–passive transition current density and enlarges the passive regions which makes passivation easier. In highly oxidizing solutions, such as boiling nitric acid, the presence of intermetallic phases and the segregation of solute atoms at grain boundaries also play a very important role. Annealed, nonsensitized stainless steels,[2,45] as well as pure nickel[23] and nickel alloys[3] have been found to undergo IGC in these highly oxidizing environments. **Figure 9** gives a schematic summary of an anodic polarization curve for austenitic stainless steel in sulfuric acid with superimposed potential regions of the ASTM A 262[10] tests. The oxalic acid test is not shown, and would be located at a potential higher than the Huey test.

IGC tests for nickel alloys are summarized in **Table 3**. These tests are restricted to a range of alloy groups comprising alloy 600, alloy 625, alloy 800, alloy 825, and alloy C-276. A list of susceptible materials is given in the corresponding standard practices. Evaluation of the IGC susceptibility is carried out by visual examination after bending (ISO 9400 – Method B),[28] or with weight loss measurements. Additional metallographic examinations are generally recommended. It is also recommended to compare differences in the measurements of weight loss to material reference conditions, since variations in composition and thermomechanical process history may exert a large effect on determined corrosion rates. This is also reflected in differences in the recommended exposure durations, which vary with alloy composition (ASTM G 28).[35]

2.06.3.2 Microstructure Screening

Metallographic etching procedures can be used for rapid first principle assessments of IGC susceptibilities. This requires knowledge of individual microstructure constituents attacked by the metallographic etching procedure and, more importantly, whether there exists a relationship between these constituents and the real IGC behavior. For instance, a typical test for aluminum–magnesium alloys (5XXX) is the orthophosphoric acid etch. This etch highlights intermetallic β-phase (Al_3Mg_2) precipitates, and has been used to gain information for exfoliation testing.[42]

2.06.3.3 Degree of Sensitization Testing

For characterizing the degree of sensitization, (i) the length of grain boundaries with chromium depletion, (ii) the width of the depleted zone, and (iii) the

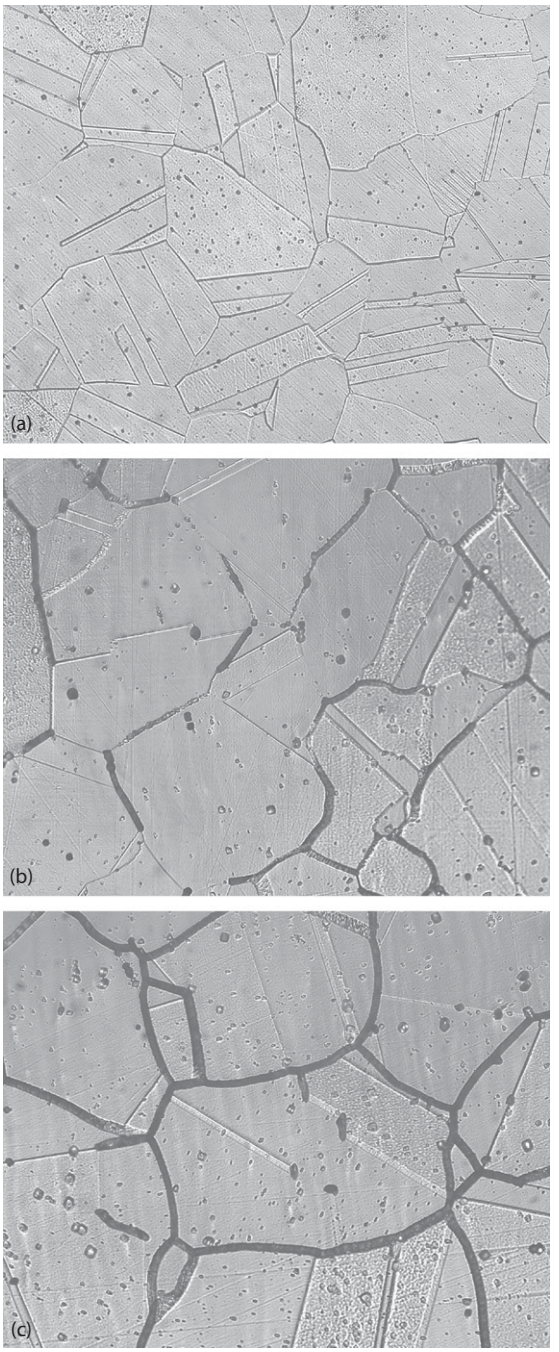


Figure 8 Optical micrographs of a Type 304 austenitic stainless steel after IGC screening in 10 wt% oxalic acid. Micrograph (a) shows a stepped, (b) a dual, and (c) a ditched microstructure.

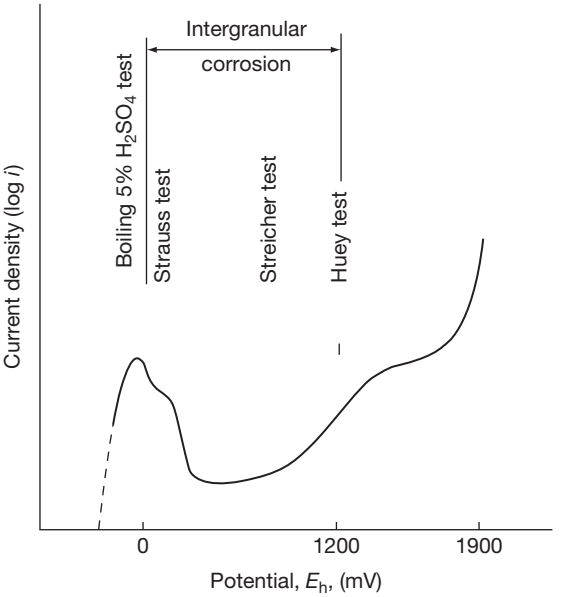


Figure 9 Schematic polarization curve of an austenitic stainless steel in deaerated sulfuric acid with superimposed potential regimes for the ASTM A 262 test methods. Potentials are given versus the normal hydrogen electrode (NHE, E_h). Reproduced from Osozawa, K.; Bohnenkamp, K.; Engell, H.-J. *Corros. Sci.* **1966**, *6*, 421–433.

minimum chromium concentration should be taken into consideration. For this purpose, a number of electrochemical test methods to quantify the degree of sensitization in stainless steels and nickel-base alloys are available.[20] Electrochemical tests are typically considered to be more sensitive for measuring the whole spectrum of sensitization, including lightly to severely sensitized materials. A saturation of the discriminating power, however, has been reported at high degrees of sensitization.[46] IGC tests based on weight loss or surface fracture measurements (i.e., **Tables 2** and **3**) are more sensitive to medium and high degrees of sensitization, but can generally not discriminate low to medium degrees of sensitization.

Three electrochemical test methods are typically used for the assessments of the degree of sensitization, as summarized in **Table 4**. These test comprise (i) the electrochemical reactivation (EPR)[37] test, (ii) the double-loop electrochemical potentiokinetic reactivation (DL-EPR)[47] test, and (iii) Cihal's DL-EPR[31] test. All tests are based on similar principles, but differ in the way the data are analyzed. **Figure 10(a)** shows EPR curves for a sensitized and non-sensitized stainless steel and **Figure 10(b)** shows a DL-EPR curve for a sensitized stainless steel with characteristic parameters. All three methods for quantifying the degree of sensitization are based on measuring the reactivation current density ($i_{reactivation}$) in a solution of sulfuric acid and thiocyanate. The thiocyanate functions as a

Table 4 Methods for detecting and quantifying the degree of sensitization in stainless steels and nickel alloys[31,37,47]

Method	*Test solution*	*Materials*	*Assessments*
ASTM G 108 (2004) -EPR Test-	0.5 M Sulfuric acid + 0.01 M KSCN	Type 304/Type 304L	$Q_{reactivation}$ and $i_{reactivation}$ ($Q_{reactivation}$/GB area)
ISO 12732 (2006) -Cihal's DL-EPR Method-	0.5–2.5 M Sulfuric acid + 0.001–0.05 M KSCN	Stainless steels Nickel alloys	Ratio = ($i_{reactivation}/i_{anodic}$) Ratio = ($i_{reactivation}$/GB area)/($i_{anodic}$/Specimen area) Ratio = ($Q_{reactivation}$/GB area)/(Q_{anodic}/Specimen area)
DL-EPR	0.5 M Sulfuric acid + 0.01 M KSCN	Type 304	Ratio = ($i_{reactivation}$)/(i_{anodic})

GB = Grain boundary, $i_{reactivation}$ = Maximum reactivation current density, $Q_{reactivation}$ = Reactivation charge density, Q_{anodic} = Anodic charge density.

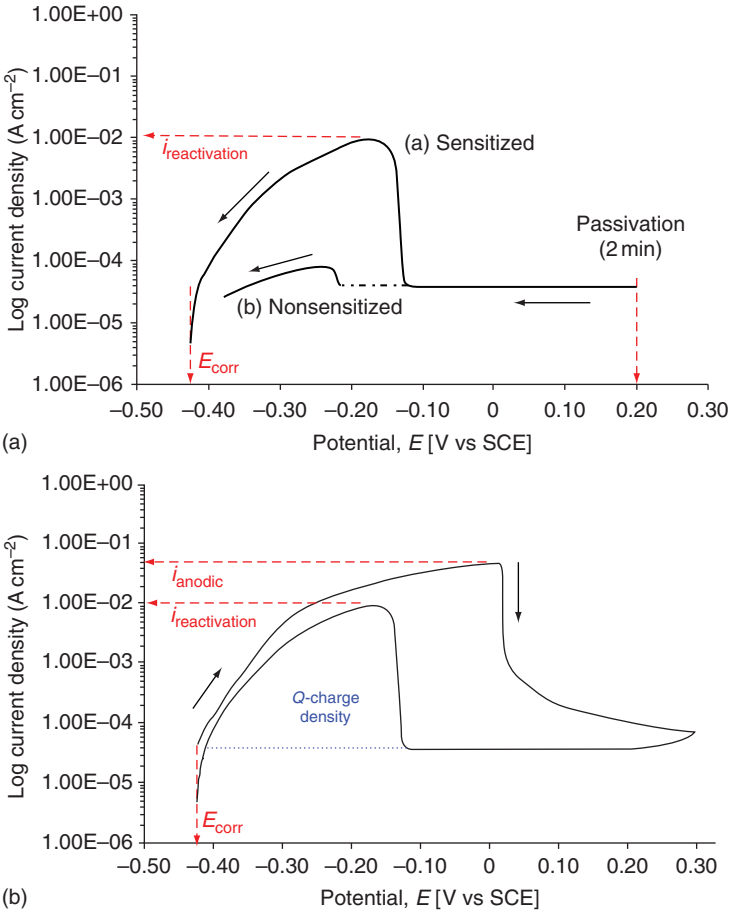


Figure 10 (a) Schematic diagram of an EPR test for sensitized (curve a) and non-sensitized (curve b) stainless steel. (b) Schematic diagram of a DL-EPR test with characteristic parameters. (E_{Corr} = corrosion potential, i_{anodic} = maximum anodic current density, $i_{reactivation}$ = reactivation current density, Q = charge density.)

depassivator, and increases the anodic dissolution during the anodic sweep, as well as the grain boundary attack during the reactivation sweep.[47] On sensitized material, the breakdown of the passive film on chromium depleted areas leads to the reactivation peak, but on nonsensitized material, the passive film remains virtually intact, and either no or only a small increase in reactivation current density is observed. The maximum height ($i_{reactivation}$ – current density) and area (Q–charge density) under the reactivation curve is a quantitative measure of the degree of sensitization (**Figure 10**).

The EPR test is a standard test method (ASTM G 108),[37] often referred to as the single-loop EPR test (SL-EPR). Evaluation of the degree of sensitization is carried out after an initial passivation treatment ($+0.2\,V_{SCE}$ for 2 min), by measuring the maximum reactivation current ($i_{reactivation}$) or charge density (Q). The reactivation charge density can also be normalized to the grain size, and gives a measure of attacked grain boundary area. After testing, the specimen has to be examined under an optical microscope, to determine whether the measured reactivation current is derived from chromium depleted regions at the grain boundaries, and not from intragranular pitting. A major drawback of this test is its sensitivity to material composition, thermomechanical process history, surface finish, and nonmetallic inclusions.[37]

An advanced version of the EPR tests is the double-loop technique (DL-EPR)[47] for Type 304 stainless steels. This test has frequently been used due to its ease of application. The degree of sensitization is determined with the maximum reactivation current density ($i_{reactivation}$) divided by the maximum anodic current density (i_{anodic}), without assessment of the grain boundary area. For an improved correlation between results derived from the DL-EPR and the EPR test, however, it is recommend to normalize the reactivation current density by the grain boundary area.[47] The DL-EPR test is not sensitive to the surface finish of the specimen, since the anodic sweep causes a general dissolution of the test surface.

A modified version of the DL-EPR test, based on Cihal's method, has been incorporated as a standard test method (ISO 12732).[31] This test is applicable for a broad range of stainless steels, including supergrades, as well as for certain nickel-base alloys. For this purpose, the sulfuric acid and thiocyanate concentrations can be tailored (**Table 4**). The assessment of Type 304 stainless steel is carried out with the same test parameters and solution concentrations as used for the EPR and DL-EPR tests. Cihal's test is assessed with the ratio of the reactivation current density, divided by the grain boundary area, and the maximum anodic current density, divided by specimen area. This accounts for a general dissolution of the surface during the anodic sweep, in relation to the dissolution of the sensitized grain boundary area during the reactivation sweep. The same assessment can also be carried out by using the charge densities instead of current densities. For comparing sensitized materials with similar grain sizes, the ratio between the maximum reactivation current density ($i_{reactivation}$) and the maximum anodic current density (i_{anodic}) can also be employed, similar to the DL-EPR test.[47] A comparison of typical values for the EPR, DL-EPR, and Cihal's test is given in corresponding literature sources.[31,47]

Acknowledgments

The author would like to thank Dr. T.J. Marrow, Dr. P.M. Wood and Prof. S.B. Lyon for their support and valuable feedback during the preparation of this manuscript.

References

1. Cihal, V. *Intergranular Corrosion of Steels and Alloys*; Elsevier, 1984.
2. Joshi, A.; Stein, D. F. *Corrosion* **1972**, *28*(9), 321–330.
3. Kasparova, O. V. *Protect. Met.* **2000**, *36*(6), 524–532.
4. Henthone, M. *STP 516: Localized Corrosion – Cause of Metal Failure*; ASTM, 1972; pp 66–119.
5. Buchheit, R. G.; Boger, R. K.; Carroll, M. C.; Leard, R. M.; Paglia, C.; Searls, J. L. *J. Miner. Met. Mater. Soc.* **2001**, *53*(7), 29–36.
6. Dayal, R. K.; Parvathavarthini, N.; Raj, B. *Int. Mater. Rev.* **2005**, *50*(3), 129–155.
7. Randle, V. *Acta Metall. Mater.* **1994**, *42*(6), 1769–1784.
8. Wolf, D. In *Materials Interfaces – Atomic-Level Structure and Properties*; Wolf, D., Yip, S., Eds.; Chapman & Hall, 1992; pp 1–57.
9. ASTM E 112: Standard Test Method for Determining Average Grain Size; 1996.
10. ASTM A 262: Standard Practices for Detecting Susceptibility to Intergranular Attack in Austenitic Stainless Steel; 2001.
11. Sutton, A. P.; Balluffi, R. W. *Interfaces in Crystalline Materials*; Oxford University Press, 1995.
12. Wolf, D.; Merkle, K. L. In *Materials Interfaces: Atomic-Level Structure and Properties*; Wolf, D., Yip, S., Eds.; Chapman & Hall, 1992; pp 87–150.
13. Bennet, B. W.; Pickering, H. W. *Metall. Trans. A* **1985**, *18*, 1117–1124.
14. Lifka, B. W.; Sprowls, D. O. *ASTM STP 516: Localized Corrosion – Cause of Metal Failure*; ASTM, 1972; pp 120–144.
15. Davis, J. R. *Corrosion of Aluminum and Aluminum Alloys*; ASM International, 1999.
16. Marshall, P. *Austenitic Stainless Steels – Microstructure and Mechanical Properties*; Elsevier, 1984.
17. Streicher, M. A. In *ASTM STP 656: Intergranular Corrosion of Stainless Alloys*; Steigerwald, R. F., Ed.; ASTM, 1978; pp 3–84.
18. Frankenthal, R. P.; Pickering, H. W. *J. Electrochem. Soc.* **1973**, *120*(1), 23–26.
19. Lula, R. A.; Davis, J. A. *ASTM STP 656: Intergranular Corrosion of Stainless Alloys*; ASTM, 1978; pp 233–247.
20. Cihal, V.; Stefec, R. *Electrochim. Acta* **2001**, *46*, 3867–3877.
21. Devine, T. M. *Acta Metall.* **1988**, *36*(6), 1491–1501.
22. Devine, T. M.; Drummond, B. J. *Corrosion* **1981**, *37*(2), 104–115.
23. Palumbo, G.; Aust, K. T. *Acta Metall. Mater.* **1990**, *38*(11), 2343–2352.

24. Uhlig, H. H.; Duquette, D. J. *Corros. Sci.* **1969**, *9*, 557–560.
25. Mieluch, J.; Smialowski, M. *Corros. Sci.* **1964**, *4*, 237–243.
26. Zhang, X. G. *Corrosion and Electrochemistry of Zinc*; Springer, 1996.
27. Nielsen, R. In *Ullmann's Encyclopedia of Industrial Chemistry*, 5th ed.; Elvers, B., Hawkins, S., Eds.; VCH, 1996; Vol. 28A.
28. ISO 9400: Nickel-Based Alloys-Resistance to Intergranular Corrosion 1996.
29. ISO 3651-1: Determination of Resistance to Intergranular Corrosion of Stainless Steels: Part 1: Austenitic and Ferritic–Austenitic (Duplex) Stainless Steels – Corrosion Test in Nitric Acid Medium by Measurement of Loss in Mass (Huey Test) 1998.
30. ISO 3651-2: Determination of Resistance to Intergranular Corrosion of Stainless Steels: Part 2: Ferritic, Austenitic, and Ferritic–Austenitic (Duplex) Stainless Steels – Corrosion Test in Media Containing Sulfuric Acid 1998.
31. ISO 12732: Corrosion of Metals and Alloys – Electrochemical Potentiokinetic Reactivation Measurement using the Double Loop Method (based on Cihal's Method) 2006.
32. ISO 15329: Corrosion of Metals and Alloys – Anodic Test for Evaluation of Intergranular Corrosion Susceptibility of Heat-Treatable Aluminum Alloys 2008.
33. ISO 11846: Corrosion of Metals and Alloys – Determination of Resistance to Intergranular Corrosion of Solution Heat-Treatable Aluminum Alloys 2008.
34. ASTM G 34: Standard Test Method for Exfoliation Corrosion Susceptibility in 2XXX and 7XXX Series Aluminum Alloys (Exco Test) 2001.
35. ASTM G 28: Standard Test Methods of Detecting Susceptibility to Intergranular Corrosion in Wrought, Nickel Rich, Chromium-Bearing Alloys 2002.
36. ASTM G 110: Standard Practice for Evaluating Intergranular Corrosion Resistance of Heat Treatable Aluminum Alloys by Immersion in Sodium Chloride + Hydrogen Peroxide Solution 2003.
37. ASTM G 108: Standard Test Method for Electrochemical Reactivation (EPR) for Detecting Sensitization of AISI 304 and 304L Stainless Steel 2004.
38. ASTM G 67: Standard Test Method for Determining the Susceptibility to Intergranular Corrosion of 5XXX Series Aluminum Alloys by Mass Loss After Exposure to Nitric Acid (NAMLT Test) 2004.
39. ASTM A 763: Standard Practice for Detecting Susceptibility to Intergranular Attack in Ferritic Stainless Steel 2004.
40. ASTM G 31: Standard Practice for Laboratory Immersion Corrosion Testing of Metals 2004.
41. ASTM G 66: Standard Test Method for Visual Examination of Exfoliation Corrosion Susceptibility of 5XXX Series Aluminum Alloys (ASSET Test) 2005.
42. Sprowls, D. O.; Walsh, J. D.; Shumaker, M. B. *ASTM STP 516: Localized Corrosion – Cause of Metal Failure*; ASTM, 1972; pp 38–65.
43. Craig, H. L., Jr. *ASTM STP 516: Localized Corrosion-Cause of Metal Failure*; ASTM, 1972; pp 17–37.
44. Osozawa, K.; Engell, H.-J. *Corros. Sci.* **1966**, *6*, 389–393.
45. Osozawa, K.; Bohnenkamp, K.; Engell, H.-J. *Corros. Sci.* **1966**, *6*, 421–433.
46. Clarke, W. L.; Cowan, R. L.; Walker, W. L. In *ASTM STP 656: Intergranular Corrosion of Stainless Alloys*; Steigerwald, R. F., Ed.; ASTM, 1978; pp 99–132.
47. Majidi, A. P.; Streicher, M. A. *Corrosion* **1984**, *40*(11), 584–593.
48. Trillo, E. A.; Beltran, R.; Maldonado, J. G.; Romero, R. J.; Murr, L. E.; Fisher, W. W.; Advani, A. H. *Mater. Char.* **1995**, *35*, 99–112.
49. Bruemmer, S. M.; Was, G. S. *J. Nucl. Mater.* **1994**, *216*, 348–363.

2.07 Galvanic Corrosion

H. P. Hack
176 Via Dante, Arnold, MD 21012, USA

Abbreviations

C.D. Current density
FS Full scale
IR Potential drop due to a current across a resistance
LCS Low carbon steel
MRE Meals ready to eat
SS Stainless steel

Symbols

A Cross-sectional area
B_a Anodic Tafel slope
B_c Cathodic Tafel slope
dE/di Local slope of the polarization curve
E_A Potential of material A
E_{AB} Coupled potential of materials A and B
E_B Potential of material B
E Potential
E_{corr} Corrosion potential
E_{couple} Coupled potential
$E(H^+/H_2)$ Potential of the reaction between hydrogen ions and hydrogen gas
$E(M/M^+)$ Potential of the reaction between metal M with its ions
$E(N/N^+)$ Potential of the reaction between metal N with its ions
E_{pp} Primary passivation potential
i Current density

i_0 **(H_2/M)** Exchange current density for the reaction of hydrogen gas on metal M
i_0 **(H_2/N)** Exchange current density for the reaction of hydrogen gas on metal N
i_0 **(M)** Exchange current density for the reaction of metal M with its ions
i_0 **(N)** Exchange current density for the reaction of metal N with its ions
i_{cc} Critical current density
i_{corr} **(couple)** Current exchanged between all anode and cathode reactions
i_{corr}**(M)** Corrosion rate per unit area of metal M
i_{corr} **(N)** Corrosion rate per unit area of metal N
i_{corr} Corrosion current density
I_{corr} Total corrosion current
$I_{measured}$ Measured galvanic current
i_p Passive current density
I Current
I_{AB} Coupled current between materials A and B
I_{anodic} Total anodic current
$I_{applied}$ Total applied current
$I_{cathodic}$ Total cathodic current
I_{total} Total current
L Length
n Number of electrons exchanged in a reaction
R Resistance
W Wagner number
ρ Resistivity

2.07.1 Introduction

Galvanic corrosion, also known as bimetallic corrosion, is a common mode of corrosion failure that is, for the most part, entirely preventable by proper corrosion design. It is the aim of this chapter to provide the reader with the knowledge and data to aid in recognizing this form of corrosion when it occurs and making the right design decisions to prevent it.

The ASTM International Committee G01 on Corrosion of Metals defines galvanic corrosion as "accelerated corrosion of a metal because of an electrical contact with a more noble metal or nonmetallic conductor in a corrosive environment."[1] NACE International defines galvanic corrosion as "corrosion associated with the current resulting from an electrical coupling of dissimilar electrodes in an electrolyte."[2] International Organization for Standards (ISO) defines galvanic corrosion as "Corrosion due to the action of a corrosion cell [a short-circuited combination of different electrodes connected in series with an ionic conductor]."[3] Other organizations that have definitions for galvanic corrosion include The Electrochemical Society, the American Water Works Association, The American Chemical Society, and others. What these definitions all have in common is that galvanic corrosion involves the electrical interaction of at least two different metals or nonmetallic conductors in an environment that accelerates the corrosion of at least one of them. The common factors are dissimilar metals, electrical contact, and a conductive electrolyte in contact with them. If any of these factors is absent, galvanic corrosion cannot occur.

Galvanic corrosion accelerates the normal corrosion of a metal in an electrolyte. Even without galvanic corrosion, metals may suffer from uniform corrosion, crevice corrosion, pitting, or other forms of corrosion. Galvanic corrosion has an accelerating effect on these other forms of corrosion and, in some cases, causes a form of corrosion to occur that would otherwise not be seen.

2.07.2 Electrochemical Theory

When a metal or nonmetallic conductor is immersed in a conductive solution, or an electrolyte, a series of reactions takes place that can cause corrosion. The corrosion reaction itself involves a change in the charge of the metal atoms from zero in the metal to a positive value for metal ions in the solution. Because this is a positive change in charge the reaction is called an anodic reaction and can be written as:

$$M \rightarrow M^{n+} + ne^-$$

where M = some metal atom or ion, and n is a number typically between 1 and 3 depending on the metal and electrolyte.

Since this reaction generates free electrons and since charge neutrality must be maintained, these electrons must be used up in one or more other reactions where charge is reduced in order for corrosion to proceed. These are called cathodic reactions, and the most common ones found in aqueous environments are:

$$2H^+ + 2e^- \rightarrow H_2(g) \text{ in acids}$$
$$2H_2O + 4e^- + O_2(\text{dissolved}) \rightarrow 4OH^- \text{in neutral or alkaline electrolytes}$$

The progression of the anodic and cathodic reactions, the flow of electrons, and the ionic motion in the electrolyte needed to complete the current flow are

shown in **Figure 1** for a piece of iron corroding in neutral water.

Corrosion cannot proceed unless all of the electrons generated by the anodic reaction are used up by a cathodic reaction. The balance between these two reactions leads to each metal, alloy, or nonmetallic conductor equilibrating at an electrochemical potential that is unique to that material and electrolyte, called the corrosion potential. For a given electrolyte, the corrosion potential of a number of materials can be listed together. Such a listing, where the materials are arranged according to ascending or descending corrosion potential, is called a galvanic series. If two materials with different corrosion potentials are immersed in that electrolyte and electrically coupled, there is a driving force for a current, called the galvanic current, to flow between them. It is this flow of current that increases corrosion of the material with the more negative potential, called the anode, while suppressing the corrosion of the more positive material, called the cathode.

The most common galvanic series is the one shown in **Figure 2** for metals in seawater. Metals with more negative corrosion potentials are said to be anodic compared to metals with more positive, or cathodic, potentials in a galvanic series.

In this case, the corrosion potentials are given as ranges, since seawater composition can vary and since corrosion potentials can be influenced by the presence of corrosion products on the surface of the corroding metal. Some metals show two bands of corrosion potentials, one in white and the other in black. These are metals that can exist in two different states in seawater, a passive state during which they are not corroding and an active state when they are corroding. Each state is associated with a different range of corrosion potentials. This galvanic series is only applicable in seawater at a specific temperature range and flow rate, and should not be applied to other environments.

Galvanic series are useful to predict which metal in a pair of electrically connected materials in an electrolyte, called a galvanic couple, will tend to have its corrosion rate accelerated by the electrical contact. The fact that there is a driving force for current to flow between the dissimilar materials and

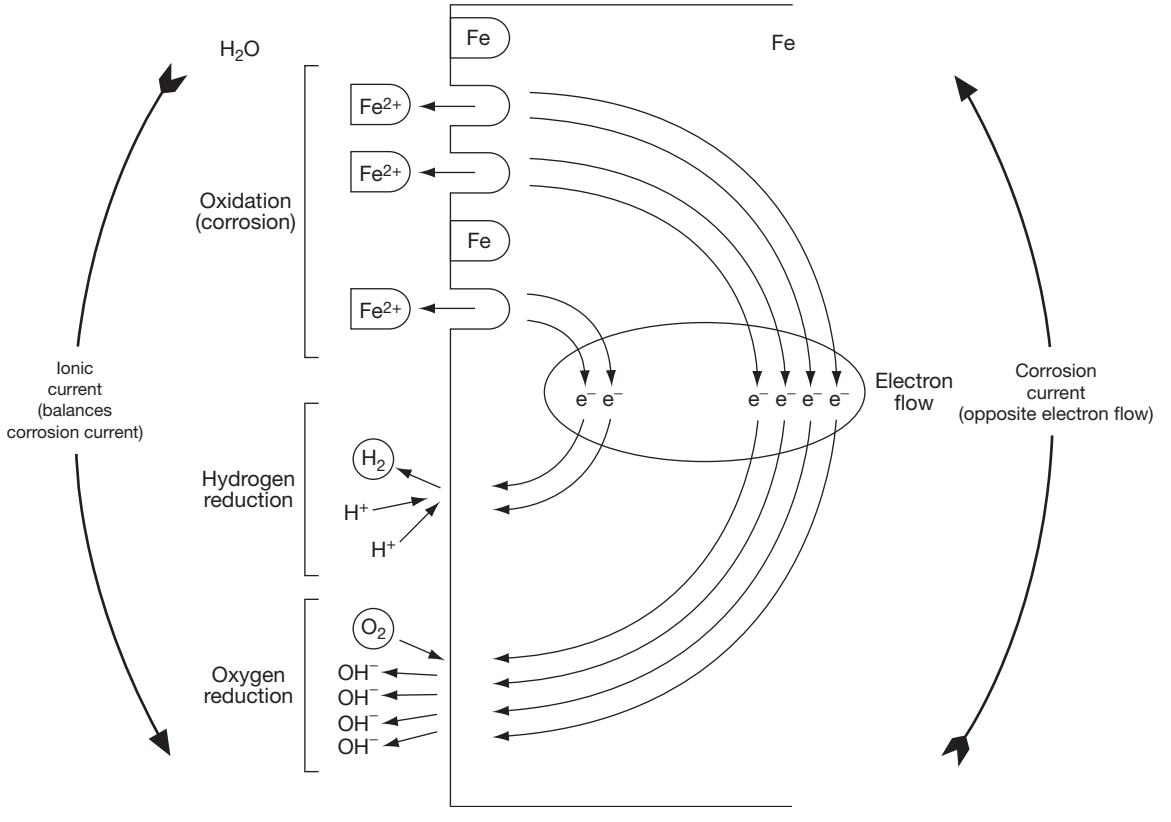


Figure 1 Iron corroding in water. Reproduced from Hack, H. P. *Corrosion Testing Made Easy*; NACE International: Houston, TX, 1993; Volume 2.

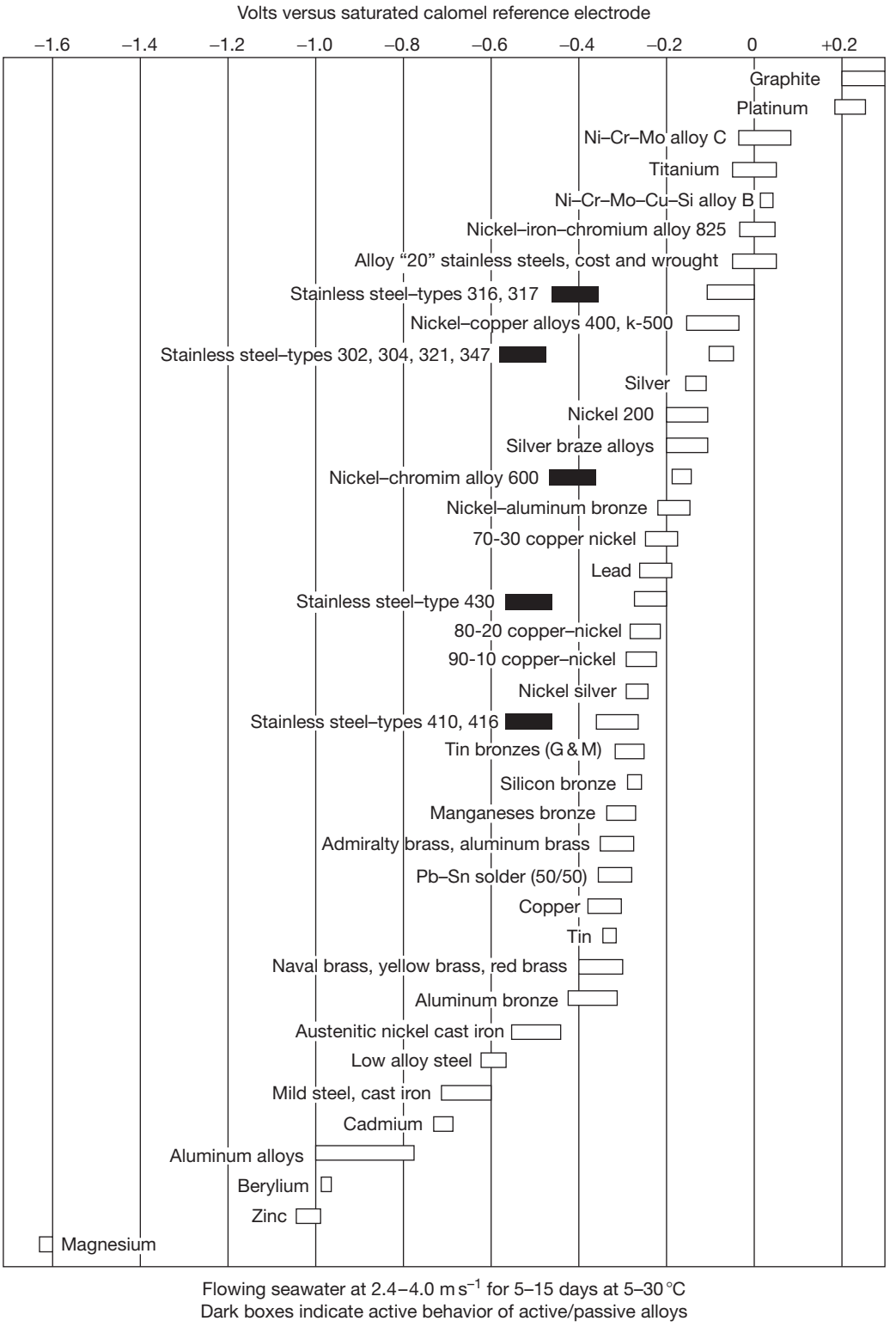


Figure 2 Galvanic series for metals in seawater with moderate flow. Reproduced from Standard G82: Guide for Development and Use of a Galvanic Series for Predicting Galvanic Corrosion Performance, *ASTM Book of Standards*; ASTM International: West Conshohocken, PA, 2006.

for the corrosion of the anode in the couple to increase does not mean, however, that the magnitude of this current and resulting corrosion will be significant.

When two dissimilar materials with dissimilar corrosion potentials are coupled in an electrolyte, they will each be forced to a common potential somewhere between the two individual corrosion potentials. When coupled, materials with similar corrosion potentials have no driving force for causing galvanic corrosion. However, materials far apart in potential will not necessarily result in higher galvanic corrosion rates than materials close together in potential. The magnitude of current flow and the increase in corrosion rate due to this coupling are determined by the tendency of each of the materials in the couples to deliver current when its potential is forced to change, called polarization. If a material delivers little current for a large change in potential, it is said to have a large polarization resistance, while if it can deliver a large current for a small enforced potential change, it is said to have a small polarization resistance. This is illustrated by plotting the current produced as a function of applied potential, called a polarization curve. Examples of polarization curves for metals with high and low polarization resistances are shown in **Figure 3**. In this figure, when the cathode has low polarization resistance and the anode has high resistance, the couple is said to be under cathodic control. When the reverse is the case, the couple is said to be under anodic control, and when the polarization resistances of the anode and cathode reactions are similar the couple is said to be under mixed control.

When two materials are coupled in an electrolyte, the galvanic current flowing between them is determined by a superposition of the cathodic polarization curve of the anodic metal with the anodic polarization curve of the cathodic material, shown in **Figure 4**. In this figure, $I_{measured}$ is the galvanic current that flows through the metallic path between the two metals, in this case copper and iron.

The magnitude of the galvanic current flowing in a galvanic couple gives a qualitative feel for the amount of increased corrosion of the anode, but it is not a good quantitative measure. To understand why this is so, it is necessary to look at the magnitude of the anodic and cathodic reactions that are occurring on each metal in the couple, both before the metals are electrically connected and afterwards. To understand corrosion rates in a galvanic couple, it is necessary to look at a plot of individual reaction

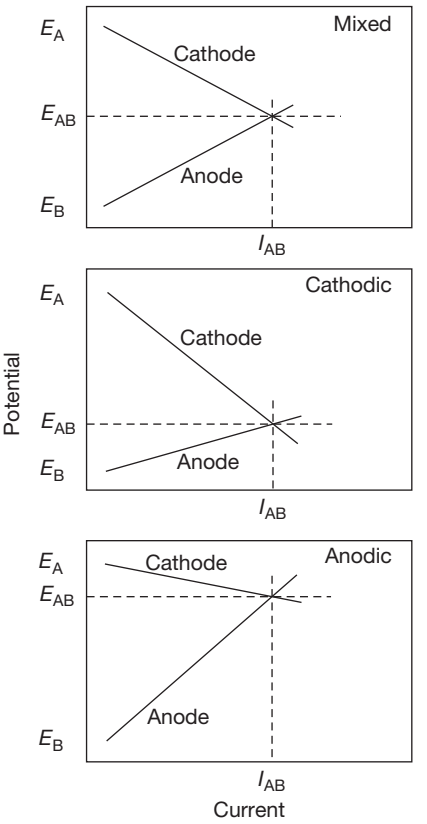


Figure 3 Effect of polarization resistance on anode and cathode reactions. Reproduced from Standard G82: Guide for Development and Use of a Galvanic Series for Predicting Galvanic Corrosion Performance, *ASTM Book of Standards*; ASTM International: West Conshohocken, PA, 2006.

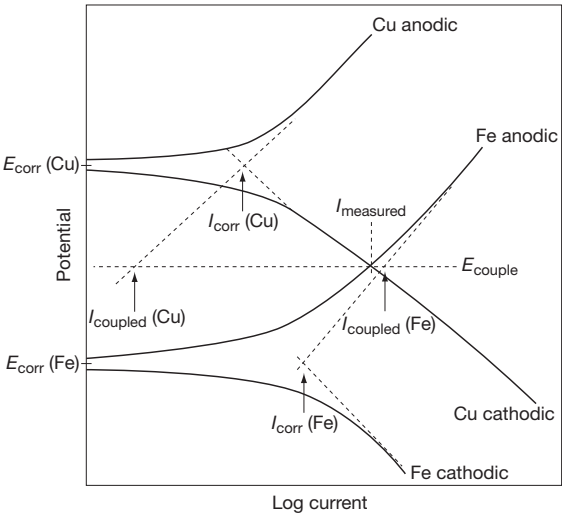


Figure 4 Superposition of polarization curves to determine galvanic current. Reproduced from Hack, H. P. *Corrosion Testing Made Easy*; NACE International: Houston, TX, 1993; Volume 2.

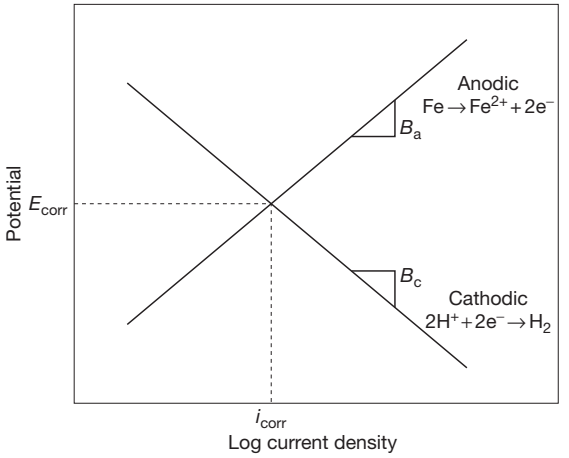


Figure 5 Evans diagram for a corroding metal. Reproduced from Hack, H. P. *Corrosion Testing Made Easy*; NACE International: Houston, TX, 1993; Volume 2.

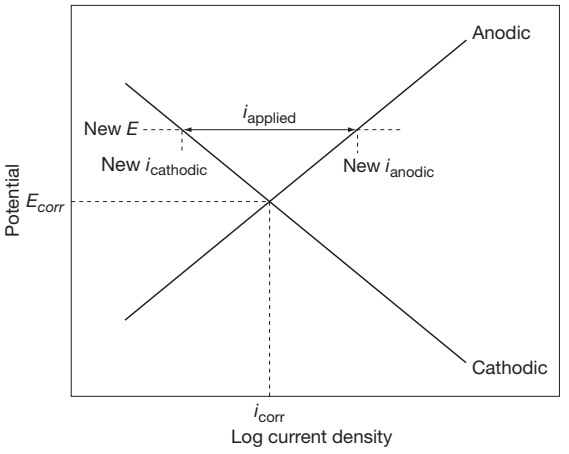


Figure 6 Relationship between a forced potential shift and applied current. Reproduced from Hack, H. P. *Corrosion Testing Made Easy*; NACE International: Houston, TX, 1993; Volume 2.

rates as a function of potential, called an Evans diagram, for each of the materials in the couple before they are coupled (**Figure 5**). This figure shows the cathodic reaction to be hydrogen production, but any cathodic reaction, such as oxygen reduction, would behave similarly. The corrosion rate for iron in this diagram is i_{corr} and the corrosion potential is E_{corr}.

The rate of the anodic reaction increases as the potential becomes more positive, while the rate of the cathodic reaction increases as the potential becomes more negative. The net current flowing to or from the metal is the difference between these two currents, or zero at the point where the two curves cross. Applying a current will force a potential shift to occur, as shown in the Evans diagram in **Figure 6**.

When two materials are coupled, the Evans diagrams for them are shown in **Figure 7**.

For two materials in electrical contact, the sum of all anodic reactions must equal the sum of all cathodic reactions. Where the anodic sum crosses the cathodic sum determines the potential of the galvanic couple, called the coupled potential. Once this potential is known, the rates of all reactions on both materials are read from the Evans diagram for each reaction at that coupled potential. This coupled potential can also be obtained from polarization curves, as shown in **Figure 4**. **Figure 8** shows the polarization curves on an Evans diagram so that the relationship can be seen. The galvanic current from the polarization curves is the actual measurable current that flows between the two materials, but is not necessarily directly related to the corrosion rate of the anode.

In practice, it is difficult to determine the exact position of all reaction curves on all materials, particularly in complex environments like seawater. For this reason, it is more common to determine corrosion rates for galvanic corrosion by first determining the coupled potential using polarization curves, and then testing the anode material by holding it at that potential and measuring the actual corrosion rate experienced.

All the Evans diagrams and polarization curves for galvanic couples shown above have used either current density or total current on the horizontal axis, assuming that the exposed surface areas of the anode and cathode metals are the same. The effect of different exposed surface areas of the anode and cathode materials can be examined by realizing that the total current is just the current density multiplied by the exposed surface area. Larger exposed areas will shift the polarization curves to the right for that material. The effect of this on coupled potential and galvanic current is shown in **Figure 9**.

This leads to the conclusion that to minimize galvanic corrosion small cathode–anode area ratios are required, or conversely the larger this ratio, the larger is the galvanic corrosion. This is sometimes called the catchment area principle. For this reason, to control galvanic corrosion it is best to paint the noncorroding cathode, not the corroding anode, contrary to what might be inferred at first glance.

Polarization does not always take the form of straight lines on Evans diagrams. Real polarization

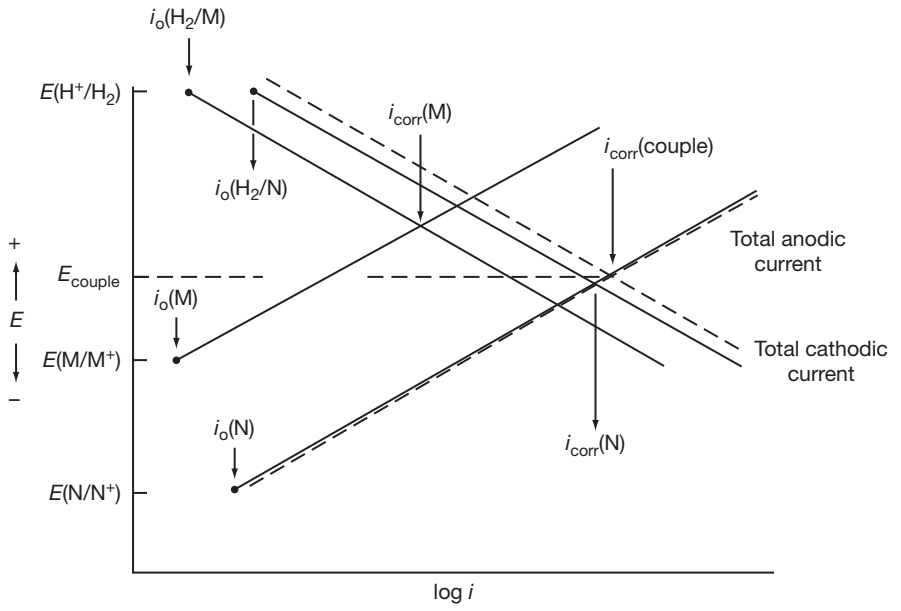


Figure 7 Evans diagram for two materials in electrical contact. Reproduced from Standard G82: Guide for Development and Use of a Galvanic Series for Predicting Galvanic Corrosion Performance, *ASTM Book of Standards*; ASTM International: West Conshohocken, PA, 2006.

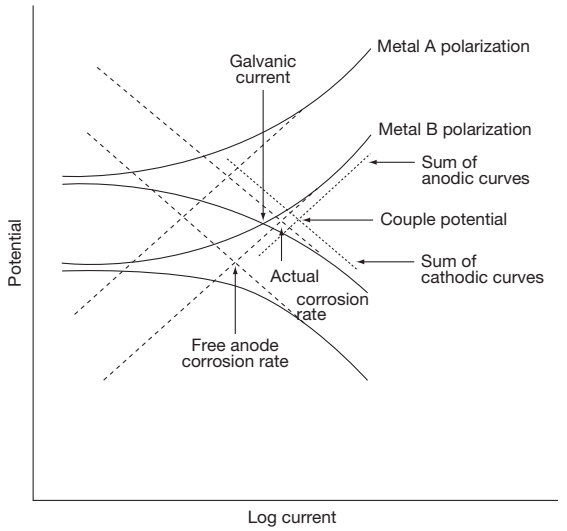


Figure 8 Polarization curves on an Evans diagram. Reproduced from Hack, H. P. *Corrosion Testing Made Easy*; NACE International: Houston, TX, 1993; Volume 2.

curves are affected by the formation of corrosion products and the ability of reactants and reaction products to move to and from the surface where the reactions take place. This leads to various polarization curve shapes, including some that curve back on themselves, as shown in **Figure 10**.

Galvanic reactions with materials having these types of curves can lead to some unusual corrosion consequences, including having several different stable corrosion potentials for a given galvanic couple, as seen in **Figure 11**.

If more than two materials with different corrosion potentials are electrically connected, they can be analyzed using the same principles as when two materials are analyzed, that is, using Evans diagrams. The anode and cathode reaction curves for all three materials are normalized to the proper surface areas and then put on the same diagram, together with the total anodic and total cathodic curves. Then the intersection point of latter curves defines the coupled potential. From this coupled potential the corrosion rate of all three materials can be determined.

Galvanic corrosion can also be affected by electrical or ionic resistances between the two coupled materials. Electrical resistance can come about if one of the materials is a poor electrical conductor or if the contact between them is not made properly. Ionic resistance occurs because electrolytes are not perfect conductors, so that high electrolyte resistivity coupled with a large distance between the anode and cathode materials can lead to significant resistance between the anode and cathode. Electrical and ionic resistance are treated the same when determining their effect on a galvanic couple. The total circuit resistance is the product of the galvanic current and the various resistances, frequently called *IR* drop. The effect of *IR* drop is seen in **Figure 12**.

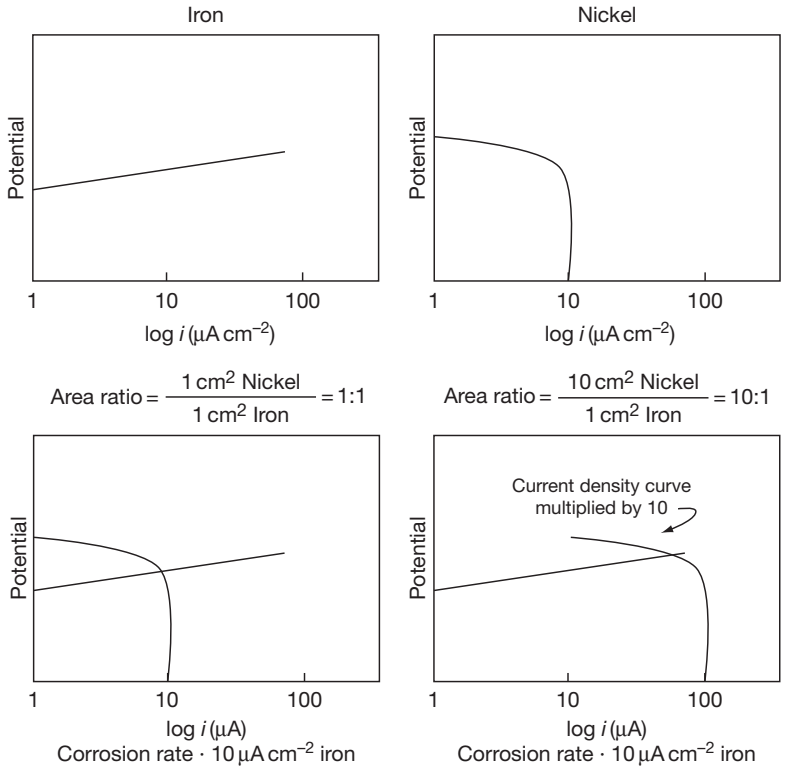


Figure 9 Effect of area ratio on superimposed polarization curves. Reproduced from Hack, H. P. *Corrosion Testing Made Easy*; NACE International: Houston, TX, 1993; Volume 2.

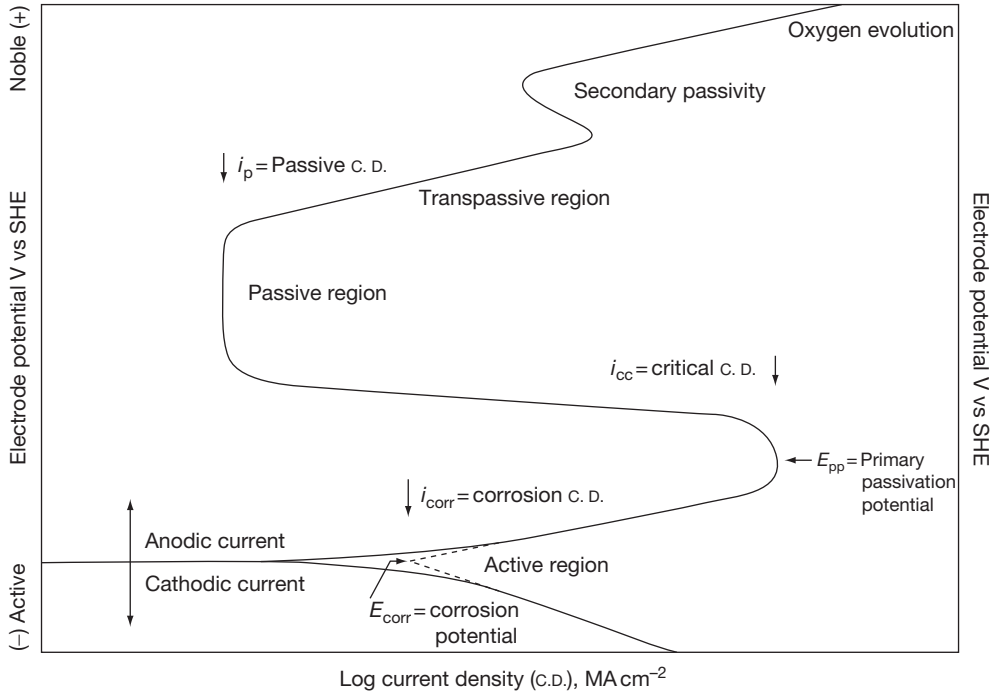


Figure 10 Polarization curve for passive-film-forming metal such as 430 stainless steel in sulfuric acid. Standard G3: Practice for Conventions Applicable to Electrochemical Measurements in Corrosion Testing, *ASTM Book of Standards*; ASTM International: West Conshohocken, PA, 2006.

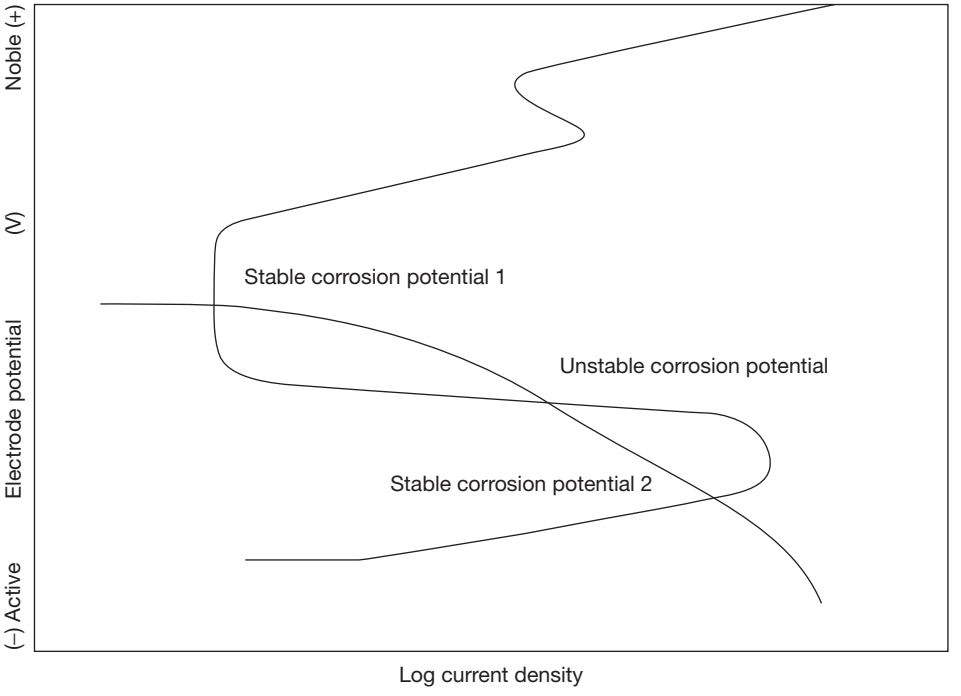


Figure 11 Triple crossing of anodic and cathodic polarization curves leading to two stable corrosion potentials.

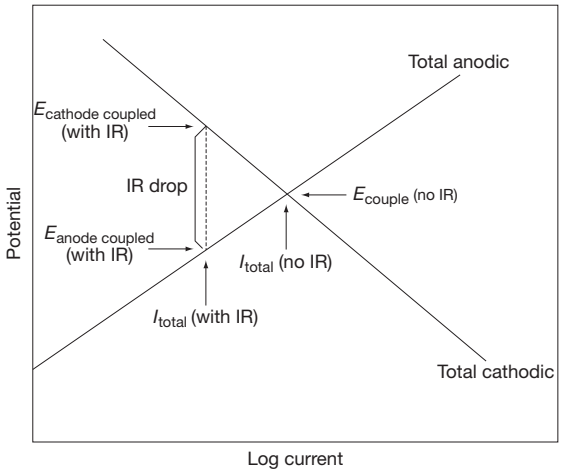


Figure 12 *IR* drop leading to anode and cathode being at different potentials. Reproduced from Hack, H. P. *Corrosion Testing Made Easy*; NACE International: Houston, TX, 1993; Volume 2.

IR drop leads to the anodic and cathodic materials being at different potentials, with the associated lowering of corrosion rate of the anode from the value if there were no resistance. Low galvanic currents lead to low *IR* drop, so this effect is more pronounced on large structures. Also, larger distances between anode and cathode lead to larger *IR* drops, and therefore lower galvanic corrosion rates. For this reason, in galvanic couples that are large or where electrolyte resistance is high, galvanic corrosion rates are highest on the anode closest to the cathode and drop off as distance between anode and cathode increases. This falloff in corrosion rate with distance is one good method for identifying galvanic corrosion, although if conductivity is high or distances are small this dropoff effect may not be observed. This can be seen in Figure 13.

2.07.3 Examples of Galvanic Corrosion

One typical example of galvanic corrosion is the use of dissimilar metal fasteners. When the fastener is anodic compared to the structure that it is fastening, the anode-to-cathode ratio is large and the fasteners will corrode quickly. When the fastener is cathodic compared to the structure the anode-to-cathode ratio is small and the corrosion acceleration of the structure around the fasteners is present but minimal as long as the electrolyte conductivity is high, as seen in Figure 14.

Galvanic corrosion also commonly occurs in the attachment of dissimilar metal piping. Here,

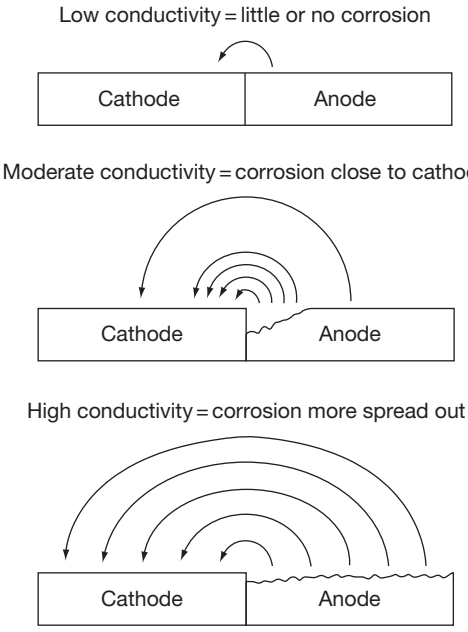


Figure 13 Concentration of corrosion near cathode as a function of conductivity. Reproduced from Hack, H. P. *Corrosion Testing Made Easy*; NACE International: Houston, TX, 1993; Volume 2.

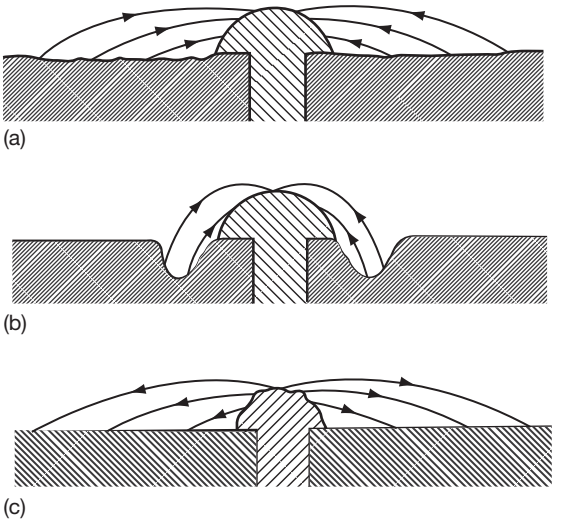


Figure 14 Corrosion of fasteners and base plate in relation to the anode.[6] (a) Base plate is the anode in high-conductivity water; (b) base plate is the anode in low-conductivity water; (c) fastener is the anode.

corrosion of the more anodic pipe will be concentrated at the junction with the cathode.

Structures made from dissimilar metals are sometimes joined in soils. This can be inadvertent, such as when a section of pipe is replaced with a new piece of pipe, not realizing that the new pipe is frequently cathodic to the old pipe. The new pipe will therefore cause the old pipe to corrode faster.

Sometimes in a mixed-metal structure one part that corrodes quickly will be 'upgraded' to a more corrosion-resistant material. Steel may be replaced with stainless steel, stainless steel with nickel alloys, etc. The more corrosion-resistant material is frequently cathodic to the other materials in the structure, causing galvanic corrosion of the structure close to the parts that were replaced. This makes it look like the corrosion problem is moving around the structure, while it is in fact due to lack of attention to galvanic corrosion in design of the replacement material.

Tube–tubesheet heat exchangers frequently use different materials for the tubes than for the tubesheet, support structure, or shell. This can lead to galvanic corrosion of any of these parts, depending on the particular material combinations chosen.

2.07.4 Benefits of Galvanic Corrosion

Galvanic corrosion can be beneficial. The potential and current between dissimilar metals is used in a sensor for time of wetness in atmospheric corrosion testing.[7] When rainwater or dew is present on the sensor, a galvanic couple exists, whereas when the moisture is not present the galvanic couple is not active. Similar sensors can be buried inside lap joints or other hard-to-access areas and can be used to determine whether sufficient moisture is present for corrosion to occur where it cannot be seen.

Special materials that can generate high galvanic corrosion currents have been prepared in powder form. When wetted, these powders result in galvanic corrosion rates that are so high that significant heat is generated. This is how meals for the U.S. Armed Forces, called MREs (meals ready to eat), are heated in the field: just add water to get a hot meal.

There are other benefits to high galvanic corrosion rates. Special corrodible pins, consisting of a hollow anode with a cathode on the inside, have been produced in a way that the failure time for the pin is well characterized. Such pins are used in seawater to keep crab and lobster pot buoys submerged to prevent poaching for a given time period, after which the buoys are released by corrosion of the pin and the pots can be retrieved by their owners. A similar system is used to cause military sonobuoys,

used to detect enemy submarines, to sink after their useful lifetime is exhausted. A corrodible plug is used for this purpose to allow flooding of the buoy after a known period of time.

Galvanic corrosion is the basis for the operation of most batteries. The galvanic current produced by the couple is put to good use in our toys, cell phones, and other electronic hardware.

Finally, the cathode in a galvanic couple frequently has its corrosion rate decreased as a result of the couple. This is the basis for sacrificial anode cathodic protection used widely in seawater and soils to protect steel structures. By coupling the steel to a piece of anodic material such as zinc, aluminum, or magnesium, the accelerated corrosion of the anode causes corrosion of the steel to be reduced. The anode sacrifices itself to prevent corrosion of the steel and is periodically replaced as it is consumed.

2.07.5 Recognizing Galvanic Corrosion

There is no absolutely certain method of determining that corrosion seen on a structure is actually galvanic corrosion, since the action of galvanic corrosion is usually to accelerate the type of corrosion normally experienced by the anode material. There are clues, however, that can help to make the determination. Galvanic corrosion always involves two materials with different corrosion potentials electrically connected in a common electrolyte. Therefore, if two materials with different corrosion potentials are present in the system and they are electrically connected in a common electrolyte, this is a strong clue that the corrosion that is observed is galvanically accelerated. This is particularly true if the anodic material has a small surface area compared to the cathodic material and the two different materials are close together. If both materials are corroding, then this corrosion is unlikely to be galvanically accelerated.

There is one good indicator for galvanic corrosion, however. If the anodic material corrodes the fastest near the cathodic material, and if the amount of corrosion lessens as distance increases from the cathode, galvanic corrosion is the likely cause. For example, if a flange connects two dissimilar metal pipes and the metal with the more negative corrosion potential is corroding a lot near the flange and less far away from the flange, galvanic corrosion is indicated. Thickness loss of the anodic metal near dissimilar metal lap joints or fasteners is likely to be due to galvanic corrosion. Some examples of galvanic corrosion are shown in Figures 15–18.

2.07.6 Predicting Galvanic Corrosion

When designing a structure it is important to take into account galvanic corrosion either by preventing it or by predicting where it will occur and how much metal will be lost over time so that structure lifetime can be determined and maintenance activities can be planned accordingly. The methods used to predict galvanic corrosion include experience with similar materials and geometries in similar environments, use of exposure data in similar environments, use of laboratory test data, computer modeling, physical scale

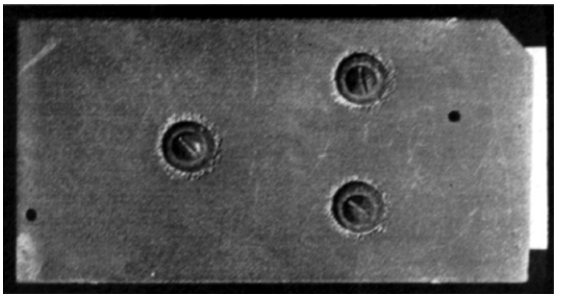


Figure 15 Corrosion of an anodic plate due to cathodic screws. Reproduced from Copson, H. R. In *Corrosion Resistance of Metals and Alloys*, 2nd ed.; LaQue, F. L., Copson, H. R., Eds.; Reinhold Publishing: New York, NY, 1963; p 36, Chapter 1.

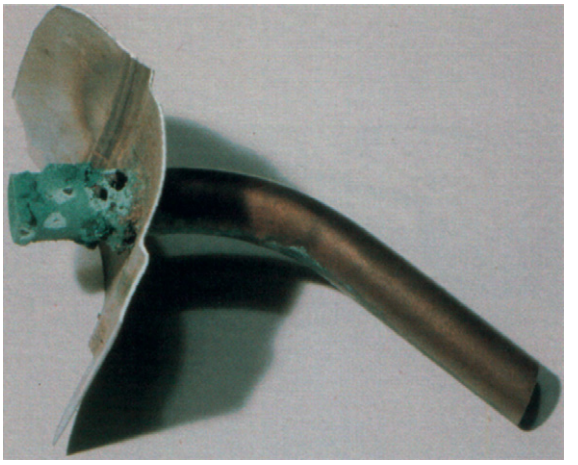


Figure 16 Corrosion of copper pipe in a stainless-steel hot water cylinder. Reproduced from Francis, R. *Galvanic Corrosion: A Practical Guide for Engineers*; NACE: Houston, TX, 2001.

Figure 17 Corrosion of a carbon steel weld in piping carrying partially deaerated seawater. Reproduced from Francis, R. *Galvanic Corrosion: A Practical Guide for Engineers*; NACE: Houston, TX, 2001.

modeling, and numerical calculations such as Wagner number. Each of these will be discussed below.

2.07.6.1 Experience

The best way to predict galvanic corrosion is to rely on experience with similar materials in similar geometries under similar environmental conditions to those of the structure being predicted. If you are designing a power plant condenser, experience with other power plant condensers is most useful, whereas if you are designing a metal roof system for a building, experience with similar roofing systems is the most valuable information to use. Locating data for the exact materials, the exact geometry, and the same environment can be difficult or impossible, in which case prediction must be based on another method.

2.07.6.2 Exposure Data

When the exact geometry cannot be found, exposure data from similar materials in a similar environment to that of interest can be used instead. The geometrical effects of area ratio and proximity of dissimilar materials can usually be accounted for in a qualitative fashion, while computer modeling can sometimes be used to get more quantitative predictions. Besides selecting data from the same materials as those of interest, the most important test variable to consider is the environment. Galvanic corrosion can be very different in atmospheric environments, immersion in seawater or other chemicals, soils, inside concrete (such as reinforcing bar), etc. Even for atmospheric exposure, many variables can affect corrosion such as industrial versus rural, proximity to the sea, rainfall, humidity, amount of sunshine, and geometrical effects such as angle of exposure and shielding by neighboring structures.

2.07.6.3 Laboratory Data

Small-scale exposures of specimens in the laboratory can sometimes yield valuable information needed to predict galvanic corrosion. Laboratory tests can generate data on corrosion potentials, polarization behavior, and galvanic currents for specific geometries. Information regarding how to measure corrosion potentials and generate a galvanic series in a specific environment is given in a standard guide from ASTM International.[5] There are probably as many ways to generate polarization curves as there are investigators, but a good place to start is by generating either a potentiodynamic scan or a series of potentiostatic tests. Potentiostatic tests can also give a measure of the amount of corrosion of a metal as a function of its potential. Methods for conducting these tests are beyond the scope of this chapter.

Once corrosion potentials are known for the materials of interest in the environment of interest, classification of the material as an anode or a cathode can be made. Superposition of a cathodic polarization curve of the anode metal with an anodic polarization curve of the cathode material, with suitable scaling for relative surface areas, can generate information on the potential of the couple. Assuming no *IR* drop, this potential is the potential of both of the materials in the couple. Using potentiostatic corrosion rate data at this potential, the corrosion rate of the anode can be determined. This process of predicting galvanic corrosion rate from laboratory data is shown in **Figure 18**.

2.07.6.4 Computer Modeling

Computer modeling, specifically boundary element modeling and to a lesser extent finite element modeling, can be used to predict the amount and distribution of galvanic corrosion on a complex structure. The computer requires input of a detailed geometrical

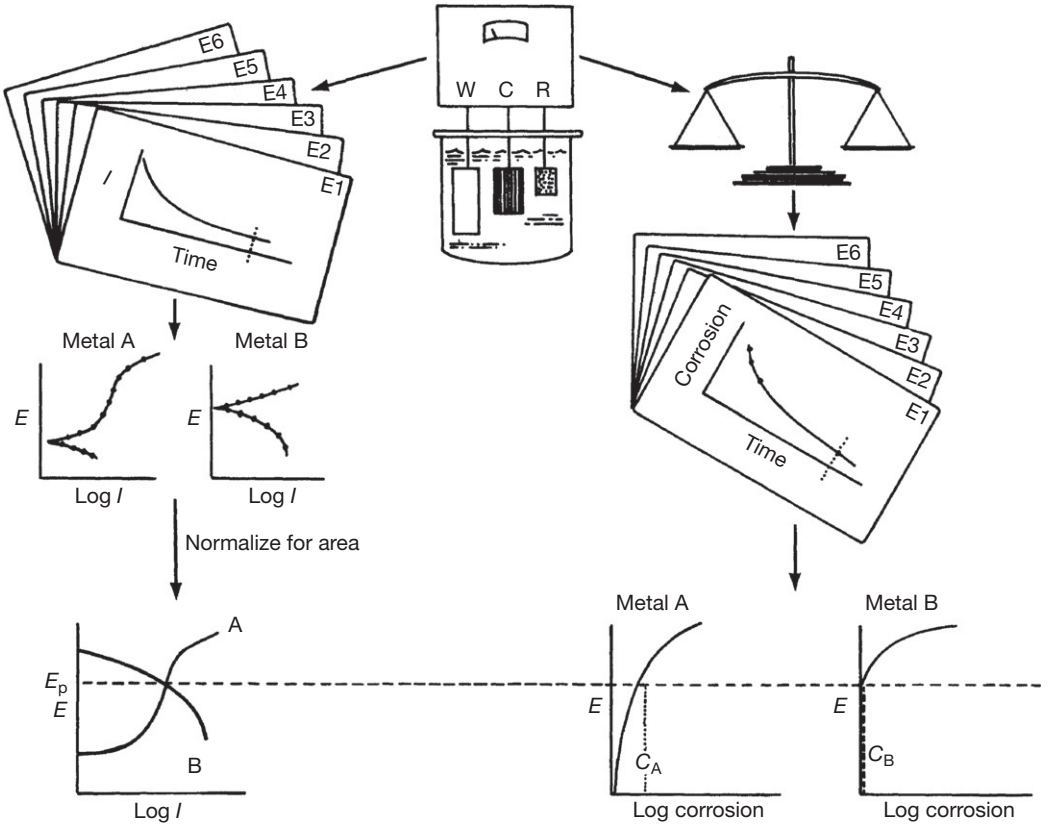


Figure 18 Predicting galvanic corrosion rate from potentiostatic data. Reproduced from Hack, H. P. *Corrosion Testing Made Easy*; NACE International: Houston, TX, 1993; Volume 2.

model of all surfaces exposed to the environment, the ionic conductivity of the environment, and the polarization curves for all of the materials involved. The latter is the most difficult data to obtain, since in many real situations polarization curves vary as a function of time and potential. Nevertheless, some excellent modeling has been conducted for large structures in seawater, mostly related to sacrificial anode cathodic protection (the opposite of galvanic corrosion) but to a lesser extent to galvanic corrosion.[8–21] Examples are shown in **Figures 19 and 20**.

2.07.6.5 Physical Scale Modeling

Physical scale modeling of complex structures has also been used to predict the amount and distribution of galvanic corrosion and cathodic protection.[22–25] Known by some investigators as dimension and conductivity scaling (DACS), proper scale modeling of corrosion processes involves increasing the resistivity of the environment by a factor equal to the dimensional scale factor. For example, a 1:10 scale model must be tested in an environment that has 10 times the resistivity of the full-scale environment, as seen in **Figure 21**.

2.07.6.6 Wagner Number

The Wagner number is the ratio of the polarization resistance of the materials of interest to the resistivity of the environment, expressed as the slope of a polarization curve.[12,26,27] It can be expressed as

$$W = (\mathrm{d}E/\mathrm{d}i)/R \qquad [1]$$

where $\mathrm{d}E/\mathrm{d}i$ is the local slope of the polarization curve of the metal of interest and R is the resistivity of the electrolyte.

If the Wagner number is large, galvanic current will flow far from the interface between anode and cathode, whereas if the Wagner number is small galvanic current will be more localized near the joint. Thus, large Wagner numbers are indicative of uniform distribution of galvanic corrosion, while

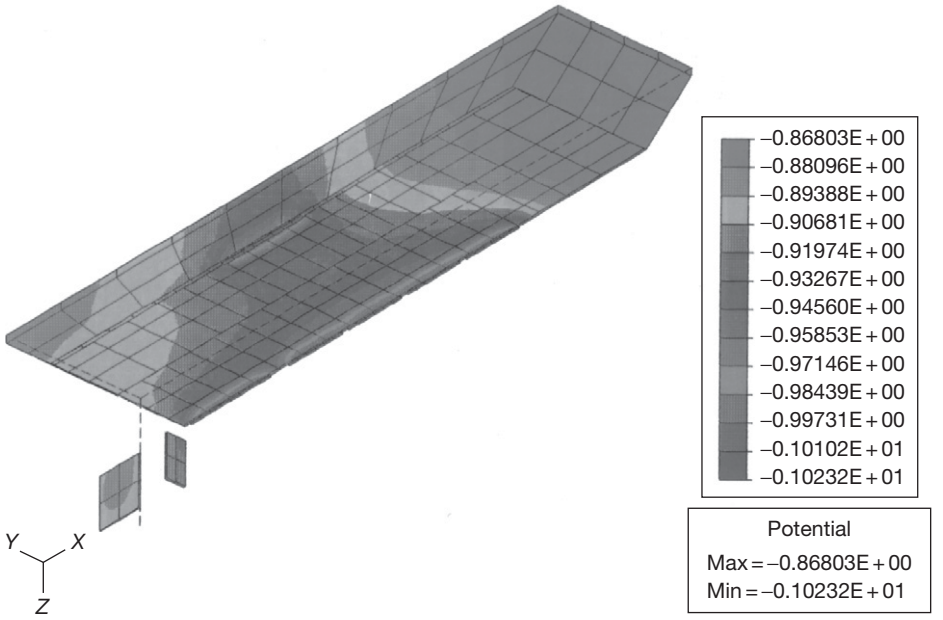


Figure 19 Computer-predicted potential distribution on a 48-ft barge with sacrificial anodes (half model). Reproduced from Hack, H. P.; Janeczko, R. M. Verification of the boundary element modelling technique for cathodic protection of large ship structures, CARDIVNSWC-TR-61–93/02, Naval Surface Warfare Center Carderock Division, Bethesda, MD, December 1993.

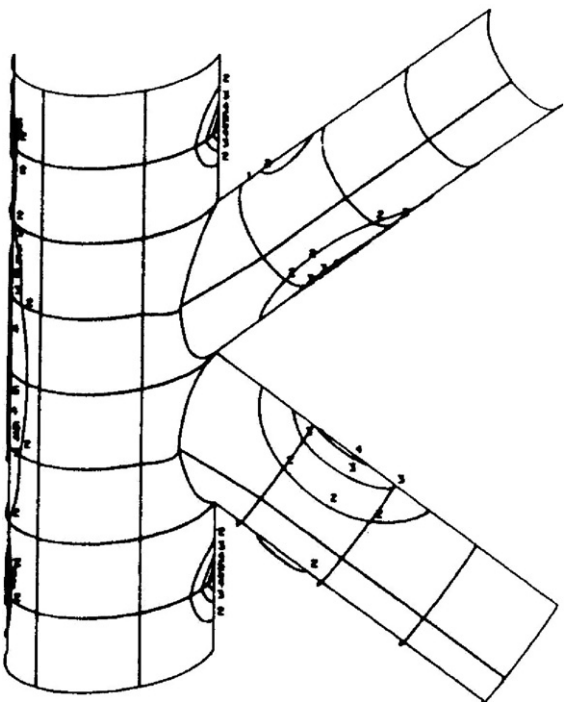


Figure 20 Computer prediction of potential profiles on cathodically protected K joint on offshore oil platform. Reproduced from Adey, R. A.; Niku, S. M. In *Galvanic Corrosion*; Hack, H. P. Ed.; ASTM International: West Conshohocken, PA, 1988; ASTM STP 978, pp 96–117.

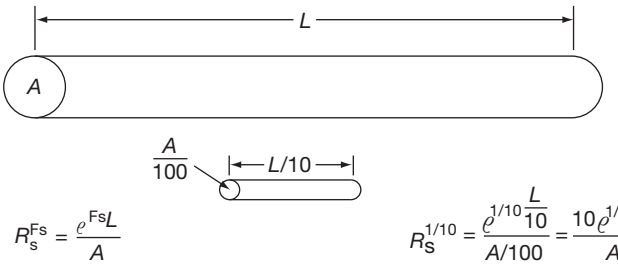
small Wagner numbers indicate localized galvanic corrosion. The effect of Wagner number is shown in Figure 22.

2.07.7 Prevention

There are several ways to minimize or prevent galvanic corrosion. These include design changes, electrical isolation, electrical insulation, exclusion of the environment, increasing *IR* drop, anode-to-cathode area ratio control, corrosion inhibitors, cathodic protection, and counter-current devices. Each of these corrosion control methods seeks to reduce or eliminate one of the requirements for galvanic corrosion to occur: materials with dissimilar corrosion potentials, metallic pathway, and ionic pathway.

Design changes can be made to minimize or eliminate contact of dissimilar metals. One such change is the use of a galvanic strip, shown in Figure 23. Use of this strip forces the dissimilar metal contact to occur in a metallurgically bonded transition material where water has limited access.

Electrical isolation eliminates galvanic corrosion by removing the metallic pathway between the dissimilar materials. This is usually accomplished with nonconductive spacers, gaskets, washers, and sleeves.



Potential drop:

$$E_s = IR_s = iAR_s$$

For $E_s^{FS} = E_s^{1/10}$ and $i^{FS} = i^{1/10}$:

$$AR_s^{FS} = \frac{A}{100} R_s{}^{1/10}$$

$$\rho^{1/10} = 10\rho^{FS}$$

Figure 21 Calculations to scale dimensions and conductivity for a galvanic couple.

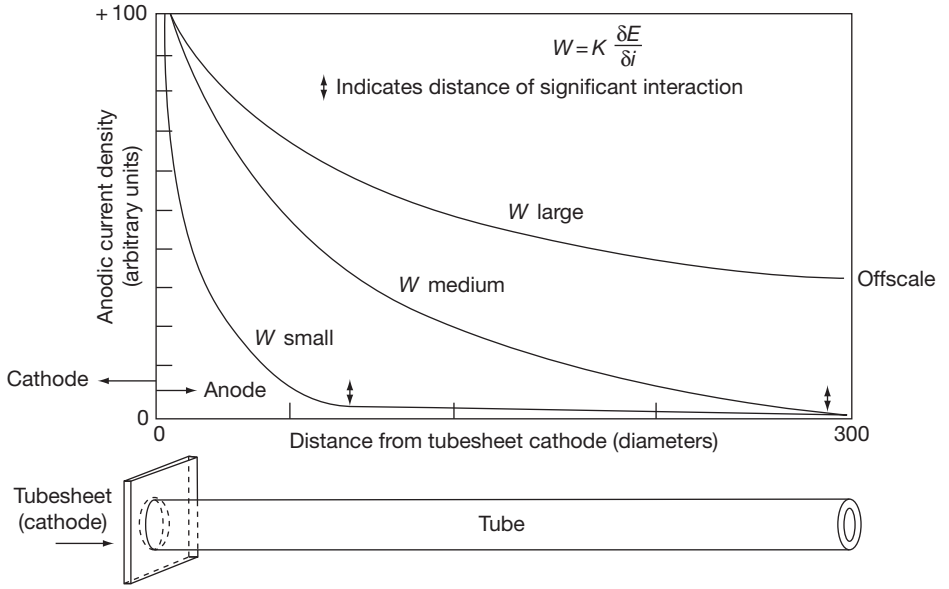


Figure 22 Effect of Wagner number on galvanic current distribution down an anodic tube. Reproduced from Scully, J. R.; Hack, H. P. In *Galvanic Corrosion*; Hack, H. P., ED., ASTM International: West Conshohocken, PA, 1988; ASTM STP 978, pp 136–157.

Typical methods for providing electrical isolation in a bolted joint are illustrated in **Figure 24**.

Commercial isolation devices are also available for pipelines from several manufacturers. The material that is used for isolation must be electrically nonconductive, so carbon-bearing rubbers cannot be used. It does not matter whether the isolation material absorbs electrolyte, so it can be porous. Typical isolation materials include Acetal, glass-reinforced plastic, Mylar, and polyethylene. Many nonmetallic bearing materials provide electrical isolation automatically for rotating equipment. Providing electrical isolation for complex structures is extremely difficult. Many isolator pieces are usually used, and unintentional shorts are difficult to control. Such unintentional shorts include pipe hangers, ground wires for electrical equipment, machining burrs, and inadvertent movable equipment like screwdrivers or steel wool pads. When providing electrical isolation it is imperative to check that electrical conductivity has been eliminated after assembly and periodically thereafter. Such checks must be done with the equipment dry.

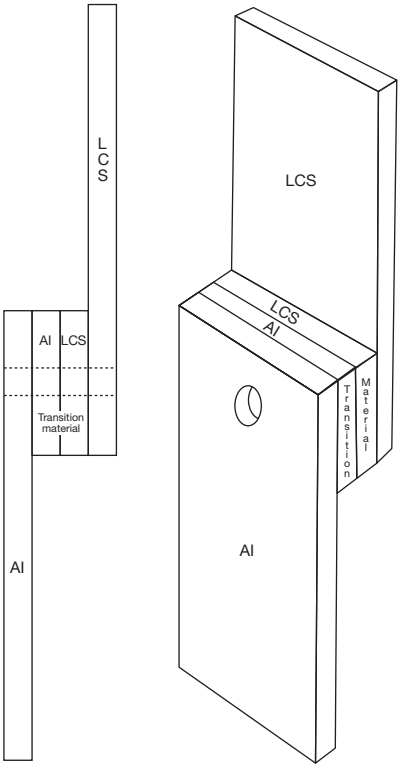


Figure 23 Use of a bimetallic transition material to minimize galvanic corrosion between low carbon steel (LCS) and aluminum. Reproduced from Baboian, R.; Haynes, G.; Turcotte, R. In *Galvanic Corrosion*; Hack, H. P., Ed.; ASTM International: West Conshohocken, PA, 1988; ASTM STP 978, pp 249–259.

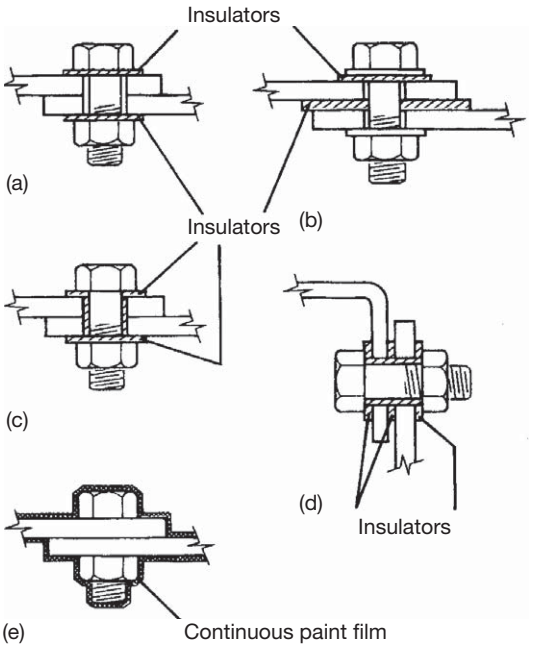


Figure 24 Designs for insulating joints to avoid corrosion likely if the fastener is anodic to the plate[1]: (a) bolt and plates are of different alloys (not ideal; see text); (b) the bolt is the same alloy as the lower plate; (c) same as (a) with insulating sleeve on bolt shank; (d) bolt and plates are all of different alloys; (e) suitable only for atmospheric or occasionally wetted surface. Courtesy of HMSO. Reproduced from Francis, R. *Galvanic Corrosion: A Practical Guide for Engineers*; NACE: Houston, TX, 2001.

Insulation is different in that the intent is not to remove the metallic pathway but instead to provide spacing between the anode and cathode to increase the resistance of the ionic pathway. This is commonly used in high-resistivity environments such as atmospheric exposure, where a small increase in spacing will have a large effect on ionic resistance. For this reason, plastic spacers are frequently used between stainless steel fasteners and aluminum siding material. Another application of this principle is the use of nonmetallic spool pieces to separate dissimilar metals in piping systems. Although the dissimilar metals may still be in metallic contact through supporting structure, the spool piece provides ionic resistance which lowers the amount of galvanic corrosion, and also causes the corrosion to be less localized on the anode pipe near the cathode.[28]

Galvanic corrosion can also be stopped by excluding the environment from the anode, the cathode, or both. This has sometimes been done successfully with the use of paint; however, paint systems will develop defects over time which must be accounted for. Water-excluding greases, silicone sealants, and heavy nonmetallic linings will work for various periods of time, until they wash out or deteriorate enough to absorb electrolyte. It is safest to exclude the environment from the cathode, since this will eliminate the driving force for the galvanic corrosion.

Anything that increases the resistance of the ionic path between anode and cathode will reduce the rate of galvanic corrosion. Reducing electrolyte conductivity, increasing anode-to-cathode spacing, and providing nonmetallic barriers in the electrolyte between the anode and cathode are effective corrosion control methods.

Since the magnitude of galvanic corrosion is a function of the ratio of the wetted surface areas of the cathode to the anode, reducing this ratio will reduce galvanic corrosion. Masking of the cathode with paint can be highly effective, whereas painting of the anode alone must never be done. Area ratio control is most effectively done during the design phase of a structure. For example, fasteners should

be cathodic to the materials they fasten. When putting stainless steel fasteners through aluminum for seawater use, there are surface treatments that will make even this small area of noble material appear even smaller by masking the corrosion potential of the fastener. Such treatments include galvanizing, cadmium plating, and the use of a proprietary fastener treatment called Alumizite.

Corrosion inhibitors can control galvanic corrosion either by acting directly to prevent corrosion of the anode, called anodic inhibitors, or by increasing the polarizability of the cathode to lower its driving force on the anode, called cathodic inhibitors. Inhibitors that do both are called mixed inhibitors. If an anodic inhibitor is used, the coupled potential will shift in the cathodic direction, whereas a cathodic inhibitor will cause the coupled potential to shift in the anodic direction. Automobile cooling systems contain a variety of materials, including steel, copper alloys, and aluminum, all in direct metallic contact in a common electrolyte, the antifreeze/coolant. For this reason, companies that produce antifreeze/coolants have developed a suite of corrosion inhibitors that are highly effective at controlling galvanic corrosion in automobile engines, provided they are changed out every few years to replenish the inhibitors.

Cathodic protection of a structure will eliminate galvanic corrosion since it provides electrodes that are anodic to both materials in the galvanic couple. This is true for both sacrificial and impressed current cathodic protection systems. The one caution regarding this is that the normal corrosion of aluminum in seawater cannot be prevented by cathodic protection, and overprotection will cause considerable damage to the aluminum. Therefore, when trying to prevent galvanic corrosion of aluminum caused by contact with other materials, just enough cathodic protection should be used to polarize the other materials to the corrosion potential of the aluminum. The cathodic protection system must have an ionic path to all cathodes in the galvanic couple to be effective. Since cathodic protection cannot reach into faying surfaces, it is usually good practice to use nonconductive, nonporous spacers between dissimilar metals so protected, not to provide isolation but to force any potential galvanic currents to travel outside of the faying surface area where they can be mitigated by the cathodic protection system. A typical insulation scheme for faying surfaces uses two layers of MIL-I-24391 tape[29] with a total thickness of 0.43 mm (0.017 in.).

Counter-current devices are relatively new to the field of galvanic corrosion control. Their principle is that they generate an *IR* drop in the electrolyte counter to the *IR* drop generated between the anode and cathode in a galvanic couple. This effectively prevents current from flowing in the electrolyte, stopping the galvanic corrosion. To date, one such device, called a bielectrode, has been the subject of research[30–33] and patents.[34] The galvanic couple geometry that can be controlled with this technology must be well controlled, and control of the counter current is the most difficult part of designing such a system. The principle of a counter-current device is shown in **Figure 25**.

2.07.8 Testing

Galvanic corrosion testing is a complex subject that is discussed in detail elsewhere.[4,35–39] This chapter provides only a summary.

2.07.8.1 General Principles

In general, when conducting tests for galvanic corrosion, the closer the test comes to the actual situation, the more accurate will be the results. The environment must be accurately modeled, with the amount of any constituents that might affect the reactions accurately reproduced and the amount of fluid flow reproduced as well. The materials tested should accurately reflect the materials of interest in composition, thermal processing, and surface films or treatments. If materials are to be tested as a couple, the geometry should reflect as closely as possible that being studied. Finally, galvanic corrosion, like any other form of corrosion, can be strongly time-dependent. Tests must be conducted over sufficient duration to ensure that the results can accurately predict corrosion behavior over the service life anticipated.

2.07.8.2 Testing Hardware

The most accurate test is one in which the actual hardware is exposed in the actual environment for the anticipated service life. In practice this seldom can be done, so that approximations may be required. The considerations for hardware testing are discussed elsewhere.[35]

2.07.8.3 Making a Galvanic Series

If a galvanic series is to be constructed, it is important that the materials and environment accurately reflect those of interest. In many environments, corrosion

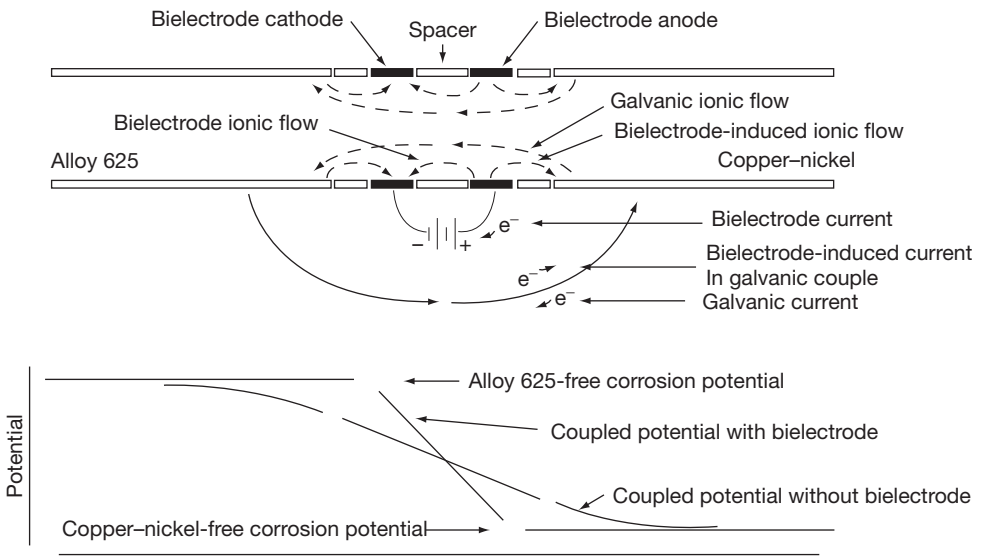


Figure 25 Principle of operation of counter-current device. Reproduced from Shifler, D. A. Advanced measures to control galvanic corrosion in piping systems, CARDIVNSWC-TR-61–99–18 Naval Surface Warfare Center Carderock Division, Bethesda, MD, September 1999.

potential varies with time, so that the length of time of the measurement must be long enough to allow measurement of the full range of potentials. These, and other considerations for developing a galvanic series, are discussed in more detail elsewhere.[5]

2.07.8.4 Determining Polarization

Polarization curves may be developed in a variety of ways, depending on the corrosion system being modeled. Potentiodynamic polarization curves are generated by holding a specimen at a given potential and gradually sweeping the potential in one direction while measuring the current required. If the starting potential is the corrosion potential, then an anodic polarization curve is generated by sweeping in the positive direction, while a cathodic polarization curve is generated by sweeping in the negative direction. Sometimes scans are started at a cathodic potential and swept in the positive direction, generating both anodic and cathodic branches of the polarization curve in a single scan. Sometimes the specimens may be pre-exposed before starting the scan either freely corroding or at a potential indicative of the likely coupled potential. The speed of the potential scan is important, with slower usually giving better results. Details of generating polarization curves are discussed elsewhere.[4,37,40–44]

Polarization curves can also be generated potentiostatically by taking a series of specimens and holding each at a different constant potential while measuring current as a function of time. This better approximates the potential behavior of a galvanic couple than a potentiodynamic scan, and current stability over time is easily obtained, but the technique is very labor and time intensive and is seldom used in practice. Details of how to perform this type of testing are discussed elsewhere.[4,42,45,46]

Actual galvanic couples of the materials of interest may be exposed in the proper environment while the coupled potential and galvanic current are measured as a function of time. Although the exact geometry is usually not modeled in this type of test, typically the anode-to-cathode area ratio is modeled by adjusting specimen size. This type of testing yields galvanic current information which may not be directly relatable to the magnitude of galvanic corrosion, but if the anode and cathode corrosion potentials are more than 120 mV apart, the anodic reactions on the coupled cathode are almost completely suppressed and the cathodic reactions on the anode are almost completely suppressed, so that the galvanic current will give a rough measure of the anode corrosion rate.[4] Exposure testing of coupled specimens is discussed elsewhere.[4,35]

2.07.8.5 Determining *IR* Drop

The above testing usually does not take into account *IR* drop. *IR* drop can be measured during an exposure of galvanic couples either by using a small

reference electrode or Luggin–Haber capillary and moving it near the various surfaces, or by rapid interruption of the galvanic current if this current travels through an external wire followed by immediate measurement of potential. Typical interruption times are on the order of milliseconds. The various methods of measuring *IR* drop are described elsewhere.[47]

2.07.8.6 Determining Corrosion Rate

Once corrosion current is known, corrosion rate can be calculated using Faraday's law.[48] If the anode and cathode corrosion potentials are more than 120 mV apart, the current that can be used in Faraday's law calculations is the galvanic current. If not, Evans diagrams must be used to determine reaction currents that can be put into Faraday's law to get corrosion rates of the anode and cathode materials.

An easier method for obtaining corrosion rates is to determine specimen mass of the anode metal before and after an exposure in a galvanic couple or at a potential indicative of the coupled potential. This corrosion rate should be compared to that of a freely corroding specimen of the anode metal to determine the amount of acceleration due to the presence of the couple.

Some highly localized corrosion phenomena, such as pitting and crevice corrosion, cannot be easily quantified by mass loss measurements. In those cases, other methods besides corrosion rate must be used to determine the amount of acceleration of the corrosion by the galvanic couple.[49]

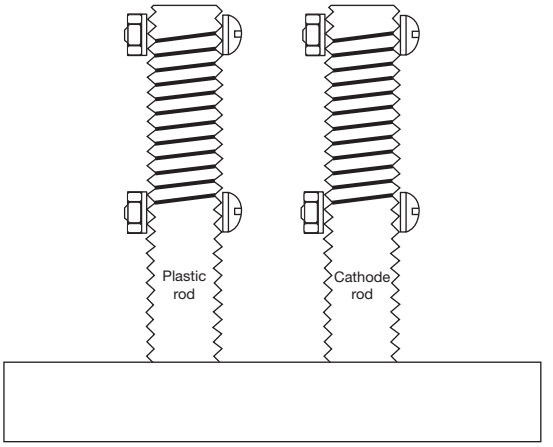


Figure 26 Wire-on-bolt or CLIMAT test specimens. Reproduced from Standard G116: Practice for Conducting Wire-on-Bolt Test for Atmospheric Galvanic Corrosion, *ASTM Book of Standards*; ASTM International: West Conshohocken, PA, 2006.

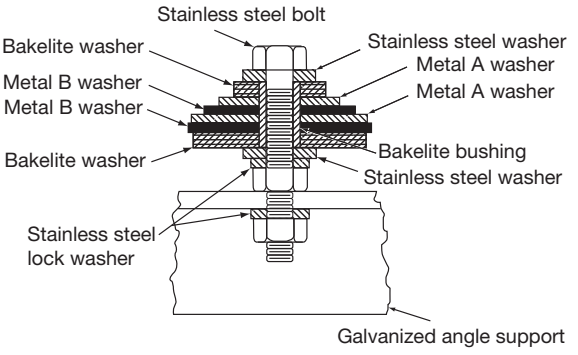


Figure 27 Washer atmospheric test specimen. Reproduced from Standard G149 (withdrawn 2004): Practice for Conducting the Washer Test for Atmospheric Galvanic Corrosion, *ASTM Book of Standards*; ASTM International: West Conshohocken, PA, 2006.

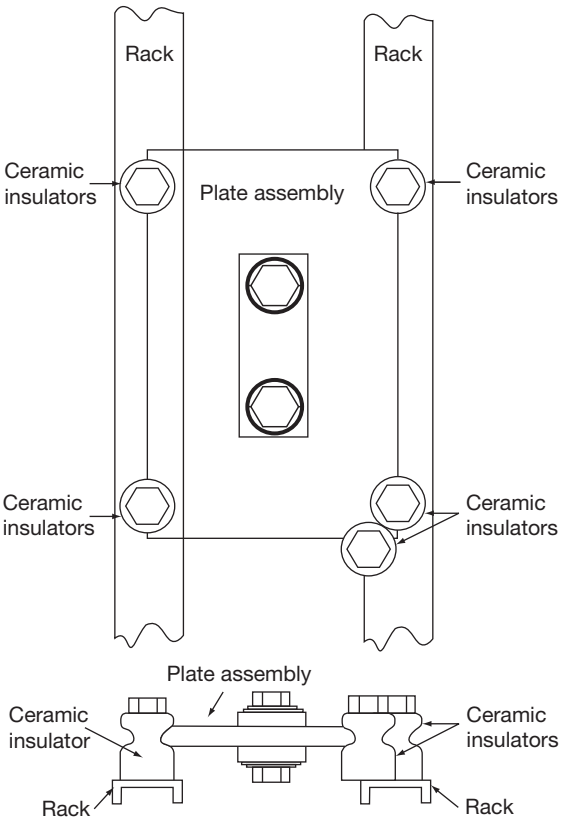


Figure 28 ISO atmospheric corrosion test using plate material. Reproduced from Hack, H. P. *Corrosion Testing Made Easy*; NACE International: Houston, TX, 1993; Volume 2.

2.07.8.7 Accelerated Tests

Accelerated testing to predict galvanic corrosion is not normally productive. Any attempt at accelerating the corrosion has a significant chance of changing the corrosion mechanism, invalidating the data that is collected. It is far more productive to allow the galvanic corrosion to proceed at its normal pace and use accurate current and mass loss measuring techniques to get data in a shorter time period.

2.07.8.8 Atmospheric Testing

There are several standard tests for galvanic corrosion in the atmosphere. The most popular of these is called the wire-on-bolt, or CLIMAT, test[50,51] (**Figure 26**). This test consists of wrapping a wire made from the anode material around a threaded rod made from the cathode material. If the anode and cathode are not known beforehand, two tests are run with the materials reversed. A control is usually run consisting of the anode wire wrapped around a nonconductive rod so that the non-galvanically accelerated corrosion of the anode material can be compared with its corrosion in the galvanic couple. This test is fast, typically taking 30 days, and sensitive, and only has the disadvantage that the materials must be available in the appropriate wire or threaded rod forms.

For materials that cannot be made into the right form for the wire-on-bolt test, another test using stacked washers of the anode and cathode material alternating can be used.[52] Although this test is not as sensitive and usually requires longer exposures, it is

Fastener selection-above water

Fastener	C steel[a]	Aluminum[a]	Copper-based alloy[a]	Stainless-steel alloy[a]	Alloy 400[a]
Steel, Zn/Cd	G	R	R	R	R
Aluminum	Y	Y	R	R	R
Copper alloy	G	R[b]	Y	R	R
Type 316 SS	G	Y[b]	Y	Y	Y
Alloy 400	G	R[b]	G	G	G
Ni–Cr–Mo	G	Y[b]	G	G	G
MP35N	G	Y[b]	G	G	G
Titanium	G	Y[b]	G	G	G

[a]Material to be fastened.
[b]Loss of aluminum around fastener has caused structural failure in coastal buildings, but stainless-steel fasteners are widely used.
G = generally satisfactory. Y = may be satisfactory; detailed information needed. R = avoid.

(a)

Fastener	Graphite composite[a]	Wood[a]	Concrete[a]	FRP[a]	Rubber[a]
Steel, Zn/Cd	R	Y	G	Y	Y
Aluminum	R	Y	Y	Y	Y
Copper alloy	R	G	G	G	G
Type 316 SS	R	Y	G	G	G
Alloy 400	R	G	G	G	G
Ni–Cr–Mo	?	G	G	G	G
MP35N	?	G	G	G	G
Titanium	?	G	G	G	G

[a]Material to be fastened.
G = generally satisfactory. Y = may be satisfactory; detailed information needed. R = avoid.
? = very limited information available; suggest caution.

(b)

Figure 29 Fastener selection – above water. Reproduced from Ross, R. W., Jr.; Tuthil, A. H. *Master. Perform* **1990**, *29*(4), 65–69.

Corroding metal	Coupled metal: Aluminum	Anodized Al	Zinc	Carbon steel	Weathering steel	Tin	Lead	Copper	Nickel	Stainless steel	Chromium
Magnesium											
Aluminum											
Anodized aluminum											
Zinc											
Carbon steel											
Weathering steel											
Tin											
Lead											
Copper											
Nickel											
Stainless steel											

■, no risk of galvanic corrosion; ■, some increase in corrosion; ■, large increase in corrosion; ☐, no data.

(a)

Corroding metal	Coupled metal: Aluminum	Anodized Al	Zinc	Carbon steel	Weathering steel	Tin	Lead	Copper	Nickel	Stainless steel	Chromium
Magnesium											
Aluminum											
Anodized aluminum											
Zinc											
Carbon steel											
Weathering steel											
Tin											
Lead											
Copper											
Nickel											
Stainless steel											

■, no risk of galvanic corrosion; ■, some increase in corrosion; ■, large increase in corrosion; ☐, no data.

(b)

Corroding metal	Coupled metal: Aluminum	Anodized Al	Zinc	Carbon steel	Weathering steel	Tin	Lead	Copper	Nickel	Sta8inlss steel	Chromium
Magnesium											
Aluminum											
Anodized aluminum											
Zinc											
Carbon steel											
Weathering steel											
Tin											
Lead											
Copper											
Nickel											
Stainless steel											

■, no risk of galvanic corrosion; ■, some increase in corrosion; ■, large increase in corrosion; ☐, no data.

(c)

Figure 30 Risk of galvanic corrosion in an (a) urban atmosphere, (b) rural atmosphere and (c) marine atmosphere for some common metals. Reproduced from Francis, R. *Galvanic Corrosion: A Practical Guide for Engineers*; NACE: Houston, TX, 2001.

Fastener	C steel[a]	Aluminum[a]	Copper-based alloy[a]	Stainless-steel alloy[a]	Alloy 400[a]
Steel, Zn/Cd	Y	R[b]	R	R	R
Aluminum	R	Y	R	R	R
Copper alloy	G	R	Y	R	R
Type 316 SS	G	R[b]	Y	R	R
Alloy 400	G	R[b]	G	Y	Y
Ni–Cr–Mo	G	R[b]	G	G	Y
MP35N	G	R[b]	G	G	G
Titanium	G	R[b]	G	G	G

[a]Material to be fastened.
[b]Loss of aluminum around fastener is as serious as loss of fastener.
G = generally satisfactory. Y = may be satisfactory; detailed information needed. R = avoid

Fastener	Graphite composite[a]	Wood[a]	Concrete[a]	FRP[a]	Rubber[a]
Steel, Zn/Cd	R	R	Y	Y	Y
Aluminum	R	Y	R	R	R
Copper alloy	R	G	G	G	G
Type 316 SS	R	R	Y[b]	R[c]	R
Alloy 400	R	G	G	G	G
Ni–Cr–Mo	?	G	G	G	G
MP35N	?	G	G	G	G
Titanium	?	G	G	G	G

[a]Material to be fastened.
[b]The alkalinity of concrete provides protection for small areas of type 316 extending beyond the concrete.
[c]Some small boat manufacturers report good performance for type 316 bolting in FRP hulls below waterline when bolts are packed with water-repellent lubricant and recessed in FRP.
G = generally satisfactory. Y = may be satisfactory, detailed information needed. R = avoid.
? = very limited information available; suggest caution.

Figure 31 Fastener selection – below water. Reproduced from Ross, R. W., Jr.; Tuthil, A. H. *Master. Perform* **1990**, *29*(4), 65–69.

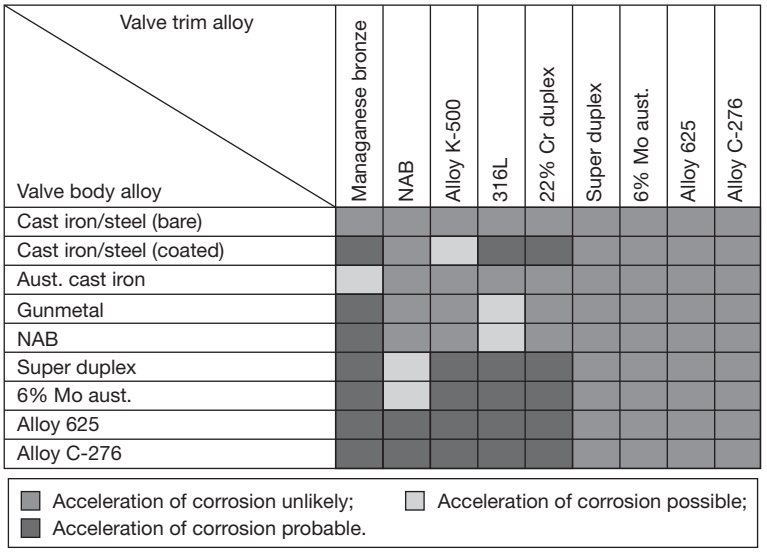


Figure 32 Alloys suitable for valve trim in ambient temperature seawater. Reproduced from Francis, R. *Galvanic Corrosion: A Practical Guide for Engineers*; NACE: Houston, TX, 2001.

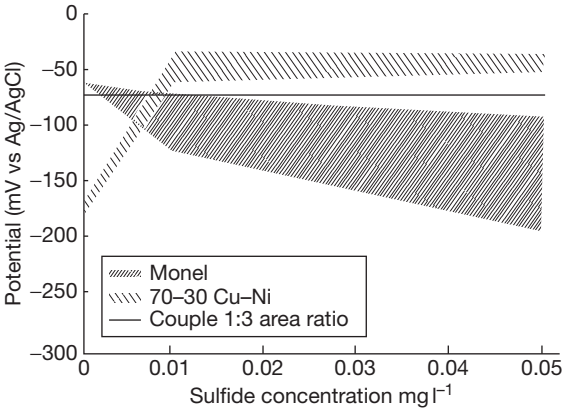


Figure 33 Potential reversal of copper–nickel piping alloys due to sulfide pollution.

Table 1 Weight loss of iron for Fe/M bimetallic couples in 1% NaCl[a] [58]

M	*Weight loss of iron (mg)*	*Weight loss of M (mg)*	*Difference in standard electrode potentials (V)*[b]
Copper	183.1	0.0	+0.785
Nickel	181.1	0.7	+0.19
Tin	171.1	2.5	+0.30
Lead	183.2	3.6	+0.31
Aluminum	9.8	105.9	−1.23
Cadmium	0.4	307.9	+0.04
Zinc	0.4	688.0	−0.32
Magnesium	0.0	3104.0	−1.90

[a]Weight loss data after Bauer and Vogel.[5]
[b]The + sign indicates that iron has a more negative standard electrode potential than the second metal.

not necessary to know before hand which material is the anode. Although it has many advantages, this test is no longer covered by a current standard in ASTM. The test specimen is shown in Figure 27.

Sometimes deterioration in properties such as tensile strength or modulus due to galvanic corrosion is needed. For this, a plate test was designed and standardized by ISO.[53] In this test, a plate of the anode material has a strip of the cathode material affixed to it. After exposure, the anode plate can be cut into tensile specimens to test for deterioration of physical properties. Like the wire-on-bolt test, if the anode is not known beforehand, two tests are run with the materials reversed. Like the washer test, longer exposures are required but the plate form of the material is usually easier to get (Figure 28).

When performing atmospheric testing there are variables that must be considered. The amount of galvanic corrosion will be a function of the atmosphere, its rainfall acidity, chloride content, and time of wetness, which is a function of rainfall, amount of sunshine, prevailing winds, and other variables. Corrosion can be affected by the angle of exposure of specimens, sheltering from wind and sunlight, and height above ground, among others. Corrosion in the atmosphere is also a function of season; so long tests, or multiple short tests over at least a year, are required to get a proper indication of corrosion performance.[54]

2.07.9 Standards

There are many standards used in evaluating and preventing galvanic corrosion. A good reference for these is published by ASTM International,[38] and only a summary is given here. Standards fall into a variety of categories: those that relate to corrosion testing in general such as standards on how to clean specimens after test, general galvanic corrosion test guidelines, specific galvanic corrosion tests such as the CLIMAT test, electrochemical test standards whose results can be used in galvanic corrosion studies such as potentiodynamic polarization tests, and standards for protection methods which can be applied to galvanic corrosion such as cathodic protection standards. These standards have been developed by ASTM International, NACE International, the ISO, the European standards agency (CEN), and numerous countries' standards organizations. Listing of these standards and what they can be used for is beyond the scope of this chapter, and the reader is directed to ASTM International.[38]

2.07.10 Useful Galvanic Corrosion Data

The following data may prove useful for estimating rates of galvanic corrosion.

2.07.10.1 Atmospheric Corrosion

Figure 29 contains useful data showing galvanic corrosion performance of fasteners in the atmosphere. Figure 30 shows galvanic corrosion performance of a large number of material combinations in urban, rural, and marine atmospheres.

Table 2 Degree of corrosion at bimetallic contacts[a] [59]

Metal considered	*Contact metal*						
	1	*2*	*3*	*4*	*5*	*6*	*7*
	Gold, platinum, rhodium, silver	*Monel, Inconel, nickel–molybdenum alloys*	*Cupronickels silver solder, aluminum bronzes, tin bronzes, gunmetals*	*Copper brasses, 'nickel silvers'*	*Nickel*	*Lead, tin and soft solders*	*Steel and cast Iron*
1. Gold, platinum, rhodium, silver	–	A	A	A	A	A	A
2. Monel, Inconel, nickel–molybdenum alloys	B	–	A	A	A	A	A
3. Cupronickels, silver solder, aluminum bronzes, tin bronzes, gun metals	C(*k*)	B or C	–	A	A	A	A
4. Copper, brasses, 'nickel silvers'	C(*k*)	B or C	B or C(*g*)	–	B or C	B or C (*p*)	A
5. Nickel	C	B	A	A	–	A	A
6. Lead, tin and soft solders	C	B or C(*t*)	B or C(*q*)	B or C(*q*)	B	–	A or C(*r*)
7. Steel and cast Iron (*a*) (*f*) (*w*)	C	C	C	C	C(*k*)	C(*k*)	–
8. Cadmium (*u*)	C	C	C	C	C	B	C
9. Zinc (*u*)	C	C	C	C	C	B	C
10. Magnesium and magnesium alloys (chromated) (*b*) (*a*)	D	D	D	D	D	C	D
11. *Austenitic* Fe–18Cr–8Ni	A	A	A	A	A	A	A
12. Stainless steel Fe–18Cr–2Ni	C	A or C(*s*)	A or C(*s*)	A or C(*s*)	A	A	
13. 13% Cr	C	C	C	C	B or C	A	A
14. Chromium	A	A	A	A	A	A	A
15. Titanium	A	A	A	A	A	A	A
16. Aluminum and aluminum alloys (*n*) (*a*) (*w*)	D	C	D(*e*)	D(*e*)	C(*k*)	B or C	B or C

[a]Based on data provided by members of the I.S.M.R.C. Corrosion and Electrodeposition Committee and others, and arranged by Mrs. V. E. Rance.

A. The corrosion of the 'metal considered' is not increased by the 'contact metal.'

B. The corrosion of the 'metal considered' may be slightly increased by the 'contact metal.'

C. The corrosion of the 'metal considered' may be markedly increased by the 'contact metal.' (Acceleration is likely to occur only when the metal becomes wet by moisture containing an electrolyte, e.g., salt, acid, combination products. In ships, acceleration may be expected to occur under in-board conditions, since salinity and condensation are frequently present. Under less severe conditions the acceleration may be slight or negligible.)

D. When moisture is present, this combination is inadvisable, even in mild conditions, without adequate protective measures.

2.07.10.2 Seawater Immersion

Figure 31 contains data showing galvanic corrosion performance of fasteners underwater. Figure 32 shows galvanic corrosion performance of a number of alloys used for valve trim in seawater. Figure 33 shows how small variations in composition of seawater caused by pollution can actually reverse the relative galvanic corrosion performance of certain alloys, reversing the anode and the cathode in a galvanic couple.

Table 3 Degree of corrosion at bimetallic contacts[63] (Continued)

8	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
Cadmium	*Zinc*	*Magnesium and magnesium alloys (chromated)*	*Stainless*			*Chromium*	*Titanium*	*Aluminum and aluminum alloys*
			Austenitic Fe–18Cr–2Ni	*Fe–18Cr–2Ni*	*13% Cr*			
A	A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A(*x*)	A
A	A	A	B or C	B	A	B or C	B or C	A(*e*)
A	A	A	B or C	B or C	A	B or C	B or C	A(*e*)
A	A	A	B or C	B or C	A	B or C	B or C	A
A	A or C(*r*)	A	B or C	B or C	B or C	B or C	B or C	A
A(*m*)	A(*m*) (*l*)	A	C	C	C	C(*k*)	C	A(*m*)
–	A	A	C	C	C	C	C	B
B	–	A	C	C	C	C	C	C(*i*)
B or C	B or C	–	C	C	C	C	C	B or C(*c*)
A	A	A	(*v*)	A	A	A	A	A
A	A	A	A	(*v*)	A	A	(*o*)	A
A	A	A	C	C	(*v*)	C	C	A
A	A	A	A	A	A	–	A	A
A	A	A	A	A	A	A	–	A
A	A	A(*c*) (*h*)	B or C	B or C	B or C	B or C(*d*)	C	(*v*)

Table 4 Total corrosion rate acceleration factors due to dissimilar metal coupling at 1:1 area ratio in flowing seawater (mean flow rate about 1–2 m s^{-1}) [59]

Coupled metal (wrought form)	*Uncoupled corrosion rate*[a]	*Acceleration factor*[b] *due to dissimilar metal coupling with:*	
		Titanium	*Mild steel*
Zinc	0.05	4	10
SIC aluminum	0.008	30	60
Mild steel.	0.15	2	1
Lead	0.01	3	3 MS
2% aluminum brass	0.01	3	3 MS
10% aluminum bronze	0.02	3	3 MS
Nickel aluminum bronze	0.015	1	3 MS
Copper	0.03	6	3 MS
90/10 cupro-nickel (1% Fe)	0.02	3	3 MS
Monel 400	0.005	2	3 MS
Stainless steel type 316	0.005	2	3 MS

[a]Corrosion rates relate to general corrosion only and are average rates obtained over about one year's exposure.
[b]Acceleration factors quoted for coupled metal corroding unless indicated by suffix (MS) for mild steel corroding.

Quantitative data is presented in, **Table 1–9**, **Table 1** shows weight loss data for iron in saltwater when coupled to various alloys. **Table 2** is an extensive list showing how various alloys perform when galvanically coupled to each other in various environments. **Table 3** is a continuation of **Table 2** with more alloy combinations. **Table 4** gives the amount of galvanic corrosion of various alloys when they are

coupled to titanium or mild steel. Table 5–7 can help to predict galvanic corrosion performance in seawater heat exchangers. Table 5 shows galvanically accelerated crevice corrosion when Monel K-500 is coupled to a super duplex stainless steel. Table 6 presents galvanic corrosion information for Muntz metal when coupled to other seawater heat exchanger alloys. Table 7 shows the amount of galvanic corrosion that can occur in heat exchangers when aluminum bronze tubesheets are coupled to stainless steel or titanium tubes. Other useful data for predicting galvanic corrosion of Naval seawater valve materials can be found in Hack.[56] Table 8 is useful for comparing galvanic corrosion performance of a variety of different alloys used in seawater when galvanically coupled to each other. Finally, Table 9 shows the effect of galvanically coupling a large variety of different materials to a titanium alloy at various area ratios and flow rates in seawater.

Data given in Hack[57] contain extensive polarization data in seawater at various flow rates that can be used for predicting galvanic corrosion between some common materials used in seawater at any area ratio. This information is also useful as boundary conditions for computerized boundary element analysis for prediction of the amount and spatial distribution of galvanic corrosion and cathodic protection.

Table 5 Maximum depth of crevice attack for couples of alloy K-500 and super duplex stainless steel[60]

Couple	*Alloy*	*Maximum pit depth (mm)*	
		Washer/plate	*Plate/plate*
1	K-500	0.26	0.33
	K-500	0.12	0.30
2	Super duplex	0.00	0.00
	Super duplex	0.00	0.00
3	Super duplex	0.00	0.00
	K-500	0.32	0.35
4	Super duplex	0.00	0.00
	K-500	0.42	0.44
5	Super duplex	0.00	0.00
	K-500	0.52	0.50

Couples were at a 1:1 area ratio and exposed to natural sea-water at ambient temperature for 84 days.

Table 6 Corrosion rate of Muntz metal coupled to various condenser tube alloys[60]

Tube alloy	*Corrosion rate (mm year^{-1})*	
	Inlet end	*Outlet end*
None	0.225	0.225
90/10 Cu-Ni	0.846	0.279
6% Mo austenitic	2.80	2.31
Titanium	6.96	6.99

This information is taken from Fu and Chow.[16]

2.07.10.3 Other Environments

Finding existing galvanic corrosion data in environments other than the atmosphere or seawater can be challenging, and frequently such data do not exist and must be created by running tests. Figures 34 and 35 show galvanic series generated in carbon dioxide or hydrogen sulfide environments. Scattered data are also contained in some of the general corrosion texts or guides.[59,63–66]

Table 7 Effect of temperature on the galvanic corrosion of aluminum bronze (alloy D) coupled to 6% Mo austenitic stainless steel or titanium heat exchanger tubes[60]

Alloy	*Tube end*	*Corrosion rate (mm year^{-1})*			
		Seawater at 22 °C	*Brackish water at 22 °C*	*Seawater at 6 °C*	*Brackish water at 11 °C*
6% Mo austenitic	Inlet	0.633	0.117	0.048	0.097
	Outlet	0.634	0.150	0.087	0.224
Titanium	Inlet	1.86	0.201	0.059	0.201
	Outlet	3.03	0.373	0.049	0.226

This information is taken from Ref. 16.

Table 8 Galvanic corrosion in seawater at 1:1 area ratio

Alloy Inducing galvanic effect	*Seawater Velocity, ft/s*	*Corrosion rate (mpy) measured on sample of*					
		Ti-6-4	*Inconel 625*	*Monel*	*70/30 Cu–Ni*	*Ni–Al Bronze*	*G Bronze*
None (Freely Corroding)	3	0.2	0.1	3.5	1.4	3.7	6.3
	13	0.1	0.1	0.6	6.0	7.0	9.3
Ti-6-4	3		0.1	3.2	3.8	11.3	21.4
	13		0.1	0.6	3.0	10.0	12.9
Inconel 625	3	0.2		3.5	19.6	32.4	37.8
	13	0.1		0.7	4.6	10.5	15.1
Monel	3	NIL	0.1		8.1	20.4	26.9
	13	0.1	NIL		4.1	11.7	12.2
70/30 Cu–Ni	3	0.2	NIL	1.5		4.1	4.9
	13	NIL	NIL	0.6		5.6	12.4
Ni–Al Bronze	3	0.1	NIL	0.9	1.3		4.4
	13	0.1	0.1	0.5	1.9		17.5
G Bronze	3	NIL	NIL	0.5	1.4	2.7	
	13	0.2	0.2	0.4	1.9	3.3	

Table 9 Summary of corrosion performance of materials coupled to Ti-6211[66]

Alloy	*Acceleration factors*							
	<0.01 m s^{-1}				*0.5 m s^{-1}*			
	1:1		*10:1*		*1:1*		*10:1*	
	General	*Pitting*	*General*	*Pitting*	*General*	*Pitting*	*General*	*Pitting*
Ni–Al bronze	2.3	1.1	10.0	1.4	0.9*	2.6	0.9	4.9+
Gun metal	7.1	1.6	38.1	3.8	3.4	1.2	6.8	1.4
Superston 40	2.9	1.0	10.0	1.0	3.4	1.0	6.0	–
CA706	17.0	–	34.7	–	6.0	–	11.2	–
CA715	3.3	45.7+	13.5	12.1	1.5	1.8	3.7	3.8
CA719	3.5	0.8	18.2	3.0	0.9	1.3	1.1	1.6
Inconel 625	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Hastelloy C	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Monel 400	0.9	3.6+	3.2	3.6+	0.8	0.9	1.0	10.0+
Monel K-500	1.3	1.6	3.6	2.6	0.7	0.5	2.1	2.2
Mild steel	1.7	2.7	5.1	5.7	1.8	0.7	23.5	23.2+
HY-80	1.9	4.2	5.7	6.4	2.0	1.2	13.0	14.6+
HY-130	1.8	2.8	5.6	17.3	1.8	2.4	21.1	42.1+
Ni-resist 1B	2.1	–	10.0	–	3.3	–	27.6	–
Stainless steel 304	2.2	0.3	6.0	0.2	0.9	1.5	0.5	1.8
Stainless steel 316	1.0	0.0	1.0	1.2	0.4	0.3	0.0	0.1
17-4 PH	29.7	6.5+	100.0	8.8+	0.8	1.1	16.3	1.9+
Carpenter 20 Cb-3	1.0	–	1.0	–	0.0	0.0	0.0	0.4
Nitronic 50	1.0	1.0	1.0	1.0	1.0	**	1.0	0.0
Nitronic 60	1.0	0.5	–	9.7	1.0	0.3	0.4	0.5
Aluminum 5456-H117	3.3	1.9	17.4	2.0	12.7	8.3	254.3	35.0+
Stellite	1.0	1.0	–	–	1.0	–	1.0	–
Ti-40	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Ti–3Al–2.5V	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Ti–6Al–4V	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

*Factors less than 1.0 are the result of variability in behavior and do not represent sacrificial cathodic protection by the Ti-6211.
**No data.

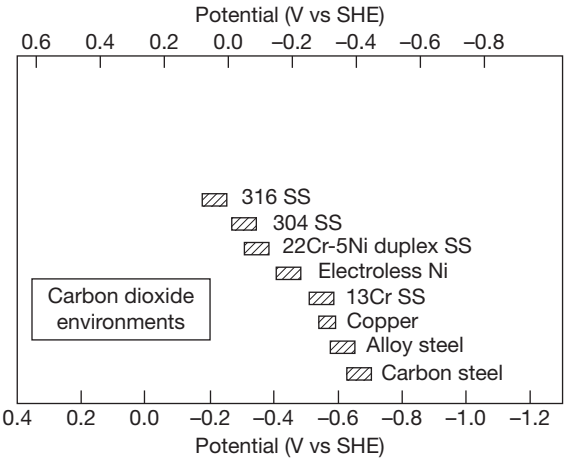


Figure 34 Corrosion potential chart for metals and alloys in carbon dioxide environments giving the relationship of their free corrosion potentials. CO_2 partial pressure = 10–100 kPa, chloride = 0.5–20 g l^{-1}, and temperature = 20–30 °C. Reproduced from Efird, K. D. In *Galvanic Corrosion*; Hack, H. P., Ed.; ASTM International: West Conshohocken, PA, 1988; ASTM STP 978, pp 260–282.

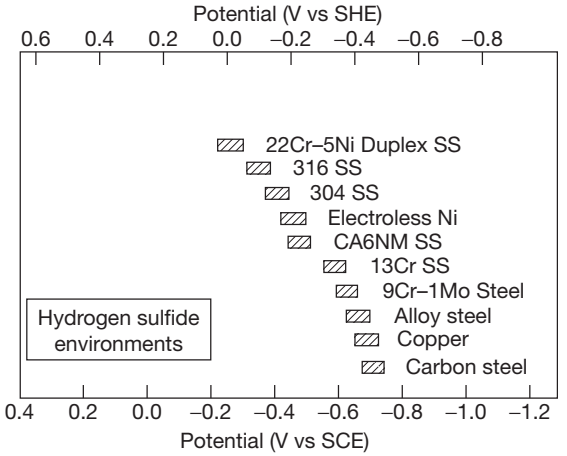


Figure 35 Corrosion potential chart of metals and alloys in hydrogen sulfide environments giving the relationships of their free corrosion potentials. H_2S partial pressure = 1 atm, 5% NaCl + 0.5% CH_3COOH, and temperature = 20–60 °C. Reproduced from Efird, K. D. In *Galvanic Corrosion*; Hack, H. P., Ed.; ASTM International: West Conshohocken, PA, 1988; ASTM STP 978, pp 260–282.

References

1. Standard G15: Terminology Relating to Corrosion and Corrosion Testing *ASTM Book of Standards*; ASTM International: West Conshohocken, PA, 2006.
2. *Standard Terminology for Corrosion, NACE Publications Style Manual*, 2nd ed., NACE International: Houston, TX, 1989.
3. ISO 8044: Corrosion of Metals and Alloys – Vocabulary The International Organization for Standardization: Geneva, 1989.
4. Hack, H. P. *Corrosion Testing Made Easy;* NACE International: Houston, TX, 1993; Volume 2.
5. Standard G82: Guide for Development and Use of a Galvanic Series for Predicting Galvanic Corrosion Performance, *ASTM Book of Standards*; ASTM International: West Conshohocken, PA, 2006.
6. Francis, R. *Galvanic Corrosion: A Practical Guide for Engineers*; NACE: Houston, TX, 2001.
7. Standard G84: Practice for Measuring Time-of-Wetness on Surfaces Exposed to Wetting Conditions as in Atmospheric Corrosion Testing, *ASTM Book of Standards*; ASTM International: West Conshohocken, PA, 2006.
8. Astley, D. J. In *Galvanic Corrosion*; Hack, H. P., Ed.; ASTM International: West Conshohocken, PA, 1988; ASTM STP 978; pp 53–78.
9. Fu, J. W. In *Galvanic Corrosion*; Hack, H. P., Ed.; ASTM International: West Conshohocken, PA, 1988; ASTM STP 978; pp 79–85.
10. Adey, R. A.; Niku, S. M. In *Galvanic Corrosion*; Hack, H. P., Ed.; ASTM International: West Conshohocken, PA, 1988; ASTM STP 978, pp 96–117.
11. Kasper, R. G.; Crowe, C. R. In *Galvanic Corrosion*; Hack, H. P., Ed.; ASTM International: West Conshohocken, PA, 1988; ASTM STP 978, pp 118–135.
12. Scully, J. R.; Hack, H. P. In *Galvanic Corrosion*; Hack, H. P., Ed.; ASTM International: West Conshohocken, PA, 1988; ASTM STP 978, pp 136–157.
13. Hack, H. P.; Janeczko, R. M. Verification of the boundary element modelling technique for cathodic protection of large ship structures, CARDIVNSWC-TR-61–93/02, Naval Surface Warfare Center Carderock Division, Bethesda, MD, December 1993.
14. Fu, J. W. *Corrosion* **1982**, *38*(5), 295.
15. Fu, J. W. A finite element analysis of corrosion cells with mathematical as well as experimental verifications, Scientific Paper 80-IDZ-MEEIC-P1, Westinghouse Corp., Pittsburgh, PA.
16. Fu, J. W.; Chow, J. D. K. *Mater. Perform.* **1982**, *21,* 9.
17. Kasper, R. G.; April, M. G. *Corrosion* **1983**, *39*(5), 181.
18. Strommen, R.; Rodland, A. *Mater. Perform.* **1981**, *20,*15.
19. Munn, R. S. *Mater. Perform.* **1982**, *21,* 29.
20. Decarlo, E. A. In *Corrosion/82*; NACE International: Houston, TX, 1982; Paper no. 165.
21. Doig, P.; Flewitt, P. E. J. *J. Electrochem. Soc.* **1979**, 2057.
22. Clerbois, L. E.; Heintz, F.; Ijselling, F.; Rowlands, J.; Simpson, J. *Br. Corros. J* **1985**, *20*(3), 107–116.
23. Decarlo, E. Computer techniques for offshore corrosion protection, *Sea Technology*; Sep. 1981.
24. Hack, H. P. *Corrosion* **1989**, *45*(7), 601–806.
25. Tighe-Ford, D. J.; McGrath, J. N.; Hodgkiss, L. *Corros. Prevent. Control* **1985**, *32*(5), 89–91.
26. Wagner, C. *J. Electrochem. Soc.* **1951**, *98,*116.
27. Wagner, C.; Traud, W. *Z. Electrochemie Bc-Bunsenges Phys. Chem.* **1938**, 44.
28. Hack, H. P.; Wheatfall, W. L. In *Corrosion/95*; NACE International: Houston, TX, 1995; Paper no. 272.
29. Military Specification – Insulation Tape, Electrical, Plastic Pressure-Sensitive, MIL-I-24391C Amendment 1, U.S. Government Printing Office, August 1991.
30. Shifler, D. A.; Hack, H. P.; Melton, D. In *Corrosion/98*; NACE International: Houston, TX, 1999; Paper no. 706.
31. Shifler, D. A.; Melton, D.; Hack, H. P. Proceedings of the 1997 Tri-Service Conference on Corrosion, Wrightsville Beach, NC, November 1997.

32. Tossey, B. M.; Beavers, J. A. In *Corrosion/2007*; NACE International: Houston, TX, 2007; Paper no. 07244.
33. Garrity, K. Galvanic corrosion inhibiting coupling interposed between two dissimilar pipes, US Patent No. 5,739,424, 1998.
34. Control for galvanic corrosion inhibiting coupling, US Patent Application No. 11/510,394, 2006.
35. Standard G71: Guide for Conducting and Evaluating Galvanic Corrosion Tests in Electrolytes, *ASTM Book of Standards*; ASTM International: West Conshohocken, PA, 2006.
36. Hack, H. P. In *Metals Handbook – Corrosion: Fundamentals, Testing, and Protection;* ASM International: Metals Park, OH, 2003; Vol. 13A.
37. Baboian, R. In *Electrochemical Techniques for Corrosion Engineering*; Baboian, R., Ed.; NACE International: Houston, TX, 1986.
38. Baboian, R. Ed. *Corrosion Tests and Standards – Application and Interpretation*, 2nd ed.; ASTM International: West Conshohocken, PA, 2005; MNL 20.
39. Champion, F. A. *Corrosion Testing Procedures*; Chapman and Hall: London, UK, 1952.
40. Standard G 5: Standard Reference Test Method for Making Potentiostatic and Potentiodynamic Anodic Polarization Measurements; ASTM International: West Conshohocken, PA, 2006.
41. Dean, S., Jr. In *Electrochemical Techniques for Corrosion Engineering*; Baboian, R., Ed.; NACE International: Houston, TX, 1986.
42. Hack, H. P. In *Electrochemical Techniques for Corrosion Engineering*; Baboian, R., Ed.; NACE International: Houston, TX, 1986.
43. Scully, J. R. In *Electrochemical Techniques for Corrosion Engineering*; Baboian, R., Ed.; NACE International: Houston, TX, 1986.
44. Siebert, O. W. In *Electrochemical Techniques for Corrosion Engineering*; Baboian, R., Ed.; NACE International: Houston, TX, 1986.
45. Hack, H. P. In *Corrosion/83*; NACE International: Houston, TX, 1983; Paper no. 73.
46. Hack, H. P.; Scully, J. R. *Corrosion* **1986**, *42*(2), 79–90.
47. Hack, H. P.; Moran, P. J.; Scully, J. R. In *The Measurement and Correction of Electrolyte Resistance in Electrochemical Tests*; Scribner, L. L., Taylor, S. R., Eds.; ASTM International: West Conshohocken, PA, 1990; STP 1056, pp 5–26.
48. Standard G102: Standard Practice for Calculation of Corrosion Rates and Related Information from Electrochemical Measurements; ASTM International: West Conshohocken, PA; Standard G46: Guide for Examination and Evaluation of Pitting Corrosion, *ASTM Book of Standards*; ASTM International, West Conshohocken, PA, 2006.
49. Compton, K. G.; Mendizza, A. In *Symposium on Atmospheric Corrosion of Non-Ferrous Metals*; ASTM International: West Conshohocken, PA, 1955; STP 175, pp 116–125.
50. Doyle, D. P.; Goddard, H. G. *Nature* **1963**, *200*(4912), 1167–1168.
51. Standard G116: Practice for Conducting Wire-on-Bolt Test for Atmospheric Galvanic Corrosion, *ASTM Book of Standards*; ASTM International: West Conshohocken, PA, 2006.
52. Standard G149 (withdrawn 2004): Practice for Conducting the Washer Test for Atmospheric Galvanic Corrosion, *ASTM Book of Standards*; ASTM International: West Conshohocken, PA, 2006.
53. ISO 7441: Corrosion of Metals and Alloys – Determination of bimetallic corrosion in outdoor exposure corrosion tests, The International Organization for Standardization: Geneva, 1984.
54. Standard G50: Conducting Atmospheric Corrosion Tests on Metals, *ASTM Book of Standards*; ASTM International: West Conshohocken, PA, 2006.
55. Hack, H. P. *J. Testing Eval* **1980**, *8*(2), 74–79.
56. Aylor, D. M.; Hays, R. A.; Marshall, L. S. In *Corrosion 2000*; NACE International: Houston, TX, 2000; paper 00640.
57. Hack, H. P. Atlas of polarization diagrams for naval materials in seawater, Report CDNSRDC-TR-61–94/44, Naval Surface Warfare Center Carderock Division, Bethesda, MD, April 1995.
58. Bauer, O.; Vogel, O. *Mitt. MatPrufAmt. Inst. Metallforsch. Berl.* **1918**, *36,*114.
59. Shrier, L. L. Ed. *Corrosion;* Newnes-Butterworth: London, UK, 1976; Vol. 1, Chapter 1.7.
60. Gehring, G. A.; Kyle, R. J. In *Corrosion/82*; NACE International: Houston, TX, 1982; Paper no. 60.
61. Niederberger, R. B.; Gudas, J. P. The effect of seawater velocity on galvanic interactions of piping system alloys, Unpublished data, Naval Surface Warfare Center Carderock Division, Bethesda, MD, February 1979.
62. Caplan, I. L.; Hack, H. P. The galvanic effects of titanium coupled to various marine alloys in seawater, Report CDNSRDC-SME-80/54, Naval Surface Warfare Center Carderock Division, Bethesda, MD, October 1980.
63. *Guide for Engineers*; NACE International: Houston, TX, 2001.
64. Fontana, M. G. *Corrosion Engineering*; McGraw-Hill: New York, NY, 1986.
65. Uhlig, H. H. *Corrosion and Corrosion Control*; Wiley: New York, NY, 1971.
66. *Bimetallic Corrosion*; National Corrosion Service, Department of Industry, Colibri Press: England, 1982; Publication No. 14.

2.08 Environmentally Assisted Cracking Overview

R. C. Newman
Department of Chemical Engineering and Applied Chemistry, University of Toronto, Canada

Abbreviations

CF Corrosion fatigue
CGR Crack growth rate
EAC Environment-assisted cracking
HE Hydrogen embrittlement
HIC Hydrogen-induced cracking
LME Liquid metal embrittlement
SCC Stress corrosion cracking
SICC Strain-induced corrosion cracking
SSCC Sulfide stress corrosion cracking

Symbols

a Crack length
K Stress intensity factor (strictly K_I for crack-opening load)
K_c Critical stress intensity factor
K_{max} Maximum stress intensity factor (in fatigue)
K_{th} Threshold stress intensity for EAC (also called K_{ISCC})
N Number of cycles (in fatigue)
N_f Number of cycles to failure (in fatigue)
R Minimum load divided by maximum load (in fatigue)
ν Crack velocity (= da/dt)
ΔK Stress intensity range (in fatigue)
$\Delta\sigma$ Stress range (in fatigue)
ε Tensile strain
σ Tensile stress
σ_y Yield stress

2.08.1 Introduction

Environment-assisted cracking (EAC) is the gradual growth of cracks under the combined – usually simultaneous – influence of stress and environment. The stress may be applied or internal (residual). The environment may be an aqueous solution or moist atmosphere, a nonaqueous liquid such as ammonia, a gas, a liquid metal, or a solid metal at a high homologous temperature. Whenever the environment is an electrolyte, the EAC process must have a partly electrochemical (i.e., corrosion) character. This may be the dominant factor or a secondary one, depending on the details of the particular system. A simple electrochemical effect is the discharge of hydrogen ions during corrosion of steel in water; the resulting hydrogen atoms may diffuse into the steel and cause embrittlement.

When the stress is more or less static, or varies over a small range, EAC is called as stress corrosion cracking (SCC), hydrogen embrittlement (HE), or liquid metal embrittlement (LME). When a blatantly cyclic stress is present, EAC is usually called corrosion fatigue (CF), hydrogen-assisted fatigue, and so on, although some specialists argue that within this category there are separate phenomena called 'true corrosion fatigue' and 'stress corrosion fatigue' (discussed later). Figure 1, from Andresen, shows the extremes that are easily categorized as 'static' or 'fatigue' loading, and some of the possibilities in between.

Even if we are sure that a cracking phenomenon is caused by hydrogen (as in the SCC of high-strength steels), we still call it SCC if it occurs under natural

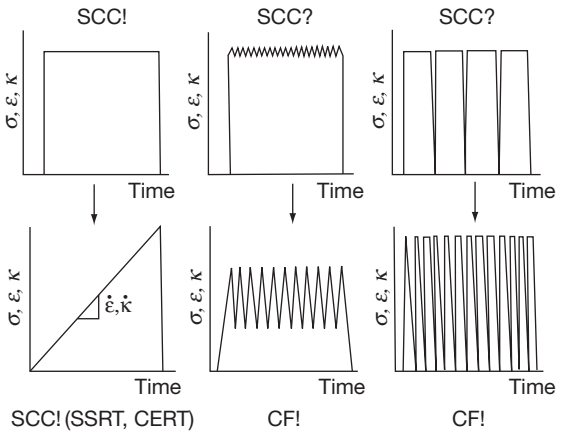


Figure 1 Illustration of the loading patterns that may occur in a component at risk of EAC, and the appropriate nomenclature. Courtesy of P.L. Andresen.

corrosion conditions. Where the distinction becomes blurred is when the steel is coated with a mildly sacrificial metal such as cadmium, and hydrogen may be absorbed through this galvanic action and not by corrosion of the steel substrate. Under full-scale cathodic protection, we would not refer to SCC, but to HE.

In principle, any EAC problem can be mitigated in one of three ways – modify the stress, modify the material, or modify the environment (which would include changing the electrode potential, either by galvanic contact or by imposing a potential). Thus, EAC presents a large number of possible variables, but also a large number of possible countermeasures. A major variable is yield strength – strong alloys are generally more susceptible to EAC, especially where hydrogen absorption is involved, and this limits their application in practice.

Many international conferences on EAC have been held since the 1940s, but a particular landmark was the Ohio State University conference in 1967.[1] This dealt mainly with SCC, but specialist proceedings are also available on CF[2] and LME,[3] recently called 'metal-induced fracture' in view of cracking observed with solid metal contact. A regular conference series on hydrogen effects in materials is held in Jackson Hole, Wyoming and has recently joined with a series pioneered in France called 'Corrosion–Deformation Interactions.'[4] The most recent conference that attempted to cover the entire field of EAC was held in Banff, Canada, in 2004,[5] and was a sequel to a 1987 conference that had a similar ambitious scope.[6] There is a vast journal literature on EAC.

The crack growth rates (CGRs) in EAC vary quite widely, and the lower limit of CGR, if there is one, has probably not been observed yet. At some point, especially at elevated temperatures, the CGR will become equal to a creep crack growth rate, at which point it may not be appropriate to talk about EAC. Creep, though, can also be affected by the environment. There is a whole area of research called creep–fatigue–oxidation interaction, and a related phenomenon that occurs under static loading, called oxygen embrittlement.[7] These phenomena are usually studied in the context of gas turbine materials, but there may be a link with SCC in hot water or steam systems.

For SCC, the CGRs tend to range from about 10^{-12} m s^{-1} (30 μm year^{-1}) to 10^{-7} m s^{-1} (9 mm day^{-1}). In the laboratory, faster rates are possible. From an industrial perspective, the lower range of rates is manageable by appropriate inspection; the highest rates are not, but they occur in certain well-known high-risk situations, such as contact of austenitic stainless steel with strong, hot chloride solutions, or contact of a susceptible alloy with mercury. Usually something has gone badly wrong when such rates occur. The fastest form of EAC is LME, which can appear almost instantaneously in certain combinations such as mercury (or liquid gallium) and a high-strength aluminum alloy.

2.08.2 Appearance of EAC

There are three common types of crack morphology in EAC: intergranular, transgranular (cleavage-like), and transgranular (fatigue-like) (**Figure 2**). When a CF surface shows 'brittle striations,' there is not much distinction from a cleavage-like SCC fracture surface, which might be obtained in the same environment by monotonic dynamic straining. There are also some cases of HE (in lower strength steels and some aluminum alloys) in which the fracture surface is dimpled, but the dimples are smaller than those that occur in a normal ductile fracture performed in air.

Classical cases of SCC (brass in aqueous ammonia; carbon steel in hot caustic or nitrate solution) were usually intergranular, but later the transgranular form became more common as new materials and technologies were developed, chiefly process industries that exposed austenitic stainless steels to the risk of SCC due to hot aqueous chloride. CF is usually transgranular, but can be intergranular at low cyclic stress (intensity) amplitudes, or in alloys with sensitive grain boundaries, such as sensitized stainless steel or

high-strength aluminum alloys. LME is usually intergranular in practice, but occurs in monocrystals of many metals (Al, Ni, etc.) in the laboratory. Hydrogen affects both intergranular and transgranular fracture. A recent review by Lynch discusses many fractographic details seen in EAC, and proposes that many EAC phenomena might be unified under an adsorption-induced plastic microfracture type of mechanism.[8]

The phenomenon of hydrogen-induced cracking of steel (HIC) occurs in pipes carrying sour oil or gas. It is not driven by preexisting external or internal stress, but by the pressure of hydrogen that is discharged as part of a corrosion reaction, enters the metal at a huge effective activity in accordance with the Nernst equation, then recombines to form high-pressure gas at defects inside the metal. It is only barely to be considered a case of EAC, and is not discussed further in this article. Importantly, though, when we increase the strength of steel used to make such pipes, we risk a transition from HIC to SSCC (sulfide stress corrosion cracking) which is a true SCC phenomenon, also caused by hydrogen, but under the influence of the stress in the pipe (**Figure 3**).

One difference that usually helps to distinguish SCC and CF is the multiplicity of cracks. CF, like ordinary fatigue, usually shows few and relatively flat cracks, whereas SCC often shows complex branched and intertwined cracks (**Figure 4**). An important factor in creating such crack patterns is the amount of local driving force (loosely, the local stress intensity factor K) that is needed to propagate the crack. If this is low, then the presence of many parallel or overlapping cracks (which reduces the Kvalue at each crack tip) is not a hindrance to the propagation of each one.

Cracking of welded or cast components cannot always be classified as 'intergranular' or 'transgranular,' and the morphologies are often complex.

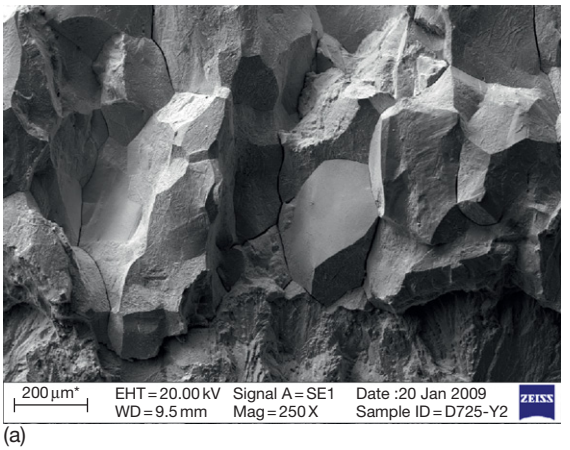


Figure 2 Fracture appearance in EAC: (a) Mainly intergranular SCC of Alloy 600, courtesy Bob Cottis, School of Materials, University of Manchester; (b) Transgranular cleavage-like SCC of Alloy 800 in 50% NaOH at 280 °C; (c) Fatigue of a nickel monocrystal showing successive regions of fracture in mercury (brittle striations), hydrogen (brittle striations), and argon (ductile striations). Reproduced from Lynch, S. P. *Acta Metall.* **1988**, *36*, 2639–2661.

2.08.3 Stress and Strain in EAC

Strong alloys are used because they need to resist stresses in service, so for these materials the main source of stress is often an externally applied load, such as

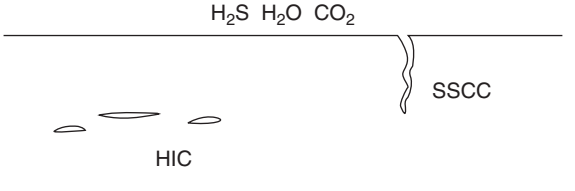


Figure 3 Distinction between HIC (hydrogen-induced cracking) and SSCC (sulfide stress corrosion cracking) as made in the oil and gas industry. HIC is not caused by the stresses present in the pipe or vessel, but by the pressure of internal hydrogen gas.

Figure 4 Illustration of the meandering path taken by typical stress corrosion cracks, indicative of a low local driving force required for crack advance (transgranular SCC of AISI type 316 stainless steel in 4 m NaOH at 280 °C).

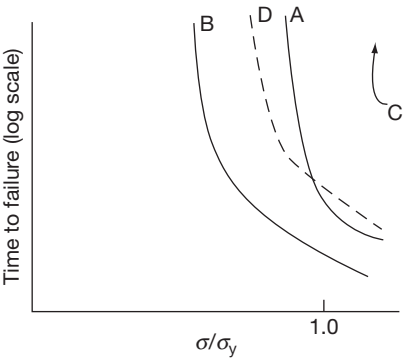


Figure 5 Schematic behavior of carbon steel pipeline material subject to external SCC with and without low-amplitude cyclic loading. In high-pH carbonate-bicarbonate, the time to failure decreases from line A to line B upon cycling. In near-neutral pH bicarbonate-CO_2, there is scarcely any cracking without cycling (line C) but cracking occurs with cycling (line D).

internal pressure in a pipe. Ductile alloys such as austenitic stainless steel are not much used for major load-bearing structure, so they usually crack under the influence of internal residual stresses (e.g., at welds). Cyclic stresses, if present, can arise from rotation, vibration, thermal fluctuation, intermittent operation, and so on. Superimposition of static and cyclic stresses is a major topic in fatigue, and thus also in corrosion fatigue.

A distinction between EAC and ordinary fatigue is that hardening is generally detrimental for EAC, whereas it is one of the ways to improve fatigue resistance.

Low-frequency, low-amplitude cycling can cause major enhancement of SCC under conditions where it would not be appropriate to refer to fatigue – such as one cycle per startup of a nuclear reactor, or one cycle per spring thaw for a pipeline crossing a riverbed. This reflects a general truth about SCC, which is that for the most part it is a plastic-strain-driven phenomenon (at least that is the prevailing opinion – this may change). If the load is completely static, cracking is less likely because any transient creep strain becomes exhausted, and there is nothing to drive processes such as oxide film rupture. R.N. Parkins is credited for much of the modern understanding of this issue, and for the idea that there exists a continuum of EAC from SCC through intermediate cases to true CF.[9] **Figure 5** illustrates, schematically, the effect of low-amplitude cycling on the two kinds of external pipeline cracking studied by Parkins – 'near-neutral pH' and 'carbonate–bicarbonate.'

EAC of pressurized or otherwise actively loaded equipment made of high-strength materials causes particular concern because it can lead to catastrophic failure. A crack initiates, grows at an increasing rate, then fast fracture occurs when K (or K_{max} in the case of CF) reaches K_c. On the other hand, in something like a stainless steel pipe carrying a corrodent under low pressure, where there is no meaningful K_c, the outcome of EAC is usually a leak. This can be serious, but it is usually more manageable than a fast rupture. Naturally, there are cases that fall somewhere in between, and here there may be uncertainty as to the structural outcome of an EAC incident; modeling and simulation can provide some of the answers, but some kind of full-scale component test may be required.

When fracture mechanics methods are applied to SCC, HE, or LME, a flat or 'plateau' region is observed at intermediate K values – **Figure 6**. (The author's Ph.D. advisor, G.T. Burstein, used to insist that such features should be called 'ledges,' not 'plateaux.') This shows that a limiting condition has been reached where something other than the mechanics controls the crack growth rate. This 'chemical' rate control is generally due to a rate-limiting dissolution, diffusion, or adsorption process. Also, the critical stress intensity factor K_c may be reduced, in effect, by hydrogen-producing environments; the implications of this are both serious and poorly understood. ('In effect,' because K is a static quantity and the reduction in apparent K_c value may be due to dynamics associated with the overload of the crack – that is,

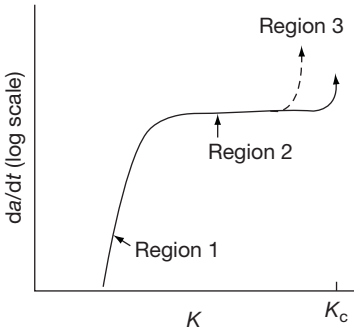


Figure 6 Schematic crack velocity versus stress intensity curve for EAC (without cyclic loading) showing the 'plateau' region at intermediate K values. Such curves are often drawn as though there is no reduction in K_c due to the environment, but in reality such reductions (dashed line) can occur and are troublesome in certain safety-related prediction procedures.

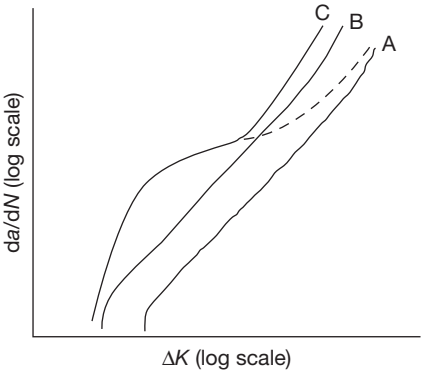


Figure 7 Fatigue crack growth curves showing some different possible effects of an environment. Line A is for an inert environment. Then, the distinction is made between 'true corrosion fatigue' (line B) and 'stress corrosion fatigue' (line C; alternatively the dashed line), which shows a plateau-like feature at intermediate ΔK values. Most corrosion engineers do not make this distinction, and most corrosion fatigue is of type C.

there is a strain rate sensitivity of the apparent reduction in K_c.)

Stress in a real component is always multiaxial to some extent, but most SCC behavior is referred to the principal tensile stress, whether residual or applied. The stress *state* is important, though, in hydrogen-affected EAC. This is because three-dimensional (3D) lattice dilation can increase the solubility of hydrogen in the metal and thus, potentially, the susceptibility to cracking. A number of authors have shown easier cracking in mode I (tension) loading than in mode III (torsion) in certain systems believed to involve hydrogen entry into the metal.[10]

In components subject to fatigue, the mean stress is rarely zero. Certain effects of the R parameter (minimum load/maximum load) are known in ordinary fatigue, but in CF these are magnified. To some extent, this can be dismissed as the superimposition of SCC (due to the mean load) on CF (due to the cyclic load), but as CF occurs in environments where SCC does not, it is an important consideration. Figure 7 indicates the distinction between 'true corrosion fatigue' and 'stress corrosion fatigue' that is made by some authors. The hump or plateau feature at intermediate ΔK values is due to SCC-like crack growth, and thus becomes more noticeable with increasing mean stress. Cyclic loading frequency is a key variable affecting crack growth in CF, in contrast to ordinary fatigue, because the corrosion process (this also applies to hydrogen-assisted fatigue) needs time to do its work on each cycle, and if the frequency is too high, the cracking behavior may revert to that seen in an inert environment. It is easy to see that all this is amenable to a mathematical modeling approach based on superposition of a time-dependent 'monotonic' or 'SCC'-like element of crack advance and a 'cyclic' contribution.

One type of SCC has its own special name. SICC (strain-induced corrosion cracking)[11] occurs in steel exposed to hot water when sufficiently large strains arise to fracture the magnetite film on the surface. This is not much different from any other case of SCC where dynamic strain is mandatory, and, indeed, CF occurs quite readily under such conditions, at least in the laboratory.

2.08.4 Test Methods for EAC

An excellent manual on SCC testing is available, written by Sedriks.[12] Testing for CF is no different from ordinary fatigue testing, except that the environment has to be contained around the sample – this of course presents a challenge when the environment is a gas or a hot, pressurized aqueous solution. Electrochemistry adds another level of complexity, especially in an autoclave.

Regarding specimen design, all EAC testing can be classified as (a) smooth-specimen testing, or (b) defected (e.g., precracked) specimen testing, and for SCC these can be further subdivided into selfloaded (e.g., bolt-loaded) and machine-loaded designs (fatigue, unless purely thermal, generally requires a testing machine). The advantage of using a machine, despite the extra complexity, is that the loading can be active; this is especially important in view of the proven role of low-level dynamic strains in SCC. Also, one can use a

test machine to do a slow strain rate test – an important accelerated method for mapping out SCC susceptibility rapidly and economically.

For SCC, HE, LME, etc. one can use precracked specimen testing to obtain the CGR, or crack velocity (v, or da/dt) as a function of the stress intensity factor K as shown in **Figure 6** – although only up to a point for ductile alloys, which undergo large-scale yielding that puts a limit on K values that can be validly quoted. This does not affect the study of crack growth at relatively low K values near the threshold (K_{th} or K_{ISCC}) below which crack growth stops altogether (it should be noted that not all authors agree with the existence of a true K_{th} for SCC). In some types of SCC, copious multibranched cracking makes K a questionable parameter, but this can be handled to some extent by machining side grooves on the specimen to keep the crack straight. The determination of da/dt is a big subject in itself – the most sophisticated laboratories generally use sensitive and stable electrical resistance (DC potential drop) methods.

Similarly, we can use precracked specimens to obtain da/dN versus ΔK curves in an environment. This is one of the most difficult of all corrosion experiments, especially if done in an autoclave.

For the most difficult types of EAC – such as the SCC of annealed stainless steels in reducing high temperature water, or the SCC of pipeline steels in near-neutral pH solutions, or HE of ordinary carbon steels in general, the occurrence of truly stable crack growth at constant K is questionable. In that case, some dynamic loading, such as fatigue loading or periodic unloading, is mandatory for stable cracking to occur. Clearly, this might represent a nomenclature problem – but generally we prefer to stay with 'SCC' (see **Figure 1**) so long as the mean stress dominates the cracking behavior. Slow strain rate testing especially exaggerates the propensity of such systems to show SCC in practice. If we believe the results of slow strain rate tests, we would never cathodically protect steel.

Data presentation is straightforward for precracked specimen testing: da/dt (crack velocity) versus K, or da/dN versus ΔK. For smooth-specimen testing, there is a multitude of types of data that can be gleaned from a test – for fatigue, the presentation is generally in the form of $\Delta\sigma$ (cyclic stress range) versus N_f (number of cycles to failure), and the effect of an environment is often to remove the 'fatigue limit' at low $\Delta\sigma$ – **Figure 8** (incidentally, such testing is normally done at quite a high frequency, so crack growth may not be much affected by corrosion – the main effect is to assist crack initiation). For SCC,

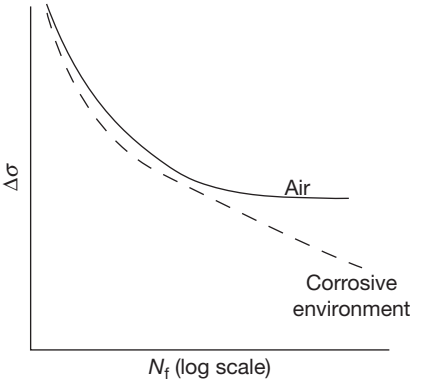


Figure 8 Disappearance of the fatigue limit in corrosion fatigue tests conducted on smooth specimens.

smooth self-loaded samples can be examined for cracks after a fixed period of time, or examined periodically and reimmersed each time (this seems dubious, but has been done, especially for long tests in hot, pressurized water), or monitored continuously by optical microscopy or acoustic emission. Multiple samples may be tested at the same condition. Ideally, the output should be a plot of time to failure against stress. Machine-loaded samples, and especially slow strain rate tests, enable a better distinction of crack initiation and crack growth – for slow strain rate, one can obtain a variety of quantities as discussed by Sedriks – strain to failure (normalized to air), reduction in area at fracture, maximum stress, and so on. Low-amplitude cyclic strains (not enough to be called fatigue) often enhance SCC susceptibility, as discussed earlier, and can be applied automatically with some modern slow strain rate machines.

2.08.5 Mechanisms of the Environmental Enhancement in EAC

A detailed discussion of crack-tip mechanisms is deferred to the SCC chapter. Broadly, the proposed mechanisms fall into several categories, which still excite controversy.

- Decohesion or enhanced plastic rupture caused by hydrogen or liquid metal adsorption or absorption.
- Slip-dissolution or slip-oxidation – that is, crack growth by a highly localized corrosion process, exacerbated by local plastic deformation.
- Film-induced substrate fracture (film-induced cleavage).
- Creep-like mechanisms, for example, 'surface mobility,' vacancy injection.

Many reviews are available – including the one by the author.[13]

In CF the possibilities are expanded, and simple anodic dissolution (without any apparent surface film) or simple passive film formation can cause an enhancement in crack initiation and da/dN. Generally CF can be seen with more subtle environmental effects than SCC.

Not much discussion has occurred lately on mechanisms of LME, but there was an intriguing study showing that in one well-known system (Ga on Al) first-principles quantum calculations could show that surface alloying (intermixing of Ga and Al in the first atomic layer) should occur.[14] This might be a general mechanism of, or rather first step in, LME. In any event, whatever the interaction is, it is extremely shallow, yet causes the fastest of all EAC phenomena. Quantum calculations are also giving insight into hydrogen effects in metals.

2.08.5.1 Occurrence of EAC in Pure Metals

CF and LME are well known to occur in pure metals, but SCC of pure metals was controversial until recently. Extremely pure iron does not crack in nitrate or caustic (and presumably other intergranular cracking environments), and intergranular SCC of pure copper could be dismissed as some kind of grain-boundary segregation effect, but in fact it should have been obvious from the earliest observations of transgranular SCC in carbon steel that pure iron would also crack transgranularly in the same (peculiar) environments – CO–CO_2–H_2O, and anhydrous liquid ammonia. The issue was laid to rest by the observation of transgranular SCC in pure copper, which has occurred now and again in practice. This is discussed further in the **Chapter 2.09, Stress Corrosion Cracking**.

2.08.6 Emerging Environments and Future Outlook

Until the advent of the slow strain rate test in the 1970s, most EAC phenomena were discovered in the field, not in the lab. Even now we can be caught out by a new environmental effect. So, as technologies develop, we need to anticipate possible EAC problems. For example, the prospects for supercritical water-cooled nuclear power plants are critically dependent on the development of new materials that are resistant to EAC in such extreme environments (and with neutron irradiation, radiolysis, etc.). Other new energy technologies such as thermochemical–electrochemical hydrogen production have unusual environments that could cause unexpected EAC problems. Finally, even the 'pure water' problem is not solved yet, because next-generation nuclear plants will have to operate for 60–100 years, and late-onset EAC is possible in the presently used materials.

We need to develop multiscale understanding of EAC, from the quantum level to the component level, before we can truly have a predictive capability. We know already that EAC can be *controlled* at all these scales – one can add some element that interacts chemically with hydrogen and mitigates its embrittlement effect, or one can induce a surface residual stress in a whole component that prevents crack initiation. In between, one can control intergranular EAC at the granular level, by inducing the formation of a high fraction of geometrically special, nonsusceptible grain boundaries – so-called grain boundary engineering. So, despite the formidable challenges presented by EAC, we have a wide range of countermeasures at many length scales.

References

1. Staehle, R. W.; Forty, A. J.; van Rooyen, D., Eds. *Fundamental Aspects of Stress-Corrosion Cracking*; NACE: Houston, TX, 1969.
2. Crooker, T. W.; Leis, B. N. *Corrosion Fatigue – Mechanics, Metallurgy, Electrochemistry and Engineering*, ASTM STP 801; ASTM: Philadelphia, 1983.
3. Kamdar, M. H., Ed. *Embrittlement by Liquid and Solid Metals*; TMS-AIME: Warrendale, PA, 1984.
4. Thompson, A. W.; Moody, N. R.; Ricker, R. E.; Was, G. S.; Jones, R. H., Eds. *Hydrogen Effects on Material Behavior and Corrosion-Deformation Interactions*; TMS-AIME: Warrendale, PA, 2003.
5. Shipilov, S. A.; Jones, R. H.; Olive, J.-M.; Rebak, R., Eds. *Environment-Induced Cracking of Materials*; Elsevier, 2008.
6. Gangloff, R. P.; Ives, M. B., Eds. *Environmental-Induced Cracking of Metals*; NACE: Houston, TX, 1990.
7. Woodford, D. A. *Energy Mater.* **2006**, *1*(1), 59–79.
8. Lynch, S. P. In Corrosion/2007; NACE: Houston, TX, 2007; Paper no. 489.
9. Parkins, R. N.; Belhimer, E.; Blanchard, W. K., Jr. *Corrosion* **1993**, *49*, 951–966.
10. Pickens, J. R.; Gordon, J. R.; Green, J. A. S. *Metall. Trans A.* **1983**, *14A*, 925–930.
11. Seifert, H. P.; Ritter, S. *J. Nucl. Mater.* **2008**, *378*, 312–326.
12. Sedriks, A. J. *Stress Corrosion Cracking Test Methods*; NACE: Houston, TX, 1989.
13. Newman, R. C. In *Stress Corrosion Cracking Mechanisms. Corrosion Mechanisms in Theory and Practice*, 2nd ed.; Marcus, P., Oudar, J., Eds. Marcel Dekker: New York, 2002; pp 399–450.
14. Stumpf, R.; Feibelman, P. J. *Phys. Rev. B* **1996**, *54*, 5145–5150.

2.10 Hydrogen Embrittlement

R. A. Cottis
Corrosion and Protection Centre, School of Materials, University of Manchester, Manchester M60 1QD, UK

Abbreviations

AIDE Adsorption-induced dislocation emission
bcc Body centered cubic
CTOD Crack tip opening displacement
fcc Face centered cubic
hcp Hexagonal close packed
HELP Hydrogen-induced localized plasticity
HEDE Hydrogen-enhanced decohesion
HIC Hydrogen-induced cracking
HISC Hydrogen-induced stress cracking
SCC Stress corrosion cracking
SCE Saturated calomel electrode
SOHIC Stress-oriented hydrogen-induced cracking
SSC Sulfide stress cracking

Symbols

a Crack length
K Stress intensity factor
K_{Ic} Critical stress intensity factor for fast fracture
K_{Iscc} Critical stress intensity factor for SCC
r Distance ahead of crack tip
σ_{th} Threshold stress for SCC

2.10.1 Introduction

There are many applications in which the yield strength of the material of construction is a limiting factor in the design. For example, a large proportion of the weight of an offshore oil production platform consists of the structure required to resist wave forces, with a relatively small proportion being the 'topside' equipment that actually does the work. If the strength of the material of the structure could be increased, this would allow the structure itself to be made lighter, with the consequence that the weight of the topside equipment could be increased. Thus, materials offering high strength-to-weight or strength-to-volume ratio are in high demand in many fields of engineering. Steels can be produced with yield strengths above 2000 MPa, and at first sight, one might expect these to have widespread application. Unfortunately, things are not that simple, and as the yield strength of materials is increased, other mechanical properties tend to decrease. In particular, the materials become more susceptible to brittle fracture, especially when assisted by environmental factors, such as hydrogen embrittlement, stress corrosion cracking, or corrosion fatigue. The related problem of hydrogen-induced blister formation or hydrogen-induced cracking tends to be more of a problem in lower-strength steels used for crude oil transmission pipelines.

It is now well established that high-strength steels are susceptible to embrittlement by dissolved hydrogen, and the majority of stress corrosion cracking (SCC) failures of these materials are attributed to hydrogen embrittlement. This should not be taken to imply that other mechanisms of SCC do not occur. However, the ease with which hydrogen can be picked up from aqueous environments is such that other SCC processes are rarely a practical problem for high-strength steels.

This document is based on a Chapter on SCC of High Strength Steels that was produced for the 3rd Edition in about 1989. It has been revised for the 4th Edition to incorporate some updated references, particularly in respect of mechanistic aspects of hydrogen embrittlement, and to provide some information on hydrogen embrittlement of materials other than carbon and low-alloy steels. Hydrogen embrittlement continues to be a major area of research, with nearly 9000 publications in the period from 1980 to Spring 2009. A number of reviews of the aspects of hydrogen embrittlement have been published in the last 25 years[1–3] to which the reader is referred for further information.

2.10.2 Terminology

As with several aspects of corrosion technology, the terminology of the effects of hydrogen on metals is somewhat confused. We can indentify two general classes of hydrogen effect:

- quasi-brittle fracture of higher strength materials that can occur with relatively low concentrations of hydrogen; we term this hydrogen embrittlement, although some processes that almost certainly occur by this mechanism are also termed SCC, such as the SCC of high-strength aluminum alloys in water or water vapor; and
- surface blistering and internal cracking at sulfide and other inclusions in lower-strength materials (primarily carbon steels) due to very high internal hydrogen fugacities, leading to hydrogen pressure-induced cracking, commonly termed hydrogen-induced cracking (HIC) (it is unfortunate that the term 'hydrogen-induced cracking' has also been used as a generic term for any cracking due to hydrogen); processes in this class are discussed separately in the subsequent **Chapter 2.11, Cracking Stimulated by Hydrogen**.

However, as is often the case in corrosion, things are not really that simple, and there is a spectrum of behaviors between these two extremes, including processes such as stress-oriented hydrogen-induced cracking (SOHIC), sulfide stress cracking (SSC) of pipeline steels, and hydrogen-induced stress cracking (HISC) of duplex stainless steels. Just to complicate the terminology further, there is the phenomenon of hydrogen cracking that is experienced at higher temperatures as a result of internal methane production by reaction between dissolved hydrogen and carbon.

2.10.3 Entry of Hydrogen into Metals

Hydrogen can enter metals, either from the gas phase, or by way of the electrochemical reduction of hydrogen-containing species, from the aqueous phase. This section is largely based on work on carbon and low-alloy steels.

2.10.3.1 Entry from the Gas Phase

Several models have been proposed for the entry of hydrogen from the gas phase, and the details of the process remain somewhat uncertain. In general terms however, the reactions involved are the adsorption of

molecular hydrogen (other gases such as hydrogen sulfide can also act as a source of hydrogen, but they will typically require water to facilitate the reactions involved and are arguably aqueous phase reactions), the dissociation of the hydrogen molecule to produce hydrogen atoms adsorbed onto the surface, and the subsequent diffusion of the adsorbed hydrogen atoms into the metal lattice:

$$H_2 \rightarrow 2H_{ads}$$

$$H_{ads} \rightleftharpoons H_{metal}$$

For perfectly clean and film-free iron surfaces, the rate-controlling step in this process may be the dissociation step.[4] This will be proportional to the partial pressure of hydrogen, and consequently so will the rate of hydrogen entry into the metal. In more realistic situations, surface films such as passive or air-formed oxide films, or the presence of gases, such as oxygen, which adsorb competitively with hydrogen, may play an important role.

2.10.3.2 Entry from the Aqueous Phase

The mechanism of electrochemical production of hydrogen on steel in aqueous solution has received much attention. It is accepted that the reaction occurs in two main stages. The first of these is the initial charge transfer step to produce an adsorbed hydrogen atom, known as the Volmer reaction. In acid solution, this involves the reduction of a hydrogen ion.

$$H_3O^+ + e^- \rightarrow H_{ads} + H_2O$$

In neutral and alkaline solution, where the concentration of hydrogen ions is very low, the reaction switches to the reduction of water molecules:

$$H_2O + e^- \rightarrow H_{ads} + OH^-$$

The second stage of the reaction to produce molecular hydrogen may occur through either of two mechanisms. In the first of these, known as chemical desorption, chemical recombination, or the Tafel reaction, two adsorbed hydrogen atoms combine to produce a hydrogen molecule:

$$H_{ads} + H_{ads} \rightarrow H_2$$

Alternatively, the adsorbed hydrogen atom may participate in a second electrochemical reaction, known as electrochemical desorption, or the Heyrovsky reaction:

$$H_{ads} + H_3O^+ + e^- \rightarrow H_2 + H_2O(\text{acid})$$

$$H_{ads} + H_2O + e^- \rightarrow H_2 + OH^-(\text{neutral or alkaline})$$

For iron, it is reasonably well established that the reaction goes by way of chemical recombination under most circumstances, although there is some evidence that electrochemical desorption may take over in very alkaline solutions or at large overpotentials.

A third reaction, which goes in parallel with the desorption reaction, is the entry of atomic hydrogen into the steel from the surface adsorbed state:

$$H_{ads} \rightarrow H_{metal}$$

In most circumstances, the kinetics of this reaction are controlled by the rate at which the hydrogen can diffuse into the underlying steel, and this reaction is essentially in equilibrium. Consequently, it is difficult to study the kinetics of this reaction. A particular situation in which this may be very important relates to the conditions at crack tips, where the hydrogen may be transported into the bulk by dislocation motion, giving rise to very high rates of hydrogen entry.

As the hydrogen entry reaction is generally in equilibrium, the hydrogen concentration just below the entry surface is directly related to the surface concentration or coverage. Consequently, the rate of entry of hydrogen into the steel is controlled by the balance between the first and second stages of the hydrogen evolution reaction, since these control the coverage of adsorbed hydrogen. In the case of chemical recombination as the second stage in the reaction, the rate of hydrogen evolution is proportional to $(H_{ads})^2$ (since two adsorbed hydrogen atoms are involved in the reaction), while the rate of hydrogen entry (or the equilibrium concentration just below the surface, which is usually the controlling factor) is proportional to (H_{ads}). Thus, as the cathodic current increases so the subsurface hydrogen concentration (and hence, the rate of hydrogen diffusion through a membrane) increases as the square root of the cathodic current (assuming that the current due to hydrogen entering the steel is small compared to that due to hydrogen evolution, as is usually the case).

However, when the second stage in the hydrogen evolution reaction is electrochemical desorption, the rate of this reaction is increased as the potential becomes more negative, and the adsorbed hydrogen concentration may remain constant or fall, according to the detailed electrochemistry. This results in curves such as that shown in **Figure 1** for steel in seawater.

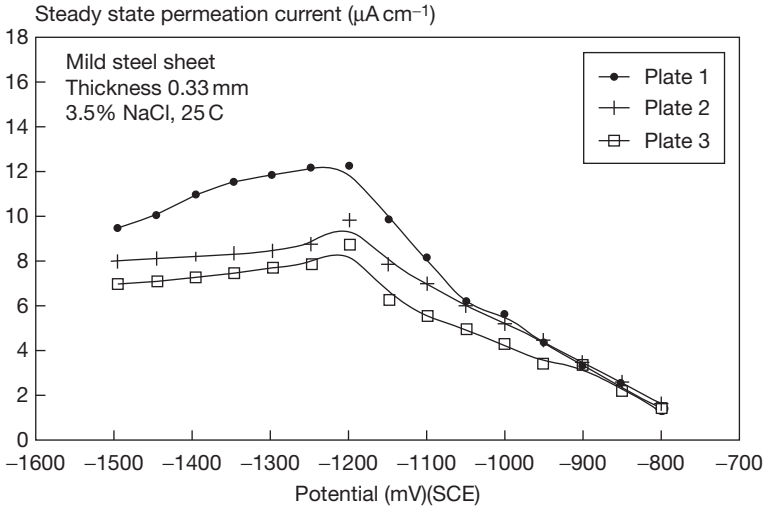


Figure 1 Hydrogen permeation current as a function of applied potential, showing effect of change in reaction mechanism below −1200 mV (SCE). Reproduced from Obuzor, U. W. PhD Thesis, UMIST, Manchester 1989.

Whether the adsorbed hydrogen is produced from the gas phase or from aqueous solution, it appears that the presence of hydrogen atoms distorts the crystal structure of the metal surface[5] and this results in a surface solubility that is higher than that of the bulk. The depth of this distortion is not clear, but it seems possible that this distorted zone may play an important part in initiating brittle fracture processes.

2.10.4 Location of Hydrogen in Steel

Hydrogen exists in metals in the monatomic form, and is commonly described as atomic hydrogen. In practice, the state of charge of the hydrogen atom is not known with any certainty, but it seems probable that it tends to acquire a slight negative charge by attracting electrons from the valence orbitals of the metal lattice. It has been suggested that this results in a weakening of the metal–metal bond which is responsible for hydrogen embrittlement.[6]

Hydrogen has a very low solubility in the iron lattice, which makes direct observation of the location of the hydrogen atom in the lattice very difficult. The hydrogen definitely occupies an interstitial site in the *bcc* iron lattice. Two such sites are normally associated with interstitial solutes in bcc structures, the tetrahedral and the octahedral sites (see **Figure 2**). Indirect evidence[7] suggests that hydrogen occupies the tetrahedral site.

In addition to interstitial sites in the lattice, hydrogen atoms are also strongly attracted to defect sites in the metal, and these are referred to as 'traps.' Trap sites include vacancies, solute atoms, dislocations, grain boundaries, voids, and nonmetallic inclusions. Of the various trap types in iron and steel, vacancies are relatively unimportant at ambient temperatures, simply because of their low concentration. Somewhat surprisingly, it appears that grain boundaries in pure iron also trap very little hydrogen, although segregation of carbon and other impurity atoms to the grain boundary can increase the tendency for trapping.[8] The more important trap sites in iron and steel appear to be phase boundaries, dislocations, voids, and inclusions.[9] Trap types may be classified in various ways, two of the more important being related to the number of hydrogen atoms which can be

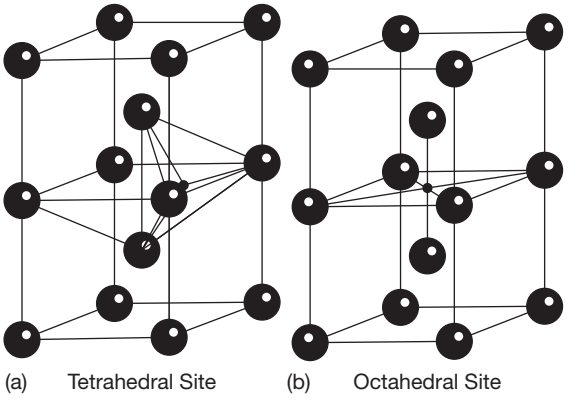


Figure 2 Octahedral and tetrahedral sites in the *bcc* lattice.

accommodated in the trap and to the binding energy of the trap. These are summarized in Table 1.

It should be noted that the terms saturable and reversible are rather loosely defined in relation to hydrogen trapping. Thus, some workers take the view that reversible and nonsaturable traps are the same (since the equilibrium hydrogen content of the trap will vary with the fugacity). Similarly, the term reversible is given a slightly different meaning from that in electrochemical reaction kinetics, where it relates to the activation energy of a reaction, rather than the overall free energy change.

There are many ways in which trapping can be studied, but the wide range of trap types and geometries make it difficult to determine the properties of specific trap sites.

Hydrogen trapped in voids consists of adsorbed hydrogen on the walls of the void, together with molecular hydrogen in the void itself. With high fugacities of hydrogen in the steel, such as can be developed in steels in contact with acidic solutions containing hydrogen sulfide, very high pressures may be developed in the void. When combined with the hydrogen embrittlement of the steel around the void, this can lead to the growth of cracks around the void. Such cracks typically develop around nonmetallic inclusions which have been flattened by rolling, giving characteristic blisters lying parallel to the rolling direction. This phenomenon is known as hydrogen-induced cracking.

2.10.5 Transport of Hydrogen in Iron and Steel

As a small interstitial atom, hydrogen diffuses rapidly in iron, the diffusion rate being of a similar order to that of solutes in aqueous solution.

Table 1 Classification of trap types

Classification	*Description*
Saturable	The number of sites for hydrogen atoms is fixed (e.g., grain boundaries, dislocations)
Nonsaturable	The number of sites for hydrogen atoms in the trap varies according to the fugacity (e.g., voids)
Reversible	The trap binding energy is relatively small, and hydrogen may escape from the trap as well as enter it
Irreversible	The trap binding energy is large, and hydrogen will not leave the trap at ambient temperature

The study of the transport of hydrogen in steel is commonly undertaken by hydrogen permeation measurements. This involves the permeation of hydrogen through a thin steel membrane (typically less than 1 mm thick). Hydrogen entry may be from the gas phase or from solution, while the flux of hydrogen through the membrane may be determined either by vacuum extraction of the gas to a suitable detector, or, somewhat more simply, by electrochemical oxidation of the hydrogen to hydrogen ions. The latter method forms the basis of the electrochemical hydrogen permeation cell developed originally by Devanathan and Stackurski[10] and illustrated in Figure 3. This has subsequently been developed into a monitoring technique by Berman *et al.*,[11] while Arup[12] has developed a small self-contained sensor using battery technology, shown in Figure 4.

In permeation measurements, the first signs of hydrogen diffusing through 1-mm steel membranes can be observed in a few minutes. The practical measurement of diffusion parameters tends to be rather unreproducible, because of the role of the various forms of trap, which tend to increase the solubility of hydrogen in the steel and thereby decrease the apparent diffusion coefficient. For pure iron at temperatures above $\sim$350 K, the activation energy for diffusion is $\sim$7.6 kJ mol^{-1}, with D_o being in the region of 1–2.5 $\times$ 10^{-7} m^2 s^{-1}. At ambient temperatures, the activation energy ranges from $\sim$4 to 7 kJ mol^{-1}, with D_o ranging from $\sim$0.5 to 1.2 $\times$ 10^{-7} m^2 s^{-1}. Diffusion coefficients at 298 K are thus $\sim$7 $\times$ 10^{-9} m^2 s^{-1}.

The above data relate to very pure iron samples with low dislocation densities. In real steels, the trapping effects result in much lower apparent diffusivities, which are dependent on the metallurgical state of the steel as well as its chemical composition. Typical values for the apparent diffusion coefficient of hydrogen in high strength alloy steel at room temperature are in the region of 10^{-11} m^2 s^{-1}.

In the context of hydrogen embrittlement, an important aspect of transport processes is the influence of plastic strain on the behavior of hydrogen in a metal. Brass and Chene[13] reviewed these effects, and concluded:

"The nature of hydrogen-deformation interactions may be complex and opposite effects can be observed during the exposure of iron or nickel base alloys to a hydrogen source:

(a) Strain may enhance hydrogen ingress in b.c.c. alloys as a consequence of a local destruction of the oxide films, while strain-induced defects favor hydrogen trapping;

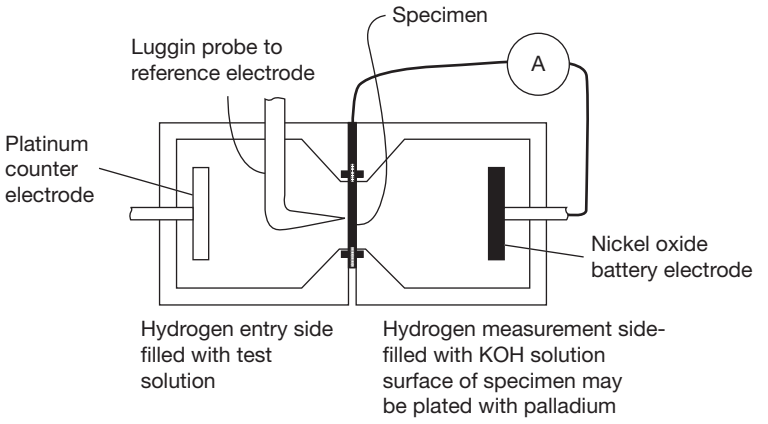


Figure 3 Hydrogen permeation cell (schematic).

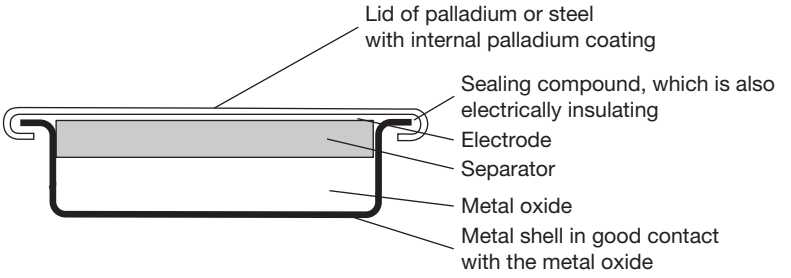


Figure 4 Permeation cell using battery technology. Reproduced from Arup, H. Proceedings of the 9th Scandinavian Corrosion Congress, 1984; p 825.

(b) Pipe diffusion may occur in the dislocation network of predeformed samples, whereas hydrogen accelerated transport by mobile dislocations is expected during plastic deformation in f.c.c. structures;
(c) Both the local plasticity and the embrittlement of these alloys are increased by hydrogen."

Using *in situ* observations in an environmental TEM, Robertson[14] has obtained further experimental confirmation of the influence of hydrogen on dislocation dynamics, and confirmed that hydrogen enhanced dislocation mobility.

2.10.6 Sources of Hydrogen

The thermodynamics of the reaction of iron with water are such that iron is always thermodynamically capable of displacing hydrogen from water, although the driving force is small at ambient temperatures, except in reasonably strong acids. Thus, iron does not suffer from serious corrosion in oxygen-free neutral or alkaline solutions. However, rates of hydrogen evolution which are insufficient to cause significant corrosion may still produce enough hydrogen to create a serious embrittlement problem for high strength steels. The hydrogen which is produced dissolves readily in the steel, and remarkably small concentrations of hydrogen (less than 1 ppm) can cause embrittlement. In addition to corrosion itself, many of the standard chemical and electrochemical treatments employed to protect steel against corrosion also tend to drive hydrogen into the steel. Some of the major sources of hydrogen are:

- *Welding* Hydrogen introduced into welds produces a particularly acute problem, as the weld and the heat-affected zone are inevitably regions of high residual stresses, contain inherent defects, and are frequently intrinsically more brittle than the parent material. Thus, it is important to minimize the introduction of hydrogen into welds, even for lower strength steels. Gas welding of steels using an oxyacetylene flame will inevitably introduce hydrogen as a result of the hydrogen-containing gases in the flame. In theory, electric arc welding,

particularly if the arc is protected from atmospheric moisture by inert gas shielding, will not introduce hydrogen. However, with normal manual metal arc welding using flux coated electrodes, it is possible for the flux coating to absorb moisture from the atmosphere, and this will react with the molten steel to produce hydrogen. For this reason, it is good practice to store coated welding electrodes in an oven in order to drive off any moisture.

- *Acid Pickling* This process is widely used for removing rust and mill scale from steel, or for removing internal scales from boilers. The objective of the process is the dissolution of iron oxides or hydroxides, or hardness scales, but at the same time, the iron will also tend to corrode in the acid, with hydrogen evolution as the cathodic reaction. This leads to hydrogen entry into the steel which may lead to hydrogen embrittlement or blister formation due to HIC.

As dissolution of the steel is normally an undesirable side-effect of the pickling process, resulting in loss of metal and wasteful consumption of acid, it is normal practice to add inhibitors (usually referred to as 'pickling restrainers') to the acid. Unfortunately, many of these work by interfering with the hydrogen recombination reaction, rather than the initial production of adsorbed hydrogen atoms. This has the effect of increasing the surface coverage by adsorbed hydrogen atoms, and consequently, the rate of entry of hydrogen into the steel is increased, even though the overall rate of hydrogen production is reduced. For this reason, it is most important that pickling restrainers for use with high strength steels should be tested for their effects on the uptake of hydrogen as well as their efficiency as inhibitors.

- *Degreasing and Cleaning* Various processes can be applied for the removal of grease and other contaminants prior to painting, electroplating, or other surface treatments. Degreasing in organic solvents (e.g., vapor degreasing) is unlikely to generate significant quantities of hydrogen unless the fluid is contaminated with water or other species, such as hydrogen chloride, capable of liberating protons. Cleaning in aqueous alkali without applied polarization is also unlikely to introduce much hydrogen, although any tendency for pitting corrosion due to chloride contamination could give rise to local problems. Cathodic cleaning, in which hydrogen is deliberately evolved on the steel is clearly undesirable.
- *Electroplating* Electroplated metal coatings provide a convenient and effective means of protecting steel against atmospheric corrosion, with zinc and cadmium being particularly useful because of their ability to provide sacrificial protection to the steel substrate at breaks in the coating. Unfortunately, all metal deposition processes require the application of potentials at which hydrogen evolution is possible (in theory it would be possible to deposit copper and more noble metals above the hydrogen evolution potential, but as this would also be a potential at which iron would dissolve anodically, this would not produce good deposits). Thus, the electrodeposition of coatings almost inevitably introduces hydrogen into the metal. The majority of electrodeposited coatings on iron are much less permeable to hydrogen than iron, the major exceptions to this being other *bcc* metals, with chromium being the only example in widespread use (although even here chromium is generally used with an undercoat of nickel or copper, which will act as a barrier). Consequently, most of the hydrogen uptake occurs in the early stages of the deposition process, before a complete coverage of the coating metal has been achieved, and the start of the deposition process is particularly critical in determining the amount of hydrogen absorbed. Clearly, the highest possible ratio of metal deposition rate to hydrogen generation rate is required in this period, and this tends to be favored by high current densities. For this reason (among others), a high current density 'flash' deposit of metal is often used as the first stage of a two-stage plating scheme. Bath composition is also important in controlling the entry of hydrogen into the steel.

Once high-strength steel components have been electroplated, it is possible (and often mandatory) to reduce the damaging effects of the hydrogen by baking at $\sim$200 °C. In part, this serves to allow hydrogen to diffuse through the coating and out to the atmosphere, but the redistribution of the remaining hydrogen within the steel, reducing damaging local high concentrations, is probably also an important part of this de-embrittlement treatment.

A very important area of use for electroplated high-strength steels is fasteners and other high stress components in aircraft. These are generally protected by cadmium plating, although the toxicity of cadmium and its compounds is giving rise to a search for alternatives. In this application, the avoidance of hydrogen embrittlement is clearly essential, and as a result, a

wide range of standard procedures and tests for cadmium plating of high-strength steels are available as civil or military specifications and codes of practice.[15]

- *Phosphating* Phosphating, which is widely used as a pretreatment for steel prior to painting, involves the controlled corrosion of the steel in acid solution, and inevitably leads to the uptake of hydrogen. By limiting the free acid in the phosphating bath, and by introducing oxidizing agents to raise the working potential, the uptake of hydrogen may be reduced.
- *Painting* Conventional paints are normally innocuous, but it seems possible that some modern water-based paints, particularly those applied by cathodic electrophoresis, may introduce significant quantities of hydrogen into the steel. Fortunately, these paints are generally stoved after application, and this is effective in removing the hydrogen, which permeates very easily through the paint film.[16] Additionally, these paints are applied primarily to lower-strength steels in motor vehicle bodies and similar applications, and as far as the author is aware, no problems have been experienced in service. However, suitable tests would be advisable if it is intended to use these processes with high-strength steels.

Paint strippers may give rise to hydrogen entry into steel, and in critical applications, such as the treatment of aircraft components, commercial paint strippers should be tested before use.

- *Corrosion in Service* Most of the modes of hydrogen entry discussed above involve a single brief charging period, and as a result are treatable, in that the hydrogen can be removed by a suitable de-embrittlement treatment (provided the embrittlement is not so severe that cracks are formed before the de-embrittlement treatment can be applied). In contrast, the entry of hydrogen due to corrosion in service is generally continuous. Thus, de-embrittlement is not feasible, and the control of hydrogen embrittlement of high strength steels presents a much more difficult problem.

In general, hydrogen will enter steel during any corrosion process involving hydrogen ion or water reduction as one of the cathodic reactions. It is frequently implied that the applied potential must be below the equilibrium potential for hydrogen evolution before hydrogen entry into steel is possible. However, for two related reasons this is not true. Firstly, the equilibrium potential for hydrogen evolution at a given pH is that potential at which protons or water molecules at the metal–solution interface are in equilibrium with molecular hydrogen at a partial pressure of one atmosphere. As the partial pressure of hydrogen is reduced, so the equilibrium potential will increase, in accordance with the Nernst equation, and even at potentials of 200 mV above the 1-atm equilibrium potential, there will be a significant concentration of hydrogen at the metal surface. Secondly, for high-strength steels, the matrix concentration of hydrogen required to cause embrittlement is very small, as the hydrogen tends to concentrate at phase boundaries and other trap sites. Thus, hydrogen embrittlement has been observed in high-strength steel in contact with gaseous hydrogen at pressures below one thousandth of an atmosphere.

A second difficulty in predicting the effect of particular corrosion conditions on the rate of hydrogen uptake by steel is the strong influence of local conditions. For passive steel, where the corrosion potential is several hundred mV above the 1-atm hydrogen equilibrium potential, one would not expect problems of hydrogen embrittlement, and this is frequently used as an argument against a hydrogen embrittlement mechanism of SCC in environments such as phosphates and carbonate/bicarbonate. However, the local environment in pits, cracks, or crevices may be very different from conditions at the free surface. In particular, acidification may occur due to metal–ion hydrolysis (especially for chromium-containing steels), and the potential in the localized corrosion cavity may be considerably more negative than that measured at the free surface. The net result is that the conditions within the cavity may be favorable for hydrogen evolution, even though the free surface conditions imply that hydrogen embrittlement is very unlikely.

In neutral saline environments, such as seawater, the rate of hydrogen entry is controlled by (among other things) the applied potential. In this case (and probably many others) the response observed, even at a smooth surface, is not as simple as might be expected, largely because of changes in the chemistry of the liquid in immediate contact with the steel. This is indicated in **Figure 5** due to Barth *et al.*,[17] which shows the rate at which hydrogen permeates through a steel membrane in response to a range of applied potentials. Note that hydrogen enters steel in aerated solution when it is anodically polarized to well above the nominal equilibrium potential for hydrogen (~ -650 mV versus SCE for this near-neutral

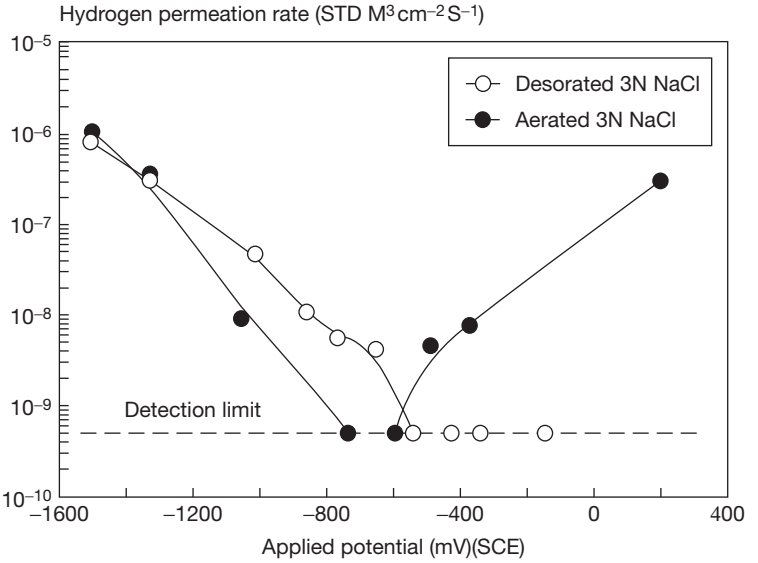


Figure 5 Hydrogen permeation current as a function of applied potential, showing effect of oxygen concentration and applied potential. Reproduced from Barth, C. F.; Troiano, A. R. Corrosion 1972, 28(7), 259–263.

solution). Similar behavior has been observed in work that has taken account of the shielding effect of a crack or crevice[18] and this has shown that the effects of applied potential can be explained on the basis of the local potential and pH in relation to the hydrogen equilibrium potential. **Figure 6** shows typical results obtained in this work, and it can be seen that the hydrogen overpotential and the rate of hydrogen entry into the steel increase with both anodic and cathodic polarization. Detailed prediction of the effect of applied potentials on the rate of hydrogen entry require very careful consideration of many factors, including the exact composition of the solution (especially with respect to the oxygen concentration), the geometry of the component (especially the presence of any cracks or crevices), and the mass transport conditions. In general terms however, when considering the performance of steel in seawater, it seems probable that slight cathodic polarization to ~−750 or −800 mV (SCE) may reduce the rate of hydrogen entry compared to free corrosion conditions, because the increased pH obtained as a result of the cathodic polarization outweighs the lowered potential. Further decrease in potential increases the rate of hydrogen entry down to ~−1200 mV (SCE), when electrochemical desorption takes over as the second stage of the hydrogen evolution reaction, and at potentials that are more negative, the rate of hydrogen entry remains roughly constant.

In deaerated neutral salt solutions, anodic polarization of chromium-free steel tends to lower the rate of hydrogen entry, as the fall in pH due to ferrous ion hydrolysis is relatively limited, and cannot compensate for the more positive potential. However, in aerated solution, the ferrous ion can be oxidized to ferric in solution by dissolved oxygen, and as ferric hydroxide is much less soluble than ferrous hydroxide, the pH can fall much lower. (This reaction is a serious problem when trying to measure the solution pH in cracks and crevices, and much of the early work which shows very acid pH values in cracks in carbon steels is now known to be incorrect because this reaction was allowed to occur between extracting the crack solution and measuring its pH). In addition to lowering the surface pH, it seems possible that the precipitated film of ferric hydroxide/oxide also presents an ohmic resistance, and may thereby allow the surface potential to become lower than that measured in the bulk solution. It seems probable that the combination of these two effects was responsible for the increase in permeation current observed for anodic polarization by Barth *et al.*[18] Many steels contain small, but significant concentrations of chromium, and it should be appreciated that chromium concentrations of ~1% can markedly increase the tendency for local acidification, owing to the much stronger tendency for chromium ion hydrolysis.

As considerations such as those discussed above have become more widely appreciated, it has become

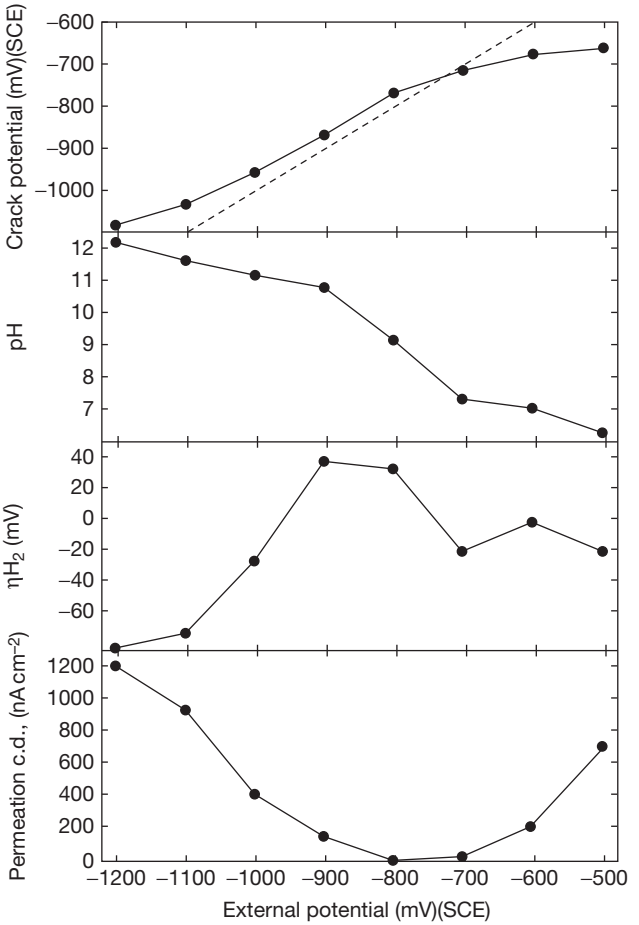


Figure 6 Potential and pH in a simulated crevice, and effect on hydrogen permeation current. Reproduced from Taqi, E. A.; Cottis, R. A. Corrosion Chemistry within Pits, Cracks and Crevices; Turnbull, A.; NPL: London, 1987; 483–494.

clear that it is quite difficult to specify aqueous environments in which hydrogen entry into steel will occur at a sufficiently low rate that the possibility of hydrogen embrittlement can be discounted. The main candidates for such an environment are strong alkalis and highly oxidizing environments, such as nitric acid or nitrates, which are free of species, such as chloride, which predispose to pitting or crevice corrosion, and high temperature environments, such as water at 250 °C (since the deleterious effects of dissolved hydrogen decrease at high temperatures).

2.10.7 Effect of Hydrogen on Mechanical Properties

2.10.7.1 Elastic Constants

While there is some evidence of small changes in the elastic properties of steel as a result of dissolved hydrogen[19] these changes are small, and of little practical consequence. This is perhaps to be expected in view of the very low solubility of hydrogen in the iron lattice and the small effect on the metal–metal bond strength.

2.10.7.2 Yield Stress

The effect of hydrogen on the yield stress of iron and steels is unpredictable. For very pure iron single crystals and polycrystals, the yield stress is frequently found to be decreased by hydrogen, but it may increase or stay the same, depending on the dislocation structure, crystal orientation, and purity of the iron.[20] Little information is available for steels.

2.10.7.3 Plastic Behavior

The effect of hydrogen on the plastic behavior of iron and steel is somewhat complex, as hydrogen may

harden or soften the material, according to its structure and the mode of slip. On the basis of work on single crystals of pure iron, Lunarska[20] has concluded that segregation of hydrogen around dislocations decreases their elastic stress fields. At room temperature (where the rate of diffusion of hydrogen is sufficiently high that it can keep up with a moving dislocation), this leads to a softening of single crystals when only one slip system is operative. The hydrogen also suppresses cross-slip of screw dislocations, and this results in increased work hardening when multiple slip systems are active. Lunarska also notes that "the presence of residuals (even at very low concentration) can drastically change all of these effects."

It is partly because of the variable effect of hydrogen (giving both softening and hardening, according to the nature of the slip) that the extrapolation of model experiments on very pure iron to predict the behavior of commercial materials is so difficult. It is further hindered by the ability of dissolved hydrogen to modify the dislocation structure of a straining material.

2.10.8 Hydrogen Embrittlement of Steels

By far the most important impact of dissolved hydrogen on the mechanical properties of steels and particularly high-strength steels is the production of apparently brittle fracture where the steel would normally behave in a ductile fashion. If a steel contains dissolved hydrogen, this can result in immediate fracture at stresses approaching the fracture stress in the absence of hydrogen, or it may result in delayed failure at lower stresses (**Figure 7**). The latter behavior is most pronounced for tensile tests on notched or precracked specimens, as the stress field around the notch or crack creates a high triaxial stress which dilates the metal lattice and tends to attract hydrogen from other parts of the specimen.

Some typical properties of common steels when exposed to environments supplying hydrogen are presented in **Table 2**. This table is presented to illustrate the order of magnitude of hydrogen embrittlement effects, and it is important to appreciate that the heat treatment and mechanical processing of a particular material, as well as the exposure conditions, can markedly affect its resistance to hydrogen embrittlement. Many of the data presented in this table have been extracted from the reviews of Sandoz[21] and McIntyre.[22] One useful fact that is indicated by the data in **Table 2** is the good hydrogen embrittlement resistance of maraging steels when used somewhat below their ultimate capability. These data also illustrate the general result that a given steel will usually have a greater susceptibility to hydrogen embrittlement as it is tempered to a higher strength.

The effect of strength of the embrittlement can be rationalized by considering the changes in the size of the plastic zone at the crack tip as a function of the yield strength. The stress ahead of a perfectly sharp crack is proportional to $1/r$, where r is the distance from the crack tip. This goes to infinity as $r \rightarrow 0$, so there must be a plastic zone ahead of the crack that has yielded. The size of this plastic zone will be

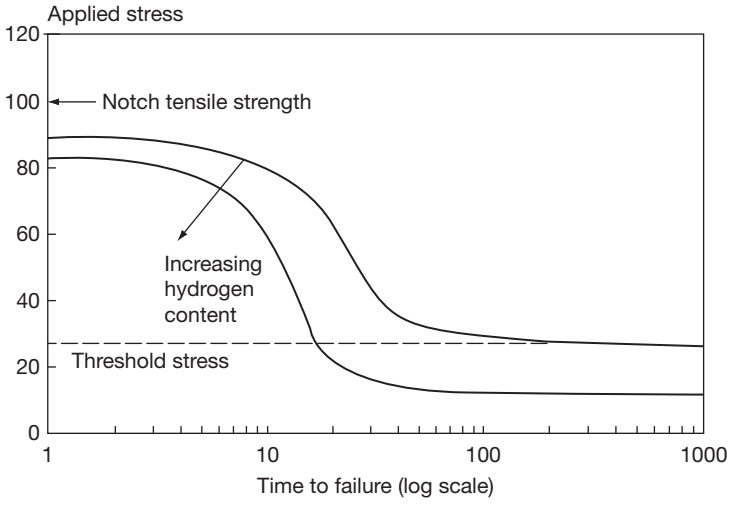


Figure 7 Typical results of static load tests on notched specimens.

Table 2 Typical properties of common high-strength steels in salt solutions

Material	*Yield stress (MPa)*	*Failure time*[a] *(h)*	K_{Ic} *(MPa√m)*	K_{ISCC} *(MPa√m)*	*Plateau velocity* ($m\,s^{-1}$)
HY130	900	NF	130–170	60–140	5×10^{-8}
300M	1600	40–150	–	115	10^{-8}
4340	1400–1800	–	–	20	$10^{-5} - 10^{-2}$
4340	<1400	–	–	30–80	
4130	1300	–	–	30	$10^{-6} - 10^{-4}$
4130	1050	–	–	120	
18Ni MS	1900	400	40–100	5–35	
18Ni MS	1500	10,000	180	40–100	
18Ni MS	1200	NF	160	120	10^{-5}

[a]The failure time presented is the approximate time to failure at an applied stress of 75% of the yield stress of the material.
MS = Maraging Steel.

approximately inversely proportional to the yield stress, and the peak stress will be equal to the yield stress. This discussion oversimplifies the stress state ahead of the crack tip, which is three dimensional and rather complex. However, the general conclusion is accurate. Thus, a higher strength material will have a smaller plastic zone with a larger stress in it, both the factors tend to increase the tendency for hydrogen to concentrate in this region and to cause fracture.

2.10.9 Hydrogen Embrittlement of Other Alloys

While the majority of work on hydrogen embrittlement is concerned with high strength steels, many other alloys have been shown to suffer embrittlement by hydrogen. We can list a few general factors that have a strong influence on typical susceptibility to hydrogen (though, like most generalizations, there will be exceptions):

- Susceptibility increases as the yield strength increases. This is an almost inevitable consequence of the reduction in the volume of the crack tip plastic zone and the increase in stress within that zone as the yield strength increases.
- Face-centered cubic (*fcc*) alloys are less susceptible than body-centered cubic (*bcc*) alloys. There are several reasons for this, (1) the greater inherent ductility of *fcc* alloys; (2) the much lower diffusion coefficient for hydrogen in *fcc* structures; and (3) the much greater solubility of hydrogen in *fcc*.
- Hexagonal close-packed (*hcp*) alloys might be expected to have an intermediate susceptibility between *fcc* and *bcc*, having lower inherent ductility than *fcc* but comparable hydrogen solubility and diffusion coefficient. However, many of the common *hcp* alloys (Ti, Mg, Zr among others) form stable hydrides, and are therefore susceptible to hydrogen embrittlement by a hydride cracking mechanism.

It is sometimes stated that *fcc* alloys (such as austenitic stainless steels) are immune from hydrogen embrittlement. However, this is an overstatement; it is more accurate to say that they are more resistant, which typically means that more hydrogen, and/or a longer time are necessary for fracture to occur. Thus, hydrogen embrittlement has been observed in nickel sheet (used in the author's laboratory as a cathode in an electrochemical machining cell), high strength aluminum alloys,[23] and austenitic[24] and duplex[25] stainless steels. There are relatively few reports of hydrogen embrittlement of copper or copper-based alloys, but this is probably due to the noble character of copper, which means that hydrogen cannot be produced at a significant concentration by corrosion of copper, rather than to any inherent immunity from embrittlement.

2.10.10 The Influence of Microstructure and Composition

The influence of the composition and metallurgical structure of steels on their susceptibility to hydrogen embrittlement have been discussed from a theoretical basis by Bernstein and Pressouyre.[6] They considered the nature of the various trap sites in the steel, both with respect to their tendency to accumulate hydrogen, and their sensitivity to fracture in the presence of hydrogen. This approach seems to offer considerable potential in respect of designing alloys for hydrogen resistance without sacrificing other properties, although with our current knowledge the methods are probably more relevant for the rationalization of results.

As a general rule, the damage due to dissolved hydrogen tends to become more severe as the strength of the steel increases. In part, this may be associated with the greater resistance of the stronger material to plastic deformation, which facilitates the transition to brittle behavior. In addition, the size of the crack tip plastic zone, which defines the size of the region to which hydrogen is attracted, is inversely proportional to the yield strength of the material. Hence, for the stronger steels, a given amount of dissolved hydrogen will be concentrated into a smaller region, and will, therefore, have a more damaging effect. However, this general effect of strength level does not explain all aspects of hydrogen embrittlement of high strength steel, and steels with the same mechanical properties may be affected by hydrogen in quite different ways as a result of their different microstructures.

A particularly important aspect of the microstructural state of the steel is the condition of the grain boundary. As noted in **Section 2.10.4** above, 'pure' grain boundaries do not act as major sites for the trapping of hydrogen, but it is clear that impurity segregation and carbide precipitation at the grain boundary may significantly modify its behavior. Thus, it is now widely recognized that there is a strong link between various forms of temper embrittlement and hydrogen embrittlement, and quite small changes in tempering treatment can give large variations in hydrogen embrittlement resistance.[26,27]

On a more positive note, it seems clear that steels can be made more resistant to the effects of hydrogen by incorporating as many strong, finely dispersed traps in the microstructure as is possible, while ensuring that there are no continuous trap sites (such as embrittled grain boundaries). This may explain the better resistance of steels hardened by cold-working compared to quenched and tempered steels.[28] Whatever the reason, this effect is particularly fortunate in the production of prestressed concrete, where cold-drawn pearlitic (or 'patented') wires give high strength levels (~1700 MPa) with good (though not complete) resistance to hydrogen embrittlement. In this particular case, it is probably also important that most of the phase boundaries in the structure lie parallel to the tensile axis, allowing them to collect large quantities of hydrogen without serious detriment to the performance of the wire. When such wires do suffer from hydrogen embrittlement, they exhibit a characteristic elongated 'green-stick' fracture.

Similarly, it seems that retained austenite may be beneficial in certain circumstances[29] probably because the austenite acts as a barrier to the diffusion of hydrogen, although in high concentrations (such as are obtained in duplex stainless steels) the austenite can also act as a crack-stopper (i.e., a ductile region in the microstructure which blunts and stops the brittle crack).

2.10.11 Theories of Hydrogen Embrittlement

Despite the major technical importance of hydrogen embrittlement, and the wealth of research work on the subject, the mechanism (or perhaps mechanisms) of hydrogen embrittlement remains uncertain. Much of the book edited by Oriani, Hirth, and Smialowski is concerned with mechanistic aspects of hydrogen embrittlement, and the reader is referred in particular to the summary by Thomson and Lin.[30]

In considering hydrogen embrittlement mechanisms, it is important to keep in mind the concentration of hydrogen in the steel, as matrix concentrations are very low, typically of the order of one atom of hydrogen for every 10^6 iron atoms. It is very difficult to see how such small amounts of hydrogen can modify fracture properties so markedly, and it must be supposed that hydrogen present in traps or possibly in surface layers (where the solubility may be markedly increased)[6] is the main cause of embrittlement. Thus, realistic mechanisms of hydrogen embrittlement will be based on the effect of hydrogen on dislocation behavior, on the effect of hydrogen at phase boundaries or grain boundaries or on the effect of hydrogen at the metal surface. These factors will be assisted by the concentration of hydrogen to the region of triaxial tensile stress ahead of the crack.

In considering the various theories, it is also apparent that many of them may be considered as alternative descriptions of essentially the same physical process, or as descriptions of parallel processes which collaborate in the failure. Thus, a complete description of hydrogen embrittlement in a given situation will almost inevitably incorporate aspects of several of the following theories.

2.10.11.1 The Pressure Theory

The earliest theory of hydrogen embrittlement was probably the planar pressure theory advanced by Zappfe[31] in 1941. This essentially proposes that the effect of hydrogen is to create very high pressures of hydrogen gas in voids and other defects within the metal thereby assisting in the fracture of the steel.

While this is an important aspect of the blistering of steel by HIC, it cannot by itself explain the hydrogen embrittlement of high strength steels, where fracture may occur in steel in equilibrium with hydrogen at very low pressures.[32] Some concentration of hydrogen may occur as a result of dislocation transport, but it is difficult to see how significant internal pressures can be generated by hydrogen entering from an external pressure of one thousandth of an atmosphere, yet this can cause hydrogen embrittlement.

2.10.11.2 Decohesion Theories

The decohesion models proposed by Troiano, Oriani, and others[33–35] suggest that the role of hydrogen is to weaken the interatomic bonds in the steel, thereby facilitating grain boundary separation or cleavage crack growth. In view of the very low hydrogen concentration in the matrix, it is necessary for some method to exist by which the hydrogen can be concentrated at the site of the fracture. For cracking along phase or grain boundaries, this can be explained in terms of the trapping of hydrogen at the phase boundary. It is a little more difficult to see how transgranular cracking can be explained; processes which have been invoked include the concentration of hydrogen at the region of triaxial tensile stress at the crack tip and local high concentrations of hydrogen being generated by reaction or adsorption at the crack tip.

2.10.11.3 Surface Energy Theories

Surface energy theories were first proposed by Petch and Stables.[36] By lowering the surface energy of the newly formed crack, the hydrogen reduces the stress intensity required for brittle fracture. As with the decohesion models, surface energy models only seem reasonable for the case of hydrogen derived from surface layers or grain boundaries, since the hydrogen adsorption must occur at the same time as the fracture event in order for the reduction in surface energy to be effective in lowering the energy required for fracture.

2.10.11.4 Hydride Formation

In some systems, notably titanium alloys, hydrogen embrittlement has been attributed to the formation and subsequent fracture of brittle hydride phases at the crack tip. This has been postulated by Gahr *et al.*[37] as a mechanism of hydrogen embrittlement of niobium. While there is little evidence for hydride formation in steels, it can be argued that hydrides would be unstable and would dissolve as soon as the crack has propagated through them. In view of recent evidence of significant structural rearrangement of steel surface containing chemisorbed hydrogen,[6] the possibility should also be considered of the induction of cleavage by a brittle surface film, similar to the film-induced cleavage model of stress corrosion cracking.[38]

2.10.11.5 Local Plasticity Theories

Various workers have suggested that hydrogen acts by reducing the stress required for dislocation motion. This follows observations by several workers (summarized by Morgan and McMahon[39]) of enhanced dislocation motion in thin films exposed to hydrogen. Similarly, Lynch[40] found very similar fractography for liquid metal embrittlement and hydrogen embrittlement, and proposed that both processes involve the facilitation of dislocation emission at the crack tip. Lynch,[3] echoing the earlier review of Birnbaum *et al.*,[2] describes the three currently favored mechanisms:

- hydrogen-enhanced localized plasticity (HELP);
- hydrogen enhanced decohesion (HEDE); and
- adsorption-induced dislocation emission (AIDE).

In the HELP mechanism, crack propagation occurs by highly localized slip due to local softening by hydrogen at the crack tip. HEDE is due to weakening of the metal–metal bond allowing brittle fracture rather than plastic deformation (essentially this is the decohesion model described above). AIDE is described by Lynch as "... based on hydrogen induced weakening of interatomic bonds (as for HEDE) but with crack growth occurring by localized slip (as for HELP)." The corrosion-enhanced plasticity model proposed by Delafoss and Magnin[41] is rather more general description of an essentially similar model to HELP, in that the enhanced plasticity may be due to corrosion, as well as hydrogen, effects.

2.10.12 Hydrogen Embrittlement Tests

There are several classes of test for hydrogen embrittlement, according to the application. Three general types of mechanical test can be identified, together with chemical and electrochemical tests intended to determine the hydrogen content of steels or the rate of entry of hydrogen from an environment.

2.10.12.1 Constant Stress Tests

The simplest test for hydrogen embrittlement involves applying a constant stress to a specimen. This can be applied both to the testing of samples which already contain hydrogen (e.g., as a result of electroplating) and to the testing of samples in environments causing hydrogen entry. The specimen may take various forms but is commonly a tensile specimen. Particularly where the test is being used to examine the effect of dissolved hydrogen which is already present in the steel, the specimen may be notched, in order to develop a region of triaxial stress at the tip of the notch that will tend to concentrate hydrogen.

When a constant stress test of a notched specimen is being used for the evaluation of samples containing hydrogen, it is commonly referred to as a sustained load test, although the terms stress rupture and static fatigue are also used in the older literature. The results of a typical sustained load test are shown in **Figure 7**. At a stress which is frequently (though not necessarily) below the notch tensile strength for hydrogen-free material, the specimen will fail instantaneously as a result of the damaging effect of the uniform hydrogen concentration in the steel. At lower stresses, delayed failure occurs as hydrogen diffuses to the region of triaxial stress at the tip of the notch. Eventually, as the stress is reduced, the quantity of hydrogen in the neighborhood of the notch is insufficient to cause failure, and a critical or threshold stress is reached. The behavior of a steel sample in this test is a function of both the material and the hydrogen concentration, with an increase in hydrogen concentration giving instantaneous failure at lower stresses, shorter times to failure, and a lower critical stress. Testing at lower temperatures will give longer failure times, since the rate-controlling process is hydrogen diffusion, but the critical stress will go down, as the intrinsic tendency of the steel to brittle rather than ductile failure is increased. At elevated temperatures, hydrogen may be lost by degassing and by trapping, and in this case, the critical stress will increase. The failure stress is also strongly influenced by the notch tip radius, with a smaller radius giving shorter times to failure and failure at lower stresses. This can be explained in terms of the size and intensity of the region of triaxial stress at the notch tip, with a sharper notch giving a smaller triaxial region, which can develop a higher hydrogen concentration in a given time. Because of the difficulty of obtaining a reproducible notch tip geometry, the sustained load test tends to give rather scattered results. Smooth specimens can be used for sustained load testing, but in this case, the defects responsible for local hydrogen concentration will typically be surface or near-surface nonmetallic inclusions, and the chance organization of these will give highly scattered results. Smooth specimen results will also tend to be nonconservative, as the majority of real components are liable to contain stress concentrators of one form or another.

In addition to testing for the effects of preexisting hydrogen in the steel, sustained load tests can also be used to test for the effect of hydrogen entering from the environment. Both smooth and notched specimens can be used. At short times, the results obtained will differ from those for steels containing dissolved hydrogen, since time will be required for the hydrogen to enter the steel. However, this time can be very short, since hydrogen can diffuse ~10 μm in 1 s, and failure times of the order of seconds can be obtained for severe environments and loading conditions. For tests at ambient temperature, it is found that the critical stress is reached fairly quickly, and it is generally found that failures will occur within a month of exposure of the samples. This provides something of a problem in relating service behavior to laboratory tests, since service failures are often observed after much longer exposure times. This may be a result of variations in exposure conditions in service or it may be associated with the need to develop initiating defects by pitting corrosion or similar processes. Whatever the cause, it does lead to the conclusion that sustained load tests cannot be used to predict service life with any accuracy.

2.10.12.2 Controlled Strain-Rate Tests

Controlled strain-rate tests were first developed by Parkins[42] for the study of stress-corrosion cracking. These took the form of constant strain-rate tests (also known, perhaps more accurately, as constant extension rate tests). Since then alternative forms of test have been developed to modify the conditions under which the specimen is exposed.

The slow strain-rate test is based on the principle that stress corrosion cracking processes are normally dependent on plastic strain in the material. By extending the specimen very slowly, any stress corrosion cracking phenomena are given every possible opportunity to occur, hence, the slow strain-rate test can quickly reveal any tendency to cracking in a given metal–environment combination. A particular advantage of the slow strain-rate test is that failure of the specimen will always occur in a reasonable time, if

only by normal ductile fracture processes. Test durations clearly depend on the strain-rate and the ductility of the metal, but typical test durations for steels in hydrogen embrittling environments are ~1 week. In addition to providing a rapid indication of the possibility of SCC in a particular system, the slow strain-rate test can also be used to study the effect of material and environmental factors on the susceptibility. For example, **Figure 8** shows the effect of environment composition and tempering temperature on the reduction in area obtained for tests on a quenched and tempered steel in various aqueous solutions. The minimum in the reduction in area which can be seen for tempering temperatures of 350 °C provides evidence of the interaction between hydrogen embrittlement and temper embrittlement.

While the conventional slow strain-rate test offers many benefits, it does suffer from a tendency to overstate the susceptibility of materials to hydrogen embrittlement. Thus, structural steels of modest strength will fail even under conditions giving relatively low rates of hydrogen entry. This is because the enforced plastic deformation provides a very severe test condition, which may not be reproduced in service when relatively low stresses are applied to components. This problem is partly addressed by the sustained load test, but this may be unduly mild, as there will be no plastic strain in the specimen once the initial creep strain has died away. In order to provide a test with intermediate severity, Erlings and his co-workers[43] developed a test which exposes the specimen to a quasi-constant strain-rate, but at the same time keeps the specimen within the elastic range. This involves prestraining a slow strain-rate specimen up to the yield stress. Then the specimen is exposed to a regular cycle, the stress being reduced (typically) to 90% of the yield stress, slowly strained to 95% of the yield stress, then dropping back to 90%. This test is proposed as a part of a sequence of tests, starting with the low-cost slow strain-rate test, for the qualification of materials for service, as indicated in **Figure 9**.

Another modification to the slow strain-rate test involves the superimposition of a low amplitude sine wave ripple on the slow uniform extension (**Figure 10**). In effect, this produces higher strain-rates (which appear to be more damaging for hydrogen embrittlement), while still giving a long test duration, with adequate time for the accumulation of hydrogen in the steel.[44]

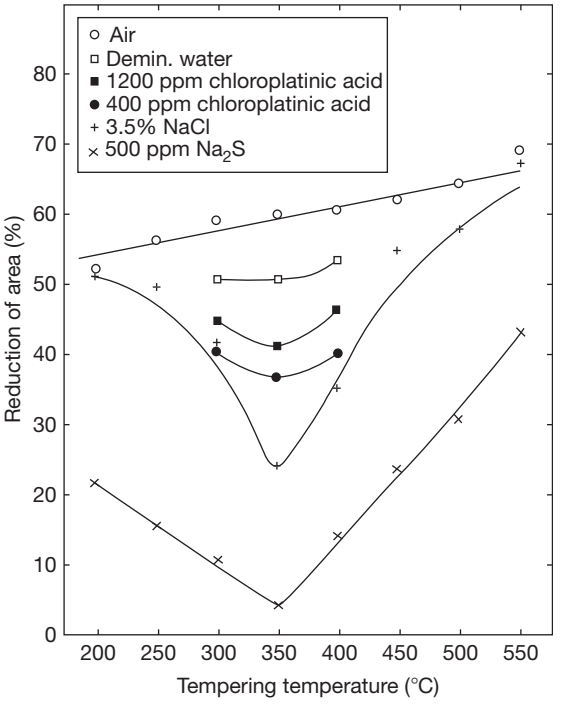


Figure 8 Variation of reduction in area as a function of tempering temperature. Reproduced from Johnston, J. W.; Cottis, R. A.; Procter, R. P. M. Electrochem. Soc. Extended Abstracts 1987, 87(2), 454–455.

2.10.12.3 Fracture Mechanics Tests

One problem of both sustained load and slow strain-rate tests is that they do not provide a means of predicting the behavior of components containing defects (other than the inherent defect associated with the notch in a sustained load test). Fracture mechanics provide a basis for such tests, and measurements of crack velocity as a function of stress intensity factor, K, are widely used. A typical graph of crack velocity as a function of K is shown in **Figure 11**. Several regions may be seen on this curve. At low stress intensity factors, no crack growth is observed until a threshold stress intensity factor, K_{ISCC} is reached. Typically, K_{ISCC} is fairly insensitive to the rate of hydrogen entry into the steel (as the crack is not growing at a significant rate, hydrogen will eventually accumulate at a sufficient concentration at the crack tip, whatever the rate of entry). As the stress intensity factor increases above K_{ISCC} the crack growth rate increases rapidly, until a plateau crack velocity is attained. The plateau velocity is a function of the rate of hydrogen entry, and the plateau is usually attributed to the availability of hydrogen at the crack tip being the rate-limiting factor. Finally, the crack velocity increases again as the stress intensity factor approaches the critical stress intensity factor for fast fracture, K_{IC}.

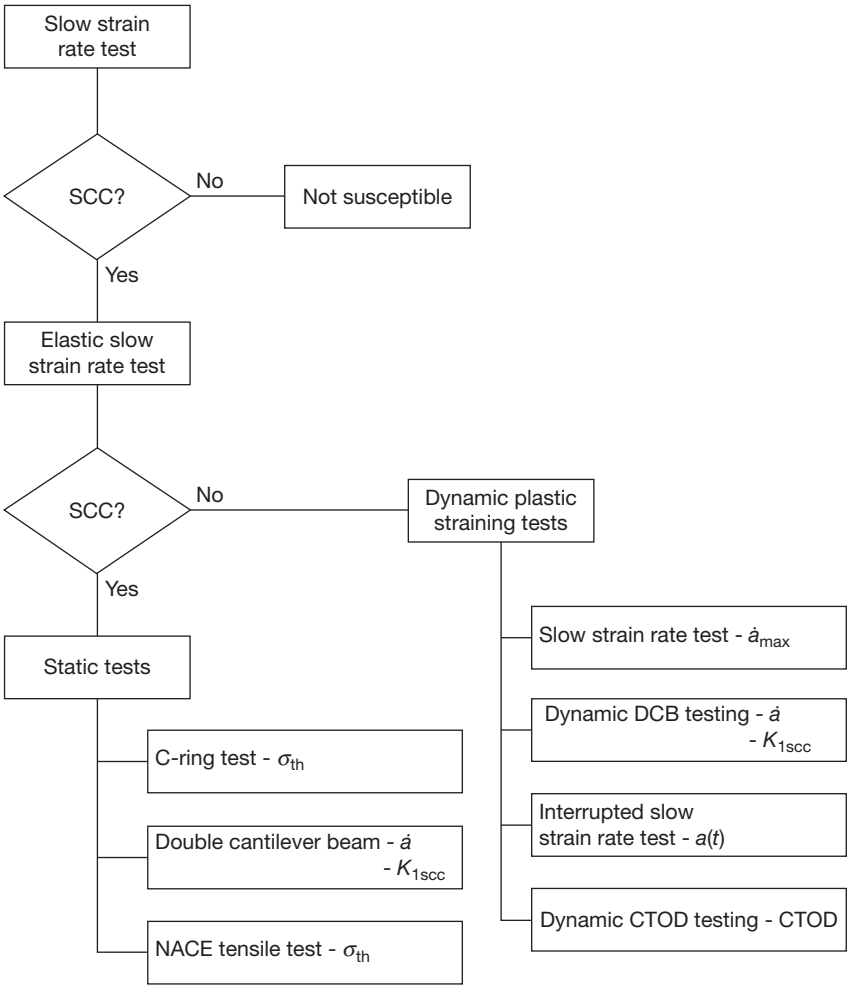


Figure 9 'Elastic' slow strain-rate test. Reproduced from Crooker, T. W.; Hauser, J. A. H A Literature Review on the Influence of Small-Amplitude Cyclic Loading on Stress Corrosion Crack Growth in Alloys, NRL Memorandum Report 5763, Naval Research Laboratory, Washington DC, April 3, 1988.

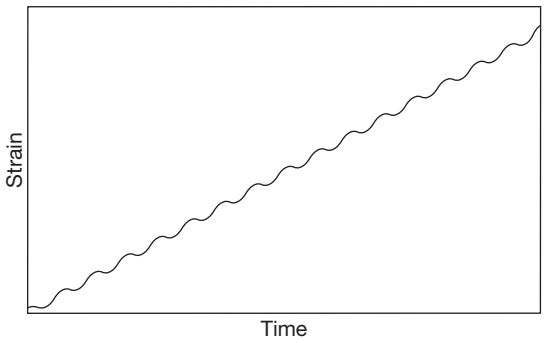


Figure 10 Strain-time history for 'Ripple' test (schematic).

In this region, the fracture consists partly of ductile tearing and partly of hydrogen embrittlement, with the hydrogen assisting the fracture of the tougher areas in the metal. The above discussion has assumed that the crack is loaded in Mode I (the crack opening mode, with a tensile stress normal to the plane of the crack). Hydrogen has relatively little effect in Modes II or III, as these generate shear stresses at the crack tip, rather than tensile stresses, and the shear behavior of steels is relatively little affected by hydrogen, presumably because dilation of the lattice at the crack tip (which does not occur in Modes II and III) is required for hydrogen accumulation.

2.10.13 Controlling Hydrogen Embrittlement in Service

Clearly there are two major ways in which hydrogen embrittlement can be controlled in service – either the hydrogen content of the steel can be reduced

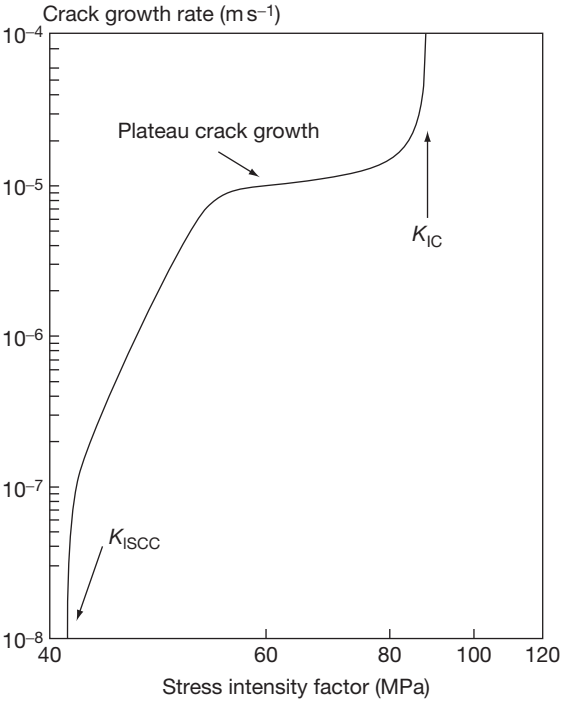


Figure 11 Stress-corrosion crack velocity as a function of applied stress intensity factor.

(by limiting the rate of entry or by baking to remove existing hydrogen), or the steel can be made more resistant to hydrogen.

2.10.13.1 Reduction of Hydrogen Content

Preexisting hydrogen in steel, introduced by electroplating or other processing, can be reduced by a suitable de-embrittlement treatment, as discussed above. A more difficult problem is the limitation of hydrogen entry in service. The feasibility of controlling hydrogen embrittlement in this way will clearly depend on the application, and in particular on the source from which the hydrogen is entering the steel. Thus, hydrogen pickup from acid pickling can be reduced by the use of appropriate inhibitors, and de-embrittlement treatments can be used after pickling. In contrast, hydrogen entry into components in an electrical generator using hydrogen as the cooling gas cannot easily be modified.

In general preventing the entry of hydrogen from the gas phase is liable to be difficult. Inhibition is possible, for example, small quantities of oxygen can markedly reduce the rate of hydrogen uptake, presumably as a result of blocking active sites on the metal surface. Similarly, ethylene gas will react with adsorbed hydrogen atoms, thereby preventing them from diffusing into the steel. However, there will be few cases in which it is permissible to modify the environment in this way (adding oxygen to hydrogen will certainly cause some concern). Organic coatings will not generally be very effective in reducing hydrogen entry, since these are usually rather permeable to hydrogen. Furthermore, such coatings are usually susceptible to local damage. Metal coatings may offer some reductions in hydrogen entry, as most close-packed metals have a significantly lower diffusion coefficient for hydrogen than iron, although this is partially compensated for by a higher solubility.

In aqueous solutions, it becomes somewhat more feasible to modify the entry of hydrogen into the steel. This can be achieved by the addition of inhibitors to the solution, by control of the electrochemical potential of the metal and by coatings.

In situations such as the acid pickling of steel or the use of steel pipes to handle sour oil streams, the use of suitable inhibitors can give a significant reduction in hydrogen entry. In this context, it is important to emphasize that the efficiency of an inhibitor in reducing hydrogen entry is not the same as its efficiency in reducing corrosion. Thus, arsenic and antimony compounds are effective inhibitors of corrosion of steel in acid, but they achieve this by inhibiting the hydrogen recombination reaction, thereby increasing the rate of hydrogen entry into the steel. In contrast, the addition of chloroplatinic acid to a solution will increase the corrosion rate by accelerating the hydrogen evolution reaction. However, a part of this acceleration arises because the hydrogen recombination reaction is catalyzed by the particles of platinum which are deposited on the surface by an exchange reaction. As a result, chloroplatinic acid acts as an inhibitor of hydrogen entry. This can be seen in **Figure 8**, where the addition of chloroplatinic acid to 3.5% sodium chloride solution has increased the reduction in area for a given heat treatment. Clearly, this is impractical for service use, but several organic inhibitors are available which are effective in reducing both corrosion and hydrogen entry into the steel.

The electrochemical potential has a direct influence on the entry of hydrogen into steel from aqueous solutions, and control of potential represents an obvious way to control hydrogen uptake. Unfortunately, the effect of potential on hydrogen entry is complex, and is strongly dependent on solution composition. For example, as has been discussed above, it has been shown[18] that anodic polarization of steel in salt

solution will give an increased rate of hydrogen entry in aerated solution, but negligible entry in deaerated solution. The situation becomes even more complex in passivating environments, especially where pitting or crevice corrosion can occur. As a generalization, for aerated neutral salt solutions the optimum potential for minimum hydrogen uptake appears to be slightly below typical free corrosion potentials, a typical value for seawater being ~−750 to −800 mV (SCE). This potential is also sufficient to reduce corrosion rates, but it gives very little margin for potential differences over the structure being protected. Potentials much below −900 mV will undoubtedly increase the rates of hydrogen entry, and in view of the hazards involved in overprotection, the deliberate application of cathodic protection to high strength steel structures is probably best avoided.

As in gaseous environments, metal coatings can reduce the rate of hydrogen entry by acting as a low permeability barrier. In addition, the coating metal may also modify the electrochemical properties at the metal–solution interface, as in the case of traces of platinum derived from chloroplatinic acid. For application in seawater and other neutral salt solutions, zinc, and cadmium probably offer the best combinations of properties, having low exchange current densities for hydrogen evolution and relatively low hydrogen permeabilities. In the event that part of the steel substrate is exposed by mechanical damage, a zinc coating has a rather low corrosion potential in seawater, and may be expected to give rapid hydrogen entry. However, cadmium gives a potential which is around that at which the minimum rate of hydrogen entry is observed, and is highly recommended as a coating material for high strength steels. Unfortunately, the toxicity of cadmium compounds precludes its use in most applications. Proposed schemes for 'cadmium replacement' should be examined carefully for their effect on hydrogen embrittlement as well as their corrosion performance.

Organic coatings such as paints are also effective in reducing the corrosion of steel, but as with inhibitors, it is not certain that this will also mean that the rate of hydrogen entry is reduced. Organic coatings are typically rather permeable to water, and it is possible for this to be reduced to hydrogen. Whether this hydrogen subsequently enters the steel will depend on the characteristics of the coating–steel interface, and little is known about this. It is recommended that suitable tests should be performed before using organic coatings for protection of high strength steels.

2.10.13.2 Increasing Resistance of Steel

For a given steel, the hardness will play a major part in determining its resistance to hydrogen embrittlement, but other factors are also significant. Thus, it is particularly important to obtain an appropriate microstructure, without any temper embrittlement or other deleterious features (see above). Where possible the use of work-hardened ferritic–pearlitic structures will probably give better performance than quenching and tempering, providing the stress is applied parallel to the working direction.

In particularly severe conditions, it may be found necessary to use austenitic steels, such as stainless steels from the 300 series. The *fcc* structure of austenite is inherently more ductile than *bcc* ferrite, and it is therefore less liable to switch to brittle modes of fracture in the presence of hydrogen. Additionally the permeability of austenitic steels is less than that of ferritic. The common commercial grades of austenitic stainless steel tend to be rather marginal in respect of the stability of the austenite phase, and they are therefore susceptible to the formation of martensite by strain-induced transformation. This martensite is susceptible to hydrogen embrittlement, and it is therefore important to ensure that steels used for hydrogen service have sufficient austenite stabilizing elements, primarily nickel. This is illustrated in **Figure 12**, where the effect of nickel content on the residual ductility in hydrogen is presented.[45] Below ~10% Ni there is a strong tendency to strain-induced martensite

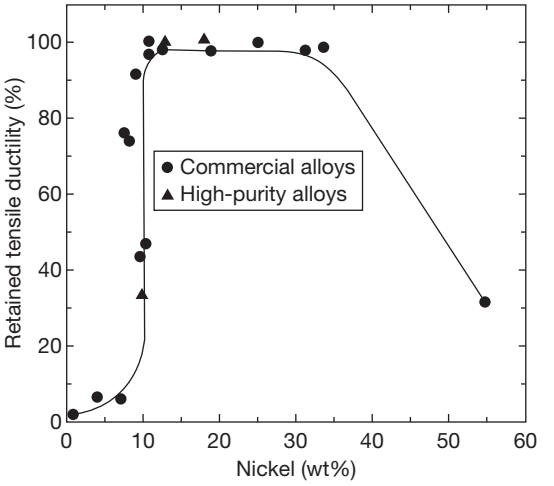


Figure 12 Effect of nickel content on residual ductility of austenitic stainless steels. Reproduced from Caskey, G. R. Third International Congress on Hydrogen and Materials, Paris, 1982; p 611–616.

formation, and low residual ductilities are observed. Once the austenite becomes fully stable however, there is hardly any loss in ductility due to hydrogen.

2.10.13.3 Prediction of Behavior

An important aspect of the use of materials in aggressive environments is the prediction of service life, such that planned maintenance procedures can be used. A typical approach that might be suggested for steels in conditions where hydrogen embrittlement can occur is to assume a maximum possible defect in the structure, usually based on the expected capabilities of the inspection system. Using a suitable relationship for the crack velocity as a function of stress intensity factor, the time to failure is calculated by integrating from the initial defect size up to that necessary for fast fracture. While this method is attractive in theory, in practice it is liable to be rather unproductive for high strength steels. The reason for this is that typical plateau crack velocities for hydrogen embrittlement are very high (between 10^{-6} and $10^{-2}\,\mathrm{m\,s^{-1}}$), and the transition from K_{ISCC} to the plateau occurs over a rather small range of stress intensity factor. Thus, unless the initial defect happens to give a stress intensity factor fractionally above K_{ISCC}, this calculation will predict either no failure or a very short life (even $10^{-6}\,\mathrm{m\,s^{-1}}$ corresponds to a crack velocity of $3.6\,\mathrm{mm\,h^{-1}}$). Currently, the understanding of the failure of components by hydrogen embrittlement after long periods in service remains somewhat limited, but in many cases, it seems probable that pitting corrosion plays an important part in the slow crack initiation process. Thus, prediction of life in hydrogen embrittlement conditions should be concerned with initiation processes in addition to the crack growth above K_{ISCC}, although there is not yet a sufficiently clear understanding of the initiation processes for this to be a productive exercise.

2.10.14 Modeling of Hydrogen Embrittlement

At a fundamental level, improvements in techniques for *ab initio* modeling at the atomic level, coupled with the continuing rapid increase in the power of high performance computers, has led to a significant improvement in our understanding of the behavior of hydrogen in metals, including aspects such as the binding energy of hydrogen to dislocations[46] and the segregation of hydrogen to grain boundaries.[47]

While it is not currently feasible to construct a completely reliable *ab initio* model of hydrogen embrittlement initiation or propagation, some moves have been made in that direction. Serebrinsky *et al.*[48] have developed a model of intergranular hydrogen embrittlement "based upon: (1) a cohesive law dependent on impurity coverage that is calculated from first principles; (2) a stress-assisted diffusion equation with appropriate boundary conditions accounting for the environment; (3) a static continuum analysis of crack growth including plasticity; and (4) the Langmuir relation determining the impurity coverage from its bulk concentration." They concluded that "The calculations reproduce the following experimental trends: (1) time to initiation and its dependence on yield strength and stress intensity factor; (2) finite crack jump at initiation; (3) intermittent crack growth; (4) stages I and II of crack growth and their dependence on yield strength; (5) the effect of the environmental impurity concentration on the threshold stress intensity factor; and (6) the effect of temperature on stage II crack velocity in the low-temperature range."

References

1. *Hydrogen Degradation of Ferrous Alloys;* Oriani, R. A., Hirth, J. P., Smialowski, M., Eds.; Noyes Publications: Park Ridge, New Jersey, 1984.
2. Birnbaum, H. K.; Robertson, I. M.; Sofronis, P.; Teter, D. 2nd International Conference on Corrosion-Deformation Interactions, CDI '96, Nice, France, 24–26 Sept 1996, pp 172–195, 1997.
3. Lynch, S. In *Corrosion/2007*; NACE International: Houston, Tx, p 55 2007; Paper 07493.
4. Pasco, R. W.; Ficalora, P. J. *Hydrogen Degradation of Ferrous Alloys*; Oriani, R. A., Hirth, J. P., Smialowski, M., Eds.; Noyes Publications: Park Ridge, New Jersey, 1984; pp 199–214.
5. Imbihl, R.; Behm, R. J.; Christmann, K.; Ertl, G.; Matsushima, T. *Surf. Sci* **1982**, *117*, 257.
6. Bernstein, I. M.; Pressouyre, G. M. *Hydrogen Degradation of Ferrous Alloys*; Oriani, R. A., Hirth, J. P., Smialowski, M., Eds.; Noyes Publications: Park Ridge, New Jersey, 1984; pp 641–685.
7. Puls, M. P. *Hydrogen Degradation of Ferrous Alloys*; Oriani, R. A., Hirth, J. P., Smialowski, M., Eds.; Noyes Publications: Park Ridge, New Jersey, 1984; pp 114–130.
8. Asaoka, T.; Lapasset, G.; Aucouturier, M.; Lacombe, P. *Corrosion* **1978**, *34*, 39.
9. Kedzierzawski, P. *Hydrogen Degradation of Ferrous Alloys*; Oriani, R. A., Hirth, J. P., Smialowski, M., Eds.; Noyes Publications: Park Ridge, New Jersey, 1984; pp 271–288.
10. Devanathan, M. A. V.; Stackurski, Z. *J. Electrochem. Soc.* **1964**, *111*(5), 619.
11. Berman, D. A.; Beck, W.; Deluccia, J. J. *Hydrogen in Metals*; Bernstein, L. M., Thompson, A. W., Eds.; ASM, 1974; pp 575–607.
12. Arup, H. *Proceedings of the 9th Scandinavian Corrosion Congress* 1984, p 825.

13. Brass, A. M.; Chene, J. *Mater. Sci. Eng. A* **1998**, *A242*, 210–221.
14. Robertson, I. M. *Eng. Fract. Mech*. **2001**, *68*, 671–692.
15. See, for example, AMS 2401C, ASTM A165, ISO 2082, BS1706 and BS3382.
16. Echizen, Y.; Takada, K. *Kinsoku Hyomen Gijutsu* **1987**, *38*(3), 113–115.
17. Barth, C. F.; Troiano, A. R. *Corrosion* **1972**, *28*(7), 259–263.
18. Taqi, E. A.; Cottis, R. A. *Corrosion Chemistry within Pits, Cracks and Crevices*; Turnbull, A., Ed.; NPL: London, 1987; pp 483–494.
19. Lunarska, E.; Zielinski, A.; Smialowski, M. *Acta Met*. **1977**, *25*, 305–308.
20. Lunarska, E. *Hydrogen Degradation of Ferrous Alloys*; Oriani, R. A., Hirth, J. P., Smialowski, M., Eds.; Noyes Publications: Park Ridge, New Jersey, 1984; pp 321–352.
21. Sandoz, G. In *Stress Corrosion Cracking in High Strength Steels and in Titanium and Aluminium Alloys*; Brown, B. F., Ed.; Naval Research Laboratory: Washington, 1972; pp 79–145.
22. McIntyre, P. *Hydrogen Degradation of Ferrous Alloys*; Oriani, R. A., Hirth, J. P., Smialowski, M., Eds.; Noyes Publications: Park Ridge, New Jersey, 1984; pp 763–798.
23. Magnin, T. *Mater. Sci. Forum* **1996**, *217–222*, 83–94.
24. Shivanyuk1, V.; Gavriljuk, V.; Foct, J. *Mater. Sci. Forum* 2007, *539–543*, 4249–4254.
25. Cassagne, T.; Busschaert, F. In *Corrosion/2005*; NACE International: Houston, Tx, 2005; p 26; Paper No 05098.
26. Hippsley, C. A. *Mater. Sci. Technol*. **1987**, *3*(11), 912–922.
27. Johnston, J. W.; Cottis, R. A.; Procter, R. P. M. *Electrochem. Soc. Extended Abstracts* **1987**, *87*(2), 454–455.
28. Ryder, D. A.; Davies, T. J.; Strecker, E. 2nd International Congress on Hydrogen in Metals; 1977; Vol. 3; Pergammon Press; Paper 3B2.
29. Kerr, R.; Guiterrez-Solana, F.; Bernstein, I. M.; Thompson, A. W. *Met. Trans. A* **1987**, *18A*, 1011–1022.
30. Thompson, R.; Lin, I. H. *Hydrogen Degradation of Ferrous Alloys*; Oriani, R. A., Hirth, J. P., Smialowski, M.; Noyes Publications: Park Ridge, New Jersey, 1984; pp 454–511.
31. Zappfe, C.; Sims, C. *Trans. AIME* **1941**, *145*, 225.
32. Oriani, R. A.; Josephic, P. H. *Acta Metall*. **1974**, *22*, 1065.
33. Troiano, A. R. *Trans. ASM* **1960**, *52*, 54.
34. Morlet, J. G.; Johnson, H. H.; Troiano, A. R. *J. Iron Steel Inst*. **1958**, *189*, 37.
35. Oriani, A. *Stress Corrosion Cracking and Hydrogen Embrittlement of Iron Base Alloys*; NACE: Houston, TX, 1973; Vol. 5, pp 351–358.
36. Petch, N. J.; Stables, P. *Nature* **1952**, *169*, 842.
37. Gahr, S.; Grossbeck, M. L.; Birnbaum, H. *Acta Met*. **1977**, *25*, 125.
38. Sieradzki, K.; Newman, R. C. *Phil. Mag. A* **1985**, *51*, 95.
39. Morgan, M. J.; McMahon, C. J., Jr. *Hydrogen Degradation of Ferrous Alloys*; Oriani, R. A., Hirth, J. P., Smialowski, M., Eds.; Noyes Publications: Park Ridge, New Jersey, 1984; pp 608–640.
40. Lynch, S. *Scripta Met*. **1979**, *13*, 1051.
41. Delafoss, D.; Magnin, T. *Eng. Fract. Mech*. **2001**, *68*, 693–729.
42. Ugiansky, G. M.; Payer, J. H. Eds. *Stress Corrosion Cracking–The Slow Strain Rate Technique*; ASTM STP, 1979; p 665.
43. Erlings, J. G.; de Groot, H. W.; Nauta, J. *Corros. Sci*. **1987**, *27*(10/11), 1153–1167.
44. Crooker, T. W.; Hauser, J. A. H. A Literature Review on the Influence of Small-Amplitude Cyclic Loading on Stress Corrosion Crack Growth in Alloys, NRL Memorandum Report 5763, Naval Research Laboratory, Washington DC, April 3, 1988.
45. Caskey, G. R. Third International Congress on Hydrogen and Materials Paris, 1982, pp 611–616.
46. Gesari, S. B.; Irigoyen, B. L.; Juan, A. *Surf. Rev. Lett*. **2005**, *12*(2), 227–232.
47. Gesari, S. B.; Pronsato, M. E.; Juan, A. *Surf. Rev. Lett*. **2002**, *9*(3 & 4), 1437–1442.
48. Serebrinsky, S.; Carter, E. A.; Ortiz, M. *J. Mech. Phys. Solids* **2004**, *52*, 2403–2430.

2.11 Cracking Stimulated by Hydrogen

C. M. Fowler
Exova Corrosion Centre, 182 Halesowen Road, Netherton, Dudley, West Midlands DY2 9PL, UK

Abbreviations

EFC European Federation of Corrosion
HIC Hydrogen induced cracking
HPIC Hydrogen pressure induced cracking
SOHIC Stress orientated hydrogen induced cracking
SSCC Sulfide stress corrosion cracking

Symbols

HAZ Heat affected zone
HV10 Vickers hardness 10 Kg load
H_2S Hydrogen sulfide
MPa Mega pascals

2.11.1 Introduction

There are several forms of environmental cracking that are either stimulated or produced by hydrogen.

Over the past decade and in previous decades, there have been a significant number of pipeline failures attributed to such mechanisms. The types of cracking and compatibility of materials will be described in this chapter.

In an ideal world, high strength steels would be fully utilized in, for example, pipelines and offshore structures, the weight advantage being very attractive. However, in general terms, as the strength of steel increases so does the susceptibility to hydrogen embrittlement. This has been treated in the previous chapter, and only one variant of this process, which is experienced in harder welds in pipeline steels, will be discussed here.

In addition to the hydrogen embrittlement of harder steels, there are also hydrogen cracking mechanisms that occur in softer steels, such as hydrogen pressure induced cracking (HPIC), and these processes are the focus of this chapter.

Hydrogen embrittlement is clearly a function of dissolved atomic hydrogen, while HPIC clearly depends on the formation of molecular hydrogen within the steel (and hence depends on the presence of very high hydrogen fugacity). However, a crack morphology now recognized and termed stress orientated hydrogen induced cracking (SOHIC) has been defined, and modern thinking suggests that the crack mechanism is driven by a combination of atomic and molecular hydrogen. This latter crack mechanism has been and is still the subject of much research.

The three types of cracking described above have all been responsible for a significant number of pipeline and pressure vessel failures.

In an effort to provide a clear picture, the three crack mechanisms will be dealt with separately.

2.11.2 Sources of Hydrogen

The general sources of hydrogen have been described in the previous chapter; here we concentrate on the

production of the very high hydrogen fugacities that are necessary for HPIC and SOHIC. In order to achieve these high fugacities, it is necessary to generate hydrogen at the metal surface at an activity that would be in equilibrium with a very high partial pressure of hydrogen (far above atmospheric pressure). The only common way in which this can be achieved is by corrosion in the presence of hydrogen sulfide (H_2S). The hydrogen sulfide may be naturally-occurring in oil and gas reservoirs (produced fluids may contain more H_2S than oil, with the highest reported H_2S content being 87%[1]) or from reservoir souring as a result of infection by sulfate reducing bacteria introduced during water injection.

Figure 1 illustrates the processes that are involved in the production of hydrogen gas at the steel surface. In most environments, the vast majority of the hydrogen is evolved as gas and the amount entering the steel is very small (typically less than 1%).

The mechanism whereby the atomic hydrogen enters the steel in great quantities depends on the poisoning of the recombination reaction in which two adsorbed hydrogen atoms combine to form a hydrogen molecule that then disperses as dissolved hydrogen or as bubbles of gas. This causes the adsorbed hydrogen concentration to increase and consequently, more hydrogen enters the steel.

While several species act as poisons for hydrogen recombination, H_2S is by far the most common source of problems, largely because of its presence in oil and gas production. It is considered that the presence of the HS^- from disassociated hydrogen sulfide is the primary poison. Thus, in the presence of hydrogen sulfide, it has been estimated that more than 90% of the atomic hydrogen produced can enter the steel substrate.

It has also been suggested that the HS^- ion acts as a catalyst for hydrogen entry, and details of the mode of action of H_2S remain to be clarified. However, the facts are that in the presence of hydrogen sulfide or other particular compounds such as cyanides, atomic hydrogen can enter steel in vast quantities, whereas under highly corrosive conditions such as corrosion in hydrochloric acid, where copious quantities of gaseous hydrogen is produced, hydrogen atoms do not enter the steel anywhere near such amounts.

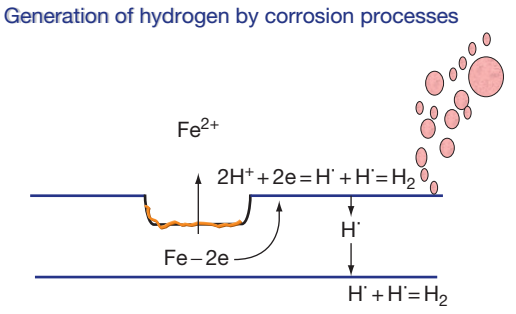


Figure 1 The corrosion process.

2.11.3 Hydrogen Pressure Induced Cracking (HPIC)

This phenomenon is also called blister cracking, stepwise cracking, or simply hydrogen induced cracking (HIC). HPIC is far more descriptive as it is the pressure of hydrogen that causes the cracking as described later.

Atomic hydrogen enters the steel substrate and will continue in theory to the other side; however, the atoms collect at sites such as metal-inclusion interfaces, voids, and other crystal defects. While hydrogen is essentially insoluble in metals in a molecular form, hydrogen molecules can form at metal-inclusion interfaces, leading to the development of internal pockets of hydrogen. Further, atoms are attracted to this site, form more hydrogen molecules, and a gas pressure builds up. This pressure can reach many MPa, and eventually, the yield strength of the steel is exceeded and cracks are initiated. As more hydrogen enters the steel and further small cracks initiate, some of the cracks join together in a shear mode. These often look like steps, hence, the name stepwise cracking, see Figures 2 and 3.

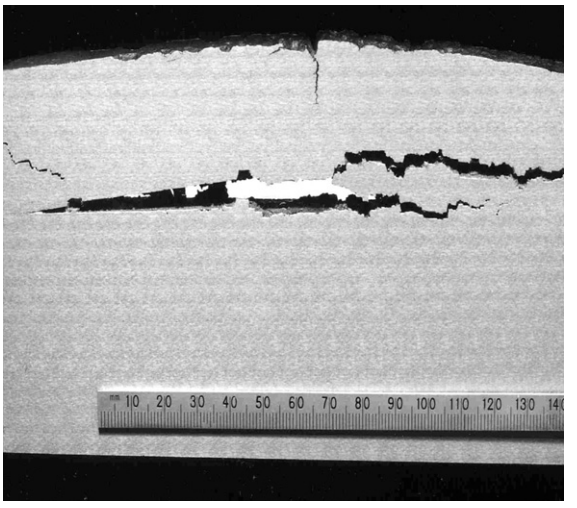


Figure 2 Example of HPIC in Plate Steel (courtesy Hay, M. Shell Canada).

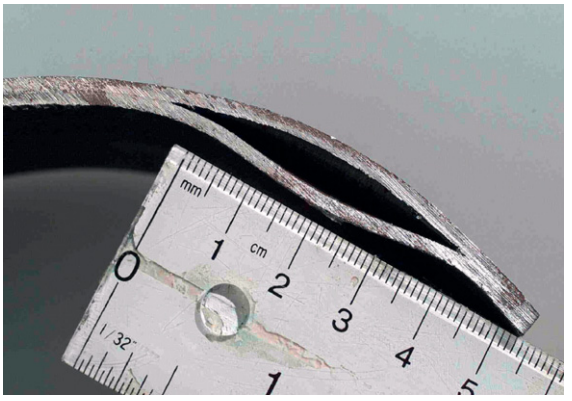


Figure 3 HPIC in pipeline steel (courtesy Hay, M. Shell Canada).

During the era of 'controlled rolled plate,' such features were not uncommon and a number of failures occurred.

Steels have now been developed with improved microstructure, achieved using calcium treatment which minimizes the inclusions and banded microstructure that are the sites for HPIC initiation. Having said that, HPIC failures are still occurring when high levels of H_2S are encountered, thus, steels for use in sour environments should be tested – see Section 2.11.6.

HPIC can occur in low strength steels with hardness levels well-below the recognized level needed for hydrogen embrittlement to occur.

2.11.4 Hydrogen Embrittlement

The first distinction to be made is that hydrogen embrittlement is an atomic hydrogen mechanism not a molecular hydrogen mechanism. Hydrogen embrittlement is a particular problem in the oil and gas industry in the context of welds, where higher hardness than the parent plate can be obtained, together with high residual stresses. Owing to the high hydrogen levels that can be achieved by the action of H_2S, failure can occur at somewhat lower levels than is normally found for hydrogen embrittlement. When it occurs in the presence of H_2S, the phenomenon is known as sulfide stress corrosion cracking (SSCC) or sulfide stress cracking (SSC).

In general terms, the higher the strength and hardness of a steel the higher the susceptibility to this form of cracking. The industry recognized upper hardness limit as Vickers HV_{10} 248 or Rockwell 'C' scale 22, see Figure 4.[2]

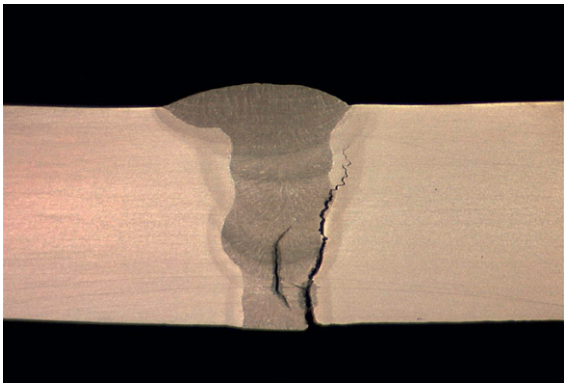


Figure 4 Hydrogen embrittlement (SSCC) of Pipeline weld.

2.11.5 Stress Oriented Hydrogen Induced Cracking (SOHIC)

The morphology of this type of crack is illustrated in Figure 5.

This type of cracking has also been termed ladder cracking or stacked array cracking.

Pargeter has presented a documentary history of SOHIC failures[3]; the evidence strongly suggests that a complex or tri-axial stress condition is required to generate such cracking, in combination with a high flux of hydrogen. It is important to note that the stress can be residual rather than applied.

The first SOHIC failures occurred in spirally welded line-pipe (Figure 6). There have been a number of catastrophic bursts.

By virtue, its manufacturing process, spirally welded pipe has a high residual stress adjacent to the weld. SOHIC cracks are often positioned adjacent to the heat-affected zone (HAZ) of the weld.

SOHIC has also been identified in failures in pressure vessels, and in one case a seamless pipeline.

2.11.5.1 SOHIC Mechanism

At the time of writing, there are a number of ongoing research programs to try to fully identify the SOHIC mechanism, and one to develop an acceptable test method.

It is generally considered that the following events occur: the steel is flooded with hydrogen from the corrosion process, there is a tri-axial stress, voids are generated, and small internal blister type cracks initiate by the HIC mechanism of molecular hydrogen being formed and trapped. These small cracks themselves generate internal stress fields which further attract atomic hydrogen, further small

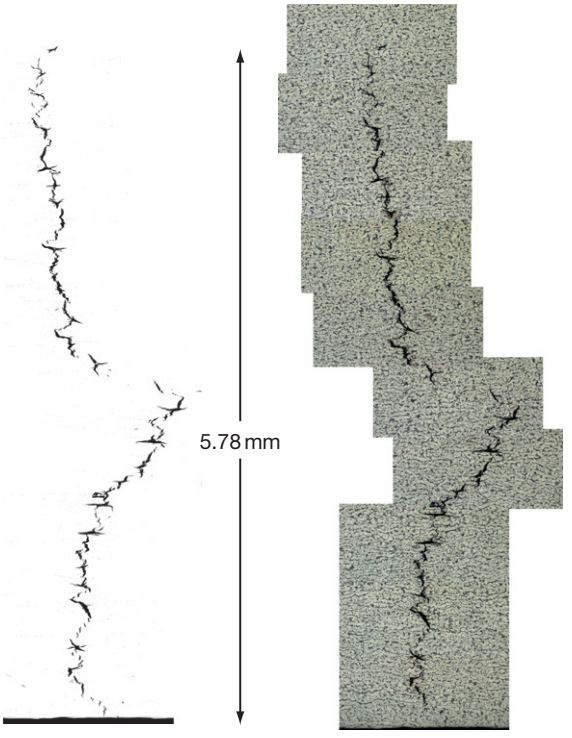


Figure 5 Stress-oriented hydrogen induced cracking (SOHIC).

Figure 6 Spirally welded pipe – SOHIC failure (courtesy Shell Canada).

cracks initiate in the through thickness direction, which heightens the stress level, atomic hydrogen then leads to cracking by processes similar to those involved in hydrogen embrittlement, the resultant cracks joining up the small HIC cracks. The eventual failure is by coalescence of the through wall features.

Some workers in the field are also of the opinion that the initial cracks could be formed around small spherical inclusions. In some instances, these have been found in the center of the crack. The definitive mechanism whereby the cracks form in a vertical array is yet to be discovered.

Current research has identified the overall importance of a tri-axial stress, and a test method that applies a complex stress field to the sample is in the late stages of refinement.

It should be noted that SOHIC is not limited to line-pipe, it has been observed in plates and fittings. Interestingly, the materials in question had hardness significantly below the level accepted for SSCC/hydrogen embrittlement resistance.

2.11.6 Testing Methods for Hydrogen Cracking

2.11.6.1 Hydrogen Pressure Induced Cracking

An industry accepted test method has been developed and used for many years, namely NACE TM0284.[4] The method is quoted in many standards such as ISO 15156[2] and EFC publication No. 16.[5]

It is important to note that ISO 15156, the industry standard for sour oil and gas production, notes that all tests should be undertaken on samples representative of the final product form.

Although stress is thought not to influence HPIC, it has been noted several times by various workers that residual strain can be detrimental, as such tests on plate have shown no failure, whereas tests in pipe made from the plate have exhibited failure when the outer strain level reached 2% or more.

The test is a simple immersion test of unstressed coupons which are exposed to H_2S saturated brine. After a set exposure time, the samples are metalographically examined for cracks.

2.11.6.2 Hydrogen Embrittlement

Considerable further work has been undertaken since Erlings' work described in the previous chapter,[6] and there are now several variations in loading technique, such as interrupted slow strain and step loading slow strain rate testing.

One of the more common loading schemes is illustrated in **Figure 7**:

Further information on this type of testing can be found in ASTM F 1624–06.[7]

Other published test methods include:

NACE TM0177[8] methods A, B, C, and D; method 'A' smooth tensile, method 'B' three point bend (now largely superseded by other configurations), method 'C' ring, method 'D' double cantilever bend.

EFC 16[5] Four Point Bend (ASTM G39)_Mostly applicable to Weld qualification.

NACE TM0198[9] Slow Strain Rate

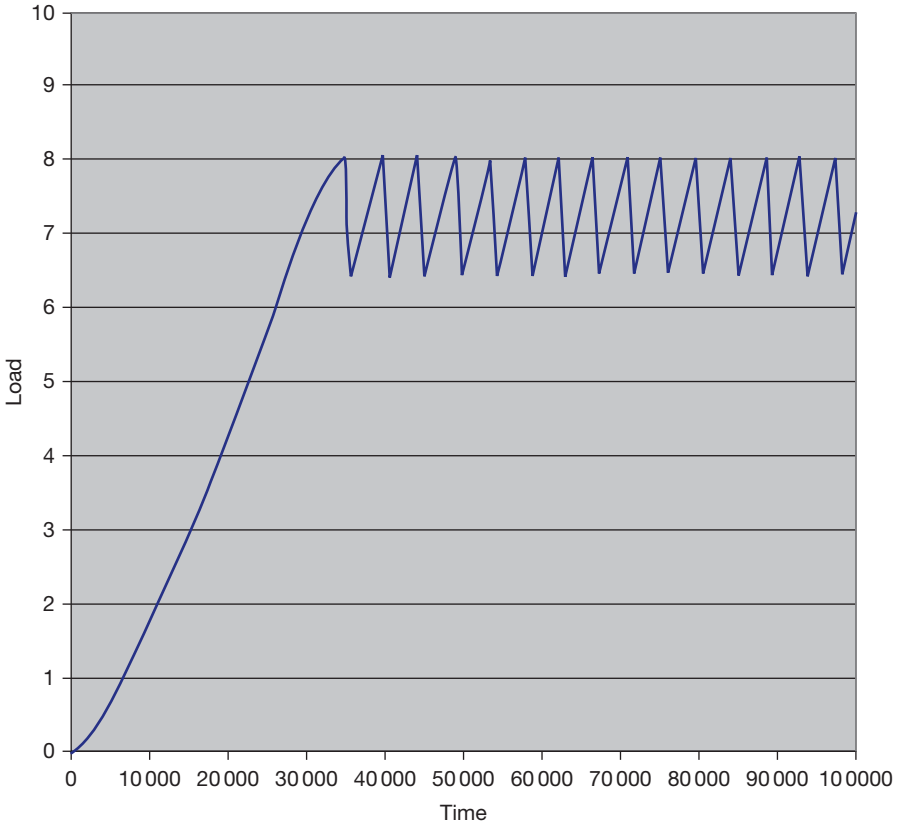


Figure 7 Load-time history for a cyclic slow strain rate test.

2.11.6.3 Testing for SOHIC

At the time of writing, there is currently no universally accepted test for SOHIC. There are two published methods, one for pressure vessel plate (NACE TM0103[10] the double beam test, not to be confused with the double cantilever beam test; under review for withdrawal) and one for line-pipe (HSE Publication OTI 95 635[11] the full ring test).

References

1. Kutney, G. *Sulfur: History, Technology, Applications and Industry*; ChemTec Publishing, 2007; p 146, ISBN 1895198372, 9781895198379.
2. ISO 15156/MR0175; NACE International: Houston, TX.
3. Pargeter, R. J. In *Corrosion 2007*; NACE International: Houston, Tx, 1975; Paper 07115.
4. NACE TM0284; NACE International: Houston, TX.
5. *Materials Requirements for Carbon and Low Alloy Steels for H_2S-Containing Environments in Oil and Gas Production EFC 16*, 2nd ed., 2002; pp 64; Maney Publishing: London, ISBN 1 90265 354 8.
6. Erlings, J. G.; de Groot, H. W.; Nauta, J. *Corros. Sci.* **1987**, *27*(10/11), 1153–1167.
7. ASTM F1624–06 *Standard Test Method for Measurement of Hydrogen Embrittlement Threshold in Steel by the Incremental Step Loading Technique*; ASTM International: Philadelphia, 2006.
8. NACE TM0177; NACE International: Houston, TX.
9. NACE TM0198; NACE International: Houston, TX.
10. NACE TM0103; NACE International: Houston, TX.
11. HSE OTI 95–635 *The Full Ring Test*; UK Health and Safety Executive – Books.

2.09 Stress Corrosion Cracking

R. C. Newman
Department of Chemical Engineering and Applied Chemistry, University of Toronto, Canada

Abbreviations

AIDE Adsorption-induced dislocation emission
BWR Boiling water reactor
C steel Carbon steel
Cl-SCC Chloride-induced SCC (of stainless steel)
EAC Environmentally assisted cracking
EBSD Electron backscattered diffraction
EDS Energy dispersive analysis of X-rays
FIB Focused ion beam
GB Grain boundary
HEDE Hydrogen-enhanced dislocation emission
HELP Hydrogen-enhanced local plasticity
HS High strength (steel)
HSLA High strength low alloy (steel)
HT High temperature (water)
HWC Hydrogen water chemistry (in BWR)
IASCC Irradiation-assisted SCC
IG Intergranular
IGA Intergranular attack
LA Low alloy (steel)
LME Liquid metal embrittlement
LT Low temperature (water)
MS Medium strength (steel)
PWR Pressurized water reactor
Q&T Quenched and tempered (steel)
RIS Radiation-induced segregation
SCC Stress corrosion cracking
SDM Slip dissolution model (for SCC)
SMM Surface mobility model
SS Stainless steel
SSRT Slow strain rate test
TEM Transmission electron microscope (or microscopy)
TG Transgranular

Symbols

a Atom diameter (in the SMM)
a_c Critical depth of pit or crack
A Atom fraction of less-noble metal atoms in an alloy
D_s Surface diffusion coefficient (in the SMM)

E Electrode potential, Young's modulus
E_m Mixed potential
E_R Repassivation potential
f Oxide-film fracture strain
F Faraday's constant
i Current density
$\bar{i}$ Mean anodic current density flowing during a time interval
i_0 Exchange current density
i_{max} Maximum anodic current density flowing on a bare metal surface
i_{net} Net current density
$K_{(I)}$ (Mode I) Stress intensity factor
K_{Ic} Critical stress intensity factor
K_{Iscc} Threshold stress intensity factor for SCC
L Diffusion distance along the crack flank (in the SMM)
L Distance ahead of the crack tip where the crack-tip strain rate is evaluated
n Number of electrons transferred per atom oxidized; inverse strain hardening coefficient
N_A Avogadro's number
Q^{-1}, f Quantities of interest in internal friction measurement
s Crack-tip strain rate
v Velocity of SCC or corrosion (or V_{scc}, V_{crack}, $V_{corrosion}$, etc.)
δ Crack-tip opening displacement according to ordinary fracture mechanics
ε_p Plastic strain
$\varepsilon(r)$ Plastic strain distribution ahead of a crack tip
Ω Molar volume of metal
σ Stress
σ_y Yield stress
τ Interval between film rupture events at a crack tip

2.09.1 Introduction and Definitions

Stress corrosion cracking (SCC) is the gradual growth of cracks under the simultaneous influence of a nominally static stress and a chemically reactive environment. The meaning of 'nominally static' is discussed in **Chapter 2.08, Environmentally Assisted Cracking Overview** – this does not exclude slow or low-amplitude dynamic loading, or periodic unloading. It is convenient to refer to SCC as occurring under 'tensile' stress, although shear stresses certainly play a major role microscopically, and some SCC testing is done under mode III loading (antiplane shear; torsion), to confirm whether or not hydrogen is playing a role in the cracking – hydrogen is supposed to be less aggressive under torsion because there is, nominally, no lattice dilation to increase the solubility of hydrogen. In this chapter, we restrict ourselves to metals, although interesting SCC phenomena occur in glasses and hot rocks, and are thought to be involved in major geologic upheavals.[1] Actually, fracture of brittle solids can be more amenable to first-principles calculations than the more complex phenomena that occur in metals, so in some ways nonmetals are at the leading edge of the subject.[2] Polymers and polymer–matrix composites also show environmentally assisted cracking (EAC) phenomena that some people would call SCC.

In metals, SCC cracks are normally multiple, branched, and either intergranular, or transgranular and cleavage-like (see **Chapter 2.08, Environmentally Assisted Cracking Overview**, for examples). Sometimes the two modes coexist, or transgranular SCC occurs simultaneously with intergranular corrosion – **Figure 1**. This fractography, when examined in sufficient detail, continues to provide important clues to the cracking mechanism, whether in the laboratory or in the field.[3,4] In particular, transgranular SCC facets are close to a particular fracture plane, for example, {110} in α-brass, but have ultrafine detail indicative of the operation of plastic processes on the nanometer scale.

SCC can be considered a 'brittle' mode of failure, in that the material fails to manifest its usual tensile strength and ductility. But SCC is always accompanied by microscopic plastic strain, and many believe that such strains, occurring dynamically at the crack tip, are a causal factor of crack growth.

It is easier to pose the question – what kinds of EAC are not SCC – than to list pedantically what is SCC. So, hydrogen-induced cracking of a high-strength steel due to gaseous hydrogen, or to excessive cathodic protection, is not SCC. But cracking of the same material in saltwater under free corrosion conditions is SCC, even if we are 100% sure it is due to hydrogen released by the corrosion process. Liquid metal embrittlement (LME) is not SCC. Cracking of high-strength steel in dry gaseous chlorine[5] is SCC, but cracking in dry gaseous H_2S might be too obviously a hydrogen effect to be SCC. Cracking of carbon steel in anhydrous liquid ammonia is certainly SCC. Cracking of alloys in air at very high temperatures where dry oxidation and creep are both possible could, arguably, be called SCC,[6] although such phenomena have other names such as oxygen embrittlement. But despite these nonaqueous

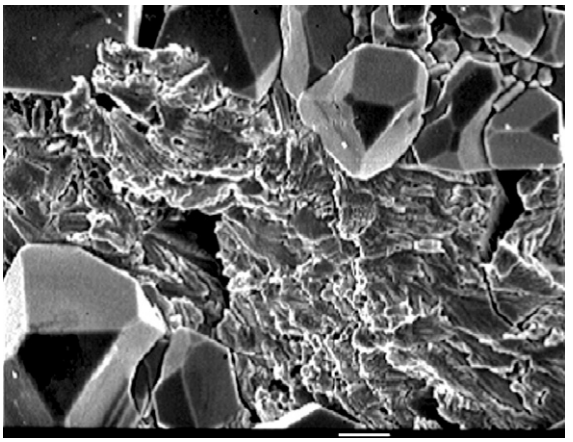


Figure 1 SCC of Monel 400 in copper sulfate solution, showing intergranular corrosion/SCC and transgranular SCC. Reproduced from Barnes, A.; Deakin, J.; Newman, R. C. *Corrosion* **2007**, *63*, 416–418.

examples, the vast majority of SCC is caused by some combination of water, oxidant, and ions or soluble molecules. And like other forms of corrosion, most SCC can be retarded or prevented by removing oxygen.

The velocity of SCC (crack growth rate) can be as high as 10^{-5} m s^{-1} or as low as 10^{-12} m s^{-1}. Common practical rates are in the range 10^{-11}–10^{-8} m s^{-1} – roughly, 1 μm to 1 mm per day. In real components, cracking is often intermittent, occurring during start-ups or other plant transients, which can make the apparent or average crack velocity very low indeed.

The consequences of SCC are quite different in low-strength and high-strength alloys. In a high-strength material, stresses are usually externally imposed, and growth of a stress corrosion crack can lead to catastrophic rupture when the fracture toughness is exceeded, whereas in low-strength materials, the dominant stresses are often internal (residual), and SCC causes leakage or gradual plastic tearing.

The history of SCC has often been told, and will be mentioned quite briefly. In this chapter, we focus on modern understanding of the various phenomena that are classified as SCC and their practical significance.

2.09.2 Metal–Environment Combinations Showing SCC; the Issue of Mechanical Difficulty

The major difference between different SCC phenomena is the severity of the mechanical conditions that are required (e.g., value of stress normalized to the yield stress or necessity of time variation in stress). Such considerations, and the notion of a spectrum of phenomena from SCC through corrosion fatigue, depending on the importance of dynamic straining, were introduced in **Chapter 2.08, Environmentally Assisted Cracking Overview**.

Table 1 gives a summary of the known SCC systems with one new feature – we try to categorize them according to the difficulty of cracking in a real component, on a scale of 1–5. So, 1 means SCC that requires only a low and not particularly dynamic tensile stress, while 5 means SCC that has only been observed under laboratory conditions and requires dynamic tensile loading (slow strain rate testing), or major cyclic loading, or alternatively an improbably oxidizing potential. Of course, there is always a first time for a phenomenon to occur in the field, but at least we can be sure that a 4 or a 5 would require unusually severe conditions, and might be classified as corrosion fatigue in any case.

An explosion in the number of known SCC combinations (material ... metallurgical condition ... environment) occurred when the slow strain rate test (SSRT) was introduced in the 1970s.[7] This test exposes a smooth tensile sample to the environment, often with electrochemical potential control, and the sample is slowly strained to failure at an engineering strain rate of typically 10^{-7}–10^{-5} s^{-1}.[8] The use of the SSRT revealed, almost for the first time, SCC processes that had not yet occurred in the field, such as the cracking of pure copper in sodium nitrite solution.[9] Some of these subsequently did occur in the field; others never have. To call the SSRT 'too severe' misses the point – it is a tool for predicting and understanding SCC. No one thinks that all SCC observed in SSRT can or will occur under realistic component loading. Sometimes, especially for hydrogen-related SCC, one can rationalize this neatly in terms of when, during a SSRT, cracks initiate and cause the load-extension curve to deviate from that measured in an inert environment – Figure 2.

Some cases of SCC are hard to predict and control because the environment that causes cracking is not well-characterized, or exists only transiently. This is the case for nuclear steam generators, where impurities present at parts per billion or trillion in the feedwater (Pb ... S) concentrate on heat-transfer surfaces under deposits[10] – Figure 3. Sometimes these cause SCC during operation, but it is increasingly thought that cracking may occur during plant

start-up, or even during plant shut-down, when oxygen is more available. Test methods that incorporate heat transfer into the test specimen are useful for the clarification of such mechanisms.[11]

The more usual test methods for SCC have been very well discussed by Sedriks, and are not given their own special section in this chapter; miscellaneous comments on test methods are inserted into

Table 1 Summary of metal–environment combinations that are susceptible to SCC

Material	*Environment*	*Exacerbating or **inhibiting** factors*	*Severity (scale 1–5) 1 highest*	*Other notes*
Carbon steel	Hot conc aq caustic O_2 Active–passive transition	GB precipitation, segregation C, P . . . Evap cond's	2	IG
Carbon steel	Warm-hot conc aq nitrate Nitrate can be the oxidant 'Transpassive' state	GB precipitation, segregation C, P . . . Evap cond's	1	IG
Carbon steel	Warm aq carbonate–bicarbonate O_2 Active–passive transition	GB precipitation, segregation C, P . . .	3	IG
Carbon steel	Anhydrous NH_3 with O_2 Ambient *T* Also NH_3/CH_3OH	***H_2O***	3	TG Also pure Fe
Carbon steel	Aq CO–CO_2 Ambient *T* O_2 not required?		3	TG Also pure Fe
Carbon steel	Aq CO_2–HCO_3^- near-neutral Ambient *T* O_2 not required?	Higher strength ***Dynamic strain required in, e.g., X65 grades***	4	TG H effect
Carbon steel	Methanol [ethanol] O_2 Ambient *T*	?	3 [ethanol 4]	IG TG with Cl^-
Carbon steel	Warm aq amine containing mixtures (NH_3?) O_2 not required?	CO_2	3	IG/TG H effect in HS
Carbon steel	HT water ($>150\,^\circ C$) O_2	Ionic impurities Cu deposition MnS inclusions	3	TG
Q&T LA steels MS	Hot water, steam Some O_2	GB segregation, e.g., P Ionic impurities	2–4	IG
Carbon steel Q&T LA steels MS	Aq CO_2–H_2S No O_2	Higher strength ***Fewer large carbides*** ***$T>60\,^\circ C$***	1–4	TG/IG H effect
HS C steel HSLA steels	Moist air. . .Water. . .Saltwater O_2	Higher strength ***$T>60\,^\circ C$*** ***Various metallurgy***	1–3	IG/TG H effect
HSLA steels HSLA steels	Dry Cl_2 Dry H_2S	***Only in ultra-HS*** ***O_2***	3 2	IG IG/TG H effect
HS carbon steel (prestressing wire)	Water O_2	Salts Lack of high pH ***Elongated microstructure favorable***	3	IG/TG
Strong SS Martensitic, etc.	Salt water O_2 Ambient *T*	H_2S Other reduced sulfur Crevices, deposits Higher strength ***$T>?$***	2	TG/IG H effect Requires loc corr

Continued

Table 1 Continued

Material	*Environment*	*Exacerbating or **inhibiting** factors*	*Severity (scale 1–5) 1 highest*	*Other notes*
Austenitic SS	Warm-hot aq Cl^- (duplex requires higher *T*)	H_2S	1	TG
Duplex SS	O_2 (trace)	Other reduced sulfur Evap cond's Crevices, deposits Mg, Ca (evap cond's) Cold work (lower Ts) Intermetallics (duplex or high austenitic) ***Ni > 12% (absolute resistance for Ni > 40%)*** ***Silicate or other inhibitors*** ***Low free P (not commercial)***		Requires loc corr
Austenitic SS Duplex SS	Hot aq caustic (needs hotter than Cl^-) Occurs without O_2 in lab Duplex less sensitive	Evap cond's	1	TG
Austenitic SS	HT water with O_2 e.g., 290 °C	Cold work Neutron irradiation Ionic impurities Crevices	2–4	IG/TG
Austenitic SS	HT water without O_2 e.g., 290 °C	Cold work Neutron irradiation Ionic impurities?	3–5	IG/TG
Austenitic SS Sensitized	HT water with O_2 e.g., 290 °C Occurs as low as 80 °C	Cold work Neutron irradiation Ionic impurities Crevices	1	IG
Austenitic SS Sensitized	HT water without O_2 e.g., 290 °C	Cold work Neutron irradiation Ionic impurities?	3–5	IG
Austenitic SS Sensitized (also Alloy 600 sensitized)	Aq solutions of reduced sulfur species – $S_2O_3^{--}$, $S_4O_6^{--}$, $FeS/O_2/H_2O$ (others?) Ambient *T*	***Excess strong-acid anions → IGA/pitting*** ***Alloyed Mo strongly beneficial***	1	IG
Austenitic or duplex weld metal	All the above	Complex; generally more resistant than annealed material; can be similar to duplex	3–5	Complex
Ni base alloys e.g., Alloy 600	HT water without O_2 e.g., 310 °C	H_2 in certain range near Ni/NiO equilibrium Cold work ***Grain boundary carbides*** ***Switch to Alloy 800 or 690***	2	IG
Ni base alloys e.g., Alloy 600	HT water with O_2 e.g., 310 °C	Cold work Neutron irradiation Ionic impurities Crevices	2	IG
Ni base alloys e.g., Alloy 600	Complex HT environments (nuclear SG crevices, etc.)	O_2 or other nonreducing condition during start-up Deposits Boiling Pb, reduced S from sulfate reduction	2	IG

Continued

Table 1 Continued

Material	*Environment*	*Exacerbating or **inhibiting** factors*	*Severity (scale 1–5) 1 highest*	*Other notes*
		Switch to Alloy 800 or 690 ***Clean deposits*** ***Design out heat-transfer crevices***		
Corrosion resistant Ni base alloys e.g., 625	Cl^-/H_2S [S] HT e.g., 200 °C No O_2	***Extreme conditions only***	3	IG
Strong Ni base alloys (γ')	HT water LT aq solutions	Highest strength normally most susceptible.	2–3	IG/TG H effect at low T
Monel 400 (70Ni–30Cu)	Hot water, or aq sulfate esp with Cu^{++}	Coldwork	3–4	IG/TG
Monel 500 (high strength)	Sulfide containing seawater	Highest strength normally most susceptible.	3–4	IG
Pure or nearly pure Cu	Aq ammonia Ambient *T*	Thin-layer exposure High P content Cold work	3–4	IG
Pure Cu	Aq nitrite O_2 (possibly not essential) Ambient *T*	***Needs dynamic strain*** (?always)	3–4	TG
Pure Cu	Aq acetate and others inc complex industrial Ambient *T* O_2	Cold work	4–5	IG [TG]
Pure Cu	Warm aq sulfide No O_2	? ***Only one report***	5	IG
Pure Cu	Hot (e.g., 300 °C) dry halogens	***Only shown in dynamic-load testing***	4	IG
α Brass	Aq ammonia O_2 Ambient *T* Also amines	Thin-layer exposure pH ~ 7 or ~ 11 Cold work	1	IG/TG
α Brass	Aq nitrite Ambient *T* O_2 (possibly not essential)	Synergistic effect of other anions e.g., MoO_4^{--} ***Difficult without dynamic strain?***	2–3	TG
α Brass	Aq sulfate, nitrate ($CuSO_4$, $Cu(NO_3)_2$) Ambient *T*		3	TG
α Brass	Other aq – acetate, chlorate	***Difficult without dynamic strain***	4	IG/TG
Al bronze	Aq ammonia O_2 Ambient *T*	***Borderline Al content for SCC***	3	IG/TG
β Brass 'Mn bronze'	Water, aq solutions Ambient *T*		3	TG
Other Cu base inc strong alloys (Cu–Ti, Cu–Be)	Aq ammonia O_2 Ambient *T*		3	IG/TG
Strong Ti alloys	Salt water O_2 Ambient *T* (separate phenomena at HT)	Sharp preexisting flaws	2–4	TG
Ti and simple alloys	Acids, methanol Ambient *T*	***H_2O (for methanol)*** ***In acids, etc. requires severe mechanics***	2–4	TG H effect (probably)

Continued

Table 1 Continued

Material	*Environment*	*Exacerbating or **inhibiting** factors*	*Severity (scale 1–5) 1 highest*	*Other notes*
Ti, Zr and alloys	Fuming or v hot HNO_3 Liq N_2O_4		3–4	TG
Ti, Zr and alloys	Dry halogens	***Severe mechanics required***	3–4	TG
Ti, Zr and alloys	HT water Cooling after hydrogen absorption at temperature	High pH aggressive for Zr H absorption	3–4	TG Hydride cracking
Strong Al alloys 7000, 2000, 8000	Aq chloride (water only in extreme cases) Ambient *T* O_2	Under- or peak ageing Cu alloying generally ***Overageing*** ***Flat grain structure in sheet***	2	IG
Medium strength 7000Al (weldable)	Aq chloride Ambient *T* O_2	Weld toe cracking	3	IG
Medium strength 5000Al	Aq chloride Ambient *T* O_2	GB intermetallic precipitation	3–4	IG
Various Al	Organics e.g., chlorinated	Poorly understood – also other reactive metals	3–4	IG
MgAl alloys	Aq chloride Ambient *T*	***Requires fairly severe mechanics***	3–4	TG
Au alloys – AuCu, AuAg – also Pt alloys	Aq acids (also some neutral solutions) Strong oxidant or complexant required for AuAg	***Requires dealloying – stops at parting limit***	1	IG TG in single crystals

C steel, Carbon steel; GB, Grain boundary; HS, High strength (steel); HSLA, High strength low alloy (steel); HT, High temperature (water); IG, Intergranular; IGA, Intergranular attack; LA, Low alloy (steel); LT, Low temperature (water); MS, Medium strength (steel); Q&T, Quenched and tempered (steel); SCC, Stress corrosion cracking; SS, Stainless steel; TG, Transgranular.

the text as appropriate. For completeness, we note that test methods can be roughly classified as follows:

- Type of specimen
  Smooth … defected (notch … precrack).
- Severity of loading
  Static (constant load … constant deflection) … Dynamic (monotonic … slow cyclic).
- Loading mode
  Tension … bending … torsion … multiaxial.
- Environment exposure
  Simple immersion … electrochemical control … thin layer/evaporating/heat transfer … gaseous environment.

2.09.3 Role of Metallurgy in SCC

2.09.3.1 Summary of Metallurgical Factors in SCC

Not all SCC requires a special metallurgical condition – the transgranular cracking of pure copper in sodium nitrite solution, of pure iron in anhydrous liquid ammonia or $CO–CO_2–H_2O$, or of pure zirconium in hot fuming nitric acid, is a property of the pure metal lattice. But generally, there is some exacerbating metallurgical factor, such as:

- Unfavorable solid solution composition, allowing selective dissolution of one or more less-noble alloy components (dealloying – see **Chapter 2.05, Dealloying**).
- Presence of grain boundaries, especially those with grain-boundary segregation of solutes or impurities.
- Presence of reactive phases at grain boundaries.
- Presence of solute-depleted regions adjacent to grain boundaries.
- Cold work and hardening in general.
- Features tending to localize strain – dynamic strain ageing, intense deformation bands, dislocation 'channeling' in neutron-irradiated materials.

Conversely, metallurgy may mitigate the possibility of SCC, as in duplex stainless steels, where the different electrochemical properties of the ferrite and austenite phases dictate different potential dependences of SCC

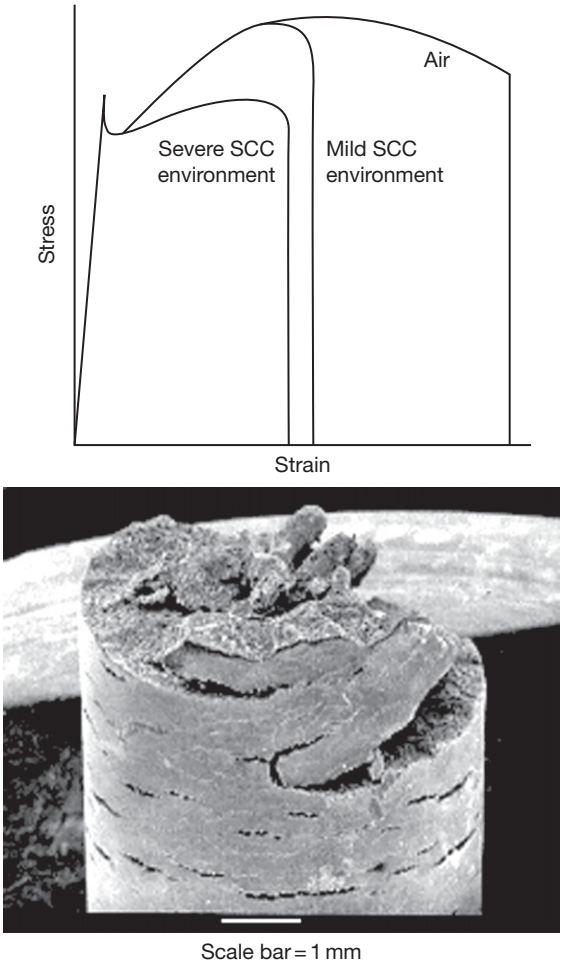


Figure 2 Schematic stress–strain curves obtained on carbon steel using the slow strain rate test: in air, in an environment causing severe SCC (e.g., anhydrous ammonia–methanol – see micrograph, courtesy W. Zheng), in an environment where cracking requires severe straining of the sample, to the point that necking has begun before cracks appear (e.g., aqueous CO_2 solution).

for each phase – and since the alloy can only have one corrosion potential, this is not (usually) a value where both phases can crack rapidly – Figure 4. This topic was discussed in some detail in another review,[12] so will not be elaborated here. Instead, we focus on a more recent development of SCC mitigation by so-called grain boundary engineering.

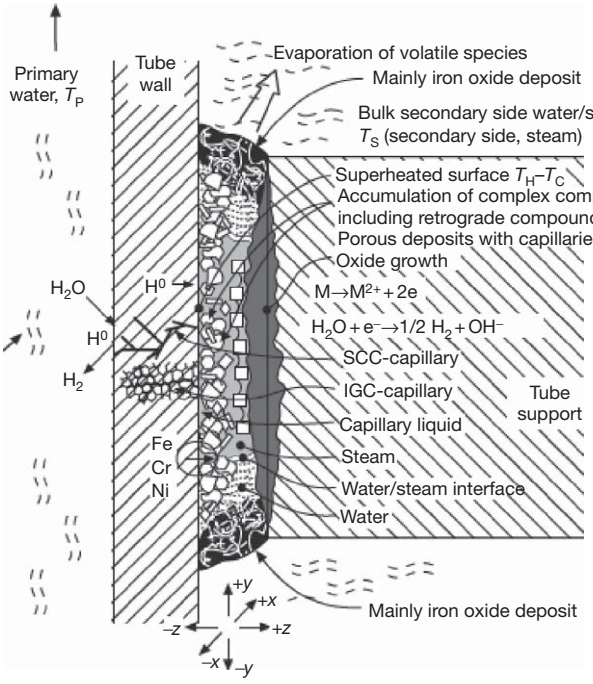
2.09.3.2 Solid Solution Composition

There is persuasive evidence that SCC of many solid-solution alloys involves dealloying (see **Chapter 2.05, Dealloying**). In the 1950s, H.R. Copson produced his famous 'Copson curve' for time to failure as a function of nickel content in chloride-induced SCC (Cl-SCC) of austenitic stainless steels, or duplex at the lowest Ni contents.[13] Around the same time, L. Graf produced very similar composition dependences for noble-metal alloys (CuAu, AgAu) where gold played (on a dealloying hypothesis) the same role as nickel.[14] SCC vanishes near the parting limit for dealloying, which is ~45 at.% of more-noble metal. The similarity was noticed by Latanision and Staehle,[15] who proposed that nickel was enriched on the crack walls and protected them in lieu of a passive film, along the same lines as Graf (nowadays, we would suggest a more direct causal role of dealloying at the crack tip – see Section 2.09.10). This very attractive hypothesis was later abandoned, apparently because the new surface analytical technique of Auger electron spectroscopy failed to detect nickel enrichment on fracture surfaces (Staehle, private communication). In fact, there is nickel enrichment, but it is subtle, and there is simultaneous dissolution of the alloy components.[16] Importantly, the state of the crack walls in Cl-SCC is precisely that of active crevice corrosion, so one need not be distracted by considerations of passive films.[17] There is some evidence, from cross-sectional TEM examination, for dealloying at the crack tip.[18]

The Copson and Graf curves are compared in Figure 5, adapted from Newman.

In strong, hot caustic solutions, dealloying of stainless steel is much more blatant than in chloride.[19] A dealloyed layer formed in a caustic environment that induces SCC is shown in Figure 6. Also, the dependence of caustic SCC on nickel content is similar to that of Cl-SCC – for example, SCC occurs readily in type 310 stainless steel (20% Ni), but requires more extreme conditions in Alloy 800 (33% Ni) – but when it does occur, it is spectacular – see Figure 2b in **Chapter 2.08, Environmentally Assisted Cracking Overview**. The dealloying of 310SS and its ability to inject a crack into the ductile substrate are shown in Figure 7.

Latanision and Staehle[15] noted that only a small Ni content, ca. 1%, was required for SCC, and at first sight this seems to favor the Graf type of mechanism, rather than one requiring an actual dealloyed layer of definable mechanical properties. But alloys with low contents of more-noble metal can form discrete, nanoporous dealloyed layers at low dealloying overpotentials, because the layer is not pure more-noble metal – it is an alloy. This was clearly shown in Figure 6.[19] Type 316 stainless steel (12% Ni) formed a dealloyed layer with about equal amounts of Ni and residual Fe. Figure 8 shows a better example of



Cl	(Cl^-, HCl)
SO_x	(retrograde compounds plus $S^{6+} \rightarrow S^{4+} \rightarrow S^{2+} \rightarrow S^0 \rightarrow S^{2-}$)
SiO_x	(SiO_2, complex compounds)
AlO_x	(Al_2O_3, complex compounds)
Cu	(Cu^0, Cu^{2+}, CuO)
Pb	(Pb^0, PbO_x^{+y})
Na, Ca, Mg	(complex compounds)
Na_2HPO_4	(retrograde compounds plus H_3PO_4)
B, Ti, Zn	(inhibitors)
O_2	(O_2, H_2O, compounds)
H_2	(H_2, H^+, H^0)
N_2H_4	(N_2H_4, NH_3, N_2)
C	(CO_3^{2-}, organic)
N	(NO_x, organic)
Fe, Cr, Ni	(Fe^{2+}, Cr^{3+}, Ni^{2+}, complex compounds)

Figure 3 Complex chemistry of a nuclear steam generator and location of corrosion phenomena. Adapted from Staehle, R. W.; Gorman, J. A. *Corrosion* **2003**, *59*, 931–994; *60*, 5–63; **2004**, *60*, 115–180.

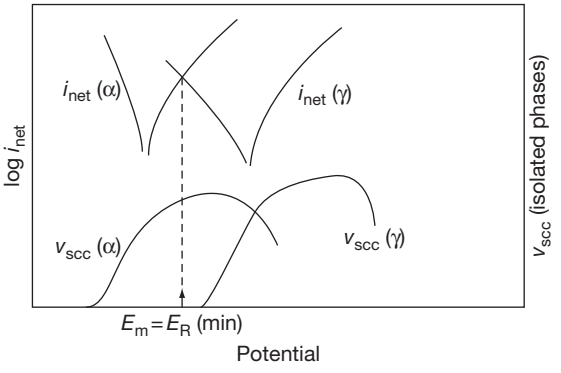


Figure 4 Reason for the greater resistance of duplex than austenitic stainless steel to chloride-induced SCC – both phases cannot (usually) be at their optimal cracking potential simultaneously, and under mild conditions, there can only be one mixed potential (E_m), and one phase is commonly below the minimum possible value of its repassivation potential (E_R) by Tsujikawa's criterion. The case illustrated would give very restrained SCC because the γ phase does not crack at all. Reproduced from Newman, R. C. Stress Corrosion Cracking Mechanisms. In *Corrosion Mechanisms in Theory and Practice*, 2nd ed; Marcus, P., Oudar, J., Eds.; Marcel Dekker: New York, 2002; pp 399–450.

95Ag–5Au (at.%) alloy. So, one cannot rule out a causal role of an actual dealloyed layer, even at 1% of noble metal, despite assertions to the contrary that appear now and again in the literature.[20,21]

As mentioned in **Chapter 2.05, Dealloying**, not all binary alloy systems show the parting limit (55 at.% less-noble element) seen in binary gold alloys – SCC and superficial dealloying of CuZn or CuAl alloy monocrystals in cuprous ammonia solutions starts at 20 at.% less-noble element, or even a little lower for the CuAl system – **Figure 9**. This has been identified as being at, or close to, a site percolation threshold for the fcc lattice.[22] Departures from this threshold could be due to like-with-like atomic clustering in the solid solution. Intergranular SCC in ordinary polycrystals persists at lower Zn contents, but transgranular SCC vanishes.[23] Evidently, the intergranular variant does not require an actual dealloyed layer, which is reasonable, as pure copper also shows intergranular SCC in ammonia solutions.

Other alloy systems where dealloying should be considered as a cause of, or contributing factor to, SCC include Monel (70Ni–30Cu)[24] and cupronickel. It is likely, though unproven, that some systems may switch from the '55%' to the '20%' threshold at high temperatures such as 300 °C. As a topical example, there is some evidence that Alloy 690 (Ni–29Cr–8Fe) shows dealloying in hot caustic environments, despite the sum of the Cr and Fe contents being well below 55%.[25]

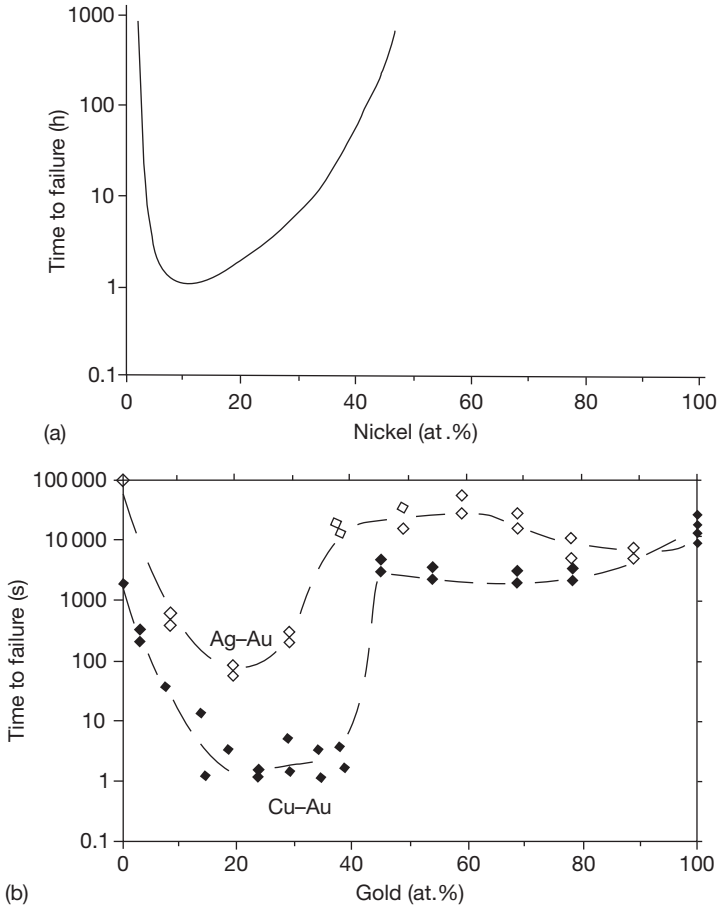


Figure 5 (a) The 'Copson curve'[13] for chloride-induced SCC of stainless steels; (b) the corresponding behavior of noble metal alloys in aqua regia, Reproduced from Newman, R. C. Stress Corrosion Cracking Mechanisms. In *Corrosion Mechanisms in Theory and Practice*, 2nd ed; Marcus, P., Oudar, J., Eds.; Marcel Dekker: New York, 2002; pp 399–450.

2.09.3.3 Grain Boundaries, without Precipitation

Since the advent of modern analytical TEM and sample preparation techniques such as focused ion beam (FIB), few have gone back to the classical SCC systems and questioned what is at the grain boundaries, and whether SCC is dependent on a second phase at the boundary, or on some more subtle chemical effect such as equilibrium segregation of alloying elements or impurities. Regarding the case of carbon steel in hot caustic or nitrate solutions, it has been known from the time of H.H. Uhlig[26] that SCC does not occur, or anyway is much more difficult, in 'pure' iron. But it is not clear whether carbide precipitation, or merely monolayer segregation of carbon or some element, is required for SCC. Modern techniques, including atom-probe field ion microscopy provide some answers, but also raise questions.[27] Most likely, in the case of carbon as the only impurity, fine carbides are required, but we do know that phosphorus and even molybdenum can induce SCC by pure segregation, without carbides (although carbides make matters worse).[28] In a related study, it was shown that an amorphous Fe–P alloy, simulating the grain boundary composition of a segregated alloy, showed enhanced anodic activity in the right potential range to explain the SCC observations.[29] This research was inspired by a classic study on steam-turbine rotor steel by J. F. Newman.[30]

The literature on grain boundary segregation and SCC has become quite large. For example, sulfur is a strong grain boundary segregant in nickel, and causes intergranular corrosion and/or SCC, depending on the conditions.[31] Magnesium segregation has been

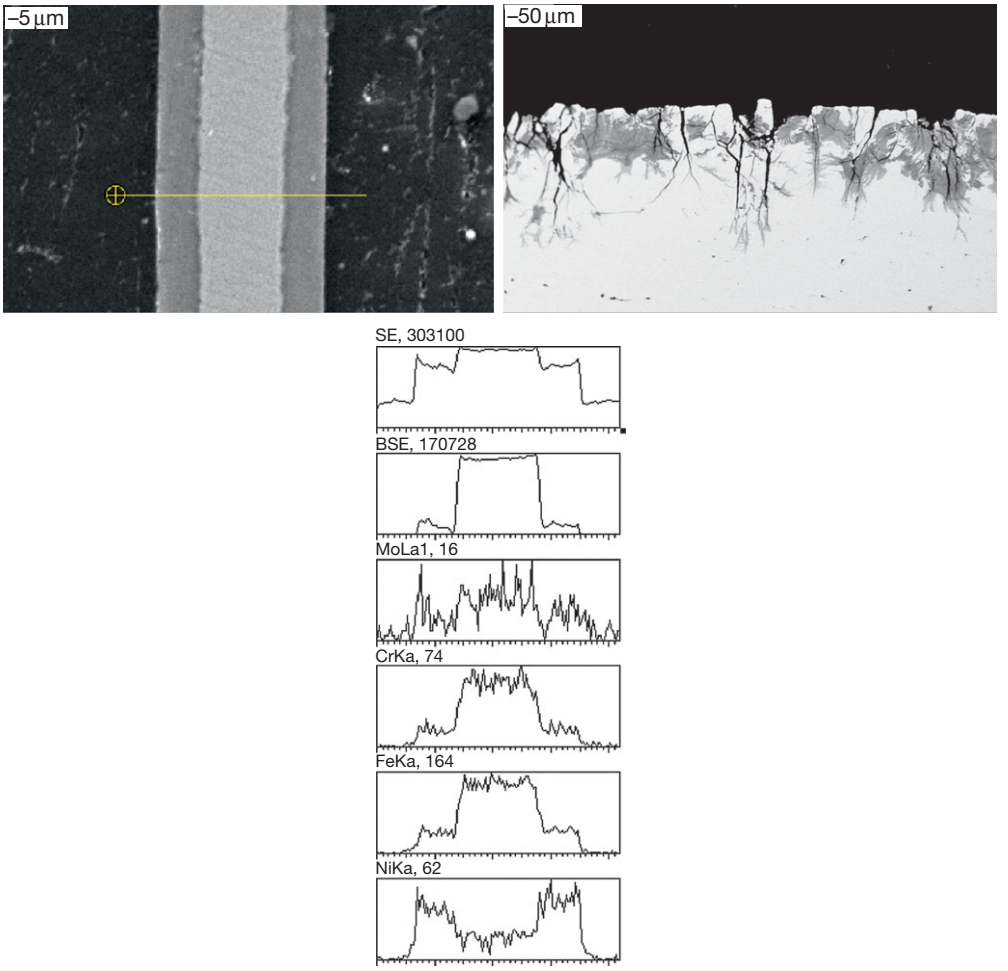


Figure 6 Secondary electron image and matching line profile across a 316SS foil exposed to 50% NaOH at 140 °C for 20 h, and SCC occurring in the same condition. Reproduced from Deakin, J.; Dong, Z.; Lynch, B.; Newman, R. C. *Corros. Sci.* **2004**, *46*, 2117–2133.

implicated in SCC of Al–Mg and Al–Mg–Zn alloys.[32] But it would not be sensible to conclude that all intergranular cracking (in the absence of precipitation) is caused by segregation. For example, there is no evidence that this is the case in solid-solution alloys like brass (in aqueous ammonia) or AuCu, AuAg, etc. – in these dealloying systems, it is more likely that the dealloyed layer is simply weaker at the grain boundary – not surprisingly, because it is known that in dealloying, individual grains retain their gross orientation, so a grain boundary is an awkward site for the nanoporosity-formation process to negotiate, leaving inevitable weakness (lack of connection of nano-ligaments across the grain boundary)[33] – **Figure 10**.

Once we consider temperatures in the 300–330 °C range, as in nuclear power plants, intergranular SCC may start to merge with low-temperature creep fracture, and clean boundaries may actually be more susceptible than those with precipitation. Hydrogen from corrosion may be facilitating such fractures. Arioka *et al.*[34] claimed common evidence for creep-like phenomena including cavities and grain-boundary nickel enrichment in tests of types 304 and 316 steels in both air (at 450–550 °C) and simulated pressurized water reactor (PWR) water (at 360 °C). We await conclusive evidence on the latter, but this is potentially a very important advance. Recently, the same group has argued that intergranular SCC of carbon steel can occur by a hydrogen-assisted creep-like mechanism near reactor operating temperatures,[35] a phenomenon of interest to operators of Canadian CANDU plants, which have carbon–steel feeder pipes. It has also been shown that creep

occurs in Alloy 600 (Ni–16Cr–9Fe) at 360 °C and can be enhanced by hydrogen absorbed from the environment,[36] although a definite role of such creep in SCC has not been established. Again, carbide precipitation at grain boundaries is beneficial. Ingress of oxygen into grain boundaries is probably the key factor in this case, and is discussed later.

Under neutron irradiation at water reactor operating temperatures, stainless steels and nickel base alloys become harder owing to atom displacement and the formation of dislocation loops, and eventually, can suffer from irradiation-assisted SCC (IASCC).[37] Not only the increased hardness, but also radiation-induced segregation (RIS) of alloying elements might contribute to the increased susceptibility to SCC. Vacancies created by irradiation migrate to grain boundaries, and the atomic mobilities are such that Ni is enriched and Cr depleted near these boundaries.[38] This is distinct from ordinary equilibrium segregation and is much wider than a single monolayer. Such phenomena are conveniently studied using proton irradiation, which induces less radioactivity than neutrons, and enables SCC testing to be done in an ordinary laboratory with comparable material properties to neutron-irradiated material.[39] Using such methods, it was shown that light annealing removed some small defects and suppressed SCC without changing significantly the hardness or the RIS.[40] The likely reason is that the small defects were promoting intense strain localization at grain boundaries by a process called dislocation channeling, and suggests that this – not segregation or hardness *per se* – was the primary cause of SCC.[41] Nevertheless, by adding oversized solutes such as Zr, RIS could be reduced, and this did reduce the susceptibility to SCC.[42]

The strain localization induced by neutron irradiation is the most extreme example of a general principle that SCC is more likely – and will occur under milder loading conditions – the more localized is the deformation at the crack tip. This can be illustrated as shown in **Figure 11**.

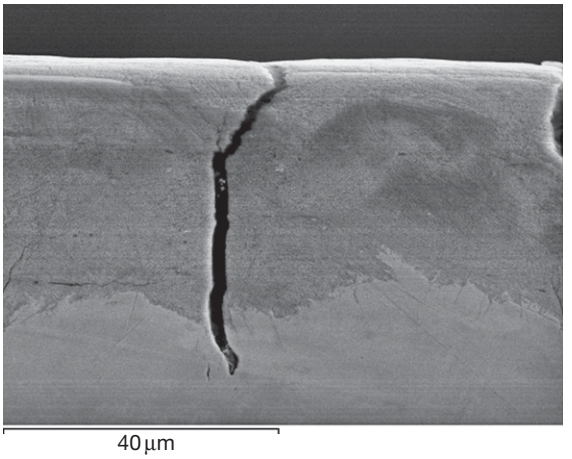


Figure 7 Dealloying of type 310 stainless steel in 50% NaOH at 280 °C with hydrogen overpressure, showing possible crack injection into the unattacked substrate upon subsequent bending in air. Courtesy of Z.L. Coull, University of Toronto.

2.09.3.4 Grain Boundary Precipitation

The most notorious example of SCC due to grain boundary precipitation is that of weld-sensitized

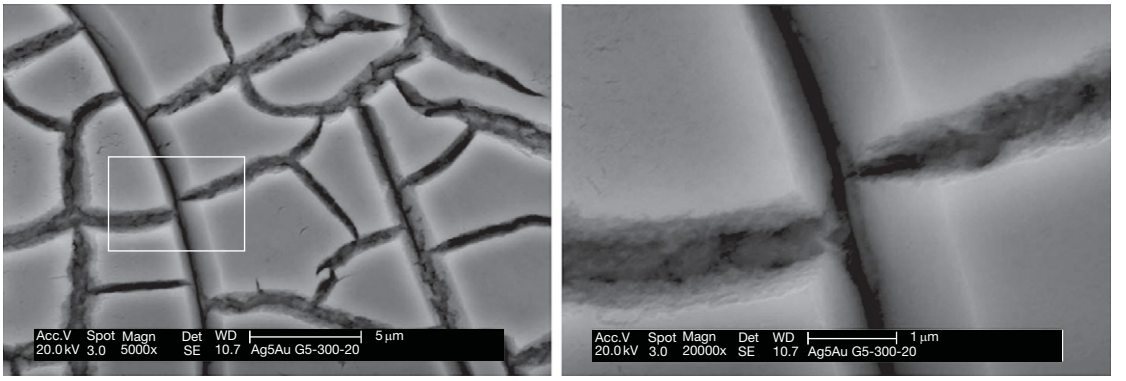


Figure 8 Dealloyed layer formed on an unstressed 95Ag–5Au (a/o) alloy in 0.77 M $HClO_4$ at 20 °C. The layer was grown by applying an anodic current density of 5 mA cm^{-2} for 300 s. We can see that a well-defined nanoporous metal layer is formed, even at this low Au content, and obviously must have a very high residual Ag content. The propensity of dealloying to induce tensile stress by lattice shrinkage is also evident – this is part of the reason why alloys crack under such mild mechanical conditions compared with pure metals. Courtesy of N.A. Senior, PhD thesis, University of Manchester, 2004.

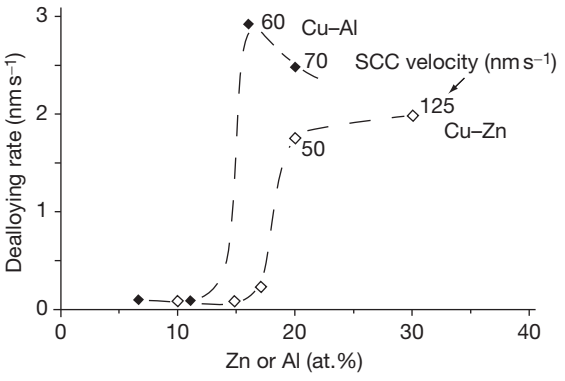


Figure 9 SCC and dealloying of Cu–Zn and Cu–Al monocrystals in cuprous ammonia solution as a function of solute content. Reproduced from Newman, R. C. Stress Corrosion Cracking Mechanisms. In *Corrosion Mechanisms in Theory and Practice*, 2nd ed; Marcus, P., Oudar, J., Eds.; Marcel Dekker: New York, 2002; pp 399–450.

austenitic stainless steel (type 304SS) in slightly oxidizing hot water in boiling water reactor (BWR) nuclear power plants.[43,44] The cost of this cracking to the industry, in terms of plant availability, was massive through the 1970s and 1980s (it is interesting to recall that even this cracking was not detected in the laboratory until the introduction of slow strain rate testing). Chromium depletion adjacent to the grain boundary carbides (generally $Cr_{23}C_6$) localizes corrosion and causes the cracking. This topic has probably been studied 'to death,' and cracking is well under control in those plants that still have 304SS – hydrogen is added to the water in stoichiometric excess compared with oxygen (and other oxidants produced by radiation, principally hydrogen peroxide), and noble metals are plated out on the plant surfaces to catalyze the hydrogen oxidation half-reaction and, thus, ensure that the conditions remain net-reducing.[45]

Although sensitized stainless steel is rather 'old hat,' and scarcely exists in modern plants that use low-carbon or stabilized austenitic steels, new insights are still being produced in studies that use low-temperature SCC of sensitized material as a model system to explore fundamental geometric features of SCC.[46] This kind of SCC, which generally occurs in solutions of partially reduced sulfur species (between sulfate and sulfide in oxidation state), is easy to study and amenable to incorporation into advanced tools such as synchrotron light sources. Originally, it was called 'polythionic acid SCC'[47] because such acids (chiefly tetrathionic acid, $H_2S_4O_6$) were believed to be produced in refinery equipment by the reaction of iron sulfide scales with oxygen and water. Later both 304SS[48] and Alloy 600 (Ni–16Cr–9Fe)[49,50] were shown to be susceptible to SCC in dilute, aerated thiosulfate ($S_2O_3^{2-}$) solutions. Thiosulfate was also the only soluble aggressive species detected in one study of the reaction of FeS with aerated water.[51]

Apart from sensitized stainless steels, the main materials where grain-boundary precipitation is implicated in SCC are aluminum alloys, and especially those of high strength (2000 series, 7000 series, 8000 (Al–Li) series) used in aircraft. SCC can occur in lower-strength 5000 series (Al–Mg–X) alloys[52] when reactive precipitates are present at grain boundaries, such as Al_3Mg_2, the β phase in the binary Al–Mg system, and in other conditions when the stresses are severe. So this is different from SCC of sensitized stainless steel, where it is the solute-depleted zone, not the particle, that is sensitive to corrosion. At the other extreme is the SCC of binary Al–Cu alloys aged to produce Cu-depleted zones adjacent to grain boundaries,[53] where the Cu-depleted material dissolves faster, at a given potential, than the matrix (there are doubts as to whether this should really be called SCC, as nearly all grain boundaries are attacked – it is more like a strain-accelerated intergranular corrosion). Regarding the highest-strength Al–Zn–Mg–Cu (7000 series) alloys, the roles of reactive particles and solute-depleted zones have been controversial, but it appears – currently – that the reactive $MgZn_2$ particles are implicated. Such an explanation provides a neat rationalization of the protective effect against SCC of overaging such alloys, as is done for all aircraft construction (the so-called T7 temper). More copper enters the precipitates with overaging, and reduces their excess reactivity compared with the matrix[54] – **Figure 12**.

Recently discovered examples of precipitation effects include the presence of boron, presumably as nickel boride, at the grain boundaries of Monel 400, leading to intergranular corrosion and SCC.[55] However, it is doubtful that such precipitates are necessary in the general case of intergranular corrosion of Monel[24] – they merely accelerate it by providing a nickel-rich path.

2.09.3.5 Yield Strength, Including Cold Work Effects

Most SCC is exacerbated by high yield strength, and there must be more than one reason for this behavior – perhaps four or five – depending on the alloy

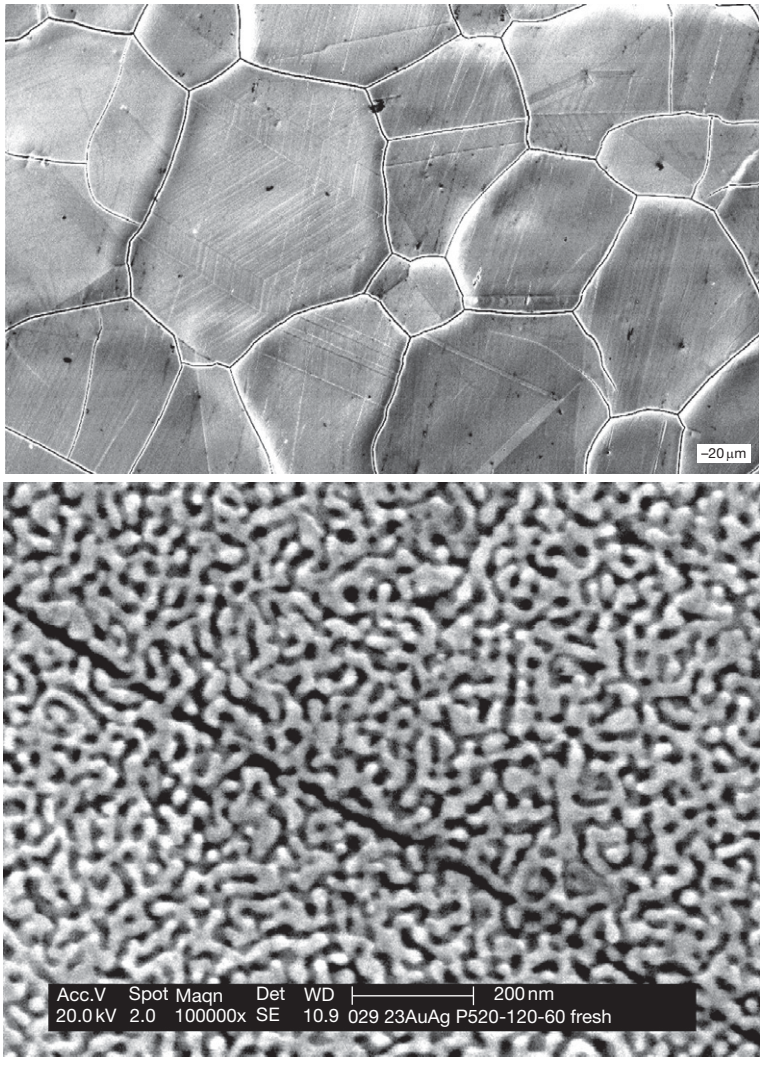


Figure 10 Spontaneous intergranular fracture of dealloyed AuAg, and the explanation in terms of limited connection of nano-ligaments across grain boundaries. Reproduced from Senior, N. A.; Newman, R. C. *Nanotechnology* **2006**, *17*, 2311–2316.

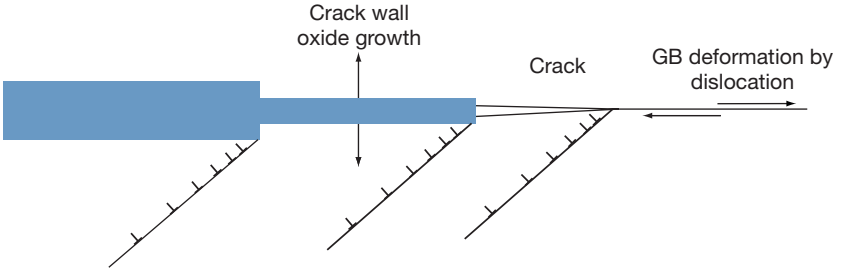


Figure 11 Enhancement of intergranular SCC by strong shear band intersection with grain boundaries, as seen in cold worked and especially neutron-irradiated austenitic steels. Adapted from Was, G. S. Proceedings of the 11th International Conference on Environmental Degradation of Materials in Nuclear Power Systems – Water Reactors, American Nuclear Society, La Grange Park, IL, 2004; pp 965–985.

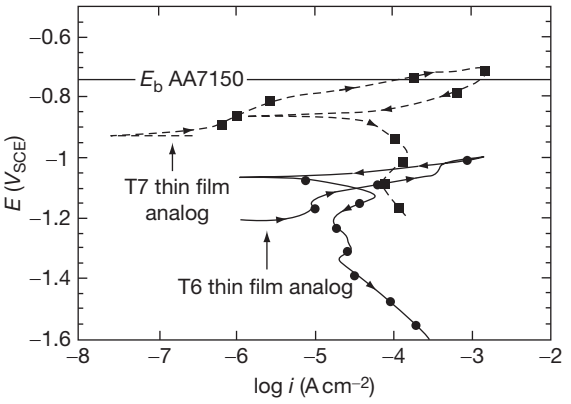


Table 3
Composition of AA7150-T6 constituents measured by analytical TEM (five replicate measurements were made on a single sample)

Element	Matrix	SDZ	Precipitate
Al	92	95	74
Mg	4.6	3.8	11
Zn	2.5	0.8	12
Cu	0.9	0.4	3.3

Table 4
Composition of AA7150-T7 constituents measured by analytical TEM (five replicate measurements were made on a single sample)

Element	Matrix	SDZ	Precipitate
Al	94.5	96	52
Mg	3.1	2.6	17
Zn	1.3	0.8	1.5
Cu	1.1	0.3	30

Figure 12 Demonstration that the reduced anodic activity and (probably) SCC susceptibility of grain boundaries in overaged high-strength aluminum alloy 7150 is due to more copper entering the Mg_2Zn precipitates. Reproduced from Ramgopal, T.; Gouma, P. I.; Frankel, G. S. *Corrosion* **2002**, *58*, 687–697.

system. However, we can make certain generalizations. In very strong steels, exposed to ambient-temperature water (with or without salts), there is no doubt whatever that SCC is due to absorbed hydrogen. The crack path is commonly intergranular, along prior austenite grain boundaries, in tempered-martensite microstructures, but can also be transgranular, cleavage-like in these and in other, more exotic steels such as precipitation strengthened alloys.[56] If we add hydrogen sulfide to the environment, as in sour gas systems, lower-strength steels become susceptible, although ordinary carbon steel remains essentially immune to SCC (it may suffer from hydrogen pressure-induced cracking – see **Chapter 2.08, Environmentally Assisted Cracking Overview**). A theme in recent oil-and-gas material development has been the gradual (very gradual) increase in the yield strength of low-alloy steel that can be employed in sour-gas applications.[57] Stainless steels based on Fe–13Cr (–Ni. . .–Mo. . .) have also been developed for mildly sour systems.

There is a class of SCC problem in carbon steels that involves apparently innocuous near-neutral-pH water, yet occurs with alarming frequency on the external surfaces of high-pressure gas transmission pipelines.[58–60] Usually the problem can be traced to some dynamic loading, such as freezing and thawing, and it is thought that SCC cannot occur under perfectly static loading, consistent with a hydrogen effect in low-strength steel, but stronger steels than the usual X-65 grade, such as X-80, seem not to require this 'fatigue' component; if such steels are installed in the field, the risk of SCC may be increased (but there is much we still do not understand about this phenomenon).

Since we know that H_2S increases the percentage of discharged hydrogen that enters the steel, its special severity clearly reflects a correlation between SCC and the concentration of atomic hydrogen that develops at some special site near the crack tip. Exactly what this site is, and what the hydrogen does when it gets there, are still under debate, but the general correlation is secure; we give a recent example of such reasoning.[61] *Ab initio* quantum-chemistry modeling is starting to illuminate some old questions such as – how does hydrogen interact with dislocations?[62] The current view seems to be that hydrogen can either promote or hinder dislocation motion, depending on factors such as strain rate and presence of other interstitials, as well as the crystal structure of the material, presence of interfaces, etc. Spectacular studies of hydrogen-dislocation interactions have been done using *in situ* transmission electron microscopy.[63] Also the growing interest in metals as storage media in hydrogen technology has provided complementary information.[64]

To quantify the SCC of high-strength steels, fracture mechanics techniques are indispensable. Generally speaking, K_{Ic} falls gradually with increasing yield strength, while K_{Iscc} falls steeply above a certain yield strength value – **Figure 13**.

High-strength steel wire (prestressing steel) is more resistant to SCC than other material forms because it has a highly elongated grain structure, and an intergranular crack has to take a tortuous path to cross the wire.[65] However, failures have occurred in such

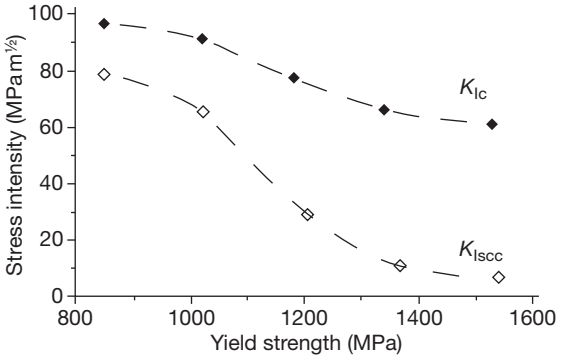


Figure 13 Classical dependence of fracture toughness (K_{Ic}) and SCC threshold in seawater (K_{Iscc}) on yield strength, for old-fashioned ultra-high strength AISI 4340 steel. Reproduced from Newman, R. C. Stress Corrosion Cracking Mechanisms. In *Corrosion Mechanisms in Theory and Practice*, 2nd ed; Marcus, P., Oudar, J., Eds.; Marcel Dekker: New York, 2002; pp 399–450.

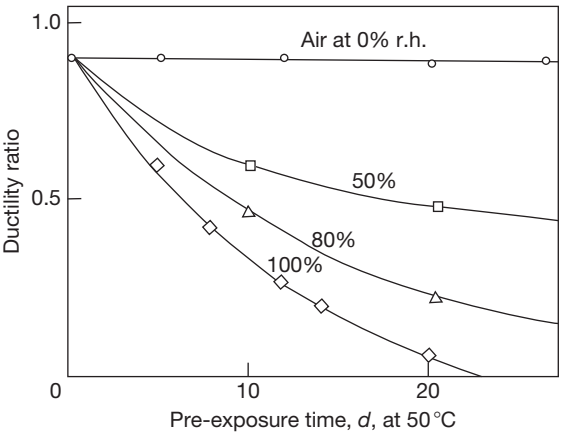


Figure 14 The preexposure embrittlement phenomenon in AlZnMg alloys. Reproduced from Scamans, G. M.; Alani, R.; Swann, P. R. *Corros. Sci.* **1976**, *16*, 443–459.

material. A similar effect protects sheet made from high-strength 2000-series (Al–Cu) alloys.

The effects of yield strength on SCC become murky when we depart from high-strength steels. For aluminum alloys, maximum sensitivity to SCC is commonly seen in the underaged or peak-aged condition, and the beneficial effect of overaging (in the most sensitive AlZnMgCu systems) is thought to involve alloying of Cu into the $MgZn_2$ precipitates at the grain boundaries. So, the root cause of the beneficial effect is related to the corrosion resistance – that is, anodic kinetics – of the grain-boundary phase, but hydrogen may be the actual cracking agent. Often *preexposure embrittlement* is observed – the material will crack in air or an inert environment after precorrosion[66] – **Figure 14**. Usually this is due to hydrogen, but a further complication is that cracking can even occur without water exposure (especially at slightly elevated temperature) by a creep-like mechanism[67] which sometimes involves a solid-metal-embrittlement type of action of internal lead particles.[68] Generally speaking, in these materials, the effect of yield strength is subordinate to microstructural considerations.[69] A similar picture presents itself when we consider β-titanium (bcc-titanium) alloys[70–72] – there is no doubt that hydrogen can embrittle the material, but the corrosion reaction has to be considered as well and the effect of yield strength *per se* cannot be divorced from microstructure. Incidentally, the SCC of titanium alloys has an exciting history – in the 1960s and early 1970s a high-strength α−β alloy was under consideration for the US supersonic passenger airliner project, but the project was abandoned – partly because of an energy crisis, and partly because of unanticipated SCC of the alloy when precracked.[73] The Apollo program was also plagued by SCC incidents, one involving methanol, which was one of a number of exotic environments found to cause SCC during that period.[74] The history of SCC from 1965 to 1990, including such practical experience, was reviewed by Procter and Newman.[75]

Solid-solution alloys that are not known to be susceptible to hydrogen also show deleterious effects of higher yield strength, and in particular cold work, on SCC. This is highly topical in the nuclear power industry, not only because certain components are inherently coldworked during manufacture (e.g., steam-generator tubes are straightened) but also because surface damage and surface condition generally are known to influence, critically, the initiation of SCC. So, whereas early failures of 304SS piping in BWR systems were confined to weld-sensitized material, nonsensitized but coldworked material began to show failures later on; such cracking will even occur in reducing conditions with hydrogen overpressure,[76] where it is associated with intense strain localization at grain boundaries as revealed by cross-sectional TEM.[77] Such localization, which we discussed earlier in the context of neutron irradiation, is probably fundamental to SCC and to the effect of yield strength, so we may formulate a rule along the following lines – *other things being equal, SCC will occur under milder loading conditions, the more localized is the plasticity near the crack tip.*

Using advanced nanoscale analytical techniques, it has been shown (Lozano Perez, unpublished) that

cold worked austenitic stainless steel exposed to PWR primary water undergoes a kind of internal grain boundary Cr oxidation and elemental Ni enrichment quite similar to what occurs in Alloy 600 in the annealed condition. Apparently the shear bands or twin deformation bands intersecting the grain boundary enhance oxidation along the grain boundary as well as along themselves. So, one could speculate that cold work removes the Fe-induced barrier to internal oxidation of Cr, unifying the mechanism of low-potential SCC in the Ni- and Fe-rich materials.

It should be mentioned that there is a related but different interpretation of the effect of cold work, or of yield strength generally[76,78] – this can only be understood by referring to details of the slip-dissolution model (SDM) for SCC, so we defer the discussion to that section.

Reverting to higher-strength steels for a moment, a problematic phenomenon in power generation is the SCC of turbine components made of quenched and tempered low-alloy steels. This does occur more easily in stronger material, but does not seem to be a hydrogen effect, and occurs at elevated temperatures where hydrogen does not affect the properties of steel because it cannot be localized for long enough near the crack tip.[79]

2.09.4 Crack Nucleation, Crack Coalescence, 3D Crack Morphology, Effect of Distributed Grain Boundary Character

It is fundamental to SCC that multiple cracks, or at least branched cracks, always occur. Whenever that happens, the nearby crack segments mutually shield each other from the local stress, so that we can say each crack has an effective K_I value that is less than the nominal one. An important stage in the life of a crack is the point where it becomes connected along the free surface for a distance similar to the specimen or component thickness. This process of crack coalescence was studied by Parkins using mainly intergranular carbonate–bicarbonate SCC of steel[80,81] – **Figure 15** – and is also of key importance in the other type of external pipeline cracking – so-called 'near-neutral pH,' transgranular SCC[82] – **Figure 16**. Akashi and others have created detailed conceptual models, suitable for statistical treatment, in which there is a long delay in the formation of a free-running crack[83] – **Figure 17**.

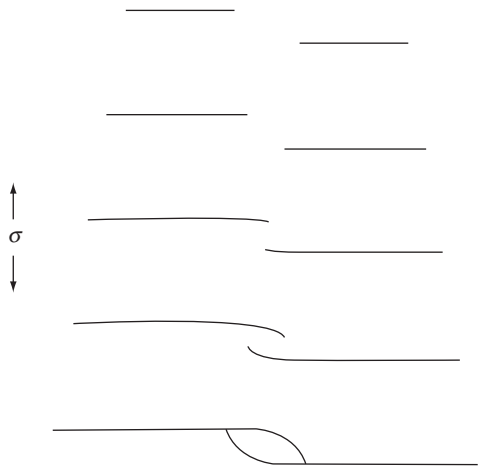


Figure 15 Coalescence of SCC on a surface (increasing time from top to bottom). Adapted from Parkins, R. N.; Singh, P. M. *Corrosion* **1990**, *46*, 485–499.

The crack coalescence phenomenon does not in itself imply anything distributed about the grain boundary properties – simply that only so many cracks per unit area can nucleate and grow for any distance. But recently the idea of grain boundary engineering has become a popular way to try to suppress, or at least understand, the progress of SCC in industrial alloys.[46,84–87] The development of software for automatic indexing of backscattered electron diffraction (EBSD) patterns was the key technical advance that made this possible. Recently this has been combined with 3D tomographic X-ray imaging in a synchrotron light source – **Figures 18 and 19**. The idea of grain boundary engineering is to create as many 'special' grain boundaries as possible within the microstructure by thermomechanical treatment. Annealing twins have special boundaries, but do not in themselves obstruct crack growth to any great extent; the crack easily bypasses the long boundaries. The boundaries that are most important are those that have a particular periodicity of registry between the lattices on either side, not only twin boundaries but also those of higher-order periodicity.[88] Such boundaries are less likely to have segregation of impurities, or detrimental forms of precipitation, and form ductile bridges that are believed – in one theory[46] – to shield the crack tip from part of the applied nominal stress intensity. Under sensitization treatment, they have delayed chromium carbide precipitation. Ordinary industrial alloys already contain special grain boundaries, often in highly variable fractions, and the challenge is to

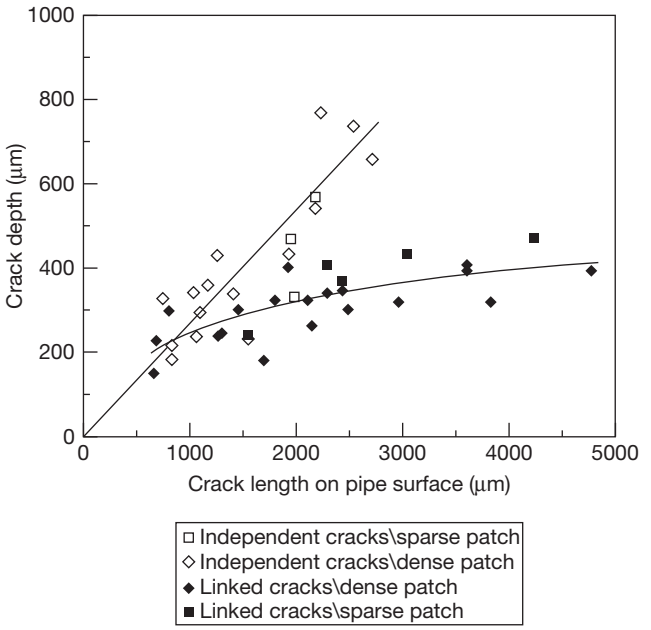


Figure 16 Behavior of crack colonies on a pipeline subject to near-neutral-pH SCC. Reproduced from Chen, W.; King, F.; Vokes, E. *Corrosion* **2002**, *58*, 267–275.

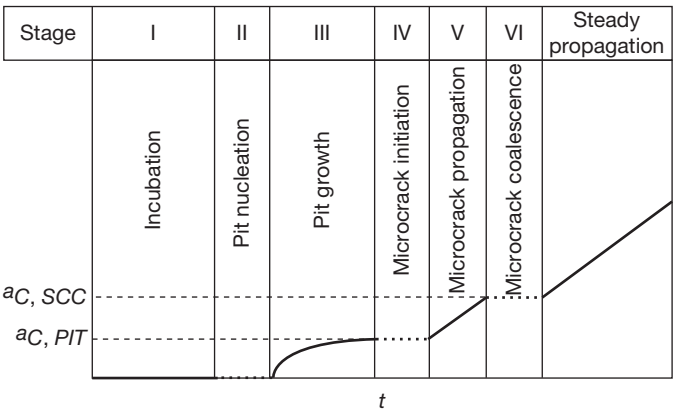


Figure 17 Life history of SCC, according to Akashi and Nakayama. Reproduced from Akashi, M.; Nakayama, G. *Effects of the Environment on the Initiation of Crack Growth*; ASTM STP 1298, Philadelphia, PA, 1997; pp 150–164.

increase this fraction whilst maintaining a reproducible, flexible, and economical processing route. A drawback is sometimes that the thermal treatments required to obtain high fractions of special boundaries have to be done at lower than normal temperature, which carries certain hazards such as deleterious precipitation reactions.

EBSD is not only able to index a microstructure – it also gives information on plastic strains and their localization.[89] An important result is that 'special' grain boundaries are soft – they do not sustain dislocation pile-ups. This is a further reason why they are resistant to SCC, which is greatly exacerbated by strain localization as we already discussed.

When SCC occurs in a microstructure with distributed grain boundary properties, the crack initially develops a ramified 3D morphology with uncracked ligaments. There are slightly different ways to approach the problem of cracking resistance in such a geometry. The simplest approach is based on percolation theory[90]; under slow strain rate conditions, where ductile ligaments will always be torn

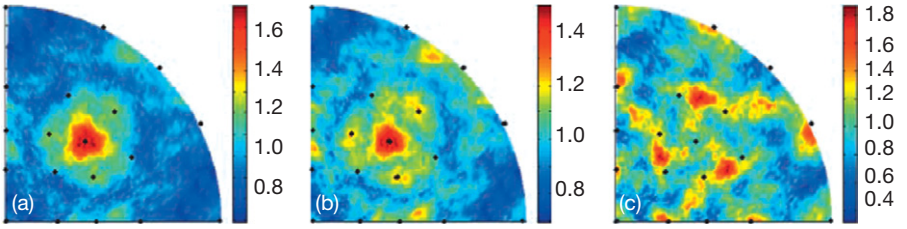


Figure 18 Capability of EBSD and tomographic imaging to identify the crack path in 3D for intergranular SCC and identify the boundaries that do, or do not, crack. Reproduced from King, A.; Johnson, G.; Engelberg, D.; Ludwig, W.; Marrow, J. *Science* **2008**, *321*, 382–385.

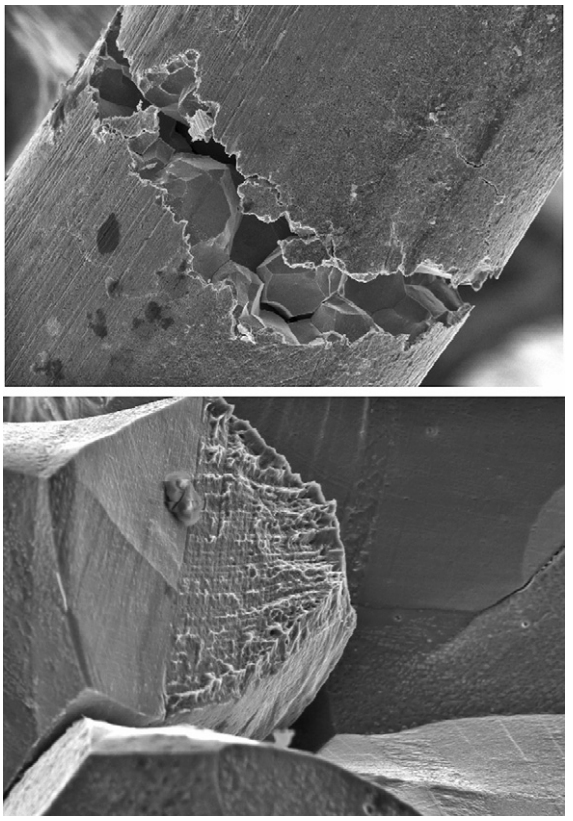


Figure 19 Appearance of the cracked sample for the kind of experiment shown in **Figure 18**, and illustration of the presence of ductile ligaments on the fracture surface. Courtesy of D.L. Engelberg.

apart, macroscopic SCC starts to occur when a certain fraction of grain boundary segments are susceptible and can form a connected 3D network (where connection is allowed at edges as well as faces). This percolation threshold is only ~23%, and such a microstructure would never develop a macroscopic crack under constant load conditions. The same authors noted that for fractions of susceptible boundaries greater than 89%, no ligaments should form and SCC should become 'completely brittle.' Later the actual grain-boundary Cr contents were taken into account in such an approach.[91] A more sophisticated approach entails restricting the definition of 'connection' to grain faces, and incorporating stress via the inability of the crack to propagate in unfavorable directions[84,92] – **Figure 20** – but is still essentially a modified percolation concept. In contrast, the model developed in Manchester[93–96] has elements of percolation, but places great emphasis on the shielding effect of uncracked ligaments, by actually calculating the magnitude of such shielding.

Of course, the physics of granular networks can be taken further than this – such networks and their behavior under fracture, dielectric breakdown, etc. have been a favorite topic in statistical physics.[97]

2.09.5 Crack Chemistry in SCC

A crack is, literally, a crevice, so we do not need to reiterate concepts that can be found in **Chapter 2.03, Crevice Corrosion**. Reviews of crack chemistry are available,[98] and a second conference 'CCPCC2' was held in 2009 to mark the twenty-fifth anniversary of a landmark conference at the National Physical Laboratory in 1984.[99]

An important aspect of SCC that relates to local chemistry is the evolution of cracks from pits or other localized corrosion sites. There is a mechanics aspect, in that a pit creates a stress concentration, but often the local solution chemistry is the dominant factor. For low-alloy steels, a recent review addressed both the chemistry and mechanics aspects,[100] but the premier example of such a sequence occurs in the chloride-induced SCC (Cl-SCC) of austenitic or duplex stainless steels. This is less well-known than it might be, because much of the published discussion is in obscure sources. **Figure 21** shows the early stage of Cl-SCC from a site of crevice corrosion.

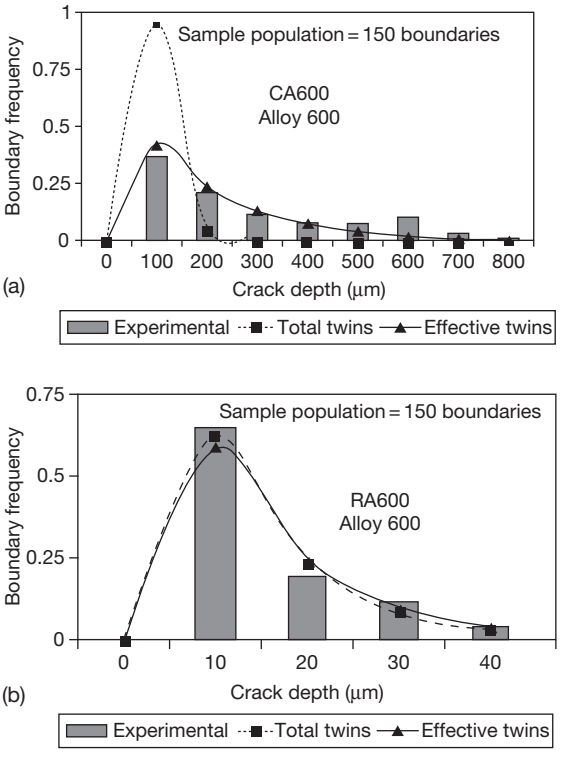


Figure 20 Grain boundary engineering of Alloy 600 against SCC, compared with a model based on connectivity of grain faces (CA – normal; RA – engineered). Reproduced from Lehockey, E. M.; Brennenstuhl, A. M.; Thompson, I. *Corros. Sci.* **2004**, *46*, 2383–2404.

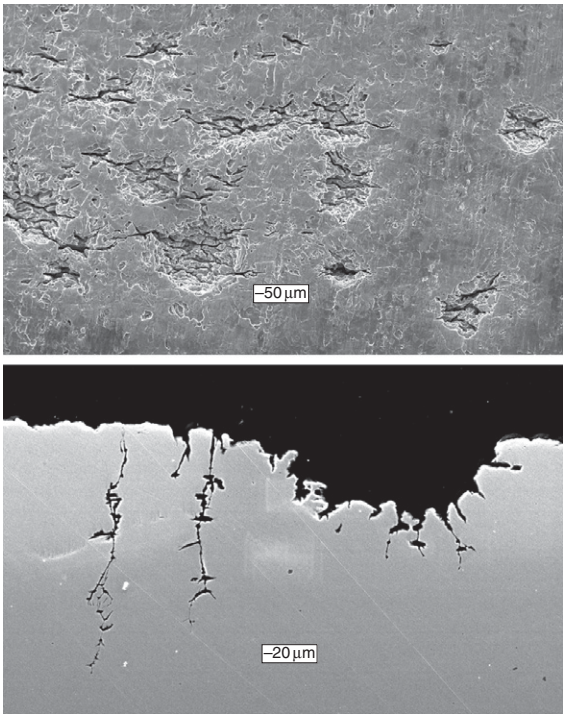


Figure 21 Early stages of Cl–SCC of austenitic stainless steel at low temperature (80 °C), starting from a corroding crevice. Courtesy of S. Wang, PhD thesis, UMIST, 2002.

Whenever SCC occurs in stainless steel in an aqueous halide environment, the local condition in the crack is similar to that in an active, strongly acidified, but slowly dissolving (because very tight) crevice. This was stated most clearly by Tsujikawa,[18,101] who first argued that the lower critical potential for chloride-induced SCC (Cl-SCC) of austenitic stainless steel was simply the repassivation potential (sometimes called protection potential) for crevice corrosion. This immediately replaced fanciful notions such as the critical potential being due to the onset of adsorption of some special species at the crack tip,[102] which are still taken seriously in some textbooks. Given that the propagation rate or velocity ($v_{corrosion}$) of localized corrosion varies with the geometry and potential because of diffusional or ohmic control of the dissolution rate, we can readily imagine that certain conditions would favor localized corrosion over SCC, or vice versa, even at a given temperature. Then when we factor in temperature, the apparent activation energy that applies to the crack velocity (v_{crack}) is higher than that for $v_{corrosion}$. Obviously, SCC can only occur if $v_{crack}(T, \text{alloy}) > v_{corrosion}(T, \text{alloy}, \text{potential})$ – **Figure 22**.

Such a framework can explain the commonly observed critical temperature for Cl-SCC of basic austenitic grades (304[L], 316[L]), as indicated in **Figure 23**. For all but the most extremely slow (i.e., blocked or stifled) localized corrosion, the slowest sustained rates of penetration tend to lie in the range of a few to 10 mm year^{-1}. This is also the order of magnitude of v_{crack} at 50 or 60 °C. But in order for a crack to initiate from a small pit that is growing at 100 or even 1000 mm year^{-1}, the temperature must be much higher, so that the condition $v_{crack} > v_{corrosion}$ is maintained. Thus, when Cl-SCC occurs at very high temperatures, it generally initiates from small pits, but at the lowest temperatures, it initiates from underdeposit corrosion sites – or sometimes in tight metal–metal or metal–polymer crevices. SCC of 304SS or 316SS can even be made to occur at room temperature by depositing $MgCl_2$ (or sea salt, which amounts to the same thing) on the steel, and controlling the relative humidity at a

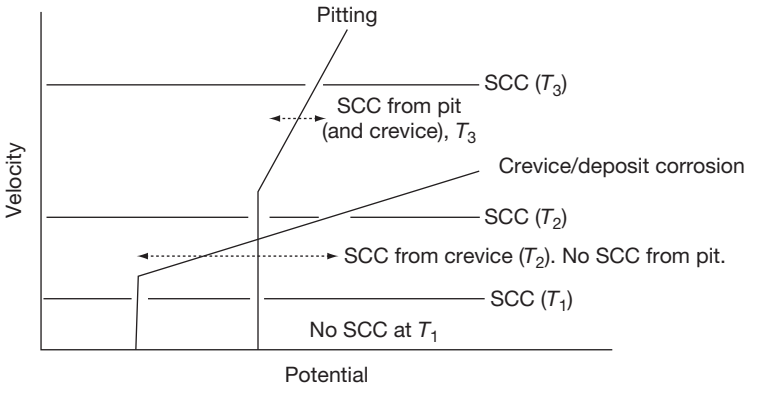


Figure 22 Potential dependence of SCC and localized corrosion in austenitic stainless steel, based on the work of Tsujikawa.

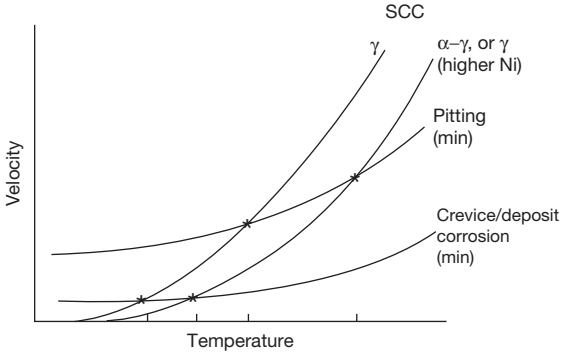


Figure 23 Origin of the 'critical' temperature for Cl–SCC of austenitic or duplex stainless steel, based on the work of Tsujikawa.

value in the 30–35% range where the surface remains wetted by nearly-saturated $MgCl_2$ solution.[103,104] Recently authors have rediscovered this kind of effect[105] – and the role of $MgCl_2$ (for the sea salt case) has become mildly controversial.[106,107] But the Japanese data from the 1980s were so convincing that there does not appear to be much room for doubt as to the role of Mg. The true application limits of duplex stainless steels are made apparent when such testing is used – a temperature limit of 100–110 °C seems to emerge,[106] contrary to the optimistic assessments made by alloy manufacturers when testing was only being done in NaCl solutions. In further work on this issue, the role of carnotite ($KMgCl_3$) would repay investigation. It is unlikely that the cases of easy SCC at low temperatures are associated with some Na-rich solution; more likely is that evaporation experiments done to date have not gone to a high enough level of concentration (80 times at least). But it is also possible that the solids precipitated from seawater have a crevice effect, exacerbating SCC, that is absent when one starts from a binary mixture of NaCl and $MgCl_2$.

Tsujikawa's analysis shows very clearly that the atomistic mechanism of Cl-SCC is not slip-dissolution, but some kind of embrittlement – most likely film-induced cleavage.

The SCC velocity v_{crack} is affected strongly by nickel alloying, which lowers v_{crack} while leaving $v_{corrosion}$ relatively unaffected – thus, as we move from types 304 or 316 stainless steel to type 310 (20% Ni) and eventually Alloy 800 (33% Ni), a higher and higher temperature is required for SCC. Eventually, somewhere above the nickel content of Alloy 800, SCC ceases altogether. This is envisaged – in Toronto anyway – as a parting limit for dealloying (see **Chapter 2.05, Dealloying**).

The situation in aluminum alloys is more complex, because the more susceptible grades are not solid solutions, and some phases are exceptionally reactive (Al_2LiMg, $MgZn_2$, Mg_3Al_2). Even pure water or water vapor can attack such phases,[66] and does so in a peak-aged, high-strength AlZnMgCu alloy such as the 7075 grade, although salt water does so more rapidly, and causes more rapid SCC. As mentioned previously, the beneficial effect of overaging on SCC resistance of AlZnMgCu alloys is thought to be due to more copper being incorporated into $MgZn_2$ precipitates, thus ennobling them. In underaged AlCu alloys, localized corrosion attacks Cu-depleted regions adjacent to the grain boundaries

(this is hardly SCC at all, more like intergranular corrosion). So the generalization that SCC is associated with localized corrosion, for passive alloys in halide solutions, remains intact, but has exceptions.

A key question that always arises when considering crack chemistry is – does the crack tip 'feel' the externally applied or naturally established electrode potential? Much ink has been spilled on this topic! Clearly, when the crack walls are quite active, it is difficult for the external potential to be thrown down the crack. But for relatively passive systems, such as titanium alloys, external polarization of the crack tip is possible. Importantly, even if there is a large IR potential drop down the crack, and the crack-tip potential is only slightly elevated from its local mixed or corrosion potential, there is a change in crack chemistry in accordance with the diffusion-migration balance for the nonreacting ions such as Cl^-. At room temperature, this is one order enrichment of Cl^- for every 59 mV (2.3RT/F) of potential drop down the crack.[108] For sulfate, it should be one order for every 29.5 mV. This must be a major factor in the SCC of turbine disk steels in dilute condensate.[79] Similarly, crack chemistry also plays a role in the numerous SCC phenomena that occur in high temperature water, especially in the somewhat oxygenated water that is present in a BWR water circuit.[44] If the environment inside the crack were truly pure water, practically no 'local cell' action could develop between the inside and outside of the crack, even if oxygen is present outside. In practice, there can be traces of sulfate derived from ion-exchange resin beads, and possibly dissolution of sulfide inclusions in the metal. The crack concentrates these anions by an electromigration process as described above, to the extent that metal cations from the alloy are soluble under the given crack conditions to balance their charge. The electrode potential at the crack tip may be only slightly shifted positive of its open-circuit value, but the open-circuit potential difference between the inside and outside drives the anion accumulation. When oxygen is removed, the potentials inside and outside the crack approach each other and the anion accumulation is less.

A specialized type of crack chemistry consideration occurs in ammonia-SCC of brass.[109] If we immerse a stressed brass sample in dilute – say 1 M – aqueous ammonia in a beaker open to the air, and wait, after a while we see the solution turning blue – this is the cuprammonium II complex ion, $Cu(NH_3)_4^{2+}$. Shortly after that, the sample starts to crack. The cathodic reaction is the reduction of the Cu II ammonia complex to the Cu I complex:

$$Cu(NH_3)_4^{++} + e \rightarrow Cu(NH_3)_2^{+} + 2NH_3$$

which may, depending on the conditions, reprecipitate as a black tarnish film consisting mainly of Cu_2O with some oxidized zinc.

Deep inside the crack, there are no Cu II ions – the crack walls efficiently scavenge these and reduce them to Cu I (so the old notion that adsorption of these ions causes cracking[102] is untenable). The environment at the crack tip is really an ammoniacal Cu I solution (containing zinc), and is very nearly in equilibrium with the copper in the brass. Indeed, SCC occurs in a simulation of such a solution,[110,111] dealloying can be demonstrated electrochemically, and a convincing correlation between such dealloying and SCC was obtained for monocrystals of different zinc contents.[22] Both dealloying and SCC ceased below 20% Zn (intergranular SCC may persist to lower Zn contents) – see **Chapter 2.05, Dealloying**.

2.09.6 Role of the Environment in SCC

An important clue to the role of the environment came as soon as fracture mechanics techniques started to be used for SCC investigations: nearly all SCC shows a 'plateau' crack velocity at intermediate K_I values – see **Chapter 2.08, Environmentally Assisted Cracking Overview**. We have to take care that the extent of this region is not exaggerated in studies of ductile alloys,[112] where K_I becomes invalid because of large-scale yielding at the crack tip, but it is definitely a real phenomenon. So what does that tell us? It tells us that the mechanics are not controlling the crack velocity. So if the rate controlling process is not the mechanics, it must be something 'chemical.' The crack velocity could be controlled by:

- The rate of oxidation or dissolution at the crack tip (which in turn can be controlled by the local environment in the crack and the local – not free-surface – potential).
- The rate of hydrogen adsorption or diffusion.
- The rate of some surface diffusion process of metal atoms or ions, and possibly others.

For SCC that occurs in an electrolyte, electrochemical considerations dominate our thinking. There are about six distinct electrochemical circumstances where SCC occurs – **Figure 24**. They are:

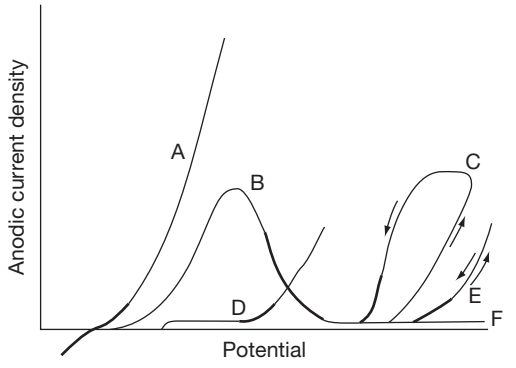


Figure 24 Regions of potential (bold lines) where SCC is likely in given electrochemical circumstances – (A) Active system, near the corrosion potential (strong alloy, hydrogen effect). (B) Active–passive transition. (C) Localized corrosion system near the repassivation potential. (D) Dealloying system near the critical potential. (E) Mild transpassivity. (F) No special reactivity, but high temperature or other exacerbating factor on film growth.

- A more-or-less actively corroding state, where SCC is entirely or predominantly due to hydrogen absorption – for example, high-strength steel in salt water.
- A compromised or borderline passive state, where SCC seems to be due to some anodic process at the crack tip – for example, carbon steel in hot caustic, nitrate, or carbonate–bicarbonate.[113]
- A state of slow localized corrosion (slow because otherwise it would be faster than the cracking) in a halide solution – for example, austenitic stainless steel in hot chloride.[16]
- A state of dealloying (see **Chapter 2.05, Dealloying**) – for example, brass in ammonia solution.[22] This does not exclude a role of dealloying in some of the other categories.
- A state of mild transpassivity, for example, due to Cr VI dissolution in a stainless alloy.
- A stable filmed state, but where the film growth kinetics, especially at deformation features, are such as to allow a certain rate of cracking despite little or no dissolution – for example, cold-worked austenitic stainless steel in hot water.[76]

Also we have to consider various hard-to-classify systems where SCC occurs, but not much seems to happen electrochemically – for example, transgranular cracking of carbon steel in anhydrous liquid ammonia or ammonia–methanol,[114,115] or intergranular cracking of nickel alloy 600 in reducing hot water.[116] These are probably exotic forms of embrittlement, caused by nitrogen[115] and oxygen[117–119] respectively, rather than hydrogen. It was also suggested, more speculatively, that transgranular SCC of carbon steel in $CO–CO_2–H_2O$[120] could be due to carbon embrittlement,[121] and that CO might be involved in SCC of carbon steel in oxygenated alcohols.[122]

Most, though not all, of these SCC schemes can be summarized as follows – SCC may occur, given the right metallurgy and stress, when the metal is passive, but not too passive. Most SCC requires some interaction between the environment and bare metal, and a passive film that forms almost instantaneously screens that interaction. Whether the actual crack growth occurs by purely anodic means, by hydrogen embrittlement, or by some other mechanism, is not always clear; Parkins' concept of the 'stress corrosion spectrum' – hydrogen effects at one end, dissolution at the other, remains valid today.[123,124] As a prelude to the next sections, we show '1980s' and '2000s' SCC spectra in **Figure 25**.

2.09.7 Anodic SCC Models

Since the 1940s, a popular scheme for the growth of SCC has been the accumulation of small dissolution or oxidation events at the crack tip, each caused by the rupture of some partially protective surface film – **Figure 26**. A popular term, probably coined by R.W. Staehle, is 'slip-dissolution,' recognizing that not all (or perhaps not any) of the oxidized metal remains actually dissolved in the small volume of solution inside the crack.

In the SDM, the plateau crack velocity is explained by postulating that above a certain K_I value, the crack tip is essentially bare the whole time, and dissolves or oxidizes at a limiting rate that is set by the local chemistry and potential in the crack. The actual value of the plateau velocity is then given by:

$$v = \frac{i_{max}\Omega}{nF} \qquad [1]$$

where v is the crack velocity, i_{max} the maximum anodic current density flowing on the bare metal surface, Ω the molar volume of the metal, F is Faraday's constant, and n is the number of electrons transferred per metal atom dissolved or oxidized.

In the 1970s and 1980s, many publications showed agreement of plateau velocities with eqn [1], where i_{max} was determined by various transient electrochemical methods, including rapid tensile straining,[125] scratching,[126] potential stepping,[127] or rapid potential scanning.[113] Parkins played a great role in

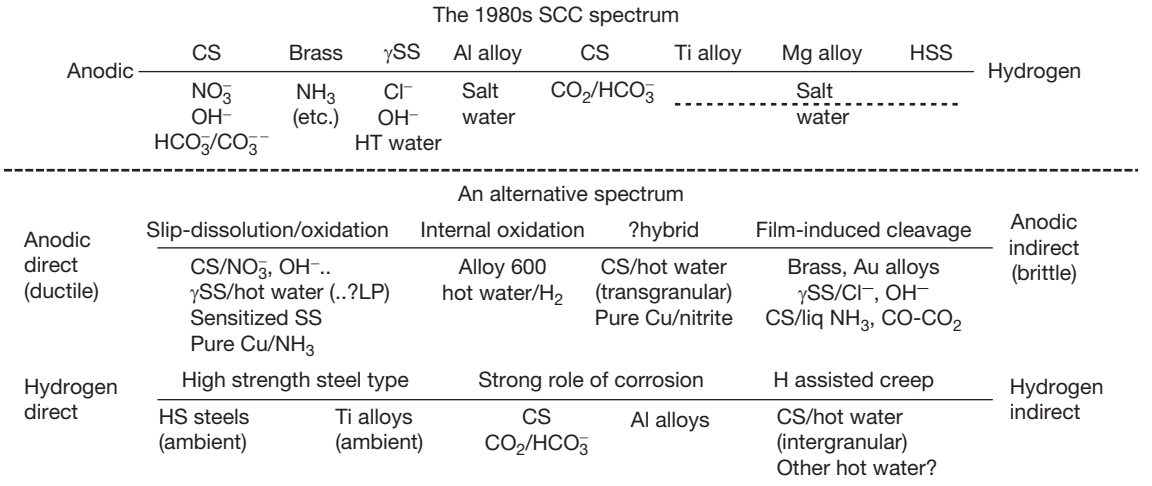


Figure 25 Old and new SCC 'spectra.'

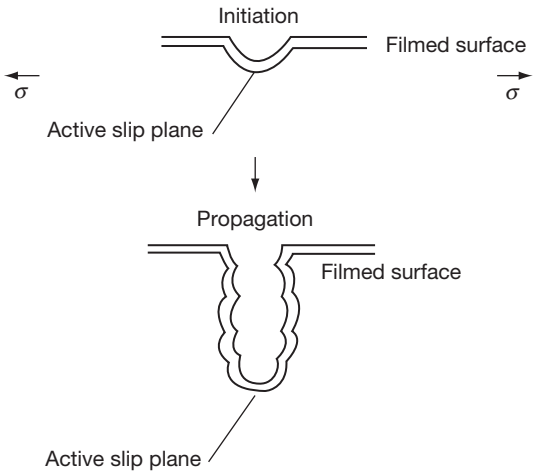


Figure 26 The slip-dissolution (or oxidation) model, after Staehle.

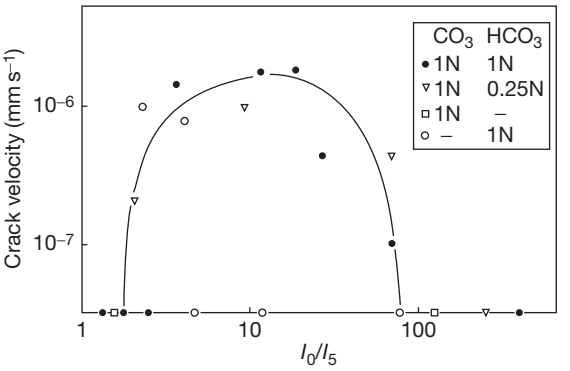


Figure 27 Analysis of current decays after potential stepping from a reducing potential to the given potential, for carbon steel in aqueous carbonate–bicarbonate, showing correlation with SCC velocity (I_0 = peak current; I_5 = current after 5 s). Modified from Parkins, R. N. *Corros. Sci.* **1980**, *20*, 147–166.

evaluating and comparing these techniques in carbon steel systems – **Figures 27 and 28**. But throughout, there were concerns that the model was being overstretched beyond carbon steel, with such views emanating especially from specialists in hydrogen embrittlement. Application of the SDM to SCC of Al–Mg alloys was certainly controversial.[128] Currently, it appears that the SDM may be valid for intergranular SCC of carbon steel in caustic, nitrate, or carbonate–bicarbonate solutions, and for a number of stainless alloys in hot water. Transgranular SCC in general does not seem to fit well with the SDM – according to Sieradzki, growth of a sharp transgranular crack is not possible under static loading,[129] and there are kinetic difficulties in many systems, such as brass in nitrite solution.[130] In general, it is probably safe to say that the SDM may apply to a number of cases of intergranular SCC in passivating environments without chloride-induced localized corrosion. It certainly does not apply to – say – chloride-induced SCC (Cl-SCC) of austenitic stainless steel, where the condition of the crack walls and tip is precisely that of an actively corroding crevice with a low pH and no possibility of very high local anodic current densities.[16,131] Recent pronouncements to the effect that the SDM does explain Cl-SCC[132] did not take into account the crack chemistry or realistic current densities that could flow at the crack tip when the metal is already in the active state.

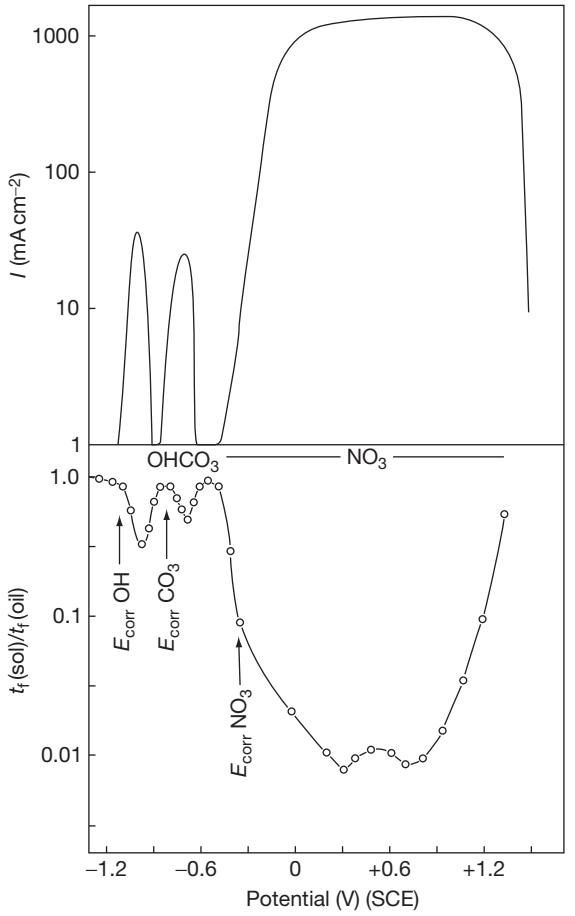


Figure 28 Similar correlation to **Figure 27**, but using the difference between the current densities measured in fast and slow potential scans. Modified from Parkins, R. N. *Corros. Sci.* **1980**, *20*, 147–166.

Below the plateau crack velocity, the SDM postulates that partial repassivation of the metal occurs between film rupture events at the crack tip – **Figure 29**. So instead of i_{max}, we write $\bar{i}$ as the mean anodic current density at the crack tip; this is given by:

$$\bar{i} = \frac{1}{\tau}\int_0^{\tau} i(t)\ \mathrm{d}t \qquad [2]$$

where $i(t)$ is the anodic current density transient on the bare metal surface, and τ is the interval between film rupture events. The crack velocity is now:

$$v = \frac{\bar{i}\Omega}{nF} = \frac{\Omega}{nF\tau}\int_0^{\tau} i(t)\ \mathrm{d}t \qquad [3]$$

If we write the crack-tip strain rate as s (for the moment we suppose this is applied externally), and the strain

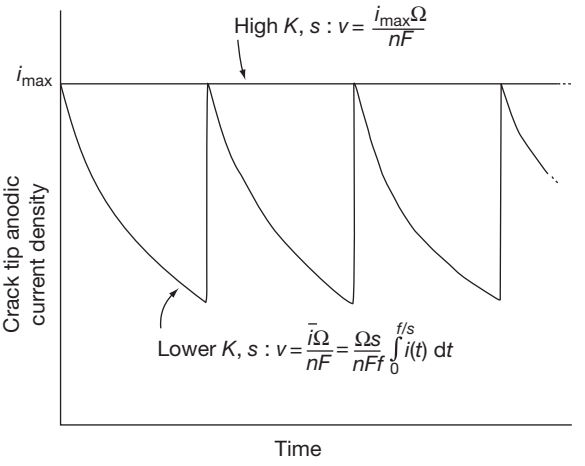


Figure 29 The variation of anodic current density at the crack tip with time, as envisaged by the slip-dissolution model.

to failure of the oxide film as f, then $\tau = f/s$, and

$$v = \frac{\Omega s}{nFf}\int_0^{f/s} i(t)\ \mathrm{d}t \qquad [4]$$

This treatment begs a number of questions, including – is the process really stepwise like this, or can it be considered as a more continuous process of stretching of a partially formed, perhaps gel-like film? Probably it is stepwise locally, but along the crack front at any given moment, there will be elements of surface that are at all possible stages of the cycle.

It is fundamental to the SDM that dynamic plasticity is occurring at the crack tip. Sometimes this is supplied by external means, such as thermal stresses that accumulate during the start-up of a BWR. But what about purely static loading – where does the dynamic strain come from in that case? In part, this can be due to low-temperature creep, but most authors have appealed to the crack growth itself as the source of the dynamic strain. Ahead of a crack, there is a plastic strain distribution $\varepsilon(r)$ where r is distance from the crack tip – **Figure 30**. Since ε falls off with increasing r, any corrosion or oxidation process that extends the crack must cause the whole strain distribution to move in the direction of crack growth, and this causes a strain transient ahead of the crack. This scheme was first clearly described by Vermilyea[133] and has been developed recently by Shoji[134] using a more modern strain distribution that had performed well in direct experimental tests on a ferritic material[135] – **Figure 31**. This possible, successful, though much-criticized, form is due to

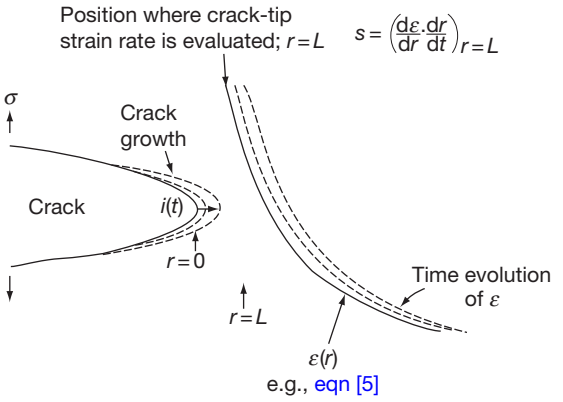


Figure 30 Basis for growth of SCC by slip-dissolution under static load – the crack growth itself provides dynamic plastic strain.

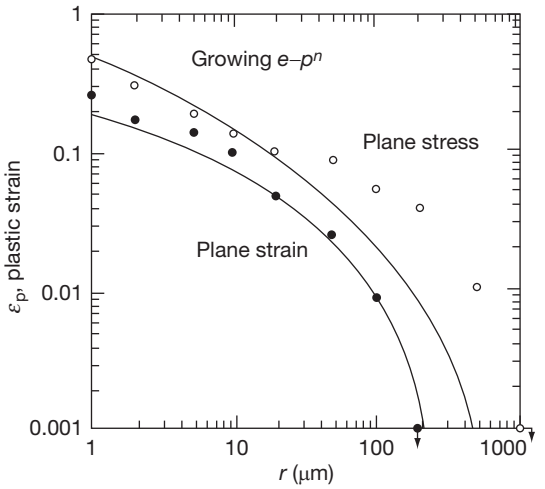


Figure 31 Comparison of experimental plastic strain distributions ahead of a growing crack with two reasonable models (for FeSi monocrystal samples). Reproduced from Gerberich, W. W.; Davidson, D. L.; Kaczorowski, M. J. *Mech. Phys. Solids* **1990**, *38*, 87–113.

Gao and Kwang[136] – they gave the strain ahead of the crack tip as:

$$\varepsilon_{ct} = \beta\left(\frac{\sigma_y}{E}\right)\left\{\ln\left[\left(\frac{\lambda}{r}\right)\left(\frac{K}{\sigma_y}\right)^2\right]\right\}^{n/(n-1)} \qquad [5]$$

where r is distance ahead of the crack, and n is the inverse strain hardening coefficient; β and λ are dimensionless.

Recently this picture has been challenged in some of its details, and its status remains unclear. According to Sieradzki,[129] it can only work if the corrosion is strongly directional, leading to the idea that it is – perhaps – only valid for intergranular SCC where there is an active path of some kind. M.M. Hall also criticizes the Vermilyea–Andresen–Shoji approach in many respects and has used a modified creep cracking approach to develop a different type of 'SDM'.[137–139] Generally, Shoji's variant of the SDM is very attractive and makes a good playing field to explore the dependence of SCC on various parameters, but the last word has not been said on this subject using such continuum approaches.

When we discussed cold work effects in SCC, discussion of the approach taken in the SDM was deferred to this section. Essentially, the SDM proposes that a stronger material has a sharper crack. Now this has become a controversial issue, not least because stress corrosion cracks are much sharper than one would naively expect from the equations of ordinary elastic–plastic fracture mechanics, such as the one that gives the crack-tip opening displacement (δ) as a function of stress intensity factor, modulus, and flow stress:

$$\delta = \frac{K_1^2}{E\sigma_0} \qquad [6]$$

which gives values for δ on the order of microns rather than the experimentally observed nanometers.[140]

But, allowing the basic assumption of the sharper crack, the SDM postulates, following Vermilyea, that the average dissolution (oxidation) rate at the crack tip (allowing for film rupture and partial repassivation) and the time-averaged plastic strain rate at and ahead of the crack tip mutually cause and determine each other. Some might call this a 'circular' argument, but it is not, in principle. So, we develop two equations – one giving the strain transient that results from a particular corrosion transient, and the other being the converse. Solving these simultaneous equations gives the crack velocity as a function of the stress intensity factor K_I, with one major adjustable parameter – the cutoff distance ahead of the crack tip at which we calculate the plastic strain rate for the purposes of the calculation (the applicable strain distribution does not allow for calculation of the strain or strain rate exactly at the crack tip). Now, we can see that if the material is stronger, and the crack sharper, the corrosion transient at the crack tip does not have to penetrate the material so deeply to create a given strain rate (other things being equal), so when we solve the equations we get a faster crack velocity. The present author developed Shoji's equations, identified some critical issues, and made

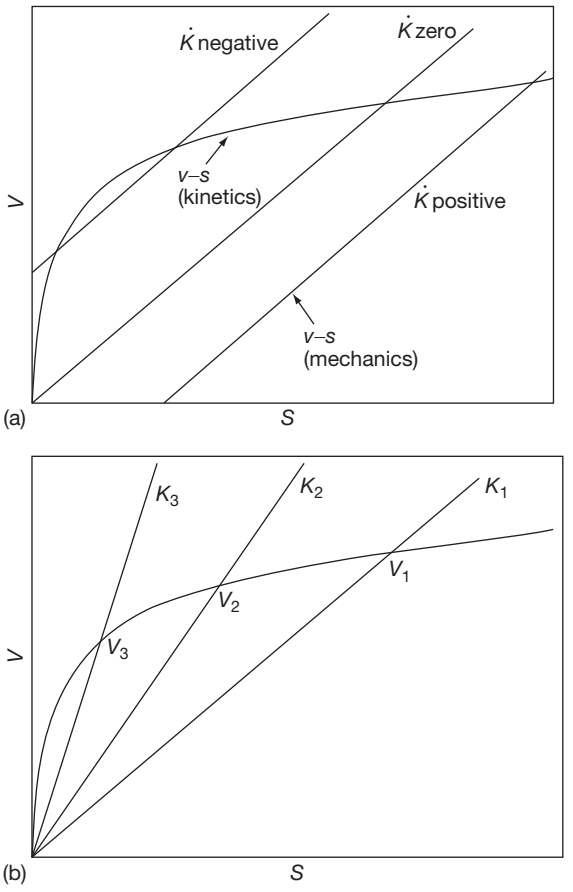


Figure 32 (a) Graphical depiction of how solutions to the equations of the slip-dissolution model lie at the intersection of two relationships between crack velocity and crack-tip strain rate. The possibility of two crack velocities satisfying the equations for negative loading rate is illustrated. (b) Illustration of how this graphical formulation of the slip-dissolution model leads to the creation of a *V–K* curve. Reproduced from Newman, R. C.; Healey, C. *Corros. Sci.* **2007**, *49*, 4040–4050.

some new predictions such as the existence of dual steady states under certain loading conditions[141] – Figure 32.

2.09.7.1 Interlude – Ultra-Sharp Cracks

Spectacular work has been done on the geometry, micromechanics, and microchemistry of SCC by Bruemmer and Thomas at the Pacific Northwest National Laboratory, which directly addresses this issue of crack sharpness as well as many other microscopic mechanistic issues, including strain localization by shear bands, alloy enrichment and depletion ahead of the crack, oxygen penetration, and porosity development ahead of the crack, as well as issues connected with neutron irradiation and impurity concentration (Pb. . .S).

Several figures are reproduced from these authors' recent work with their original captions in Figures 33–36.

2.09.8 Chemistry Considerations for SCC by Slip-Dissolution

When unalloyed carbon steel cracks in nitrate, caustic, or carbonate–bicarbonate solution, the conditions are always such as to promote the formation – perhaps fleetingly – of soluble complexes. In caustic solutions, the soluble species is $Fe(OH)_4^{2-}$; in carbonate–bicarbonate, it is $Fe(CO_3)_2^{2-}$ [142] (or similar bicarbonate complex[143]), and in nitrate it is probably an unknown complex that forms with some reduction product of nitrate. The present author used to believe that nitrate SCC was happening near the active–passive transition, like caustic or carbonate–bicarbonate SCC, but actually, the potentials are very oxidizing, in what could be called a transpassive region[113] – at least when the solution is hot and concentrated. But iron does not normally suffer from transpassive dissolution. What must be happening is that normally, nitrate contacting an iron surface is reduced so far (e.g., to N_2 or NH_4^+) that it becomes harmless, but that in a certain range of potentials, it is reduced to something that forms a soluble complex with iron. We do not know what that is – it could be an NO species. 'Transpassive' corrosion also occurs in zirconium in strong nitric acid, and must be due to a similar complexation effect.[144]

As Parkins has often shown, intergranular SCC of carbon steel is preceded by mild intergranular corrosion, and to the extent that fine carbides are to blame, this can be rationalized by their diminished passivation ability.[145] Similarly, where phosphorus segregation is involved, insights can be obtained by studying an amorphous iron phosphorus alloy[29] or simply an iron phosphide, as a rough simulation of the reacting material at the grain boundary. In the case of sensitized stainless steels or nickel base alloys, a range of alloys with reduced Cr contents can be prepared for transient electrochemical experiments.[126]

2.09.9 More Comments on Hydrogen Embrittlement as a Mechanism of SCC

Some of the oldest controversies in SCC research concern the role of hydrogen embrittlement. R.N.

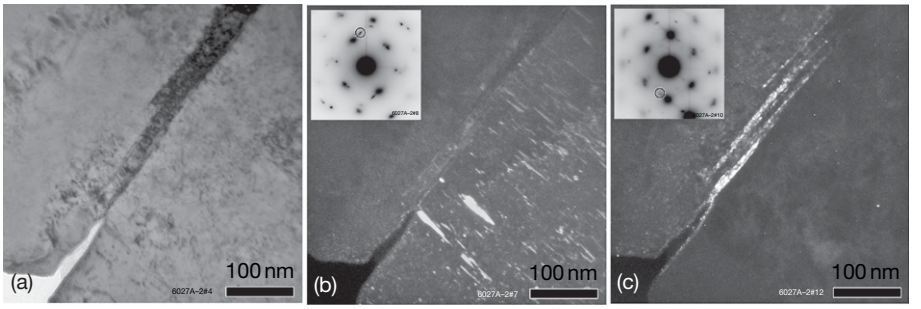


Figure 33 Crack tip region in CW316LSS sample tested in BWR HWC: (a) TEM brightfield image showing shear bands at tip. (b) $< 111 >$ Relrod darkfield image highlighting faults (one of four variants) in grain at right. (c) Relrod darkfield image highlighting faults along grain boundary in grain at left. Inset diffraction patterns show relrod reflections used for imaging. Commentary – these images show extraordinary martensite-like shear transformations ahead of cracks in cold-worked stainless steel, and vividly illustrate the strain localization that is responsible for much of the severity of SCC in cold worked material. Adapted from Toloczko, M. B.; Andresen, P. L.; Bruemmer, S. M. SCC Crack Growth of Cold-Worked 316LSS in BWR Oxidizing and Hydrogen Water Chemistry Conditions, Proceedings of the 13th International Conference Environmental Degradation of Materials in Nuclear Power Systems – Water Reactors, Canadian Nuclear Society, 2007; Paper 141.

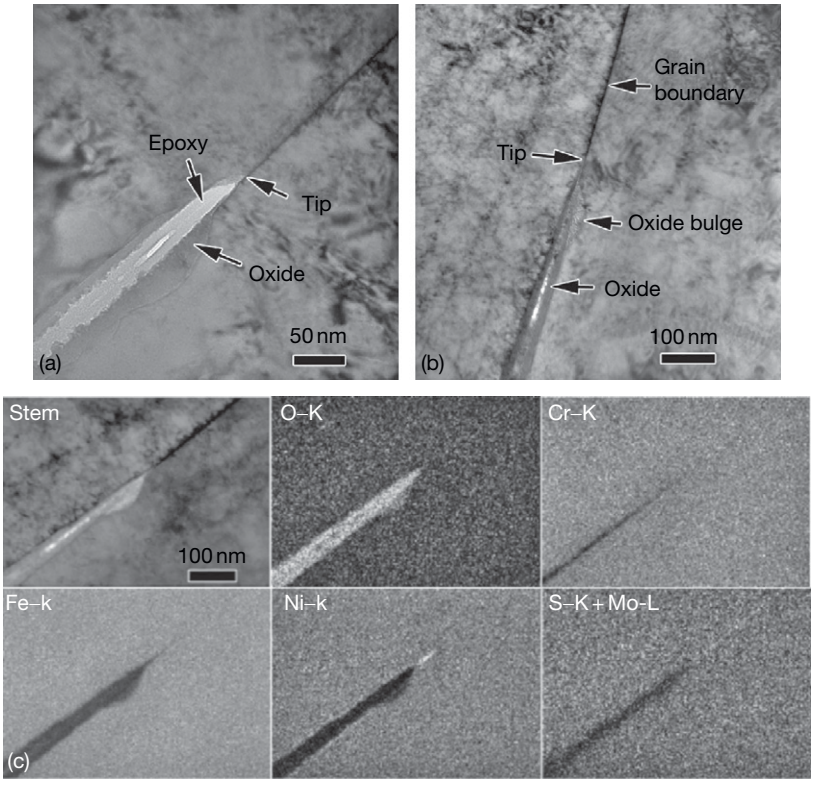


Figure 34 Open (a) and oxide-filled (b) crack tips in a BWR core shroud sample. EDS x-ray maps of oxide-filled tip (c) show Fe/Ni loss in oxide, Ni concentrated ahead of tip, and Mo segregation along the leading grain boundary. Commentary – this is neutron-irradiated material from plant, so there could be complex processes at play, including atomic mobility that leads to Ni enrichment ahead of the crack, and Mo segregation. Adapted from Thomas, L. E.; Edwards, D. J.; Asano, K.; Ooki, S.; Bruemmer, S. M. Crack-Tip Characteristics in BWR Service Components, Proceedings of the 13th International Conference Environmental Degradation of Materials in Nuclear Power Systems – Water Reactors, Canadian Nuclear Society, 2007; Paper 143.

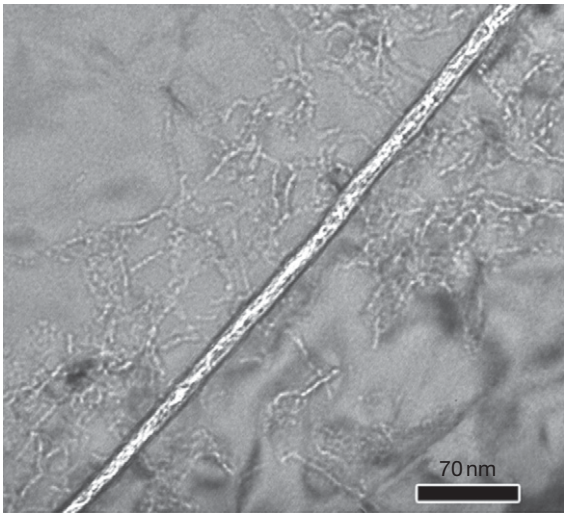


Figure 35 TEM underfocus image showing oxide-filled IG crack and adjacent tunnel-like structure along walls in Ringhals-4 alloy 182 sample. Commentary – the striking feature is the apparent oxidation along dislocation lines either side of the crack. Adapted from Thomas, L. E.; Bruemmer, S. M. High-Resolution Analytical Microscopy of Environmental Cracks in Alloy 182 Weldments, 11th International Conference on Environmental Degradation of Materials in Nuclear Power Systems – Water Reactors, American Nuclear Society, 2003; p 1212.

Parkins used to refer to a stress corrosion 'spectrum' that had SCC of high-strength steel at one end and SCC of ordinary carbon steel at the other, with various aluminum and titanium systems near the middle.[123] Recently titanium alloys seem to have arrived firmly somewhere near the hydrogen (high-strength steel) end,[70] although clearly the discharge of hydrogen ions depends on electrochemical factors like the repassivation rate of the metal at the crack tip. Nowadays, no one seems to believe in Theodore Beck's 'electrochemical knife' (SDM) that he proposed in the 1960s and early 1970s.[146] A general difficulty with aluminum and titanium alloys is that hydrogen is always generated in copious quantities at the crack tip, but this does not mean it always plays an important role in cracking. Probably these mechanisms will continue to be considered somewhat hybrid in nature. A similar comment could be made about martensitic stainless steels in ambient chloride solutions, or carbon steel in near-neutral-pH bicarbonate solutions.[147]

It was once thought that high-strength low-alloy steels immersed in salt water had an acidic crack chemistry, and that this assisted hydrogen discharge in the crack,[148] but later it appeared that this was mistaken, at least for steels without chromium.[149,150] The original experiments suffered from air-oxidation of Fe (II) to Fe (III), causing acidification. But crack depth also affects the crack chemistry (deep cracks tend to self-corrode and thus have less or no tendency to acidify, even with Cr alloying), and it was sometimes reported in the 1980s that short cracks grew faster than long ones.[151] The pH changes in such short cracks depend very sensitively on the external corrosion potential (which depends on alloying content), and no generality of the 'chemical short crack effect' was ever established. Essentially, we know that acidification occurs, with Cr alloying, if we anodically polarize the material from its natural corrosion potential. So under free corrosion conditions, there has to be both Cr alloying and some other element (Mo, Ni for example) that ennobles the anodic reaction on the free surface – then, if we are (un)lucky, we can see the chemical short crack effect. Marageing steels, which contain large amounts of Ni, are good candidates for such behavior.

There is no space here for a detailed discussion of hydrogen embrittlement mechanisms, but the subject has been very well-documented, not only in journal literature but in a series of conferences.[152] To the nonspecialist, theories of hydrogen effects seem very confusing and mutually contradictory – some theories postulate a kind of hardening, leading to decohesion of the lattice or of a grain boundary,[153] while others postulate a local softening leading to highly localized plastic rupture.[154,155] Strain rate affects the predominant mechanism. Quantum mechanics modeling gives insights into the interaction of hydrogen with other solutes,[62] with dislocations, and with grain boundaries. Two types of softening effect are proposed in the literature – Lynch[152] sees this as a surface or very near-surface effect attributed to a lifting of a compressive surface stress by hydrogen adsorption (or very local absorption), while the 'HELP' (hydrogen enhanced local plasticity) mechanism[63] postulates that internal hydrogen reduces the repulsion between dislocations in their slip plane – **Figure 37**.

2.09.10 Film-Induced Substrate Fracture ('Film-Induced Cleavage')

It is recognized that certain cases of SCC are too rapid to be due to any version of the SDM, and that some of them occur under conditions where hydrogen absorption can be ruled out. Many of these have in common that they occur in relatively concentrated solid

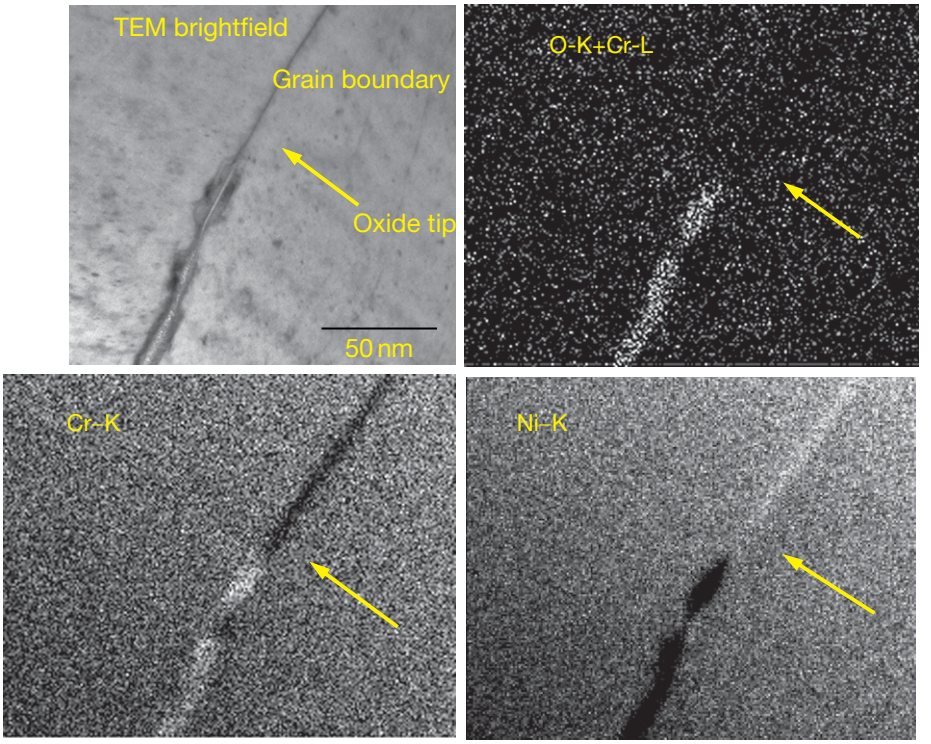


Figure 36 EDS maps showing the leading edge of IG attack ahead of primary-side IGSCC in an Alloy 600 steam generator tube are presented. Cr is highly enriched in the tip oxide and Ni is enriched along the grain boundary ahead of the tip. Commentary – these depletions and enrichments need to be studied in the context of internal oxidation and other models. Adapted from Bruemmer, S. M.; Thomas, L. E. Crack-Tip Examinations of Primary-Water Stress Corrosion Cracking in Alloy 600, Proceedings of Fontevraud 6 International Symposium on Contributions of Materials Investigations to Improve the Safety and Performance of Light-Water Reactors, French Nuclear Energy Society, 2006; p 603.

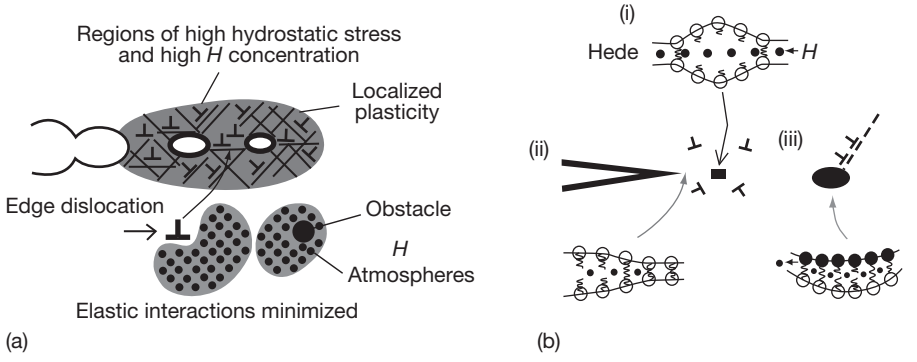


Figure 37 Plastic hydrogen fracture mechanisms, according to Lynch. Reproduced from Lynch, S. P. In Corrosion/2007; NACE: Houston, TX, 2007; Paper No 489.

solutions – brass, copper–gold or silver–gold alloys, austenitic stainless steel.[18] A natural idea is that such rapid cracking could be due to dealloying (see **Chapter 2.05, Dealloying**), but not just to a repetitive process of formation and fracture of a dealloyed layer – the rapid cracking requires that the fracture of the dealloyed layer injects a brittle microcrack, even into a fcc substrate that is normally considered unconditionally immune to brittle fracture[156] – **Figure 38**. The original source for this concept was a remarkable study on brass by Edeleanu and Forty in the late 1950s.[157] Optical microscopy of the side

surface of an α-brass crystal immersed under stress in ammonia solution showed what appeared to be regular brittle crack jumps of a few microns, each starting and ending at a slip trace. The authors stopped short of stating that the fcc lattice was undergoing a true cleavage process, as this was considered impossible even then. They proposed that the material between slip bands was embrittled by short-range order, and that corrosion – perhaps selective – at the slip band was required to reveal this intrinsic brittleness in the material. Their observations were confirmed by the group of E.N. Pugh.[158] Pugh introduced a clever method to detect discontinuities in crack growth, based on the use of periodic load pulsing to mark the fracture surfaces, and deduced that indeed the cracks were proceeding by a series of jumps. Acoustic emission was used as supporting evidence. Pugh did not adopt the short-range order suggestion, but did consider for some time that hydrogen might be causing the effect; he later rejected this, paving the way for the acceptance of a cleavage process initiated by dealloying.[159] Sieradzki and Newman presented arguments based on dislocation dynamics that went some way toward establishing a feasible framework for such effects.[160,161] This was considered a bizarre suggestion by many, but recent insights into the extraordinary mechanical properties of dealloyed layers (near-theoretical strength in compression,[162] size-scale-dependent elastic modulus,[163] surface stress-driven bending[164]) are reawakening interest in the possibility that such crack injection can occur.[165] Pugh considered that intergranular cracking in brass was a continuous process, but actually the strongest evidence for the film-induced fracture hypothesis comes from studies of intergranular SCC in noble-metal alloys[166–168] – **Figure 39**. Crack jumps into a silver–gold substrate of 50–100 µm can easily be obtained from preformed dealloyed layers. A key experiment is to dealloy under an applied potential, then step the potential to a value where little or no further

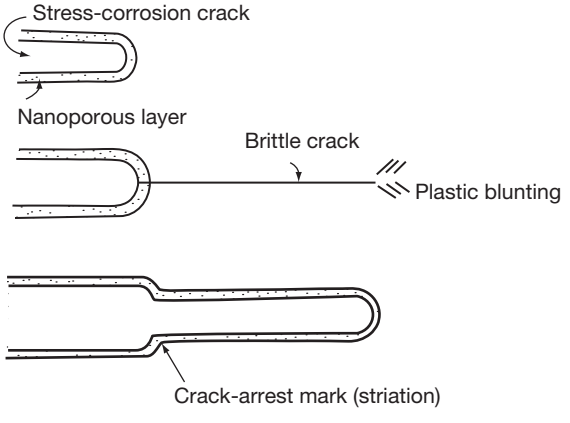


Figure 38 The film-induced cleavage concept, for a system showing dealloying.

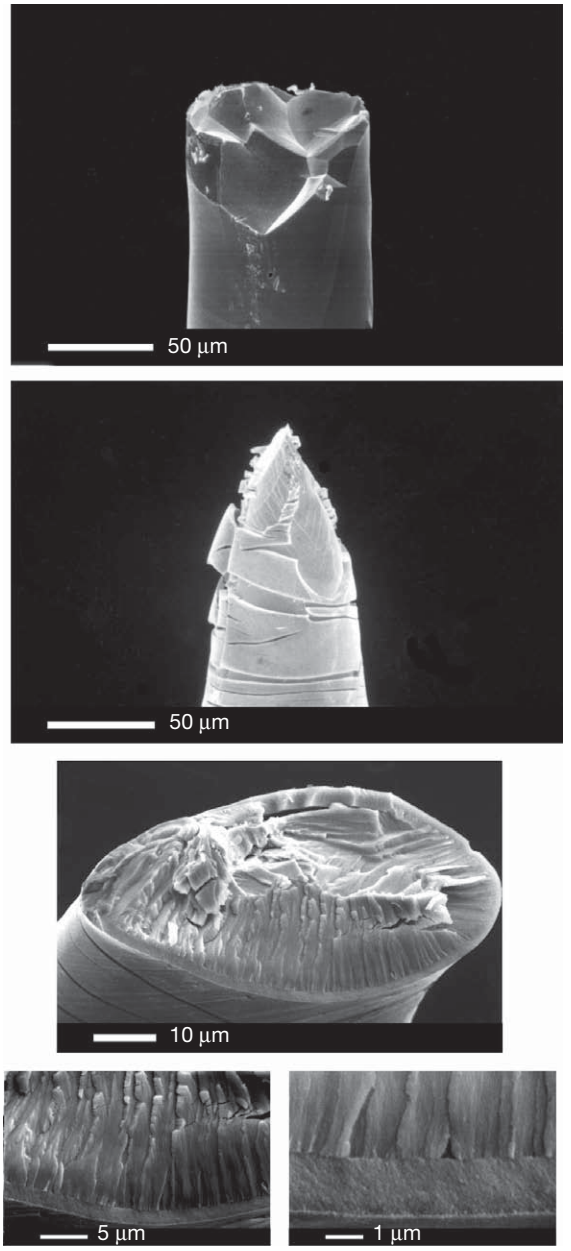


Figure 39 Brittle fracture of AuAg – intergranular or transgranular – obtained after surface dealloying; the middle picture illustrates the effect of ageing the dealloyed layer, which destroys its ability to inject a brittle crack by coarsening its nanoporosity. Reproduced from Barnes, A.; Senior, N. A.; Newman, R. C. *Metall. Trans. A* **2009**, *40*, 58–68.

faradaic reaction can occur, then load the sample to failure. This also showed good promise for transgranular fracture,[166] although more work remains to be done in that area. In both cases, ageing of the dealloyed layers destroys their ability to inject a crack – probably owing to coarsening of their porosity by surface diffusion.

Clearly, the crack injection phenomenon relies on the special properties of dealloyed layers, including their intrinsic brittleness. A key aspect is the absence of a crystallographic interface – dealloyed layers retain the same lattice orientation that they had before dealloying, so the crack can cross into the substrate without meeting any dislocation structures. Provided it is injected at a velocity close to the velocity of sound in the crystal, it can continue to grow for some distance even in a fcc material, as discussed by Sieradzki and Newman.[158]

2.09.11 Surface Mobility, Vacancy Based Models, and Others

A large amount of research has been done by the group of J.R. Galvele, based around a surface mobility model (SMM) that he proposed in 1987.[169] This is a kind of creep crack growth model in which crack extension occurs by surface diffusion of metal atoms (perhaps partially oxidized or chelated – this is not always clear) away from the crack tip. The basic crack growth equation is – using Galvele's own notation:

$$\text{cpr} = \frac{D_s}{L}\left[\exp\left(\frac{\sigma \cdot a^3}{kT}\right) - 1\right] \qquad [7]$$

where cpr means crack propagation rate, D_s is the applicable surface (self) diffusivity, L is a diffusion length along the crack flank (Galvele always uses 10^{-8} m for this), σ is the opening stress at the crack tip, and a^3 is the atomic size (approximate volume of a vacancy). The crack velocity is considered to correlate with the melting point of a surface compound (which in turn correlates with surface diffusivity on a contaminated surface), or in some cases with a kind of enhanced exchange of metal atoms at the crack surface (exchange current), for which a different equation was developed[170]:

$$D_s = \frac{i_0 N_A A n^2 a^4}{6F} \qquad [8]$$

where i_0 is the exchange current density, N_A is Avogadro's number, A is the atom fraction of base (reactive) metal atoms in the alloy (this was called AF in the original paper), n is a kind of dimensionless distance or average hop size, a is the atom size, F is Faraday's constant, and the 6 arises because this is the number of neighboring sites of an atom in a close-packed plane.

The SMM does have the attractive feature that (with an atomically sharp crack, and allowing for some adjustable parameters such as stress at the crack tip) it seems to rationalize literature data on SCC velocities. Not enough independent research has been done on the concept to give a final answer as to its merits.

The chemical potential gradients could be in the right direction for crack growth in the SMM, but – according to Friedersdorf and Sieradzki[171] – only if one neglects capillary effects – tight cracks would close up, and the SMM can only work with a relatively blunt crack (as in the creep analogues that can be found in the literature). Friedersdorf and Sieradzki also stated that the change in vacancy formation energy due to stress (σa^3) had been wrongly assigned because this quantity is appropriate for the interior of the solid, not the crack-tip surface. They also argued that chemical potentials should be used, not concentrations. The crack velocities estimated by Friedersdorf and Sieradzki are many orders of magnitude lower than those suggested by Galvele, although their amended model does have the drawback that it assumes equilibrium vacancy concentrations, as pointed out by Galvele himself.[172]

One source of confusion in the published discussions of the Galvele model is that between 1987 and 1994, he shifted the location of the site where the stress was assumed to act from the crack tip surface to somewhere just ahead of the crack, thus – perhaps – justifying the use of σa^3. This is an unsatisfactory aspect of the SMM, as it introduces an arbitrary additional vacancy transport process that must occur within the solid – this is simply assumed to be rapid compared with transport along the crack flank. Friedersdorf and Sieradzki were analyzing a version of the SMM that had been changed a few months before they published their analysis, to deal with the σa^3 problem.

Our summary of the current state of the SMM is shown in **Figure 40**.

Intriguing intergranular SCC phenomena were demonstrated by the Galvele group for copper and silver, and silver alloys, in hot, dry halogen atmospheres.[173–175] The lowest temperature where cracking was studied was 200 °C. These observations were considered to be a cornerstone of the SMM, as they related directly to classic studies of enhanced surface

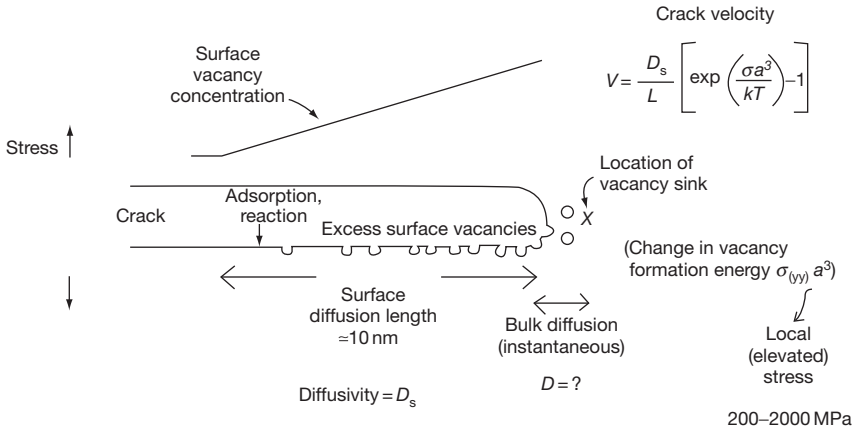


Figure 40 Current state of the Surface Mobility Model, as envisaged by the author.

mobility due to halogen adsorption.[176] Other explanations are possible, of course. Ion mobility in copper and silver halides could be high.

2.09.11.1 Vacancy Injection Models for Low-Temperature SCC

It is interesting that the moving of the vacancy sink ahead of the crack turns the SMM into a *vacancy injection* model, qualitatively similar to those proposed by other authors, including E.I. Meletis[177] and the late Denny Jones[178] who believed that he had detected interdiffusion in metallic bilayers subjected to anodic dissolution. In those models, the enhancement by stress of equilibrium vacancy concentration inside the solid was not taken into account – at least not explicitly – although it was probably part of the thought process.

One of several 'orphan' SCC models is that of Aaltonen and others in Finland, based on internal friction studies[179,180] – **Figure 41**. These results are so surprising (much more so, we would submit, than film-induced cleavage) that no other author has seriously addressed them. Essentially, these authors showed that anodic polarization of copper wires at 80 °C in a sodium nitrite solution, under conditions promoting duplex Cu(I)–Cu(II) oxide formation, produced defects that behaved like vacancies under the standard protocols used for internal friction measurement. They even claimed a reduction in shear modulus of the whole wire specimen. Later they reported an increase in creep rate under similar conditions.[181] One's immediate reaction is that even if such vacancies were to be produced, they must be confined within nanometers of the metal–film interface, so how is it possible for this nm-thick layer to dominate the internal friction response, or sample modulus, when the smallest sample dimension is 0.5 mm? Yet the results are quite striking, showing Hasiguti peaks in the anodically treated samples. One possibility (acting as devil's advocate) is that there is an autoreduction or disproportionation of a thick oxide, which produces metallic copper with a high defect density. Yet even in that case, the defected metallic layer should be a tiny fraction of the total sample dimension. But we should certainly not dismiss such observations. Metals like copper and silver may have high mobility for certain substitutional elements.

Very recently (Hanninen *et al.*, unpublished) a variety of techniques have been used to support the proposed vacancy injection model. Dislocation structures are altered near the surface during anodic polarization in the range of potentials that cause SCC of copper in nitrite. This does not mean that vacancies are penetrating the material to the entire depth to which altered behavior is observed – as shown many years ago, a near-surface effect can alter bulk mechanical properties, at least in fairly thin wire or sheet. So the relevance of vacancy injection to crack propagation may lie in the inducement of more intense localized plasticity, rather than some kind of embrittlement by the vacancies themselves.

Of course, vacancies can be involved in creep-like SCC phenomena when the temperature exceeds 300 °C, as discussed earlier.[34]

2.09.11.2 Adsorption-Induced Dislocation Emission

According to Lynch (private communication), 'surface mobility' can be harmonized with a mechanism of

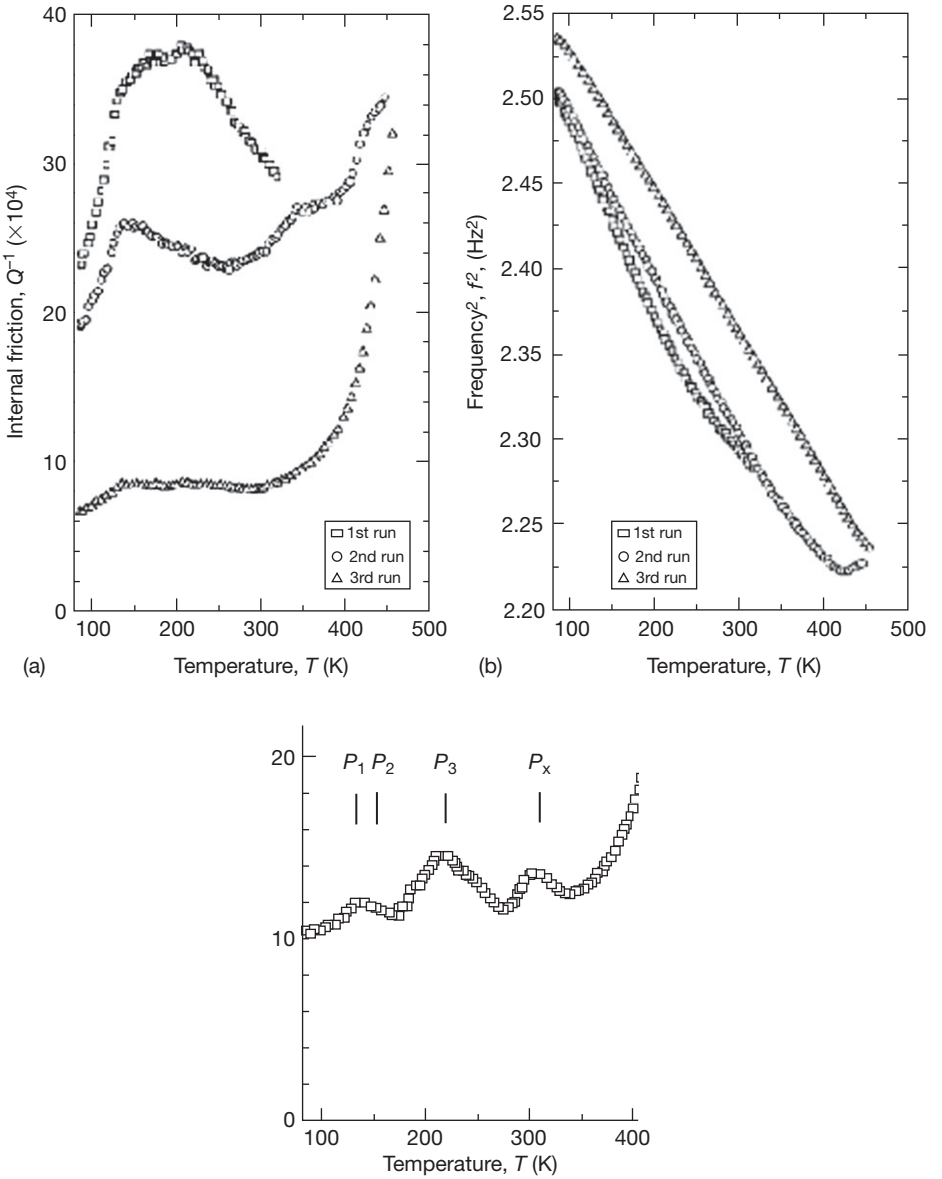


Figure 41 Upper: Temperature dependence of internal friction Q^{-1} (a) and oscillation frequency f^2 (b) for copper polarised in 0.3 M $NaNO_2$ solution for 2 h at −150 m V_{SCE} and subsequently for 10 h at +100 m V_{SEC} at 353 K (80 °C). During first run heating was performed up to 320 K, and during second and third runs up to 450 K. Lower: Location of the usual Hasiguti peaks for copper. Reproduced from Aaltonen, P.; Jagodzinski, Y.; Tarasenko, A.; Smouk, S.; Hanninen, H. *Corros. Sci.* **1998**, *40*, 903–908.

crack growth wherein adsorption of species from the environment, including hydrogen, facilitates dislocation emission from the crack tip, leading to a highly localized plastic fracture.[152] As he points out, both mechanisms involve the weakening of bonds between metal atoms, and some proportionality between the resulting crack growth rates might be expected. A strength of Lynch's proposal is that it easily encompasses liquid metal and hydrogen embrittlement, whereas these are treated a little vaguely in the SMM. Like Galvele, Lynch cites particular experiments showing very high crack velocities, to argue against dissolution or other bulk effects, but with more justification, in that the cracking of nickel monocrystals in mercury or hydrogen does seem to be a continuous process without crack jumps, whereas Galvele's studies of SCC in noble-metal alloys[182] claim continuous crack velocities of up to 20 mm s^{-1}, but are

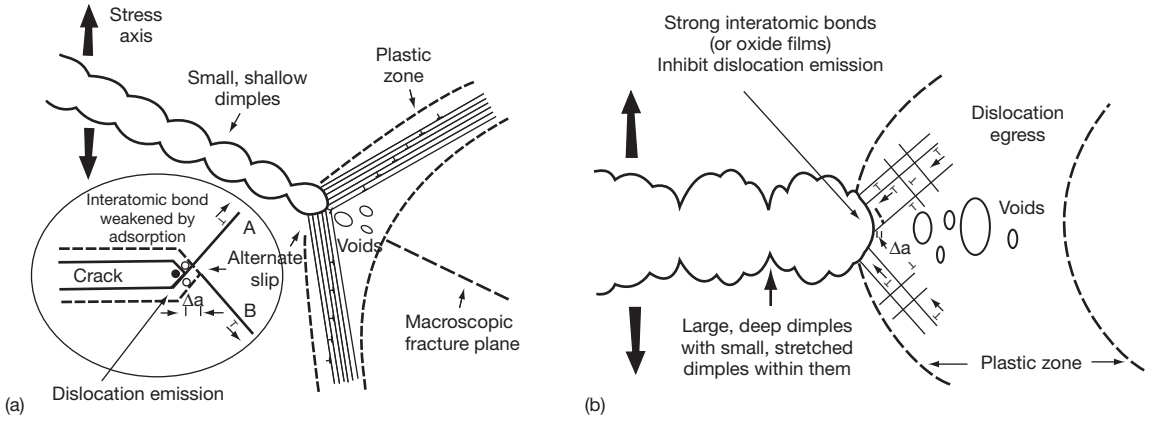


Figure 42 The AIDE mechanism. Reproduced from Lynch, S. P. In Corrosion/2007; NACE: Houston, TX, 2007; Paper No 489.

contradicted by the experimental evidence for large brittle crack jumps in such systems.[166]

From an energetic point of view, the Lynch mechanism, or adsorption induced dislocation emission (AIDE) relies on the occurrence of intense plasticity within a very small volume at the crack tip. This can lower the total energy of fracture, which is integrated over the whole deformed volume around the crack. The proposed micromechanism is shown in **Figure 42**.

A possible weakness of the AIDE mechanism used to be the insistence on a truly 2D interaction rather than any 3D reaction mechanism. Nowadays, Lynch tends to accommodate a near-surface process zone, although in the case of LME in immiscible systems, this must be very shallow. An intriguing suggestion, based on first-principles calculations for Ga on Al, is that LME systems, bulk-miscible or not, are those in which there is facile alloying of the liquid metal into the first one or two atomic layers of the substrate.[183] Since surface alloying is also known in vacuum deposition and in underpotential deposition of metals on copper and other substrates,[184,185] this raises the interesting possibility that – for example – Pb ions could cause cracking of copper or other metals by an underpotential displacement reaction followed by a LME-type cracking.

Other local softening models exist, such as that of T. Magnin, which although mainly developed for hydrogen effects[153] has also been applied to 'anodic' SCC.[186] Such softening may very well occur – the question is whether, in a given instance, it is essential for cracking. In corrosion fatigue, it is well-established that ordinary active dissolution in noble metals like copper can promote cracking,[187] and so the dissolution must be enhancing plasticity, since fatigue is essentially a plastic fracture (no evidence of brittleness is seen on such fractures). So it is not at all unreasonable to propose a similar effect in SCC, at least where dynamic straining is applied externally, as in a SSRT.

References

1. Kilburn, C. R. J.; Petley, D. N. *Geomorphology* **2003**, *54*, 21–32.
2. Ogata, S.; Shimojo, F.; Kalia, R. K.; Nakano, A.; Vashishta, P. *J. Appl. Phy.* **2004**, *95*, 5316–5323.
3. Lynch, S. P.; Moutsos, S. *J. Failure Anal. Preven.* **2006**, *6*, 54–69.
4. Lynch, S. P. *Mater. Sci. Eng. A: Struct. Mater.: Prop., Microstruct. Process.* **2007**, *468–470*, 74–80.
5. Kerns, G. E.; Staehle, R. W. *Scr. Metall.* **1972**, *6*, 1189–1194.
6. Le Calvar, M.; Scott, P. M.; Magnin, T.; Rieux, P. *Corrosion* **1998**, *54*, 101–105.
7. *Stress-Corrosion Cracking – The Slow Strain Rate Technique*; Ugiansky, G. M.; Payer, J. H. Eds. ASTM STP 665, ASTM: Philadelphia, 1977.
8. Sedriks, A. J. *Stress Corrosion Cracking Test Methods*; NACE: Houston, TX, 1989.
9. Pednekar, S. P.; Agrawal, A. K.; Chaung, H. E.; Staehle, R. W. *J. Electrochem. Soc.* **1979**, *126*, 701–702.
10. Staehle, R. W.; Gorman, J. A. *Corrosion* **2003**, *59*, 931–994; *60*, 5–63; **2004**, *60*, 115–180.
11. Poulson, B. *Corros. Sci.* **2004**, *46*, 729–753.
12. Newman, R. C. Stress Corrosion Cracking Mechanisms. In *Corrosion Mechanisms in Theory and Practice*, 2nd ed; Marcus, P., Oudar, J., Eds.; Marcel Dekker: New York, 2002; pp 399–450.
13. Copson, H. R. Effect of Composition on Stress Corrosion Cracking of Some Alloys Containing Nickel. In *Physical Metallurgy of Stress Corrosion Fracture*; Rhodin, T. N., Ed.; Interscience: New York, 1959; pp 247–269.
14. Graf, L. Stress-Corrosion Cracking in Homogeneous, Nonsupersaturated Alloys Containing Noble or No Noble Metals. In *Fundamental Aspects of Stress-Corrosion Cracking*; Staehle, R. W., Forty, A. J., van Rooyen, D., Eds.; NACE: Houston, TX, 1969; pp 187–201.

15. Latanision, R. M.; Staehle, R. W. Stress-Corrosion Cracking of Iron-Nickel-Chromium Alloys. In *Fundamental Aspects of Stress-Corrosion Cracking*; Staehle, R. W., Forty, A. J., van Rooyen, D., Eds.; NACE: Houston, TX, 1969; pp 214–307.
16. Newman, R. C.; Corderman, R. R.; Sieradzki, K. *Br. Corros. J.* **1989**, *24*, 143–148.
17. Tamaki, K.; Tsujikawa, S.; Hisamatsu, Y. Development of a New Test Method for Chloride Stress Corrosion Cracking of Stainless Steels in Dilute NaCl Solutions. In *Advances in Localized Corrosion*; Isaacs, H. S., Bertocci, U., Kruger, J., Smialowska, S., Eds.; NACE: Houston, TX, 1991; pp 207–214.
18. Nisbet, W. J.; Lorimer, G. W.; Newman, R. C. *Corros. Sci.* **1993**, *35*, 457–469.
19. Deakin, J.; Dong, Z.; Lynch, B.; Newman, R. C. *Corros. Sci.* **2004**, *46*, 2117–2133.
20. Maier, A.; Fernandez, J. S.; Galvele, J. R. *Corros. Sci.* **1995**, *37*, 1–16.
21. Serebrinsky, S. A.; Galvele, J. R. *Corros. Sci.* **2004**, *46*, 591–612.
22. Sieradzki, K.; Kim, J. S.; Cole, A. T.; Newman, R. C. *J. Electrochem. Soc.* **1987**, *134*, 1635–1639.
23. Pugh, E. N.; Craig, J. V.; Sedriks, A. J. Stress-Corrosion Cracking of Copper, Silver and Gold Alloys. In *Fundamental Aspects of Stress-Corrosion Cracking*; Staehle, R. W., Forty, A. J., van Rooyen, D., Eds.; NACE: Houston, TX, 1969; pp 118–158.
24. Barnes, A.; Deakin, J.; Newman, R. C. *Corrosion* **2007**, *63*, 416–418.
25. Staehle, R. W. Private communication.
26. Uhlig, H. H. Evaluation of Stress-Corrosion Cracking Mechanisms. In *Fundamental Aspects of Stress-Corrosion Cracking*; Staehle, R. W., Forty, A. J., van Rooyen, D., Eds.; NACE: Houston, TX, 1969; pp 86–97.
27. Atrens, A.; Wang, J. Q.; Stiller, K.; Andren, H. O. *Corros. Sci.* **2006**, *48*, 79–92.
28. Bandyopadhyay, N.; Briant, C. L. *Corrosion* **1985**, *41*, 274–280.
29. Bandyopadhyay, N.; Newman, R. C.; Sieradzki, K. Caustic SCC of an Fe–3Ni alloy, effects of alloying additions or impurities; Proceedings of the 9th International Congress on Metallic Corrosion, Toronto, 1984; Vol. 2, pp 210–215, NRC: Ottawa.
30. Harrison, R. P.; de G Jones, D.; Newman, J. F. Caustic Cracking of Temper Embrittled Steels. In *Stress-Corrosion Cracking and Hydrogen Embrittlement of Iron-Base Alloys*; Staehle, R. W., Hochmann, J., McCright, R. D., Slater, J. E., Eds.; NACE: Houston, TX, 1977; pp 659–662.
31. Danielson, M. J.; Oster, C. A.; Jones, R. H. *Corros. Sci.* **1991**, *32*, 1–21.
32. Scamans, G. M.; Holroyd, N. J. H.; Tuck, C. D. S. *Corros. Sci.* **1987**, *27*, 329–347.
33. Senior, N. A.; Newman, R. C. *Nanotechnology* **2006**, *17*, 2311–2316.
34. Arioka, K.; Yamada, T.; Terachi, T.; Miyamoto, T. *Corrosion* **2008**, *64*, 691–706.
35. Staehle, R. W. Private communication.
36. Paraventi, D. J.; Angeliu, T. M.; Was, G. S. *Corrosion* **2002**, *58*, 675–686.
37. Was, G. S. *Fundamentals of Radiation Materials Science, Metals and Alloys*; Springer: New York, 2007.
38. Faulkner, R. G. *J. Nucl. Mater.* **1997**, *251*, 270–276.
39. Was, G. S.; Busby, J. T.; Allen, T.; Kenik, E. A.; Jensson, A.; Bruemmer, S. M.; Gan, J.; Edwards, A. D.; Scott, P. M.; Andreson, P. L. *J. Nucl. Mater.* **2002**, *300*, 198–216.
40. Busby, J. T.; Sowa, M. M.; Was, G. S.; Kenik, E. A. *J. ASTM Int.*; 2004, 1.
41. Jiao, Z.; Busby, J. T.; Was, G. S. *J. Nucl. Mater.* **2007**, *361*, 218–227.
42. Hackett, M. J.; Busby, J. T.; Was, G. S. *Metall. Mater. Trans. A.* **2008**, *39*, 218–224.
43. Andresen, P. L.; Ford, F. P. *Mater. Sci. Eng. A.* **1988**, *A103*, 167–184.
44. Andresen, P. L. *Corrosion* **2008**, *64*, 439–464.
45. Kim, Y.-J.; Niedrach, L. W.; Andresen, P. L. *Corrosion* **1996**, *52*, 738–743.
46. King, A.; Johnson, G.; Engelberg, D.; Ludwig, W.; Marrow, J. *Science* **2008**, *321*, 382–385.
47. Samans, C. H. *Corrosion* **1964**, *20*, 256t–262t.
48. Isaacs, H. S.; Vyas, B.; Kendig, M. W. *Corrosion* **1982**, *38*, 130–136.
49. Newman, R. C.; Roberge, R.; Bandy, R. *Corrosion* **1983**, *39*, 386–390.
50. Newman, R. C.; Bandy, R.; Roberge, R. *Environmental Degradation of Materials in Nuclear Power Systems – Water Reactors*; NACE: Houston, TX, 1984; pp 636–647.
51. Horowitz, H. H. *Corros. Sci.* **1983**, *23*, 353–362.
52. Jones, R. H.; Gertsman, V. Y.; Vetrano, J. S.; Windisch, C. F. *Scr. Mater.* **2004**, *50*, 1355–1359.
53. deWexler, S. B.; Galvele, J. R. *J. Electrochem. Soc.* **1974**, *121*, 1271–1276.
54. Ramgopal, T.; Gouma, P. I.; Frankel, G. S. *Corrosion* **2002**, *58*, 687–697.
55. Brennenstuhl, A. unpublished research.
56. Lee, Y.; Gangloff, R. P. *Metall. Mater. Trans. A.* **2007**, *38*, 2174–2190.
57. Asahi, H.; Yagi, A.; Ueno, M. *ISIJ Int.* **1993**, *33*, 1190–1195; **1993**, *33*, 1275–1280; **1994**, *34*, 290–294.; *Corrosion* **1994**, *50*, 537–545.
58. Chu, R.; Chen, W.; Wang, S.-H.; King, F.; Jack, T. R.; Fessler, R. R. *Corrosion* **2004**, *60*, 275–283.
59. Parkins, R. N.; Blanchard, W. K., Jr.; Delanty, B. S. *Corrosion* **1994**, *50*, 394–408.
60. Parkins, R. N.; Zhou, S. *Corros. Sci.* **1997**, *39*, 159–173.
61. Kehler, B. A.; Scully, J. R. *Corrosion* **2008**, *64*, 465–477.
62. Gavriljuk, V. G.; Shivanyuk, V. N.; Shanina, B. D. *Acta Mater.* **2005**, *53*, 5017–5024.
63. Robertson, I. M. *Eng. Fract. Mech.* **2001**, *68*, 671–692.
64. Pundt, A.; Kirchheim, R. *Annu. Rev. Mater. Res.* **2006**, *36*, 555–608.
65. Toribio, J. *Eng. Fract. Mech.* **2008**, *75*, 2683–2694.
66. Scamans, G. M.; Alani, R.; Swann, P. R. *Corros. Sci.* **1976**, *16*, 443–459.
67. Behnood, N.; Cai, H.; Evans, J. T.; Holroyd, N. J. H. *Mater. Sci. Eng. A.* **1989**, *A119*, 23–32.
68. Singh, P. M.; Lewandowski, J. J.; Holroyd, N. J. H.; Evans, J. T. *Proceedings of Parkins Symposium on Fundamental Aspects of Stress Corrosion Cracking*; Meletis, E. I., Parkins, R. N., Eds.; TMS, Warrendale, PA, 1992; pp 567–583.
69. Holroyd, N. J. H.; Vasudevan, A. K.; Christodoulou, L. *Aluminum Alloys – Contemporary Research and Applications*; Academic Press, Inc.: London, 1989; pp 463–483.
70. Gaudett, M. A.; Scully, J. R. *Metall. Mater. Trans. A.* **1999**, *30A*, 65–79A.
71. Gaudett, M. A.; Scully, J. R. *Metall. Mater. Trans. A.* **2000**, *31A*, 81A–92A.
72. Teter, D. F.; Robertson, I. M.; Birnbaum, H. K. *Acta Mater.* **2001**, *49*, 4313–4323.
73. Peterson, M. H.; Brown, B. F.; Newbegin, R. L.; Groover, R. E. *Corrosion* **1967**, *23*, 142–148.
74. Meyn, D. A. *Corros. Sci.* **1967**, *7*, 721–723.

75. Newman, R. C.; Procter, R. P. M. *Br. Corros. J.* **1990**, *25*, 259–269.
76. Andresen, P. L.; Morra, M. M. *J. Nucl. Mater.* **2008**, *383*, 97–111.
77. Lozano-Perez, S.; English, C. A.; Terachi, T. In Corrosion/2008; NACE: Houston, TX, 2008; Paper No. 495.
78. Peng, Q. J.; Kwon, J.; Shoji, T. *J. Nucl. Mater.* **2004**, *324*, 52–61.
79. Turnbull, A. *Corrosion* **2008**, *64*, 420–438.
80. Parkins, R. N.; Singh, P. M. *Corrosion* **1990**, *46*, 485–499.
81. Parkins, R. N.; Belhimer, E.; Blanchard, W. K., Jr. *Corrosion* **1993**, *49*, 951–966.
82. Chen, W.; King, F.; Vokes, E. *Corrosion* **2002**, *58*, 267–275.
83. Akashi, M.; Nakayama, G. *Effects of the Environment on the Initiation of Crack Growth*; ASTM STP 1298, Philadelphia, PA, 1997; pp 150–164.
84. Cheung, C.; Erb, U.; Palumbo, G. *Mater. Sci. Eng. A.* **1994**, *185*, 39–43.
85. Schuh, C. A.; Kumar, M.; King, W. E. *Acta Mater.* **2003**, *51*, 687–700.
86. Engelberg, D. L.; Humphreys, F. J.; Marrow, T. J. *J. Microsc.* **2008**, *230*, 435–444.
87. Reed, B. W.; Kumar, M.; Minich, R. W.; Rudd, R. E. *Acta Mater.* **2008**, *56*, 3278–3289.
88. Palumbo, G.; Aust, K. T.; Lehockey, E. M.; Erb, U.; Lin, P. *Scr. Mater.* **1998**, *38*, 1685–1690.
89. Zaefferer, S.; Wright, S. I.; Raabe, D. *Metall. Mater. Trans. A.* **2008**, *39*, 374–389.
90. Wells, D. B.; Stewart, J.; Herbert, A. W.; Scott, P. M.; Williams, D. E. *Corrosion* **1989**, *45*, 649–660.
91. Gaudett, M. A.; Scully, J. R. *Metall. Trans. A.* **1994**, *25A*, 775–787.
92. Lehockey, E. M.; Brennenstuhl, A. M.; Thompson, I. *Corros. Sci.* **2004**, *46*, 2383–2404.
93. Marrow, T. J.; Babout, L.; Jivkov, A. P.; Wood, P.; Engelberg, D.; Stevens, N.; Withers, P. J.; Newman, R. C. *J. Nucl. Mater.* **2006**, *352*, 62–74.
94. Babout, L.; Marrow, T. J.; Engelberg, D.; Withers, P. J. *Mater. Sci. Technol.* **2006**, *22*, 1068–1075.
95. Jivkov, A. P.; Marrow, T. J. *Theor. Appl. Fract. Mech.* **2007**, *48*, 187–202.
96. Jivkov, A. P.; Stevens, N. P. C.; Marrow, T. J. *J. Pressure Vessel Technol.-Trans. ASME*; 2008, 130.
97. Duxbury, P. M.; McGarrity, E. S.; Holm, E. A. *Mech. Mater.* **2006**, *38*, 757–771.
98. Turnbull, A. *Corrosion* **2001**, *57*, 175–189.
99. *Corrosion Chemistry within Pits, Crevices and Cracks*; Turnbull, A. Ed. HMSO: London, 1987.
100. Turnbull, A.; Horner, D. A.; Connolly, B. J. *Eng. Fract. Mech.* **2009**, *76*, 633–640.
101. Shinohara, T.; Tsujikawa, S.; Hisamatsu, Y. *Boshoku Gijutsu (Corros. Eng.)* **1985**, *34*, 283–290.
102. Uhlig, H. H. In *Stress-Corrosion Cracking and Hydrogen Embrittlement of Iron-Base Alloys*; Staehle, R. W., Hochmann, J., McCright, R. D., Slater, J. E., Eds.; NACE: Houston, TX, 1977; pp 174–179.
103. Shoji, S.; Ohnaka, N.; Furutani, Y.; Saito, T. *Boshoku Gijutsu (Corros. Eng.)* **1986**, *35*, 559–565.
104. Shoji, S.; Ohnaka, N. *Corros. Eng.* **1989**, *38*, 111–120.
105. Drugli, J. M.; Steinsmo, U. *ACOM (Sweden)* **1998**, *4*, 7.
106. Hinds, G.; Turnbull, A. *Corros. J.* **2008**, *64*, 101–106.
107. Turnbull, A.; Zhou, S.; Nicholson, P.; Hinds, G. *Corros. J.* **2008**, *64*, 325–333.
108. Galvele, J. R. *J. Electrochem. Soc.* **1976**, *123*, 464–474.
109. Bertocci, U.; Pugh, E. N. In *Corrosion Chemistry Within Pits, Crevices and Cracks*; Turnbull, A., Ed.; HMSO: London, 1987; pp 187–198.
110. Bertocci, U.; Thomas, F. I.; Pugh, E. N. *Corrosion* **1984**, *40*, 439–440.
111. Shahrabi, T.; Newman, R. C.; Sieradzki, K. *J. Electrochem. Soc.* **1993**, *140*, 348–352.
112. Speidel, M. O. *Corrosion* **1977**, *33*, 199–203.
113. Parkins, R. N. *Corros. Sci.* **1980**, *20*, 147–166.
114. Procter, R. P. M.; Omideyi, M. V. E.; Newman, R. C.; Zheng, W. *Corros. Sci.* **1992**, *33*, 1009–1031.
115. Newman, R. C.; Zheng, W.; Procter, R. P. M. *Corros. Sci.* **1992**, *33*, 1033–1051.
116. Scott, P. M.; Combrade, P. In *ASM Handbook, Vol. 13C. Corrosion: Environments and Industries* 2006; Vol. 13C, pp 362–385.
117. Scott, P. M.; Le Calvar, M. In Proceedings of the Sixth International Symposium on Environmental Degradation of Materials in Nuclear Power Systems – Water Reactors 1993; pp 657–667.
118. Newman, R. C.; Gendron, T. S.; Scott, P. M. Proceedings of the Ninth International Symposium on Environmental Degradation of Materials in Nuclear Power Systems – Water Reactors, TMS-AIME, Warrendale, PA, 2000; pp 79–93.
119. Scenini, F.; Newman, R. C.; Cottis, R. A.; Jacko, R. J. *Corrosion* **2008**, *64*, 824–835.
120. Brown, A.; Harrison, J. T.; Wilkins, R. Electrochemical Investigations of Stress Corrosion Cracking of Plain Carbon Steel in Carbon Dioxide–Carbon Monoxide–Water System. In *Stress-Corrosion Cracking and Hydrogen Embrittlement of Iron-Base Alloys*; Staehle, R. W., Hochmann, J., McCright, R. D., Slater, J. E., Eds.; NACE: Houston, TX, 1977; pp 686–695.
121. Hannah, I. M.; Newman, R. C.; Procter, R. P. M. Environmental Cracking of C–Mn Steel in CO/CO2/H2O Solutions. In *Hydrogen Effects on Materials Behaviour*; Moody, N. R., Thompson, A. W., Eds.; TMS: Warrendale, PA, 1990; pp 965–974.
122. Newman, R. C. *Corrosion* **2008**, *64*, 819–823.
123. Parkins, R. N. *Br. Corros. J.* **1972**, *7*, 15–28.
124. Parkins, R. N. *JOM* **1992**, *44*, 12–19.
125. Diegle, R. B.; Vermilyea, D. A. *J. Electrochem. Soc.* **1975**, *122*, 180–188.
126. Newman, R. C.; Sieradzki, K. *Corros. Sci.* **1983**, *23*, 363–378.
127. Smart, N. R.; Scott, P. M.; Procter, R. P. M. *Corros. Sci.* **1990**, *30*, 877–901.
128. Ford, F. P.; Burstein, G. T.; Hoar, T. P. *J. Electrochem. Soc.* **1980**, *127*, 1325–1331.
129. Sieradzki, K. Atomistic and Micromechanical Modeling Aspects of Environment-Induced Cracking of Metals. In *Environment-Induced Cracking of Metals*; Gangloff, R. P., Ives, M. B., Eds.; NACE: Houston, TX, 1990; pp 125–137.
130. Cole, A. T.; Newman, R. C.; Sieradzki, K. *Corros. Sci.* **1988**, *28*, 109–118.
131. Newman, R. C.; Mehta, A. In *Stress–corrosion Cracking of Austenitic Alloys, Environment–Induced Cracking of Metals*; Gangloff, R. P., Ives, M. B., Eds.; NACE: Houston, TX, 1990; pp 489–510.
132. Holzleitner, S.; Kranister, W.; Mori, G.; Schmuki, P. In Corrosion/2008; NACE: Houston, TX, 2008; Paper No. 486.
133. Vermilyea, D. A. A Film Rupture Model for Stress Corrosion Cracking. In *Stress-Corrosion Cracking and Hydrogen Embrittlement of Iron-Base Alloys*; Staehle, R. W., Hochmann, J., McCright, R. D., Slater, J. E., Eds.; NACE: Houston, TX, 1977; pp 208–217.

134. Lu, Z.; Shoji, T. *J. Pressure Vessel Technol. (Trans. ASME)* **2006**, *128*, 318–327.
135. Gerberich, W. W.; Davidson, D. L.; Kaczorowski, M. *J. Mech. Phys. Solids* **1990**, *38*, 87–113.
136. Gao, Y. C.; Kwang, K. C. Elastic-plastic fields in steady crack growth in a strain-hardening material. In Proceedings of the 5th International Conference on Fracture, Cannes, March/April 1981, vol. 2, pp 669–682 (1981).
137. Hall, M. M., Jr. Crack Tip Strain Rate Equation with Applications to Hydrogen Embrittlement and Active Path Dissolution Models of Stress Corrosion Cracking. In *Environment-Induced Cracking of Materials*; Shipilov, S., Jones, R. H., Olive, J. M., Rebak, R. B., Eds.; Elsevier: Amsterdam, 2008; pp 59–68.
138. Hall, M. M., Jr. *Corros. Sci.* **2008**, *50*, 2902–2905.
139. Hall, M. M., Jr. *Corros. Sci.* **2009**, *51*, 225–233.
140. Bruemmer, S. M.; Thomas, L. E. *Surf. Interface Anal.* **2001**, *31*, 571–581.
141. Newman, R. C.; Healey, C. *Corros. Sci.* **2007**, *49*, 4040–4050.
142. Davies, D. H.; Burstein, G. T. *Corrosion* **1980**, *36*, 416–422.
143. Zhou, S.; Stack, M. M.; Newman, R. C. *Corros. Sci.* **1996**, *38*, 1071–1084.
144. Nagano, H.; Kajimura, H. *Corros. Sci.* **1996**, *38*, 781–791.
145. Green, J. A. S.; Parkins, R. N. *Corrosion* **1968**, *24*, 66–69.
146. Beck, T. R.; Grens, E. A., Jr. *J. Electrochem. Soc.* **1969**, *116*, 177–184.
147. Parkins, R. N.; Beavers, J. A. *Corrosion* **2003**, *59*, 258–273.
148. Brown, B. F.; Fujii, C. T.; Dahlberg, E. P. *J. Electrochem. Soc.* **1969**, *116*, 218–219.
149. Cottis, R. A.; Alavi, A.; Taqi, E. A. Chemical Conditions and Hydrogen Generation Within Crevices in Carbon-Manganese Steels. In *Advances in Localized Corrosion*; Isaacs, H. S., Ed.; NACE-9: Houston, TX, 1990; pp 403–406.
150. Turnbull, A.; Ferriss, D. H. *Corros. Sci.* **1987**, *27*, 1323–1350.
151. Gangloff, R. P. *Metall. Trans. A* **1985**, *16A*, 953–969.
152. Now joined with another series called Corrosion Deformation Interactions – International Conference on Hydrogen Effects on Materials Behavior and Corrosion Deformation Interactions; Moran, WY; USA; Sept. 2002; pp 22–26.
153. Katz, Y.; Tymiak, N.; Gerberich, W. W. *Eng. Fract. Mech.* **2001**, *68*, 619–646.
154. Lynch, S. P. In *Corrosion/2007*; NACE: Houston, TX, 2007; Paper No 489.
155. Delafosse, D.; Magnin, T. *Eng. Fract. Mech.* **2001**, *68*, 693–729.
156. Rice, J. R.; Thomson, R. *Phil. Mag.* **1974**, *29*, 73–97.
157. Edeleanu, C.; Forty, A. J. *Phil. Mag.* **1960**, *58*, 1029–1040.
158. Beavers, J. A.; Pugh, E. N. *Metall. Trans. A.* **1980**, *11A*, 809–820.
159. Pugh, E. N. *Corrosion* **1985**, *41*, 517–526.
160. Sieradzki, K.; Newman, R. C. *Philos. Mag. A.* **1985**, *51*, 95–132.
161. Sieradzki, K.; Newman, R. C. *J. Phys. Chem. Solids* **1987**, *48*, 1101–1113.
162. Volkert, C. A.; Lilleodden, E. T.; Kramer, D.; Weissmüller, J. *Appl. Phys. Lett.* **2006**, *89*, 061920.
163. Mathur, A.; Erlebacher, J. *Appl. Phys. Lett.* **2007**, *90*, 061910.
164. Kramer, D.; Viswanath, R. N.; Weissmüller, J. *Nano Lett.* **2004**, *4*, 793–796.
165. Weissmüller, J.; Newman, R. C.; Jin, H. J.; Hodge, A. M.; Kysar, J. W. *MRS Bull.* **2009**, *34*, 577–586. Special issue.
166. Kelly, R. G.; Shahrabi, T.; Frost, A. J.; Newman, R. C. *Metall. Trans. A.* **1991**, *22A*, 191–197.
167. Saito, M.; Smith, G. S.; Newman, R. C. *Corros. Sci.* **1993**, *35*, 411.
168. Barnes, A.; Senior, N. A.; Newman, R. C. *Metall. Trans. A.* **2009**, *40*, 58–68.
169. Galvele, J. R. *Corros. Sci.* **1987**, *27*, 1–33.
170. Galvele, J. R. *Acta* **2000**, *45*, 3537–3541.
171. Sieradzki, K.; Friedersdorf, F. J. *Corros. Sci.* **1994**, *36*, 669–675.
172. Galvele, J. R. *Corros. Sci.* **1994**, *36*, 901–910.
173. Bianchi, G. L.; Galvele, J. R. *Corros. Sci.* **1987**, *27*, 631–635.
174. Bianchi, G. L.; Galvele, J. R. *Corros. Sci.* **1993**, *34*, 1411–1422.
175. Bianchi, G. L.; Galvele, J. R. *Corros. Sci.* **1994**, *36*, 611–619.
176. Delamare, F.; Rhead, G. E. *Surf. Sci.* **1971**, *28*, 267–284.
177. Meletis, E. I.; Lian, K. *J. Mech. Behav. Mater.* **1995**, *6*, 69–84.
178. Jones, D. A.; Jankowski, A. F.; Davidson, G. A. *Metall. Mater. Trans. A.* **1997**, *28A*, 843–850.
179. Aaltonen, P.; Jagodzinski, Y.; Tarasenko, A.; Smouk, S.; Hanninen, H. *Corros. Sci.* **1998**, *40*, 903–908.
180. Jagodzinski, Y.; Aaltonen, P.; Smouk, S.; Tarasenko, A.; Hanninen, H. *J. Alloys Compd.* **2000**, *310*, 256–260.
181. Aaltonen, P., Jagodzinski, Y., Hanninen, H. Vacancy Generation in Electrochemical Oxidation/Dissolution of Copper in $NaNO_2$ Solution and its Role in SCC Mechanism. In Proceedings of International Conference on Hydrogen Effects on Materials Behavior and Corrosion Deformation Interactions, Moran, WY, USA, 22–26 Sept. 2002, pp 597–606; TMS-AIME, 2003.
182. Serebrinsky, S. A.; Galvele, J. R. *Corros. Sci.* **2004**, *46*, 591–612.
183. Stumpf, R.; Feibelman, P. J. *Phys. Rev. B.* **1996**, *54*, 5145–5150.
184. Tersoff, J. *Phys. Rev. Lett.* **1995**, *74*, 434–437.
185. Robert, S.; Gauthier, S.; Bocquet, F.; Rousset, S.; Duvault, J. L.; Klein, J. *Surf. Sci.* **1996**, *350*, 136–144.
186. Magnin, T.; Chambreuil, A.; Chateau, J. P. *Int. J. Fract.* **1996**, *79*, 147–163.
187. Yan, B.; Farrington, G. C.; Laird, C. *Acta Metall.* **1985**, *33*, 1533–1545.

2.12 Corrosion Fatigue

R. Akid
Materials & Engineering Research Institute, Sheffield Hallam University, Howard Street, Sheffield S1 1WB, UK

Abbreviations

AA2024 Aluminum Alloy 2024
ASME American Society of Mechanical Engineers
ASTM ASTM International
BWR Boiler water reactor
c.p. Commercially pure
CF Corrosion fatigue
CGRs Crack growth rates
CP Cathodic protection
DA Damage accumulation
EAC Environmentally assisted cracking
HSLA High-strength low-alloy
LEFM Linear elastic fracture mechanics
MSC Microstructurally short crack
OCP Open circuit potential
PSBs Persistent slip bands
PSC Physically short crack
RTP Roller quench temper
S–N Stress vs number of fatigue cycles

Symbols

a Defect size
***B, x*, and β** Constants
C Pit radius
d A microstructural dimension
D LEFM threshold crack growth rate
f Frequency
F Geometry factor
K_{max} Maximum stress intensity

*It is the intention in this chapter to discuss some of the more recent developments in the field of corrosion fatigue. Given the previous edition's emphasis on long crack growth, much of this chapter is devoted to an understanding of the early stages of corrosion fatigue, often termed the 'initiation' stage.

K_{min} Minimum stress intensity
K_t Stress concentration factor
m Constant
m1, m2, A, and x Experimentally determined constants
N Number of cycles
N_f Number of cycles to failure
Q Function based on crack shape
R Stress ratio
r_p Size of the plastic zone
t Time
$t_{p/c}$ Time to pit/crack transition
y Shear strain
y_{eq} Equivalent strain
ΔK Crack tip stress intensity factor range
ΔK_{th} Threshold stress intensity factor
σ Stress
σ_a Stress amplitude
σ_y Yield strength

2.12.1 Introduction and Historical Perspective

The term *fatigue* was coined in the latter half of the nineteenth century, although this type of failure had previously been recognized in the early 1800s as a different type of fracture occurring in the axles of the railway locomotives and carriages. In 1843, Rankine[1] attributed this behavior to a gradual deterioration of the material during its service life. Other views held at the time were that the metal had become tired, lost its nature, or crystallized.

Fatigue cracking[2] is a consequence of highly localized plastic deformation, where the vast majority of cracks start at a free surface. However, it is not necessary for the bulk of the grains in a piece of metal to deform plastically for it to fail by fatigue, and continued cyclic plastic deformation in one localized surface region is sufficient for fatigue failure to occur. This aspect distinguishes the problem from any other form of mechanical failure, the basic features of fatigue failure being the 'initiation' of surface microcracks and their subsequent propagation through the bulk of the material.

Corrosion fatigue is a term used to describe cracking (including both initiation and growth) in materials subject to the combined actions of a fluctuating (cyclic) stress and corrosion. Electrochemical reactions taking place at the metal–environment interface govern whether the mechanism of failure is controlled by anodic dissolution or hydrogen embrittlement. Unlike the phenomenon '*stress corrosion cracking*' where cracking is often controlled by specific combinations of material–environment, for example, brass in ammonia-containing environments, this specificity rarely applies in this type of cracking process. In fact, environments as innocuous as moist air can lead to severe reductions in the fatigue resistance of high-strength engineering alloys.[3]

Haigh[4] published possibly the earliest recorded paper on corrosion fatigue while seeking to explain the failure of towing ropes that were kept in a state of frequent vibration when exposed to sea water. Later work by McAdam[5] and Gough[6] laid the foundations for work on the subject of corrosion fatigue. McAdam's work showed the danger of adopting higher-strength alloy steels in preference to ordinary steels. It is now established that an increase in tensile strength of low corrosion-resistance alloys is associated with a decreased resistance to corrosion fatigue. McAdam's work also included the connection between pitting and cracking and showed the time-dependency effects due to corrosion by studying the effects of frequency of loading on fatigue lifetime. The influence of the environment on fatigue is assessed by plotting the relationship of stress to number of cycles to failure, the so-called (S–N) curve. **Figure 1** shows the results of S–N tests on different substrates in a 3.5% NaCl solution at open-circuit potential and open to the atmosphere. Here, it can be seen that the most significant impact of the environment is seen at stress levels close and below the in-air fatigue limit, a region that might historically have been deemed the 'crack initiation' region. As will be discussed later, it is now known that the environment has a major effect on the early stages of crack growth in the microstructural fracture mechanics regime.[7]

The electrochemical aspects of corrosion fatigue were clearly demonstrated in the studies of Evans and Simnad.[8,9] An important aspect of this work included observations showing an increase in corrosion rate with increase in applied stress (see **Section 2.12.3**). These results highlighted an increase in metal dissolution when the degree of plasticity increased. Furthermore, it was noted that when the material was subject to stresses within the elastic range, the rate of corrosion hardly differed from when the material was unstressed.

Another feature of this work included the effects of an applied cathodic current. In this case, it was found that fatigue lives were enhanced in neutral salt solutions, in which the rate of metal dissolution at potentials below the open-circuit potential was

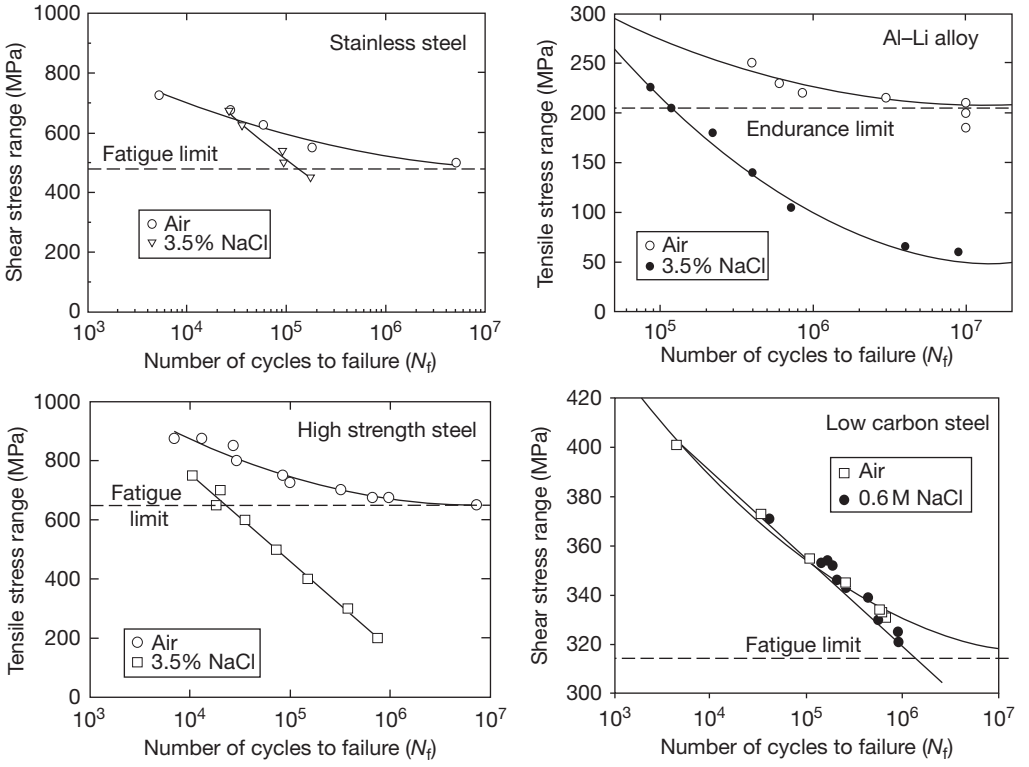


Figure 1 *S–N* curves for various alloys in 3.5% (0.6 M) NaCl. Reproduced from Miller, K. J.; Akid, R. *Proc. R. Soc. Lond. A* **1996**, *452*, 1411–1432.

negligible. For acidic solutions, a greater applied cathodic current was required for protection before an equivalent improvement in fatigue life was observed. The conclusion given was that 'while elastic deformation does not affect the chemical or electrochemical properties of the iron, deformation beyond the elastic limit (which may occur if pits produce stress intensification) alters these properties, making the iron behave like a more reactive metal. This behavior was later confirmed by measuring the growth rates of pits under different values of stress (see **Section 2.12.3.1.2**).

The following sections of this chapter will address the mechanisms and factors affecting the fatigue life of various engineering alloys operating within corrosive environments.

2.12.2 Fundamentals of Fatigue Failure

Before going on to discuss corrosion fatigue in detail, it is worthwhile briefly outlining the basic fundamentals of fatigue in the absence of an environment. Given this understanding, it is considered that the reader will better appreciate the role of the environment in the corrosion fatigue process.

Fatigue damage is responsible for around 70–80% of the rupture cases of mechanical engineering components and structures in the world; therefore, an understanding of how cracks initiate and propagate would be useful when designing these structures.

It is generally accepted that purely elastic deformations are totally reversible and do not induce damage. Furthermore, it is accepted that fatigue rupture is related to the localization of plastic deformation,[10] and during cyclic loading, the following occurs[11]:

- cyclic softening or hardening: at constant strain amplitude, the stress amplitude decreases or increases respectively,
- crack development via surface intrusion and extrusions,
- crack propagation.

The main site for crack initiation is generally the material surface, as the surface grains may deform plastically and crack more easily than other grains,

because not only are they in contact with the environment, which has a great role in fatigue damage, but they are also the only ones not totally surrounded by other grains, that is, unconstrained.

Initiation sites most frequently encountered include the material surface defects – stress concentrations near notches, corners, holes, etc. In the case of smooth specimens, without any macroscopic defects, slip bands form at the surface on activated slip systems in grains favorably oriented. These slip bands are associated with intrusions/extrusions, see **Figure 2**, which may act as crack initiation sites.

For some materials, such as body-centered cubic (bcc) alloys, initiation may be intergranular owing to shape changes of the near surface; this leads to deformation incompatibility which in turn contributes to localized deformation irreversibility at the grain boundaries. This effect has been attributed to the, thermally activated, so-called asymmetric glide of the screw dislocations in tension and compression.[10,13] Also, crack initiation very often takes place at interfaces such as inclusion/matrix or second phase/matrix.

There appears to be no real distinction between crack initiation and crack propagation, just as it is difficult to know when a slip band becomes a crack. **Figure 3** presents a schematic of the development of a crack from a free surface, while **Figure 4** relates the crack growth rate with the crack size and the relevant crack growth regime.

Figure 3 shows that a crack initiates from the free surface, usually in the largest grain along a shear slip plane at 45° to the loading axis. Plasticity induced in the grain drives the crack until a dislocation build-up occurs as the crack tip approaches the grain boundary between grains 1 and 2. As the crack tip approaches the first grain boundary (d_1), in **Figure 4**, the growth rate of the crack decreases. There are now two possible outcomes depending upon the level of the applied stress; if the stress is below the fatigue limit, the crack decelerates and may arrest, and if the stress level is above the fatigue limit, the crack grows into the next grain and can continue to propagate throughout the structure. A revised definition of the fatigue limit might be 'the stress level where a crack does not propagate beyond a defining microstructural dimension.' This dimension being one or several grains or a value related to the microstructure, for example, martensite lath size or prior austenite grain boundary size. At and around the fatigue limit, the 'short' crack growth behavior is very nonlinear. This nonlinear

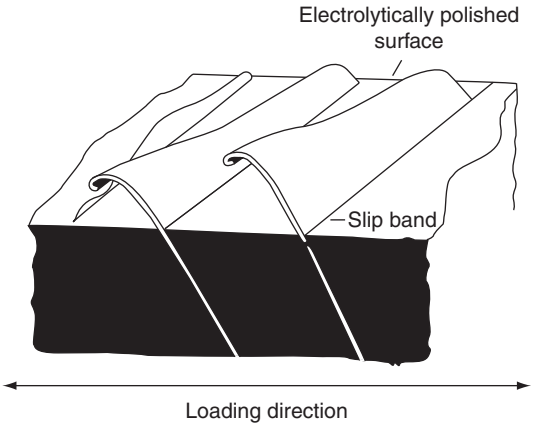


Figure 2 Formation of intrusions and extrusions at persistent slip bands (PSBs) during fatigue loading. Reproduced from Forsyth, P. J. E. *J. Inst. Metals* **1951/2**, *80*, 181.

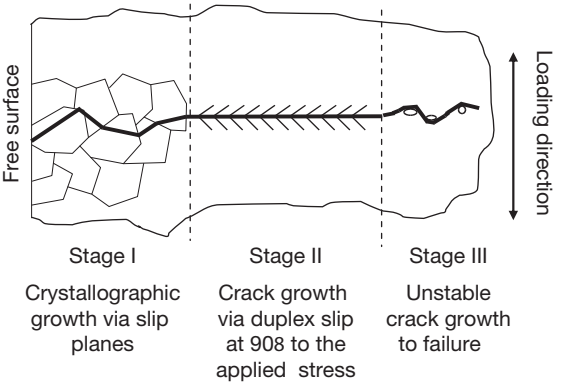


Figure 3 Schematic illustrating different fatigue crack growth stages.

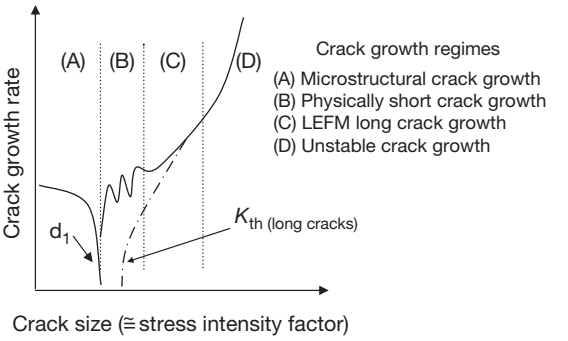


Figure 4 Fatigue crack growth rate as a function of crack size.

'short crack' growth behavior continues as the crack propagates through the microstructure. However, as the crack tip stress intensity increases with crack length, the effective barrier resistance of the grain structure diminishes until the crack growth is unaffected by the microstructure. At this point, the ratio of the crack tip plastic zone size to crack length is around 0.02, and the crack may be considered to fall within the linear elastic fracture mechanics (LEFM) regime and is described as a 'long crack' (see Section 2.12.4.2).

The first two crack growth stages (A and B) shown in Figure 4 have 'historically' been considered as the *initiation* stage, while the third regime (C) is the representative of the *propagation* stage. The variables affecting each of these stages are further discussed in Section 2.12.4.

2.12.3 Corrosion Fatigue Regimes, Mechanisms, and Modeling

The following section is aimed at addressing the different regimes that contribute to the damage processes of pit and crack development and subsequent growth of corrosion fatigue cracks. The section is divided into the growth of defects (pits and cracks) from smooth surfaces and the growth of preexisting long cracks. The examples given in this section relate predominantly to steels actively corroding within chloride environments; however, where appropriate, work illustrating other metal/environment systems is also discussed.

2.12.3.1 Initiation and Short Crack Growth

It has recently been recognized that cracks develop very early in the lifetime of a component and that the fatigue crack initiation period is considered to be a negligible phase in the fatigue failure process.[14,15] The premise is that fatigue cracks initiate below the fatigue limit but the rate of propagation is controlled by numerous factors and under some cases this rate is sufficiently small to be considered negligible, leading to infinite fatigue life. An important point here is that if the conditions change, for example, a fatigue load excursion above the fatigue limit stress, the propagation rate may change causing the crack to grow beyond a critical threshold size.

The corrosion fatigue process may be considered analogous to that of air fatigue, notably because damage occurs very early in life and the defect, pit or crack propagates at a rate governed by the loading/environmental conditions. As will be discussed later, processes such as passive film breakdown and pitting play an important role in contributing to the early stages of damage, Figure 5.

Figure 5 shows the typical development stages for air and corrosion fatigue cracking. In air, a crack develops on the surface at some defined size, a_o, often related to the surface roughness or a metallurgical feature such as an inclusion, and then grows up to the major microstructural barrier. If the stress level is not above the fatigue limit, crack arrest occurs and a nonfailure condition exists. If the stress level is just above the fatigue limit, the crack can grow beyond the major microstructural barrier and a failure condition exists. Under corrosion fatigue conditions, surface film breakdown or local microgalvanic activity occurs leading to localization of corrosion (pitting). Subject to the nature of the environment, pit growth continues until a transition to cracking occurs. Environment-assisted crack growth then ensues until failure occurs.

In terms of fatigue life, the most important regimes are (a) in air, the stage I (shear) to stage II (tensile) crack transition; Figure 3 and (b) in corrosion fatigue, the pit-crack transition. These regions are highlighted in Figure 5. At this point, the crack is propagating at its lowest rate. A similar feature is observed for corrosion fatigue loading.[7] In the case of air fatigue, the microstructural barriers impede crack growth; in corrosion fatigue, pit growth rate

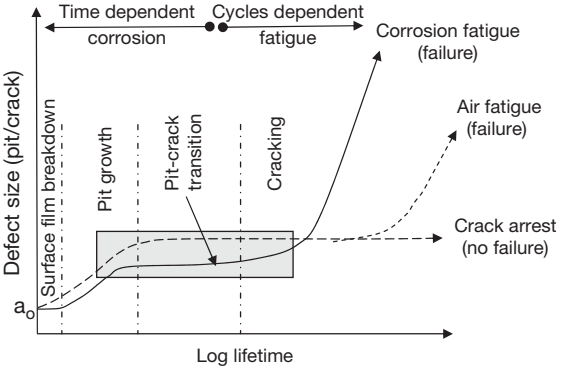


Figure 5 Schematic showing the behavior of fatigue cracks under air and corrosion fatigue conditions. Note that the stress level for the corrosion fatigue curve is less than the in-air fatigue limit of the material. Individual annotated regimes are for corrosion fatigue only. Reproduced from Akid, R. In *Effects of Environment on the Initiation of Crack Growth*; Van der Sluys, W. A., Piascik, R. S., Zawiercha, R., Eds.; ASTM STP 1298, **1997**, pp 1–17.

decreases (see Figure 8) as it is governed largely by solution chemistry and electrochemistry, which change as the pit size increases. Akid[16] undertook a comparison of these transition regimes in a high-strength steel for fatigue in air and in 0.6 M NaCl (see Figure 6). The following main conclusion can be drawn: the stage I to stage II transition in air takes place at a greater fraction of lifetime and at defect sizes larger than that of the pit/stage II transition (see shaded regions). The implication of this result is that stage II tensile cracking can be promoted by the presence of a pit and occurs earlier in the lifetime of a component subject to corrosion fatigue when compared to identical loading conditions in the absence of the environment, that is, air fatigue loading.

The term 'crack initiation' has generally been considered as that period prior to the observation of a crack. This leads to a problem, notably, 'when is a crack observable?'; hence quantification of the period of initiation becomes subjective and dependent upon the equipment being used to monitor the surface. This problem was recognized by Ford[17] who proposed that initiation, under environment-assisted cracking conditions, comprises only the time required to form a localized environment and as all surfaces contain geometrical discontinuities, this period, relative to the total lifetime, will comprise only a few percent. This is in keeping with the premise that propagation dominates the damage process.

2.12.3.1.1 Passive film breakdown and pitting

Many commercial alloys contain inclusions, and these can provide the sites for crack initiation in both air and corrosion fatigue. Inclusions are chemically different to the matrix in which they are contained and therefore will have a different electrochemical behavior to that of the matrix. Furthermore, the passive film developed over the site of an inclusion will differ primarily in its chemistry and thickness from the passive film formed on the matrix. The consequence of this is that these sites are preferentially attacked in the presence of depassivating agents such as chloride ions; for example, AlCu-type inclusions in AA2024 alloys suffer preferential chemical attack and can promote pitting.[18]

The effect of pitting on a component's fatigue life was demonstrated in the early work of McAdam,[5] and Evans and Simnad[8,9] in which it was shown that prepitting a specimen prior to air fatigue testing results in pits acting like notches, thereby lowering the fatigue limit. However, prepitting tests differ significantly from full immersion or cyclic immersion corrosion fatigue tests, which will be discussed later. Jack and Paterson[19] concluded, in their work on three turbine rotor shafts (2% Cr/Ni/Mo/V steel), that cracks initiated at corrosion pits and grew by high cycle fatigue. Metallographic examination of the rotor shafts showed that pits were widely distributed and some pits were present without cracks while all the cracks had initiated at pits. This would seem to imply that in these materials, cracking occurs after the onset of pitting and not the reverse.

Conversely, Duquette and Uhlig[20] argued that initial pit formation is not a requirement for corrosion fatigue. Evidence from their research suggested that the applied stress level determines whether cracks develop from pits or from slip bands. At high-stress levels, slip bands form quickly with insufficient time for pit formation to take place; hence slip-band

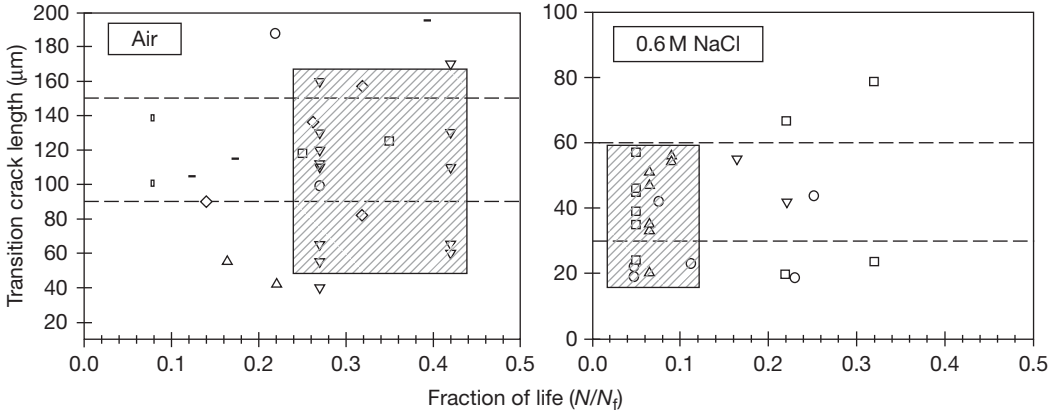


Figure 6 Comparison of stage I crack/stage II crack and pit/stage II crack transition for a high-strength carbon steel (σ_y; 1440 MPa) in air and 0.6 M NaCl. Fully reversed torsion. Reproduced from Akid, R. In *Effects of Environment on the Initiation of Crack Growth*; Van der Sluys, W. A., Piascik, R. S., Zawiercha, R., Eds.; ASTM STP 1298, **1997**, pp 1–17.

cracking dominates. It should be clarified that the above study relates to the work carried out on 0.18% C steel at pH values of 2–4. However, the development of pits or slip-band dissolution was observed to be dependent upon solution conditions as shown by Akid[21] who, using a similar steel, showed that pitting was the preferred corrosion mechanism at near-neutral pH values, while at pH 2, slip-band dissolution was the dominant process.

Magnin[22] described the effects of the stress–environment interaction on dislocations, illustrating that such interactions increase the dislocation density at the crack tip causing a more localized form of damage than would occur in the absence of the environment.

From the above discussion, it is therefore clear that substantial evidence exists for pitting to be one of the principal mechanisms for the initiation of corrosion fatigue cracks. Leis *et al.*[23] summarized some of the effects that are attributable to pitting, suggesting that pitting may accelerate the initiation process through a mechanical notch effect; pits may concentrate aggressive chemical species in the environment and that pits may serve to raise locally the stress above yield in a nominally elastically loaded material. The following section provides a summary of the influence of different variables on the development and growth of short cracks. From the results of numerous tests conducted under a wide range of conditions, a model is proposed for predicting corrosion fatigue lifetime.

Surface films and film breakdown

With the exception of some of the noble metals, the majority of engineering metals and alloys exist with a surface film that offers corrosion protection to a greater or lesser degree. Classically, this is seen in the range of alloys known as the *stainless steels*. Stainless steels contain greater than 12% Cr which, along with other alloying elements, gives rise to a mixed oxide surface film having a thickness of the order of tens of nanometers. Similarly, the corrosion resistance of Al and Al alloys is attributed to the presence of a thin aluminum oxide film.

As previously discussed, fatigue and corrosion are both highly localized surface phenomena. The nature of the metal surface, in conjunction with the chemistry of the electrolyte, therefore plays a significant role in establishing the type and degree of damage arising during corrosion fatigue. As discussed above, the inherent resistance of a material to corrosion is based upon the properties of the surface film. Breakdown of this film leads to the separation of anode and cathode sites and to the development of localized corrosion, for example, pitting. The ability to overcome this resistance to surface damage might therefore be seen as a 'primary threshold' with pit growth occurring above this threshold. A value for this threshold would be dependent upon the nature of the oxide film, notably its mechanical and chemical properties and thickness,[24] the nature of the electrolyte,[25,26] and the magnitude of applied stress to which the film is subjected.[26]

Pitting

The association of pits as precursors to crack initiation was demonstrated over 70 years ago by McAdam.[5] Recent corrosion fatigue studies[27–29] have shown that pitting is often associated with inclusions or constituent particles lying within the matrix of the alloy, for example, manganese sulfide in steel. The chemical nature, size, and distribution of such particles play an important role in the early stages of corrosion fatigue damage.

Muller[30] considered the influence of early fatigue crack growth based on crack initiation as the failure criteria. A model was proposed based on corrosion kinetics and fracture mechanics featuring three different corrosion conditions: general, pitting, and passive corrosion. In the case of pitting corrosion, the lifetime is determined by a critical pit depth from which a fatigue crack can develop. In passive corrosion conditions, the lifetime is determined by slip characteristics and repassivation kinetics leading to a critical corrosion current below which no corrosion fatigue cracks can initiate. Experimental results from crack initiation tests on chromium–nickel steels showed good correlation with the theoretical models. The model further proposed that all the three modes of corrosion are dependent upon the applied stress.

An extended discussion of the influence of stress on pitting and pit growth rates can be found from Akid.[16] In summary, it is suggested that pit growth rate is influenced by the degree of plasticity within the grains surrounding the pit and that boundary conditions exist for the effects of stress on pit growth rate, these being the onset and saturation of plastic deformation (see **Figure 7**). The pit growth rate coefficient is based upon measurements of pit size prior to the pit–crack transition. This influence of stress on pit development/growth is supported by

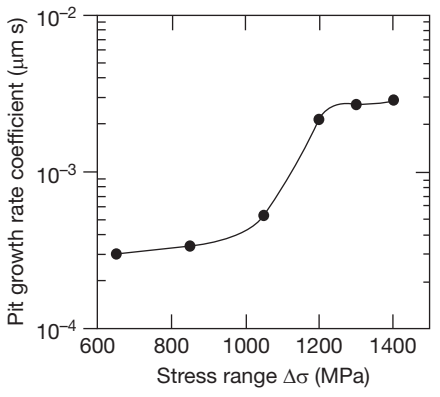


Figure 7 Influence of stress on pit growth rate coefficient. Q2N steel (0.12%C), $\sigma_y = 790$ MPa; 3 pt bend testing, $R = 0.01$; in artificial seawater. Derived from experimental data from reference. Reproduced from Wu, X. J. Ph.D. Thesis, University of Sheffield, 1995.

work recently conducted concerning the electrochemistry of deformed smooth surfaces.[32]

Pit/crack transition

It has been demonstrated above that pits act as sites for crack initiation and that pit growth rate is a function of material, environment, and loading conditions. The transitional stage from a pitting (time-dependent) dominated regime to that of a mechanical (cycles-dependent) regime is a critical stage in the corrosion fatigue process. Recent results[33] have shown that, not surprisingly, this transition stage is dependent upon local solution chemistry, stress, and time. Interestingly, during this study, which concerned the corrosion fatigue of two thermo mechanically processed steels, one containing Mo/V additions and one without, the transition times and dependencies on stress were different; however, the lifetimes were similar. This arose because the pit size at which the transition occurred differed between the two steels. Unfortunately this makes modeling of corrosion fatigue slightly more complex than would ideally be desired. However, given sufficient statistical confidence, it would be possible to develop a model based upon a 'process-competition' approach in which stages were assigned depending upon the dominant process, that is, stress-assisted pitting and dissolution-assisted cracking. Such an approach is described in Akid.[16]

An example of the pit-to-crack transition from the study by Ebera[34] is given in **Figure 8**. Here, it can be seen that a significant number of cycles elapses before a crack initiates from the pit. Ebera found that the maximum depth of the corrosion pit at the crack initiation stage is 23.5 μm in plane bending at a stress of 228 MPa ($N/N_f = 0.968$) and 17.8 μm in plane bending at a stress of 280 MPa ($N/N_f = 0.645$). However, it should be noted that the test frequency used in this study was 60 Hz which gives times for the pit-to-crack transition of 23 and 9 h, respectively for stress levels of 228 and 280 MPa.

The relationship between stress and time to the pit–crack transition for Ebera follows the trend observed by Akid[31] who correlated data for a range of steels tested in chloride environments under different loading conditions (see below) and found the following trend:

$$t_{p/c} = (A/\sigma)^C$$

where $t_{p/c}$ is the time to pit–crack transition in seconds, σ is the applied stress in MPa, and A and C are constants dependent upon material, environment, loading mode, and test frequency.

Figure 9 presents the analysis of nine different studies where the corrosion fatigue mechanism involved a pit–crack transition. The numbers in square brackets given in the figure refer to references cited in Akid.[16] These data represent steels tested under torsion, uniaxial, rotating bending and 3 pt bend loading at different frequencies within chloride environments.

Physically short crack/long crack interface

It is well established for both air and CF cracking that short cracks grow at rates many times that of 'ΔK equivalent' long cracks.[35] Subject to the limitations of calculating ΔK values for small defect sizes, **Figure 10** shows that conventional LEFM analyses are not appropriate for predicting the transition from pitting to cracking. At longer crack lengths, there is a good correlation between smooth specimen tests and LEFM test data. The data in **Figure 10** suggest that there is a greater correlation between physically short cracks and long cracks close to ΔK_{th} and that the pit-to-crack transition may be independent of the stress intensity factor. This degree of correlation between short and long crack growth rates shows the value of conducting short crack corrosion fatigue tests, notably that information on both microstructural short crack behavior and 'long crack' thresholds can be obtained using short crack test methods.

As discussed previously, in particular for passive metals/alloys, the early stages of corrosion fatigue often involve local breakdown of the passive film

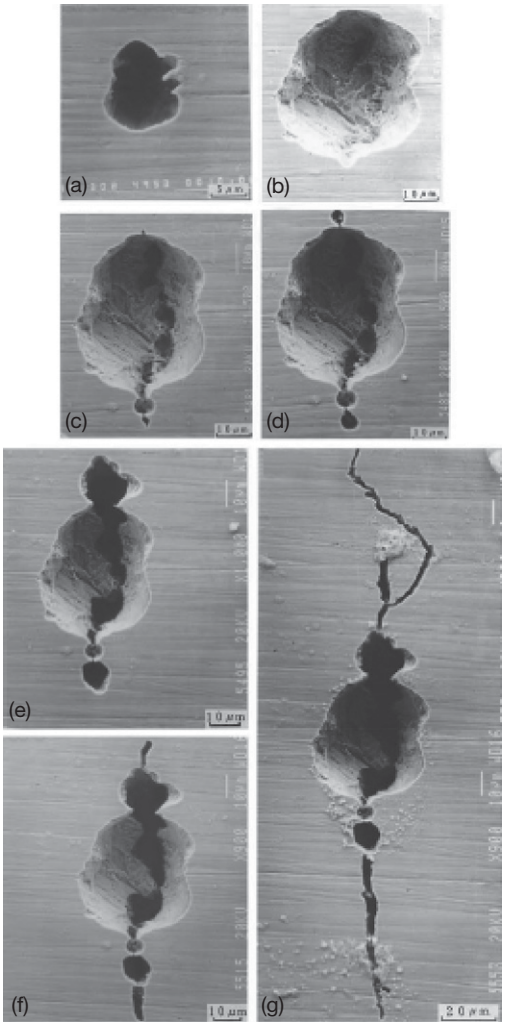


Figure 8 CF crack development of a pit: 12% Cr stainless steel; 3.5% NaCl; stress level 228 MPa (after 34); cycles elapsed, (a) 4.950×10^6 (b) 5.450×10^6 (c) 5.480×10^6 (d) 5.485×10^6 (e) 5.495×10^6 (f) 5.515×10^6 (g) 5.500×10^6.

followed by active corrosion, that is, pitting. This time dependent corrosion phase eventually gives way to crack formation and a transfer into a 'predominantly' mechanical, fatigue-cycles dominated regime. The characterization of these features, summarized in **Table 1**, is based upon numerous studies during which damage is monitored via a plastic replication technique.[28,33,36–40]

An alternative model, appropriate for slip-band cracking is given in eqn [3].

On the basis of these studies and equivalent tests conducted in an air environment, the following simple models have been proposed[37,42] which account for the individual stages identified in **Table 1**. These models are either based upon a modified in-air short crack growth model[37] or on the degree of plasticity at the crack tip,[42] see eqns [1], [2], and [3], respectively.

$$\frac{da}{dN} = A\gamma^{\alpha}(d-a) \text{ MSC} \qquad [1]$$

$$\frac{da}{dN} = B\gamma^{\beta}a - D \text{ PSC} \qquad [2]$$

where MSC and PSC represent the microstructural short crack and physical short crack regimes, respectively. A, B, α, and β are constants depending upon material and environment; γ is the shear strain, and d is a microstructural dimension, for example, grain size, a is the defect size, that is, pit size or crack length, and D is the LEFM threshold crack growth rate.

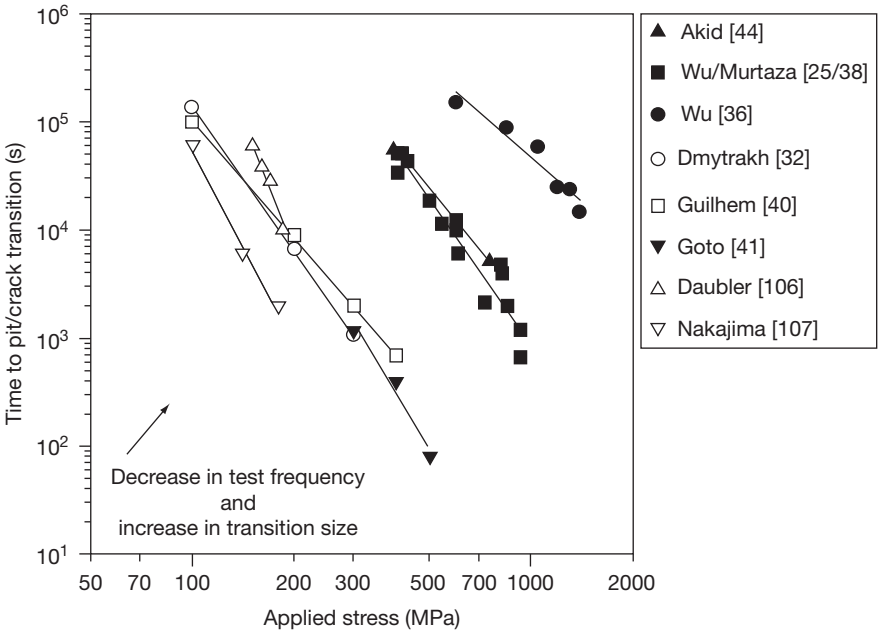


Figure 9 Relationship between time to pit–crack transition and applied stress for various steels. Reproduced from Akid, R. In *Effects of Environment on the initiation of Crack Growth*; Van der Sluys, W. A. Piascik, R. S., Zawiercha, R., Eds.; ASTM STP 1298, **1997**, pp 1–17.

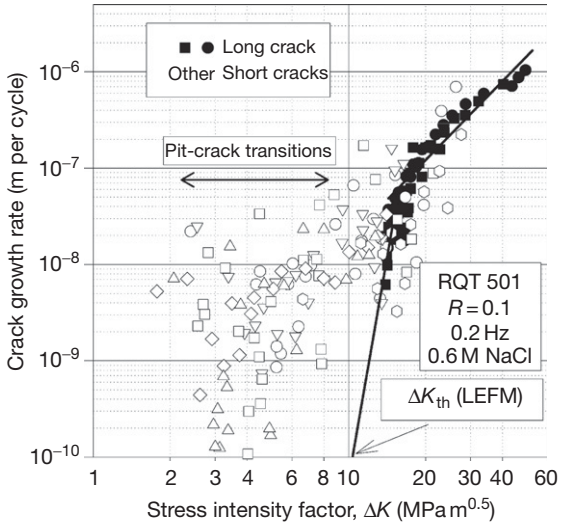


Figure 10 Comparison of short and long corrosion fatigue crack growth rates for a C-steel ($\sigma_y = 500$ MPa). Conditions: artificial seawater at OCP and uniaxial loading at a stress ratio, $R = 0.1$. Note open symbols represent data from smooth specimen short crack tests and solid symbols represent data from pre-cracked specimens. ΔK values for short cracks are based upon an average crack shape aspect ratio and geometry factor of 0.65–0.7 depending upon crack size. Reproduced from Akid, R. In *Effects of Environment on the initiation of Crack Growth*; Van der Sluys, W. A. Piascik, R. S., Zawiercha, R., Eds.; ASTM STP 1298, **1997**, pp 1–17.

Table 1 Designated damage regimes relating to the early stages of corrosion fatigue

Damage regime	*Process*	*Influential factors*
1	Pitting	Solution composition, pH, applied stress
2	Pit/crack transition	Pit shape/local pit environment, applied stress state
3	Environment-assisted short fatigue crack growth	Local crack tip chemistry, applied stress state
4	Physically short crack growth	Applied stress state (ΔK)

$$\frac{da}{dN} = C(r_p)^m \qquad [3]$$

Here, C and m are constants depending upon material and environment, and r_p is the size of the plastic zone.[31]

Where the applied stress state differs from a uniaxial loading condition, for example, fully reversed torsion or tension, a modification is made to eqns [1], [2], and [3] in which γ is replaced by γ_{eq}, an equivalent strain based upon the multiaxial loading condition in which both the principal strain and the strain normal to the shear plane are taken into account.[43,44]

While this approach offers a number of advantages in terms of the application of LEFM analyses, it does not recognize that pit growth is influenced by the microstructural state and that the transition from a pit to a crack is influenced by plasticity. This latter effect has significant implications for the application of LEFM to such small defects. This can be seen in Figure 10 which shows pit–crack transitions occurring well below the LEFM threshold.

Modeling corrosion fatigue behavior, like modeling air fatigue, has fallen into two main categories based upon the existence or otherwise on a defect. Typically, the models address either the development and growth of defects (initiation) or the propagation of defects which fall under the boundary conditions of LEFM.

2.12.3.1.2 Pitting corrosion fatigue models

As discussed in the previous section, initiation is commonly referred to as the 'nucleation stage,' that is, the beginning of defect development. Unfortunately, this can cause uncertainty with the accuracy in prediction of lifetime, as nucleation is somewhat dependent upon the first observation of a defect (crack or pit). Initiation might also be defined as 'life to a certain detectable defect size.' This subtle distinction requires that designers carefully consider the 'inspectability' or 'detectability' of a component or structure. A number of researchers have also suggested that stress intensity factor can be used to define specific stages of the CF process.[45, 46] Given that pitting corrosion is often a precursor to fatigue crack development, this section will provide details on pitting corrosion fatigue models developed to predict CF lifetime. These models are based upon the damage mechanism that consists of the formation of a pit and eventual development of a crack from the pit site. A more detailed review of this subject is given by Hagn.[47]

The models of Hoeppner,[45] Lindley,[48] Kawai,[49] Kondo,[50] and Chen[51] rely on a transition from pitting to cracking defined by a stress intensity threshold value (ΔK_{th}) or that the stress intensity associated with a pit can be defined based upon its geometry, often assuming the pit is hemispherical. Table 2 summarizes the approaches adopted by these authors.

One major problem that has not been addressed when applying a fracture mechanics approach to corrosion fatigue lifetime prediction of nondefective surfaces is the important point that LEFM analyses tend to assume that cracks do not propagate below the ΔK_{th} value. Yet it is known that cracks do initiate at pit sites below the threshold value (calculated using the dimensions of the pit), as illustrated in Figure 10 above. In considering the results shown in Figure 10, the role of crack closure should be considered. Corrosion may affect crack closure in two ways, notably: corrosion can lead to oxide film formation on the crack walls, leading to an increase in closure, or corrosion can remove material by dissolution, leading to a reduction in crack closure. Crack closure is associated more with long cracks than with short cracks because of the limited crack opening displacement of small defects.

2.12.3.2 Long Crack Growth

Differentiation of 'short' and 'long' crack growth is normally made based upon the degree of yielding or plasticity occurring at the crack tip. Knott[52] has shown that LEFM analysis is applicable when the plastic zone size (r_p) is less than one-fiftieth (1/50) of the crack length.

Table 2 Summary of pitting corrosion fatigue models

Author	*Model*	*Comments*
Hoeppner[45]	$K = 1.1\sigma\sqrt{\pi\frac{a}{Q}}$ and $t = \left(\frac{d}{c}\right)^3$	σ = applied stress, a is pit length, Q is a function based on crack shape, t is time to attain pit depth for corresponding threshold value, d is pit depth, and c is a constant dependent upon material and environment
Lindley[48]	$K_{th} = \frac{\Delta\sigma\sqrt{(\pi a)[1.13 - 0.07(a/c)^{0.5}]}}{[1 + 1.47(a/c)^{1.64}]^{0.5}}$	Pits considered as semielliptical cracks. a is a minor and c is a major axis of semielliptical crack
Kawai[49]	$\Delta\sigma = \frac{\Delta K_{all}}{F\sqrt{\pi h_{max}}}$	ΔK_{all} determined from da/dN vs. ΔK plot. F is a geometry factor and h is maximum pit depth
Kondo[50]	$\Delta K_p = 2.24\sigma_a\sqrt{\pi c\alpha/Q}$ and $c = C_p(N/f)^{1/3}$	Assumes pit as a crack, σ_a is stress amplitude, a is aspect ratio (assumed constant) and Q is a shape factor. C is pit radius, N, number of fatigue cycles, and f is cyclic frequency
Chen[51]	$\Delta K_{tr} = (1.12k_t\Delta\sigma\sqrt{\pi c_{tr}})/\Phi = \Delta K_{th}$	K_t is a stress concentration factor, c is half-pit diameter and Φ is a shape factor determined by the pit diameter and depth

For uniaxial loading, the size of the plastic zone is given by eqn [4].

$$r_p = \frac{1}{2\pi}\left[\frac{\Delta K_I}{2\sigma_y}\right] \quad [4]$$

where K_I is the stress intensity factor (MPa $\sqrt{m}$) and σ_y is the yield strength (MPa).

A bounding condition for LEFM applicability is therefore the relationship between r_p and the crack length, as previously described.

Unlike the *S–N* tests, which use smooth 'defect-free' specimens, LEFM test methods use a precracked specimen, which is subject to fatigue cycling within the desired environment. Measurements of crack length are taken at selected intervals at a specified stress ratio (minimum/maximum load). The results of this type of test are presented in a fatigue crack growth rate curve as shown in **Figure 11**. Here, LEFM long crack growth is associated with the stage II regime of the crack growth rate diagram and is often called the 'Paris' region.[53] This region is defined by the empirical relationship given in eqn [5];

$$\frac{da}{dN} = C(\Delta K)^m \quad [5]$$

where $\Delta K = K_{max} - K_{min}$, and A and m are material-, temperature-, microstructure- and stress-ratio-dependent constants. Values of m that have been reported in the literature lie typically in the ranges of values $2 \leq m \leq 5$. This equation is widely used in its integrated form to evaluate the lifetime of cracked structures knowing the stress field, the threshold and critical values of K and the constants A and m.

This empirical relation given in eqn [5] contains the self-similarity of the fatigue crack growth process. However, it is only valid within a so-called intermediate, medium-amplitude part of the fatigue kinetic diagram (stage II), as shown in **Figure 11**.

A threshold effect is observed, ΔK_0 (or ΔK_{th}), which occurs when the crack growth rate equals a value of approximately the order of one atomic spacing per cycle. Hence the plot is generally divided into three regions. When the maximum stress intensity factor becomes higher, the crack growth curve deviates from linearity (the 'Paris' region), and crack instability and rapid acceleration can occur terminating in failure (stage III).

In stage I, the crack growth rate is very low, and experimental data have shown that the crack growth in this region is greatly affected by microstructure, environment, and stress ratio.[54–57]

The self-similitude concept postulates that cracks of different lengths will have the same stress and strain fields, and hence the same crack growth rate if they are subject to the same applied ΔK. This means that cracks in specimens and structures can be directly compared. However, small-scale yielding (SYY) conditions are assumed to prevail, where SYY is dependent upon crack opening displacement and a nonhardening material. This leads to the condition that the limiting stress amplitude is less than one-third of the yield stress. More generally, for this concept to hold, low applied stress levels are required.

Substantial work on the relationship between ΔK and the crack growth rate followed on from the work of Paris and Erdogan, although it was not until the

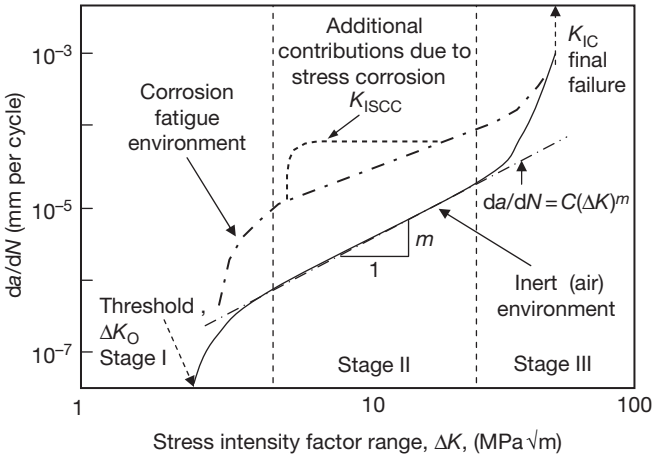


Figure 11 LEFM fatigue crack growth curves showing behaviour for air, corrosion fatigue and corrosion fatigue plus stress–corrosion interactions.

early 1970s that an international conference was held[58] to review the subject of corrosion fatigue.

As can be seen from the schematic in Figure 11, the environment generally causes an increase in the crack propagation rate for a given ΔK value. Furthermore, crack growth rate can be further augmented if the material is also susceptible to stress corrosion cracking within the test environment.

2.12.3.2.1 LEFM-based corrosion fatigue models

There are numerous models cited in the literature that attempt to quantify the contribution from both electrochemistry and the mechanical loading condition. These models may be characterized under the following categories: process superposition, process competition, and process interaction.

Wei and Landes[59] first proposed a *process superposition model*, which summed the stress corrosion component of crack growth with that of air fatigue crack growth. The model can be expressed as eqn [6].

$$(da/dN)_{cf} = (da/dN)_r + \int da/dt\ K(t)dt \qquad [6]$$

where $(da/dN)_{cf}$ is the rate of fatigue crack growth in an aggressive environment, $(da/dN)_r$ is the rate of fatigue crack growth in an inert environment, and the integral term is the environmental contribution obtained from sustained load crack growth in the same environment. This model takes no account of any synergistic interactions between corrosion and fatigue. Later refinements of this approach were predicated on the recognition that environmentally assisted crack growth is the result of sequential processes and is controlled by the slowest process in the sequence.[60]

Austen and Walker[61] proposed the *process competition model* on the basis that stress corrosion and fatigue (or true corrosion fatigue) are mutually competitive and not additive as postulated by Wei and Landes. It is assumed that the crack will propagate by the fastest available mechanism at the stress intensity operative for a given crack length. The model is given in eqn [7].

$$(da/dN)_{cf} = (da/dN)_r + (da/dt) \times 1/f \qquad [7]$$

where the first two terms are as in eqn [6] and (da/dt) $1/f$ accounts for the stress corrosion contribution.

The *process interaction model* was developed to allow interactions to occur, particularly in those systems that exhibit a degree of susceptibility to SCC.[62] This model has been used to predict corrosion fatigue data for steel and aluminum in air and saline environments. The model is given in eqn [8].

$$\left(\frac{da}{dN}\right)_e = m_1(\Delta K_{eff})^{m_2} + \int_0^{1/f} A\nu K_{eff}\alpha\ dt \qquad [8]$$

Here, ΔK_{eff} is the effective stress intensity factor, adjusted for blunting and crack branching, ν accounts for the influence of load cycle on stress corrosion rate, f is the frequency, and m_1, m_2, A, and α are experimentally determined constants.

The ability to predict fatigue lifetime relies upon accurate information on defect size (cracks or pits), shape, and orientation with respect to the applied stress and local crack tip chemistry. Panasyuk *et al.*[63] proposed the use of 'invariant diagrams' in order to account simultaneously for the stress–strain state and the electrochemical conditions at the crack tip. They showed that for a range of alloys, including steel, Al and Ti, the pH at the crack tip differed from that of the bulk solution (NaCl, pH 6.5) by up to 5 pH points, ranging from 0.8 to 3.5. Similar studies have also been conducted by other researchers.[64] In contrast to measurements made at the open circuit potential, Turnbull assessed the crack tip pH and potential changes when steel was subjected to cathodic overpolarization of −1100 mV (SCE). He found that the pH increased up to pH 13 and the potential moved 120 mV in the positive direction (i.e., towards the free corrosion potential). Furthermore, knowledge of the local stress–strain conditions at the crack tip is critical. Novel techniques are now being developed to obtain the geometrical shape of cracks and pits, for example, 3-D synchrotron X-ray imaging[65] and light-interference microscopy.[66] It should be noted that some of the early studies which used crack tip solution chemistry as a means to elucidate the crack tip mechanisms suffered from the fact that solutions oxidized before analysis could take place.

2.12.4 Variables Affecting Corrosion Fatigue

The significance of environmental effects on corrosion fatigue life has been reported by various workers. Evans and Simnad[8,9] in their two stage tests showed how removing the environment after a certain percentage of the expected life did not substantially increase the fatigue life. Rollins *et al.*[67] also concluded that the action of the corrosive environment is of

primary importance in the crack initiation stage. Duquette and Uhlig[20] suggested that a delay in crack initiation or early propagation, attributed to a reduction in dissolution by adopting a higher corrosion-resistant material, would lead to an improvement in the fatigue life. However, many of the earlier studies were devoted to assessing lifetime rather than quantifying crack growth. In this respect, relatively little attention has been paid to the anomalous crack growth rates of short cracks in corrosion fatigue even though they have been observed and known to exist for some time.[37,68] Such studies are expected to increase as environment-assisted growth of small fatigue cracks becomes increasingly recognized as an important failure mode, often dominating total life.[69] As discussed, most corrosion fatigue life prediction methods are dominated by the application and analysis of long crack data.[70] However, where defects develop from smooth 'engineering' surfaces, the application of linear elastic fracture mechanics (LEFM) becomes limited. It is therefore necessary to develop models that account for the development and growth of pits and short cracks. Furthermore, before such models can be applied, it is further necessary to consider the factors which affect the development of such small defects during the early stages.

Consideration of some of these factors is given below.

2.12.4.1 Microstructure

It is widely recognized that the 'in-air' fatigue performance of a material can be significantly influenced by varying the grain size. The principal effect of grain size on short, stage I – shear, crack growth is the introduction, ahead of the crack, of physical barriers whereby the grain boundary acts as a demarcation between two grains having different slip orientations. Therefore, by introducing a large number of barriers into the system, that is, reducing the grain size, cracks are required to orientate themselves along many different planes. The effect of this is one of crack speed deceleration, and, where the stress level is below that of the fatigue limit, crack arrest.

With respect to corrosion fatigue, there is currently no direct evidence to suggest a relationship between grain size and corrosion performance, and hence corrosion fatigue resistance. However, since the environment can influence the ability of a crack to surmount the microstructural barrier, it would not be unreasonable to suggest that corrosion moderates the influence of grain size on fatigue strength.

Furthermore, given that most engineering materials contain metallurgical heterogeneities such as intermetallic particles (IMPs) and second phase precipitates, which can give rise to sites for preferential dissolution and pitting, it might sensibly be expected that a change in corrosion fatigue response will occur with changes in alloy chemistry. IMPs are particularly influential on corrosion fatigue lifetime as they lead to multisite damage and crack initiation. Furthermore, IMP can bridge individual grains, thereby reducing the influence of microstructure on crack growth. Murtaza and Akid[38] conducted intermittent air fatigue–corrosion fatigue loading tests and showed that the dominant microstructural feature in the corrosion fatigue behavior of a high-strength spring steel was the prior austenite grain size. Individual microstructural phases can also influence the initiation behavior, for example, duplex stainless steel consists of almost equal fractions of ferrite and austenite. However, the pitting behavior of these two phases is different with pitting being preferred in the ferrite phase.[26]

2.12.4.2 Loading Frequency

The influence of loading frequency on 'long-crack' corrosion fatigue resistance is well documented, being attributed to time-dependent effects of corrosion which are ascribed to mass transport and electrochemical reaction rate limitations.[71–73]

The role of frequency on the development and growth of short fatigue cracks is, however, less well documented. Studies on smooth specimen fatigue tests[42] have shown that frequency influences short fatigue crack growth rates (see **Figure 12**). The most striking feature of this figure is that crack growth rates are enhanced at a point where the mechanical driving force is at a minimum, that is, when a crack is small and below a critical size–aspect ratio for the transfer from the stage I – shear to stage II – tensile plane. This is not an unreasonable observation as the damage process moves from a 'predominantly' time-dependent corrosion process to a mechanical, cycles-dependent fatigue mechanism. **Figure 12** shows that the differences in lifetime for tests at 10 and 0.01 Hz, which account for a reduction in life of 70%, occur during the growth of a defect below a size of 200 μm. Here, it can be seen that crack growth rates, based on measurements of surface defect size with increasing number of fatigue cycles, for 0.01 Hz are around an order of magnitude higher than at 10 Hz for the same loading conditions.

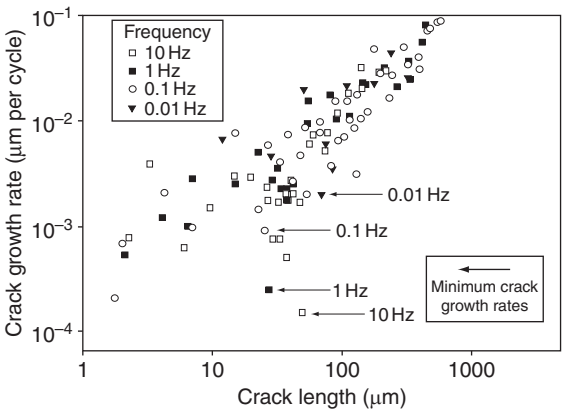


Figure 12 Influence of loading frequency on short corrosion fatigue crack growth of a low-alloy structural carbon steel in artificial seawater. Grain size approximately 10 μm. 3 Pt bend tests, stress ratio = 0.1. Reproduced from Akid, R., Wu, X. J. Structural Materials in Marine Environments. The Royal Society, London, UK. The Institute of Materials; 1994; pp 330–345, ISBN 1 86125 0029.

This accelerated growth at low frequency results in fatigue lifetimes of 5000 and 20 000 cycles for 0.01 and 10 Hz, respectively.

2.12.4.3 Solution/Electrochemical Conditions

The corrosion resistance of a metal or alloy is normally derived from its ability or inability to form a stable passive film. The majority of passive films are based upon the formation of an oxide film and therefore the nature of the environment, that is, oxidizing or reducing, has a marked affect on film formation.

Corrosion fatigue studies[7] on polished carbon steel BS4360 50D showed that neutral pH chloride solutions promoted pitting, while low pH, acidic chloride solutions promoted slip-band dissolution. These differences were attributed, respectively, to local weaknesses in the surface oxide film around inclusions, hence pitting, and total dissolution of the oxide film at low pH levels, allowing access to slip bands within the grain and accelerated corrosion at these sites. The effect of these differences, in terms of fatigue behavior, is such that multiple site damage (cracking) occurred in an acidic solution, followed by coalescence of cracks and a subsequent reduction in fatigue lifetime when compared to near-neutral solution conditions.

Where solution conditions were modified to increase the thickness of the passive film,[24] the time to pit formation was increased. A decrease in the propensity for local corrosion (pitting) can also be affected by modifying the solution conditions, that is, by using inhibitors. Such effects have been observed under corrosion fatigue conditions[25]; here, corrosion fatigue lifetimes increase on the addition of an oxidizing inhibitor, sodium nitrite ($NaNO_2$). The improved CF performance is attributed to a delay in pitting and therefore an increase in time prior to crack formation.

Where electrochemical conditions are controlled to eliminate the effects of anodic dissolution, that is, cathodic polarization (CP), it is observed that the corrosion fatigue resistance is restored towards that observed for identical tests in air. It should be noted that CP may not be effective where a component or structure contains cracks that can propagate under LEFM conditions. For 'smooth' surfaces, a nonfailure condition was observed for fatigue tests carried out below the in-air fatigue limit in NaCl with CP, compared to very early failure in the absence of CP (see **Figure 13**).[73] The effect of polarization on crack growth rate requires an understanding of the crack tip electrochemistry as this controls the dissolution, repassivation, and hydrogen evolution kinetics. The trends are also subject to the nature of the material tested, and suppression of the metal dissolution reaction through the application of a negative potential is not effective for all metals. For example, applying a negative potential to Al and its alloys can lead to corrosion because the cathodic reaction promotes the formation of a high-pH solution, and as aluminum is amphoteric it can undergo dissolution at both high- and low-pH values.

2.12.4.4 Surface Condition

As previously discussed, fatigue is the consequence of an extremely localized surface phenomenon. It is therefore not unreasonable that changes in the surface condition of a material have a marked effect on fatigue performance. Furthermore, it is well documented that fatigue lifetime decreases on increasing the surface roughness.[74] This reduction in lifetime is attributed to the presence of 'micronotches' associated with the peaks and troughs of the surface profile.

Changing the nature of the surface condition, notably stress state and microstructural condition, through the application of shot peening is a method adopted to improve fatigue performance. Here, the surface is deformed by impingement of glass or metal

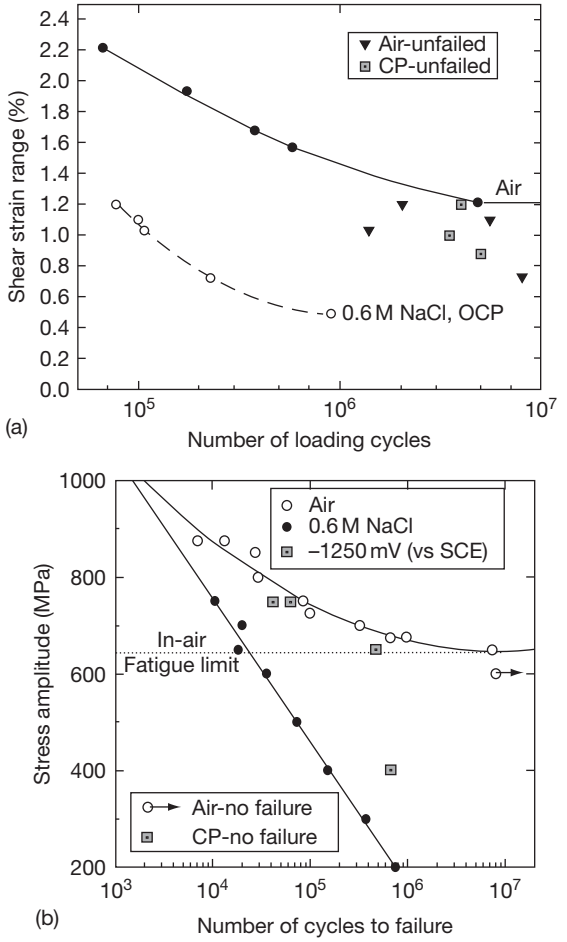


Figure 13 Influence of environment and electrochemical potential on fatigue strength; (a) fully reversed torsion loading (b) fully reversed axial loading;. High-strength carbon steel ($\sigma_y = 1100$ MPa). Reproduced from Akid, R. In *Hydrocarbon Transport and Cracking of Metals*; Turnbull, A., Ed.; The Institute of Materials, 1995; pp 50–61. Note that shear loading, as occurs in a torsion test, tends to be insensitive to hydrogen, while axial loading is more sensitive.

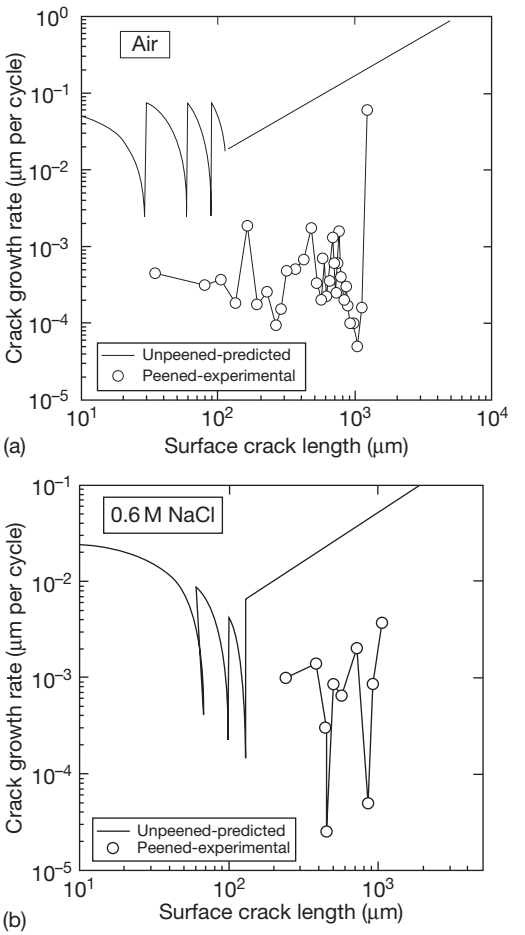


Figure 14 Influence of shot peening on short fatigue crack growth. (a) Air; $\Delta\tau = 1050$ MPa, N_f (peened) 1.4×10^6 cycles, N_f (unpeened) 114 000 cycles. (b) 0.6 M NaCl; $\Delta\tau = 820$ MPa, N_f (peened) 270 000 cycles, N_f (unpeened) 107 000 cycles. Note solid lines represent predicted growth rates based upon a crack growth model. $\Delta\tau_{fl} = 920$ MPa.

beads which impart a compressive residual stress in the surface layers of the metal. The interpretation of increased lifetime has, previously, been attributed solely to residual stress effects. However, it has been shown[7] that the resulting distortion of the microstructure also has a significant effect on the short fatigue crack growth. Figure 14 provides an example of the reduction in fatigue crack growth rate (FCGR) that occurs on the application of shot peening. Here, the FCGR under peened conditions is compared with predicted FCGR data derived from air tests conducted on unpeened samples at the same stress level.

Figure 14 further shows that FCG rates are decreased by an average of two orders of magnitude for both air and corrosion fatigue tests, respectively, and that despite testing in 0.6 M NaCl the fatigue crack growth rates of peened specimens are significantly lower than the predicted in-air growth rates of specimens peened to the same Almen intensity. Comparison of the fatigue lives, as given in the caption of Figure 14, shows that shot peening gives an improvement in fatigue lifetime of 13 and 2.5× for air and corrosion fatigue, respectively.

2.12.4.5 Stress State

Many studies have been devoted to assessing the fatigue performance of materials through the application of simple stress–life (S–N) type testing. Such an experimental approach often adopts a simple (uniaxial) loading condition and an 'inert' (air) environment. However, engineering structures and components are often subjected to nonuniform, asymmetrical loading within environments of varying aggressivity.

The implication of adopting a simple uniaxial constant-load test condition is that a crack may propagate, in shear, and fail to transfer to a fatal mode I, tensile crack. Where a multiaxial loading condition applies, both the maximum shear strain and strain normal to the maximum shear plane govern whether or not a crack will arrest (nonpropagating crack) or grow to failure.[75]

Recent multiaxial 'in-air' fatigue tests[75] have shown that, for mixed tension–torsion testing, the application of a static tensile stress, superimposed on a cyclic torsional load, significantly reduces the fatigue lifetime. Conversely, a compressive stress increases the lifetime, such effects being attributed to the relationship of the orientation of crack and stress state, and its relationship to the microstructure. Corrosion fatigue tests conducted under identical conditions[75] showed that the environment eliminates the effects of the stress state and the early stages of damage (see **Figure 15(a)**), are controlled by the environment, via a pitting mechanism, with fatigue lifetime becoming independent of biaxial stress state (see **Figure 15(b)**). As the crack size increases, environmental effects are more dominant for negative biaxial stress states; while at positive biaxial stress states, the mechanics dominate crack growth.

Additional corrosion fatigue studies concerned with block loading sequences, that is, damage accumulation (DA) loading[76] have also shown that corrosion fatigue lifetime is primarily governed by the ability to develop a small defect, that is, a pit, which transforms to a stage I, tensile crack. In this respect, there are two distinct regimes of damage; notably, a time-dependent corrosion regime, governed principally by solution chemistry, followed by a cycle dependent fatigue regime, governed by the magnitude of stress and mode of loading. Where appropriate conditions apply, that is, crack size and stress state,[73] the ability to generate hydrogen capable of being adsorbed into the metal matrix ahead of the crack tip will also play a part in controlling damage and hence lifetime.

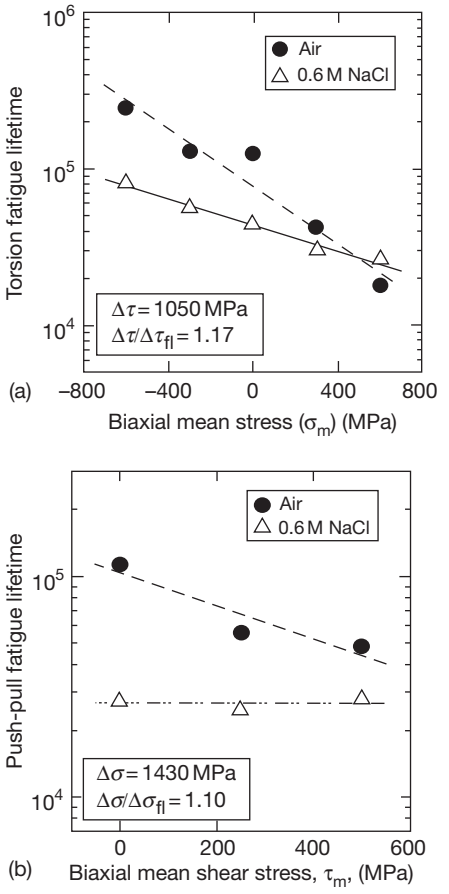


Figure 15 Influence of biaxial stress state on corrosion fatigue lifetime, (a) fully reversed torsion with tensile and compression loading, (b) fully reversed axial loading with static torsion loading. High-strength spring steel, loading frequency of 5 Hz. Reproduced from Akid, R.; Zhang, W. In *Engineering Against Fatigue*; Beynon, J. H., Brown, M. W., Lindley, T. C., Smith, R. A., Tompkins, B., Eds.; A. A. Balkema: Rotterdam, Brrokfield, 1999.

2.12.5 Corrosion Fatigue of Specific Alloys

2.12.5.1 Ferrous Alloys

Carbon steel is the highest tonnage alloy used within the engineering sector. It is, however, an alloy with limited corrosion resistance and therefore its application where both stress and corrosion may operate in combination needs careful consideration. In an attempt to extend the service life of steel

structures, carbon steels are often subjected to cathodic protection (CP), for example, in oil and gas exploration applications. In some cases, these materials are also painted to reduce the current demand from the CP system. CP reduces the operating electrode potential of the steel to typically more negative than −850 mV (vs. Ag/AgCl), that is, below its natural open circuit potential. Where CP is used, careful material selection is required in order to prevent any possible hydrogen embrittlement (HE) problems. Typically, steels having yield strengths of less than about 700 MPa are considered to be immune from HE, although this is affected by the loading conditions and hydrogen concentration.

Given that chloride is very aggressive towards ferrous based materials and that many structures are located either offshore or in coastal regions, it is not surprising that many studies have been devoted to the corrosion fatigue behavior of steels in seawater or artificial seawater environments.

The 1970 and early 1980s saw significant research activity devoted to the corrosion fatigue resistance of offshore structural steels and pressure vessel steels[55,57,61,77] and more latterly to steels having refined microstructures such as roller quench temper (RTP) grades.[33,42]

Barsom[78,79] demonstrated the effects of frequency and cyclic waveform, identifying crack growth regimes above and below the threshold stress intensity factor for stress corrosion cracking. He observed that low loading rates (negative sawtooth) were more detrimental than high loading rates (positive sawtooth) and attributed this to the time available for corrosion and crack extension during the loading part of the cycle. Scott[80] in a separate study illustrated the effects of stress ratio on crack growth rate and cathodic protection on the evolution of a SCC plateau region. The plateau region exhibits an independence of crack growth on the stress intensity value, suggesting that a rate-limiting process is operative. In this case, the rate determining step was attributed to hydrogen evolution and transport at the crack tip.

The specific mechanism(s) leading to corrosion fatigue remain(s) in debate. Gangloff[81] summarized the possible processes that lead to enhanced fatigue crack growth under CF loading suggesting the following mechanisms may be involved: hydrogen embrittlement, film rupture, dissolution, and repassivation; enhanced localized plasticity and interactions of dislocations with surface dissolution, films, or adsorbed atoms.

The dominant mechanism is controlled by both the nature of the materials and the electrochemical conditions at the crack tip. Metals with high yield strength are susceptible to HE unless the local crack tip pH values are alkaline which then lead to charge transfer controlled dissolution process governing crack growth. Wei[82] showed, for an X-70 pipeline steel in a carbonate–bicarbonate solution at pH = 9, that changing cyclic frequency and temperature led to changes in fatigue crack growth rate, indicating the important role of dissolution kinetics.

Environment also has a role to play in the cracking of a high-strength low-alloy (HSLA) steel. Recent work has been conducted by Ritter[83] who studied the CF crack growth behavior of different RPV steels over a wide range of environmental (temperature, electrochemical potential, dissolved oxygen, SO_4^{2-}, Cl^-), loading (ΔK, R) and material (S, MnS, microstructure) parameters in oxygenated or hydrogenated high temperature water, applying cyclic load tests with precracked fracture mechanics specimens. The application of hydrogenated water was always found to result in a reduction of low-frequency CF crack growth rates (CGRs) by at least one order of magnitude with respect to oxygenated conditions. A few hours after changing from oxidizing to reducing conditions, the CGRs dropped below the corresponding ASME crack growth curves.

Ford also showed the importance of passive film rupture and transient anodic dissolution for pressure vessel steels for high temperature pure-water environments.[84] Faraday's law was employed to determine the metal loss during active dissolution, the total metal loss being an integral of the current–time transient response of the metal. The contribution from anodic dissolution was shown to be strain rate and cyclic frequency dependent.

The microstructure of the material has also been shown to play an important role in the resistance to CF cracking. Using a superposition model for a high-purity FeCrNi alloy, Gao[85] assessed the role of microstructure in hydrogen-assisted corrosion fatigue crack growth. He concluded that the environmentally assisted crack growth rate in α-martensite was only twice as high as that in vacuum, while the rate through the austenite grain and twin boundaries is more than one order of magnitude faster. He suggested that the results indicated that strain-induced α-martensite formation is not a principal contributor to the environmental enhancement of corrosion fatigue crack growth and is not necessary for hydrogen embrittlement.

Mechanical factors also influence the nature of the cracking mechanism. Wu[86] studied the influence of strain rate on the low-cycle fatigue resistance of ASTM A 533B low-alloy pressure vessel steel in simulated BWR water. At strain rates above 0.001% s^{-1}, the fatigue life decreased almost linearly with decrease in the strain rate on a log–log plot with environmental effects tended to become saturated at strain rates below 0.001% s^{-1}. It was suggested that the change in strain rate leads to a change in the environment-assisted cracking (EAC) mechanism from hydrogen-induced cracking to a film-rupture–slip-dissolution-controlled process, mainly influencing the rupture rate of oxide film at the crack tip, and mass transfer and local electrochemistry processes in the system. It was also proposed that MnS inclusions play an important role in fatigue crack initiation and propagation where dissolution of the inclusions in the high temperature water environment results in H_2S formation which may enhance the hydrogen-induced cracking process.

While there is general agreement on the role of hydrogen in EAC for high- and low-strength steels in aqueous low temperature environments, it remains debated as to the exact mechanisms operative at high temperatures.

2.12.5.2 Stainless Steels

The susceptibility of stainless steels to CF is dependent upon the steel grade, notably ferritic, martensitic, austenitic, or duplex, the strength level, and environmental conditions.

Stainless steels are renowned for their superior corrosion resistance when compared to low-alloy carbon steels. However, where the prevailing environmental conditions lead to pitting corrosion the fatigue strength of the stainless steel grades is often not appreciably greater than the nonalloyed steel counterparts, assuming equivalent mechanical strength levels.

An example of the reductions in fatigue strength of a range of stainless steel grades is given in Table 3. The data represent rotating bending test results (stress ratio equal to −1) in air and for samples moistened in 3% NaCl solution.

It can be seen from Table 3 that there is a significant reduction in endurance limit of the ferritic and martensitic grades, being some 30–50% of the in-air limit. Increasing the corrosion resistance to an austenitic grade improves the CF resistance, while a combination of high mechanical strength and good corrosion resistance can be obtained by moving to a duplex grade.

Table 3 Rotating bending fatigue test results for different stainless steel grades

Steel grade	*Yield strength (0.2%, MPa)*	*Air fatigue limit (MPa)*	*CF endurance limit (MPa)*
410 (martensitic)	440	340	130
12% Cr Mo (ferritic)	55	415	140
Precipitation-hardened martensitic	1060	550	230
304 (austenitic)	326	260	230
316 (austenitic)	223	270	260
Duplex	475	450	355

Reproduced from Shrier, 3rd ed.

The nature of the CF mechanism operative in stainless steels remains in debate. While it is recognized that active dissolution takes place when the passive film is broken, it is not entirely clear whether hydrogen, generated as a result of metal ion hydrolysis, further contributes to the crack propagation rate.

Using long focal length video optical microscopy to measure hydrogen bubble diameters at the crack mouth, Olive[87] conducted *in situ* monitoring of natural cracks. A boundary condition of the hydrogen evolution inside a crack was obtained and, by considering the electrochemical reactions occurring inside a crack, a minimum volume of metal dissolved per cycle at the crack tip was proposed. No comment was made as to any synergistic effect of hydrogen adsorption and any consequential changes in embrittlement ahead of the crack tip.

Magnin[88] studied the corrosion fatigue damage mechanisms leading to crack initiation in high-purity ferritic stainless steels (Fe–26Cr–1Mo wt%) in a 3.5% NaCl solution at 300 K and an imposed anodic corrosion potential. Particular attention was paid to the influence of strain rate on both the plastic deformation mechanisms and dissolution characteristics, which govern the crack initiation process. In summary, it was found that:

- At high strain rates (about 10^{-2} s^{-1}), pencil glide induces strain localization at grain boundaries; dissolution is localized at grain boundaries; crack initiation is intergranular, and lifetimes are reduced in comparison with those in air.

- At intermediate strain rates (about $10^{-3}\,s^{-1}$), dissolution is entirely localized at grain boundaries because of pencil glide; crack initiation is intergranular, and fatigue lifetimes are shorter than at high strain rates.
- At low strain rates (less than $10^{-4}\,s^{-1}$), the amount of dissolution is lower than the critical value for corrosion fatigue damage; crack initiation is transgranular, and fatigue lifetimes are similar in air and in the corrosive solution.

Nyström[89] presented a study showing the effects of heat treatment and resulting changes in microstructural condition on the low-cycle CF resistance of 25% ferrite and 75% austenite duplex stainless steel. He and his coworker pointed out a major drawback of duplex stainless steel, being that of the susceptibility of the ferritic phase to 475 °C embrittlement; potentially limiting the use of duplex stainless steel to temperatures below 250 °C. The mechanical properties of duplex stainless steel were compared to those of an austenitic stainless steel. The results indicate that the fatigue properties of the duplex grade are little affected by annealing at 475 °C for 100 h. In addition, after annealing, the impact toughness remained high at 130 J. The austenitic grade has the same nominal fatigue crack growth threshold as the duplex grade; however, the effective crack growth threshold, related to the (intrinsic) internal resistance of the material to crack extension and the (extrinsic) level of crack closure, of the duplex grade is higher than that of the austenitic grade.

2.12.5.3 Aluminum Alloys

Aluminum alloys may be classed as either heat-treatable or nonheat-treatable, providing a range of alloys with different properties. The main industrial alloys include the Al–Cu; 2xxx series, Al–Zn–Mg 5xxx series, and Al–Mg–Zn 7xxx series. Heat treatment of these alloys produces precipitation hardening which can be controlled to alter the tensile and yield strength. This in turn affects the sensitivity of the alloys to environment sensitive cracking and localized corrosion. Like the high-strength steels, the high-strength 2xxx and 7xxx series have reduced resistance to stress corrosion cracking, hydrogen embrittlement, and corrosion fatigue. Furthermore, it has been shown that small fatigue cracks grow faster than, and below the thresholds of, long cracks in both active (air at 60% relative humidity) and inert (vacuum and dry nitrogen) environments.[90]

Of major interest to the aerospace industry is the role of intermetallic precipitates in the development of pitting and subsequent crack initiation. Inclusions may be either anodic or cathodic with respect to the surrounding matrix. In either case, a microgalvanic cell is established which leads to local corrosion. Pit-like features act as stress concentration sites where crack initiation can take place. Metal ion hydrolysis leads to the formation of a low-pH solution within the pit. In turn, this can lead to hydrogen embrittlement or advancement of the crack via metal dissolution. The heat-treatable alloys tend to be more susceptible to EAC because the heat-treatment ageing process has a marked influence on the resulting microstructure of the material. Kumai *et al.*[91] showed that ageing of the Al–Li and Al–Li–Cu alloys can lead to precipitate free zones along the grain boundaries, affecting the corrosion resistance of the alloys. In another study, it was shown that where the addition of Li promotes the formation of a passive film, the corrosion resistance of the alloy was increased.[92]

Ro *et al.*[93] showed that the LEFM-regime fatigue crack growth kinetics for an Al–Cu–Mg/Li alloy were dependent not only on ΔK and stress ratio, R but also on the factor P_{H2O}/f, namely, the ratio of water vapor pressure to cyclic frequency. It was observed that the water vapor exposure dependence of da/dN exhibits four regimes of behavior and that there is no exposure effect below a threshold P_{H2O}/f of 0.01 Pa s. At moderate water vapor exposures (0.02–2 Pa s), impeded molecular flow governs da/dN, whereas hydrogen diffusion rate limitation and surface reaction saturation intervene to reduce the da/dN dependence on P_{H2O}/f at higher exposures. A plateau growth rate response was observed above 100–1000 Pa s.

The importance of mechanical loading effects on AA7075-T651, notably periodic overloads, was shown by Chlistovsky *et al.*[94] for smooth specimens. They concluded that the effect of periodic overloads was limited at high-cyclic amplitude loads (ΔS values) between 100 and 450 MPa, but below this value fatigue life was significantly reduced (see **Figure 16**).

It was concluded that the reduction was due to premature crack initiation from surface pitting.

Further clarification of the role of overloads may be gained by considering their role in the development of short fatigue crack growth. Akid's study on damage accumulation under corrosion fatigue conditions[76] showed that cracks grow faster when subject to a low–high stress sequence in comparison

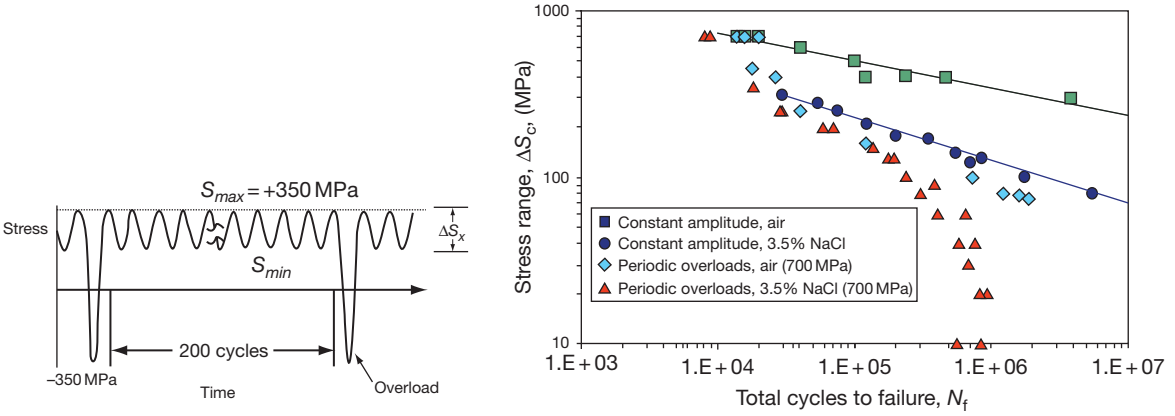


Figure 16 Periodic overload sequence and fatigue life data for 7075-T651 aluminum alloy in air, and simulated seawater. Reproduced from Chlistovsky, R. M.; Heffernan, P. J.; DuQuesnay, D. L. *Int. J. Fatigue* **2007**, 29, 1941–1949.

to that of a high–low stress sequence. This observation was rationalized on the basis that during low stress cycling (below the in-air fatigue limit), the corrosion component is more damaging than the fatigue component. Where tests are conducted at an initial high stress (above the in-air fatigue limit), the contribution from corrosion is small compared to that from fatigue. In the case of Chlistovsky's study, the maximum effect of overloads was seen for corrosion fatigue below the in-air fatigue limit. Here, corrosion can progress the defect up to the microstructural barrier where, in principle, a single overlaod can propagate the defect beyond the barrier. When the stress level is restored to below the in-air fatigue limit, the crack progresses assisted by corrosion up to the next major microstructural barrier.

The presence of preexisting corrosion pits on the fatigue behavior of notched 7075-T7351 was studied by Pao *et al.*[95] Corrosion damage was produced by immersion in 3.5 wt% NaCl solution for 336 h. They observed that pitting shortens the fatigue crack initiation life in air by a factor of two to three and decreases the fatigue crack initiation threshold, $(\Delta K/\sqrt{\rho})_{th}$, by about 50%, where ρ is the notch root radius. Fatigue cracks in polished blunt-notched specimens often initiated from large constituent particles.

2.12.5.4 Titanium Alloys

Titanium has found important uses in a range of sectors, including aerospace, offshore and biomedical because of its good fatigue strength, low specific weight, and good corrosion resistance. Its high corrosion resistance is attributed to the instantaneous formation of an inert Ti oxide surface layer. However, low wear resistance and the poor tribological properties of titanium and its alloys are of concern. This is of particular importance in the case of metal release within the body and potential poor performance where there is a combination of fretting, fatigue, and a corrosive environment that is not conducive to the repassivation of the alloy.

Titanium alloys can be classified into three main groups depending upon the microstructural condition of the alloy, namely, α, α/β, or metastable β. The microstructural condition and state of alloying provide Ti alloys with a wide range of mechanical properties, notably yield strengths up to 1400 MPa. This in turn gives a range of fatigue strength for these alloys.

Ti alloys have been shown to be susceptible to both corrosion fatigue and stress corrosion cracking.[96] As with other alloys that show this combination, there can be a marked increase in crack growth rates when the loading–defect size combination leads to stress intensity values above K_{ISCC}. These were obtained using precracked specimens, and the increase in crack growth rate has been attributed to hydrogen embrittlement. Shipilov[97] reported that, depending upon the nature of the environment used, cathodic and anodic polarization excursions may or may not accelerate or retard crack growth in VT20 and TS6 alloys. It is of interest to note that the effects of cathodic polarization, notably retardation in crack growth, can depend upon the nature of the material and the size of the crack; for these Ti alloys, crack growth is retarded, while for high strength steel, as shown in **Figure 13**, the development of pits and subsequent short fatigue cracks was retarded.

Goa showed that the fatigue response of Ti alloys in water vapor is identical to that of Al alloys[98] where crack growth rates increased with increasing water vapor pressure up to the saturation pressure. Frequency effects were also studied, and it was observed that crack growth rates increased with decreasing frequency and reached a maximum that depended upon ΔK after which growth rates decreased to those comparable with rates observed in vacuum or other inert environments.

With the recent increase in the use of Ti alloys for prosthetic devices, there has been a significant amount of research on the fatigue and corrosion behavior of alloys, such as Ti4V6Al, in actual and simulated body environments. Leinebach *et al.*[99] investigated the cyclic deformation behavior of the binary titanium alloys Ti–6Al–4V and Ti–6Al–7Nb under axial stress-controlled constant amplitude and increasing load tests as well as in loading sequences based on gait patterns from human walking. They concluded that slight differences in the materials' microstructures result in differences of the fatigue and cyclic deformation behavior. In comparison to the results of constant loading rate tests, the standard walking gait pattern tests revealed shorter lifetimes for identical maximum stresses. It was also noted that the fatigue strength was also decreased by corundum grit blasting, that is, increase in surface roughness.

Papakyriacou *et al.* similarly conducted long-term corrosion fatigue tests comparing potential skeletal implant materials, namely cold-worked c.p. niobium and commercially pure (c.p.) tantalum with titanium and Ti–6Al–7Nb alloy.[100] Constant amplitude fatigue experiments (*S–N* curves) were performed at an ultrasonic frequency of 20 kHz with two different surface structures (ground surface and blasted and shot-peened surface) in ambient air and in a corrosive fluid similar to the body fluid in the oral cavity. They recorded a decrease in fatigue limit of 5–20% at 20 million cycles and a detrimental effect of surface finish on the Ti alloys compared to the Nb and Ta candidate materials. However, while this study shows that the Ti alloys can suffer from corrosion fatigue, it should be noted that the extreme high-cyclic test frequency used is not representative of typical implant loading conditions. Given that CF is more pronounced at low test frequencies, the results obtained in this study are likely to be conservative.

Titanium alloys are also gaining increased use in aerospace applications and are used in aircraft gas turbine engine blades and discs as well as in highly stressed components such as forged wing structures and landing gear components. In this type of application, there is also the possibility that an additional damage mechanism may be operative, notably that of fretting. Fretting is the surface damage that occurs when contacting surfaces between mating bodies experience an oscillatory motion of small amplitude. The combination of this motion between two surfaces that are subject to loading is termed 'fretting fatigue.' Fretting fatigue increases the tensile and shear stresses at the contact surface producing surface defects which can act as stress concentration sites. Lietch *et al.*[101] investigated the fretting fatigue behavior of Ti–6Al–4V in ambient laboratory air and under a controlled environment consisting of synthetic seawater. Tests were performed over a wide range of the maximum axial stresses, ranging from 380 to 760 MPa, to examine both low- and high-cycle fatigue regimes. They concluded that seawater had a deleterious effect on fretting fatigue life in the low-cycle fatigue regime but improved the life in the high-cycle fatigue regime. *S–N* data showed marginal differences between the fatigue limit in air and in seawater, although there were limited data points and a high degree of scatter. The results of Leitch (Ti alloy) differ from those obtained by Shaluf,[102] working on a high-strength steel, in which Shaluf showed that fretting under corrosion fatigue conditions reduced the fatigue life at all stress levels.

2.12.6 Prevention of Corrosion Fatigue

Prevention of corrosion fatigue relies upon simple strategies such as lowering the fatigue stress level, adopting more corrosion-resistant alloys or eliminating the corrosion activity, in this case, by either forming a barrier on the surface of the substrate or by applying a negative electrode potential to suppress the anodic dissolution reaction. The former of these strategies, notably reducing the stress, is by far the easiest approach; however, this will reduce the operating loads that can be applied to a given component or structure. Consequently, this approach is rarely used.

2.12.6.1 Prevention by Barrier Coatings and Surface Treatments

Historically, barrier coatings have been used to protect substrates from reacting with the environment. Metallic and nonmetallic coatings, for example, hot-

dip galvanizing and paints offer differing benefits such as cost and ease of applicability, wear resistance, increased corrosion resistance, etc. However, there have been mixed benefits of applying coatings with respect to the enhancement of fatigue lifetime. The principle philosophy behind this approach is that the coating system will prevent contact of the corrosive media, and hence the material will essentially act as if it was under stress within a benign environment. This does, however, rely on the barrier remaining intact and not suffering failure from fatigue itself. Where coating breakdown does occur, there are several possible outcomes. With nonmetallic coatings, localized corrosion may take place at the coating failure site, thereby leading to a local stress concentration. In the case of a metallic coating, it is possible that the coating acts in a sacrificial manner and protects the underlying substrate, for example, Zn and Al coatings on steel. However, it should be pointed out that when the electrochemical potential of a coating is significantly lower than that of the substrate these types of coatings can liberate hydrogen during the dissolution process. This in turn can be adsorbed into the metal and, subject to the nature of the substrate, lead to hydrogen embrittlement. Alternatively, where a metallic coating is noble to the underlying substrate, failure of the coating can lead to accelerated localized corrosion due to galvanic action between the two dissimilar metals and subsequent crack initiation. A further consideration is that of crack initiation in the coating. This is particularly relevant for hard coatings that lack ductility, for example, chromium and nickel. The effects on fatigue life of crack initiation can be compounded if the coating is noble to the substrate, for example, Ni on steel. The initiation of a crack can lead to solution ingress to the substrate and the formation of a local galvanic cell whereby the underlying steel substrate is attacked locally leading to pitting and the formation of a stress concentrating feature.

As previously discussed, both shot peening and surface coatings can influence the fatigue performance of a material. Asquith *et al.*[103] assessed the combination of shot peening and a plasma electrolytic oxidation (PEO) treatment on the corrosion performance of a 2024 aluminum alloy. Shot peening was found to decrease the corrosion resistance, as assessed by electrochemical impedance spectroscopy, while shot peening followed by PEO improved the corrosion resistance. Marin de Camargo *et al.*[104] assessed the effects of coating residual stresses on the fatigue performance of a 7070 Al alloy. They concluded that a 17% Co thermally sprayed coating improved the axial fatigue resistance, while different anodizing treatments decreased fatigue resistance. Shot peening of anodized layers improved the fatigue life over that of nonshot-peened surfaces.

Alternative surface treatments that influence the surface metallurgy, and hence the localized corrosion performance have also been investigated. Wu *et al.*[105] evaluated the effects of laser surface melting (in air and in N_2), using an excimer laser, to improve the resistance to fatigue cracking of aluminum alloy 6013 induced by pitting corrosion. Analysis of the electrochemical impedance measurements showed that the laser-formed AlN and/or Al_2O_3 layer can greatly reduce the rate of electrochemical reaction. The corrosion fatigue lives of the air-treated and the N_2-treated specimens were found to increase by two and four times compared to the untreated specimens. Furthermore, the corrosion current for the laser-treated specimens (measured during the fatigue test) was considerably lower than that for the untreated specimens, the improvement in lifetime being considered to be due to the excellent initiation resistance to pitting corrosion of the laser-treated specimens.

2.12.6.2 Prevention by Cathodic Protection

Cathodic protection operates on the principle of lowering the potential of the metal to a value below its normal open-circuit potential. This topic is covered elsewhere in this book and therefore will not be discussed in detail. However, its relevance to corrosion fatigue is that the cathodic partial reaction can lead to the generation of atomic hydrogen which in turn can give rise to hydrogen embrittlement of high-strength alloys. A typical rule of thumb when protecting steel is to prevent the potential becoming more negative than −900 mV (Ag/AgCl) and use alloys that have a yield strength less than 700 MPa.

2.12.7 Summary

This chapter has considered the role of corrosion on the fatigue performance of a range of engineering materials. Specific attention has been given to discussing the early stages of corrosion fatigue for damage emanating at nominally 'defect' free surfaces. This stage has traditionally been referred to as 'initiation.'

Corrosion fatigue damage of alloys in aggressive solutions progresses through several stages including surface film breakdown, localized dissolution/pitting at slip bands or nonmetallic inclusion, transfer from corrosion-dominated growth to fatigue cycles-dependent cracking and subsequent transfer to, environmentally-assisted microstructurally independent cracking in a 'continuum mechanics' growth regime.

A number of factors have been shown to influence individual stages of the corrosion fatigue process; including atmospheric or solution conditions, electrochemical potential, applied stress state, loading frequency, surface stress state and microstructural condition.

The 'free corrosion' fatigue lifetime of 'smooth' specimens is considered to be dominated by the ability to grow a defect, that is, a pit or crack, beyond crack-arresting microstructural barriers inherent within the material. Given the appropriate solution and loading conditions, this early dissolution dominated regime, along with the transition of a pit to a crack, has a major effect on the fatigue lifetime of a material. The major factors, which are known to control smooth specimen lifetime under in-air fatigue loading, notably microstructure, loading mode and biaxial stress state, play a less significant role under corrosion fatigue conditions. Test frequency, however, is a dominant factor in controlling smooth specimen corrosion fatigue lifetime. Overall it is considered that the environment, in particular, anodic dissolution plays a major part in controlling fatigue lifetime for stress levels below the in-air fatigue limit.

In the case of corrosion fatigue of structures containing preexisting defects, the most important factors include solution chemistry, test frequency, load ratio, and crack tip stress intensity. Of particular concern for cracked components is hydrogen embrittlement, especially when high-strength alloys are employed in conjunction with cathodic protection. A similar problem can arise if these materials are cleaned in acidic solutions or electroplated.

The nature of the surface is especially important, as corrosion fatigue is predominantly a surface phenomenon. Controlling the surface residual stress state, surface roughness, electrode potential and applying coatings have a marked effect on performance. Coatings, in particular, can be a low-cost strategy to exclude contact between the substrate and the environment and, in effect, simulate an inert environment for the material. Caution is required where the coating is brittle and noble to the substrate as this may accelerate damage and decrease fatigue lifetime.

References

1. Rankine, W. *Proc. Inst. Civil Eng.* **1843**, *2*, 105–108.
2. Frost, N. E.; Marsh, K. J.; Pook, L. P. *Metal Fatigue*; Oxford: Clarendon Press, 1974; p 26.
3. Lisagore, W. B. Crooker, T. W.; Leis, B. N.; Eds. *Environmentally Assisted Cracking: Science and Engineering,* ASTM STP 1049, 1990.
4. Haigh, B. P. *J. Inst. Metals* **1917**, *18*, 55–77.
5. McAdam, D. J. *Proc. ASTM* **1926**, *26*, 224–254.
6. Gough, H. J. *J. Inst. Metals* **1932**, *49*, 17–92.
7. Miller, K. J.; Akid, R. *Proc. R. Soc. Lond. A* **1996**, *452*, 1411–1432.
8. Evans, U. R.; Simnad, M. T. *Proc. Roy. Soc. A* **1947**, *188A*, 372–392.
9. Simnad, M. T.; Evans, U. R. *J.I.S.I.* **1947**, *156*, 531–539.
10. Mughrabi, H. *Scr. Metall.* **1992**, *26*, 1499–1504.
11. Klesnil, M.; Lukas, P. *Fatigue of Metallic Materials*; Elsevier Scientific, 1980.
12. Forsyth, P. J. E. *J. Inst. Metals* **1951/2**, *80*, 181.
13. Magnin, T.; Driver, J.; Lepinoux, J.; Kubin, L. P. *Rev. Phys. Appl.* **1984**, *19*, 483–502.
14. Suresh, S.; Ritchie, R. O. *Int. Metals Rev.* **1984**, *29*, 445–476.
15. Miller, K. J. *Fatigue Fract. Eng. Mater. Struct.* **1993**, *16*(9), 931–940.
16. Akid, R. In *Effects of Environment on the Initiation of Crack Growth*; Van der Sluys, W. A., Piascik, R. S., Zawiercha, R., Eds.; ASTM STP 1298, **1997**, pp 1–17.
17. Ford, F. P. In *Environment-Sensitive Fracture*; Dean, S. W., Pugh, E. N., Ugiansky, G. M., Eds.; ASTM STP 821, **1984**, pp 32–51.
18. Frankel, G. S. *J. Electrochem. Soc.* **1998**, *145*(6), 2186–2198.
19. Jack, A. R.; Paterson, A. N. In *Conference of Institute of Mechanical Engineers*, London, UK **1977**; pp 75–83.
20. Duquette, D. J.; Uhlig, H. H. *Trans. ASM* **1969**, *62*, 839–845.
21. Akid, R.; Miller, K. J. *Fatigue Fract. Eng. Mater. Struct.* **1991**, *14*, 637–649.
22. Magnin, T.; Landon, J. M. *Mater. Sci. Eng.* **1995**, *76*, L7–L10.
23. Leis, B. N.; *et al. Corrosion Fatigue Crack Initiation in an Iron–Caustic System* ASTM STP 801, **1983**, pp 197–228.
24. Isaacs, H. S.; Kissel, G. *J. Elect. Soc.* **1972**, *119*, 1628–1631.
25. Wu, X. D. Ph.D. Thesis, University of Sheffield, 1995.
26. Vignal, V.; Mary, N.; Valot, C.; Oltra, R.; Coudreuse, L. *Electrochem. Solid State Lett.* **2004**, *7*(4), 1–11.
27. Cottis, R. A.; Markfield, A.; Haritopoulos, P. In *Environment Assisted Fatigue*; EGF 7 Scott, P., Ed.; Mechanical Engineering Publications: London, 1990; pp 381–393.
28. Wang, Y.; Akid, R. *Corros. J. Sci. Eng.* **1996**, *52*(2), 92–102.
29. Boukerrou, A.; Cottis, R. A. *Corros. Sci.* **1993**, *35*(1–4), 577–585.
30. Muller, M. *Met. Trans. A* **1982**, *13*, 649–655.
31. Wu, X. J. Ph.D. Thesis, University of Sheffield, 1995.
32. Dmytrakh, I. M.; Akid, R.; Miller, K. J. *Br. Corros. J.* **1997**, *32*(2), 1–7.
33. Hu, H. Ph.D. Thesis, University of Sheffield, 1997.
34. Ebera, R. *Eng. Failure Anal.* **2006**, *13*, 516–525.
35. Gangloff, R. P.; Wei, R. P. In *Proceedings of the Second Engineering Foundation International Conference – Small Fatigue Cracks*, 1986; pp 239–264.
36. Wu, X. D. Ph.D. Thesis, University of Sheffield, 1995.

37. Akid, R.; Miller, K. J. *Fatigue Fract. Eng. Mater. Struct.* **1991**, *14*, 637–649.
38. Murtaza, G.; Akid, R. *Int. J. Fatigue* **1996**, *18*(8), 557–566.
39. Zhang, W.; Akid, R. *Fatigue Fract. Eng. Mater. Struct.* **1997**, *20*(4), 547–557.
40. Guillem, F. M.Phil. Thesis, University of Sheffield, 1998.
41. Goto, M.; Nisitani, H. *Fatigue Fract. Eng. Mater. Struct.* 1992, *15*(4), 335–363.
42. Akid, R., Wu, X. J. Structural Materials in Marine Environments. The Royal Society, London, UK. The Institute of Materials; 1994; pp 330–345, ISBN 1 86125 0029.
43. Zhang, W.; Akid, R. *Fatigue Fract. Eng. Mater. Struct.* **1997**, *20*(4), 547–557.
44. Akid, R.; Wang, Y. Z.; Fernando, U. In *Corrosion-Deformation Interactions*; Magnin, T., Gras, J. M., Eds.; Les Editions de Physique Les Ulis, 1993; pp 659–670.
45. Hoeppner, D. W. In *Fatigue Mechanisms*; STP 675. Fong, J. T., Ed.; ASTM: Philadelphia, PA, 1979; p 841.
46. Wei, R. P.; Gangloff, R. P. In *Fracture Mechanics: Perspectives and Directions*; ASTM STP 1020; Wei, R. P., Gangloff, R. P., Eds.; ASTM: Philadelphia, PA, 1989; pp 233–264.
47. Hagn, L. *Mater. Sci. Eng. A* **1998**, *103*, 193.
48. Lindley, T. C.; Mclntyre, P.; Trant, P. J. *Metal. Technol.* **1982**, *9*, 135.
49. Kawai, S.; Kasai, K. *Fract. Eng. Mater. Struct.* **1985**, *8*, 115–127.
50. Kondo, Y. *Corros. Sci.* **1985**, *45*, 1.
51. Chen, G. S.; *et al. Mater. Sci. Eng. A* **1996**, *219*, 126.
52. Knott, J. F. *Fundamentals of Fracture Mechanics*; Butterworths, 1973; p 134.
53. Paris, P.; Erdogan, F. *Trans. ASME, J. Basic Eng.* **1963**, *85*, 528–534.
54. Taylor, D.; Knott, J. F. *Fatigue Eng. Mater. Struct.* **1981**, *4*, 147–155.
55. Craig, H. L., Jr. Crooker, T. W.; Hoeppner, D. W. Eds. *Corrosion Fatigue Technology,* ASTM STP 642, 1978.
56. Congelton, J.; Craig, I. H.; Olieh, R. A.; Parkins, R. N. In *Corrosion Fatigue: Mechanics, Metallurgy, Electrochemistry and Engineering*, ASTM 801; Crooker, T. W., Leis, B. N., Eds.; 1984, pp 367–389.
57. Austen, I. M.; Walker, E. F. In *Influence of Environment on Fatigue*; IMech E Conference Publications 1977-4. The Institution of Mechanical Engineers, 1977; pp 1–10.
58. Devereux, O. F.; McEvily, A. J.; Staehle, R. W. Eds. *Corrosion Fatigue: Chemistry, Mechanics and Microstructure*; NACE-2, National Association of Corrosion Engineers, 1972.
59. Wei, R. P.; Landes, J. D. *Materials. Res. Stds.* **1969**, *9*, 25.
60. Gangloff, R. P.; Wei, R. P. *Metall. Trans. A* **1977**, *8*, 1043–1053.
61. Austen, I. M.; Walker, E. F. *Fatigue* **1984**, *84*, 1457–1469.
62. Rhodes, D.; Musuva, J. K.; Radon, J. C. *Eng. Fract. Mech.* **1981**, *15*(3–4), 407–419.
63. Panasyuk, V. V.; Ratych, L. V.; Dmytrakh, I. N. *Fatigue Eng. Mater. Struct.* **1984**, *7*(1), 1–11.
64. Turnbull, A. Ed. *Corrosion Chemistry in Pits, Crevices and Cracks*; HMSO: London, 1987.
65. Nakayama, T.; Yuse, F.; Tsubokawa, Y. *Corros. Sci.* **2007**, *49*, 130–138.
66. Frantziskonis, G. N.; *et al. Eur. J. Mech. A: Solids* **2000**, *19*(2), 309–318.
67. Rollins, V.; *et al. Br. Corros. J.* **1970**, *5*, 33–40.
68. Gangloff, R. P. *Res. Mech. Lett.* **1981**, *1*, 299–306.
69. Akid, R.; Miller, K. J. In *Environmental Assisted Fatigue*; EGF 7; Scott, P., Ed.; Mechanical Engineering Publications: London, 1990; pp 415–434.
70. Craig, H. L., Jr. Crooker, T. W.; Hoeppner, D. W. Eds. *Corrosion Fatigue Technology,* ASTM STP 642, 1978.
71. Congelton, J.; Craig, I. H.; Olieh, R. A.; Parkins, R. N. In *Corrosion Fatigue: Mechanics, Metallurgy, Electrochemistry and Engineering*; Crooker, T. W., Leis, B. N., Eds.; ASTM 801, 1984, pp 367–389.
72. Wei, R. P.; Gangloff, R. P. In *Fracture Mechanics: Perspectives and Directions*; ASTM STP 1020; Wei, R. P., Gangloff, R. P., Eds.; American Society for Testing and Materials: Philadelphia, PA, 1989; pp 233–264.
73. Akid, R. In *Hydrogen Transport and Cracking of Metals*; Turnbull, A., Ed.; The Institute of Materials, 1995; pp 50–61.
74. Taylor, D.; Clancey, M. *Fatigue Fract. Eng. Mater. Struct.* **1991**, *14*(2/3), 329–336.
75. Akid, R.; Zhang, W. In *Engineering Against Fatigue*; Beynon, J. H., Brown, M. W., Lindley, T. C., Smith, R. A., Tompkins, B., Eds.; A.A.Balkema: Rotterdam, Brrokfield, 1999.
76. Akid, R. *Fatigue Fract. Eng. Mater. Struct.* **1996**, *19*(2/3), 277–285.
77. Ford, F. P. In *Environment-Sensitive Fracture*, ASTM STP 821, Dean, S. W., Pugh, E. N., Ugiansky, G. M., Eds.; 1984, pp 32–51.
78. Barsom, J. M. *Eng. Fracture Mech.* **1971**, *3*, 15–25.
79. Barsom, J. M. In NACE-2, 1972; pp 424–436.
80. Scott, P. M. In *Advances in Fracture Mechanics*; Chell, G., Ed.; Applied Science Publishers, 1981.
81. Gangloff, R. P. In *Corrosion Tests and Standards Manual: Application and Interpretation*; Baboian, R., Ed.; American Society for Testing and Materials: Philadelphia, PA, 1995.
82. Wei, R. P.; Chiou, S. *Eng. Fracture Mech.* **1992**, *41*(4), 463–473.
83. Ritter, S.; Seifert, H. P. *J. Nucl. Mater.* **2007**, *360*(2), 170–176.
84. Ford, F. P.; Hudak, S. J., Jr. In *Small Fatigue Cracks*. Ritchie & Langford Ritchie, R. O., Lankford, J., Eds.; 1986.
85. Goa, M.; Wei, R. P. *Scr. Metall. Mater.* **1993**, *28*(7), 853–856.
86. Wu, X.; Katada, Y. *Corros. Sci.* **2005**, *47*(6), 1415–1428.
87. Olive, J. M.; Cwiek, J.; Desjardins, D. *Corros. Sci.* **1999**, *41*, 1067–1078.
88. Magnin, T.; Coudreuse, L. *Mater. Sci. Eng.* **1985**, *72*, 125–134.
89. Nyström, M.; Karlsson, B. *Mater. Sci. Eng. A* **1996**, *215*, 26–38.
90. Lankford, J. *Fatigue Fract. Eng. Mater. Struct.* **1983**, *6*, 15–31.
91. Kumai, C.; *et al. Corrosion (NACE)* **1989**, *45*(4), 294–303.
92. Gui, J.; Devine, T. M. *Scr. Metall.* **1987**, *21*, 853–857.
93. Ro, Y. J.; Agnew, S. R.; Bray, G. H.; Gangloff, R. P. *Mater. Sci. Eng. A* **2007**, *468–470*, 88–97.
94. Chlistovsky, R. M.; Heffernan, P. J.; DuQuesnay, D. L. *Int. J. Fatigue* **2007**, *29*, 1941–1949.
95. Pao, P. S.; Gill, S. J.; Feng, C. R. *Scr. Mater.* **2000**, *43*, 391–396.
96. Pelloux, R. M. Proceedings of Fatigue of Materials and Structures Sherbrooke: Canada, July 1978; pp 8–26.
97. Shipilov, S. A. *Fatigue Fract. Eng. Mater. Struct.* **2001**, *25*(9), 243–259.
98. Gao, S. J.; Wei, R. P. *Mater. Sci. Eng.* **1984**, *62*, 65–78.
99. Leinenbach, C.; Schwilling, B.; Eifler, D. *Mater. Sci. Eng C* **2005**, *25*(3), 321–329.
100. Papakyriacou, M.; Mayera, H.; Pypenb, C.; Plenk, H., Jr.; Stanzl-Tschegga, S. *Int. J. Fatigue* **2000**, *22*(10), 873–886.
101. Lietch, L. C.; Lee, H.; Mall, S. *Environ. Mater. Sci. Eng. A* **2005**, *403*(1–2), 281–289.
102. Shaluf, I. M. M.Sc. Thesis, University of Sheffield, 1993.

103. Asquith, D. T.; *et al. Thin Solid Films* **2007**, *516*(2–4), 417–421.
104. Marin de Camargo, J. A.; *et al. Surf. Coat. Technol.* **2007**, *201*, 9448–9455.
105. Xu, W. L.; Yue, T. M.; Man, H. C.; Chan, C. P. *Surf. Coat. Technol.* **2006**, *200*, 5077–5086.
106. Daeubler, M. A.; Thompson, A. W.; Bernstein, I. M. *Metall Trans* **1991**, *22A*, 513–519.
107. Nakajima, N.; Tokaji, K. *Fatigue Fract. Eng. Mater. Struct.* **1995**, *18*, 345–351.

2.14 Under Film/Coating Corrosion

H. N. McMurray and G. Williams
Materials Research Centre, School of Engineering, Swansea University, Singleton Park, Swansea, SA2 8PP, UK

Abbreviations

EB Emeraldine base
EPMA Electron micro-probe analysis
ES Emeraldine salt
FFC Filiform corrosion
H-CS Camphor sulfonic acid
H_2-PP Phenyl phosphonic acid
H-pTS *para*-Toluene sulfonic acid
HDG Hot-dip galvanized steel
ICP Intrinsically conducting polymer
LB Leucoemeraldine base
PVB Polyvinyl butyral
PAni Polyaniline
SHE Standard hydrogen electrode
SIMS Secondary ion mass spectrometry
SKP Scanning kelvin probe
SKPFM Scanning kelvin probe force microscope
XPS X-ray photoelectron spectroscopy

Symbols

E_{corr} Free corrosion potential (V)
E_{intact} Free corrosion potential of an intact polymer coated metal (V)
pK_a Acid dissociation constant
x_{del} Delamination distance (μm)
t_{del} Delamination time since electrolyte contact (min)
t_i Delamination initiation period (min)
ϕ_{sc} Strontium chromate pigment volume fraction
ϕ_{pa} Polyaniline pigment volume fraction

2.14.1 Introduction

Corrosion-driven organic coating delamination, in which the more-or-less intact coating becomes physically separated from the underlying substrate, is a key failure process affecting painted, lacquered and otherwise organically coated metal products, and the subject has been reviewed elsewhere.[1–6] Here, we will attempt a brief overview and update of the topic. In so doing, we will concentrate on two processes of corrosion-driven coating delamination, which have been extensively characterized and have reasonably well-defined mechanisms, namely *cathodic delamination* and *filiform corrosion* (FFC). Both of these phenomena tend to propagate from preexisting breaks or penetrative defects in the organic coating, which act to admit species from an external electrolyte. In the case of cathodic delamination, separation of the organic coating is linked to the cathodic reaction in a localized corrosion cell. In the case of FFC, separation of the organic coating is frequently linked to the anodic reaction. As such, cathodic delamination and FFC are closely related to the phenomena of cathodic blistering and anodic blistering.[7–10] It is not our intention to further describe or discuss blistering as a separate phenomenon. Neither is it our intention to describe or discuss the processes by which ions from an external electrolyte diffuse or migrate through a nominally intact organic coating.[11–13] We will restrict our comments to situations in which electrolyte species have penetrated through a penetrating defect and describe the

corrosion processes that then ensue at the defect and beneath the organic coating.

2.14.2 Cathodic Delamination

One of the most common and therefore economically important mechanisms of coating delamination on iron (mild steel) and zinc (galvanized) surfaces involves so-called cathodic delamination.[1,3,4,14–24] This process can occur wherever the metallic substrate becomes exposed to scratches or cut edges during manufacturing or through in-service damage. Numerous studies have indicated that when iron or zinc substrates are exposed to aqueous electrolyte through penetrative defects in an organic coating, a *cathodic delamination cell* may become established. In this cell, a thin layer of electrolyte, which penetrates beneath a region of already delaminated coating, acts to couple anodic metal dissolution occurring at the coating defect to cathodic oxygen reduction occurring at the site of coating disbondment. An alkaline environment forms at the cathode, and suggested mechanisms of coating disbondment include the dissolution of an amphoteric metal oxide film at the metal–coating interface (in the case of zinc), base-catalyzed polymer degradation, base-catalyzed hydrolysis of interfacial bonds, and attack on the polymer by hydrogen peroxide and other reactive intermediates in cathodic oxygen reduction.

Koehler showed that cathodic delamination (which he called halo detachment) occurred only when group I (alkali metal) cations were available to act as counter ions to cathodically generated OH^-.[14] This is because only group I cations possess the hydrolytic stability and solubility at high pH necessary to support a strongly alkaline electrolyte. Ammonium salts do not produce cathodic delamination because the ammonium cation (NH_4^+, pK_a 9.3) decomposes to ammonia (NH_3) at high pH.[14,21] Group II (alkaline earth) salts do not produce cathodic delamination because of the limited solubility of group II hydroxides.[14,21,25] When group I cations are present, a very high local pH can be produced. In a work involving cellulose nitrate coatings on iron, Ritter and Kruger measured a pH approaching 14 beneath the delaminated coating for a natural corrosion situation with 0.05 M NaCl.[26] However, from their accompanying ellipsometric measurements, it seems most likely that actual delamination took place when the pH was ~11.

Two possible routes exist for species undergoing transport to the cathodic site, (vertically) through the coating or (horizontally) along the polymer–substrate interface. Leidheiser *et al.* showed that water transport occurs through the coating.[1,18] They also showed that oxygen transport occurs predominantly through the coating in nonpigmented coatings, with a minor contribution from interfacial transport. Water and oxygen transport are not usually the rate controlling processes in cathodic delamination.[1,18] However, in pigmented coatings, where the pigment hinders through-coating diffusion, interfacial oxygen transport can become important and rate limiting.[21] Leidheiser identified a linear relationship between cation diffusion coefficient in aqueous solution and cathodic delamination rate for a range of group I chloride salts. This suggested that cation transport occurred principally through migration along the interface. However, because delamination rates also showed a strong dependence on film thickness, through-coating transport could not be ruled out.[27] The notion of interfacial cation migration was further supported by the work of Castle and Watts.[28] They used electron probe micro-analysis (EPMA) to show that Na^+ cations superficially penetrated the delaminated coating from both bulk solution and the interface, but detected no Na^+ within the interior of the coating, thus ruling out the possibility of through-coating transport. They also found that delamination rate decreased with increasing substrate surface roughness and, therefore, increasing tortuosity of the coating–substrate interface. This finding is also consistent with the notion of interfacial cation migration as the rate controlling step.

Probably the most systematic recent work on the cathodic delamination of organic coatings from iron[29–31] and the zinc surface of electrogalvanized steel[32–34] was carried out by Stratmann *et al.* in a series of experiments designed to preclude the possibility of through-coating cation transport. They prepared their samples and used a scanning Kelvin probe (SKP) apparatus in a configuration similar to that shown in **Figure 1**. Under these circumstances, a cathodic delamination cell becomes established, and a thin layer of electrolyte ingresses beneath the delaminated coating in the manner shown schematically in the lower part of **Figure 2**. The SKP is calibrated to record the local free corrosion potential (E_{corr}) as a function of distance (x) from the edge of the organic coating defect. As delamination proceeds, $E_{corr}(x)$ profiles develop as shown schematically in the upper part of **Figure 2**. These may be approximately resolved into four characteristic regions as follows:

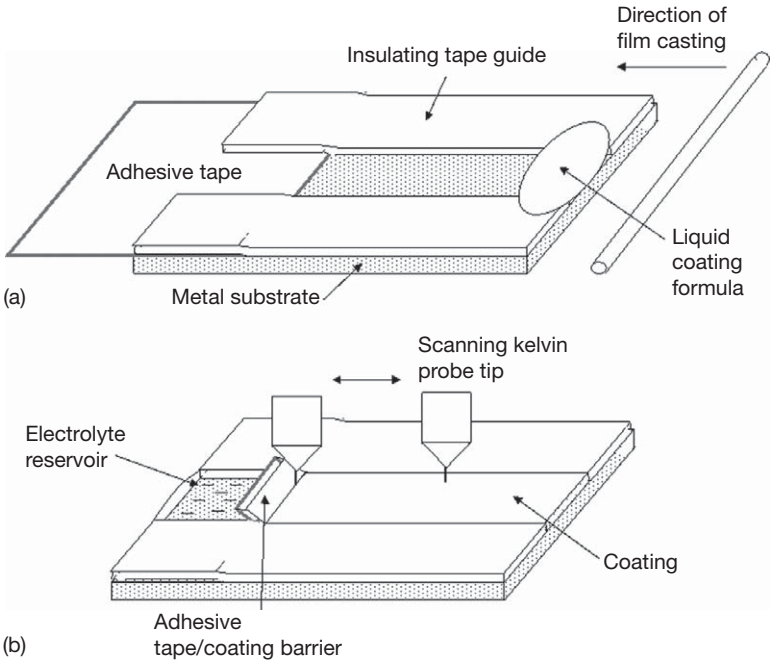


Figure 1 Schematic diagram showing (a) sample preparation and (b) experimental procedure for a typical delamination experiment.

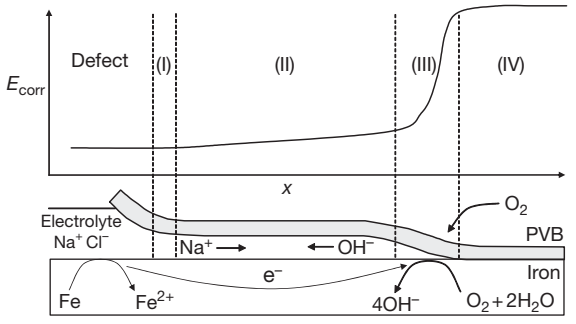


Figure 2 Schematic representation of the corrosion-driven delamination cell showing correspondence with various regions of the time-dependent $E_{corr}(x)$ profile.

1. The region immediately next to the defect where E_{corr} is low and ultimately becomes similar to that of the freely corroding metal.
2. A region of E_{corr} gradient linking region (I) with region (III) below and resulting, at least in part, from ohmic resistance to ionic current flow in the thin under-coating electrolyte layer.
3. A region of sharp transition in E_{corr} corresponding to the cathodic delamination front where coating disbondment takes place.
4. A region of as yet undelaminated polymer where oxygen can diffuse to the metal surface but electrolyte is not present and metal cannot corrode. Here, E_{corr} (or the intact potential E_{intact}) is high.

Stratmann's technique is robust and reproducible. **Figure 3** shows SKP derived $E_{corr}(x)$ profiles obtained, by ourselves, from unpigmented 30 µm polyvinyl butyral (PVB) coatings on the zinc surface of (hot dip) galvanized steel (HDG), undergoing cathodic delamination with 5% aqueous NaCl in contact with the coating defect. The location of the delamination front can be determined from the point of maximum $dE_{corr}(x)/dx$ gradient.[30] The distance (x_{del}) over which delamination has occurred is related to the time since electrolyte contact (t_{del}) by,

$$x_{del} = k_d(t_{del} - t_i)^{1/2} \quad [1]$$

where k_d is a parabolic rate constant and t_i is the initiation period – that is the time delay between electrolyte contact and the onset of cathodic delamination. Initially, the substrate–coating interface is intact, and anions and cation ingress simultaneously by diffusion. Because the coating is fully hydrated at the time of electrolyte contact, t_i was taken to be the time required for interfacial ionic concentration to rise to a level capable of supporting cathodic O_2 reduction.[30] Below a concentration threshold of 0.05 M inside the defect, delamination was not observed, implying that a minimum interfacial ionic

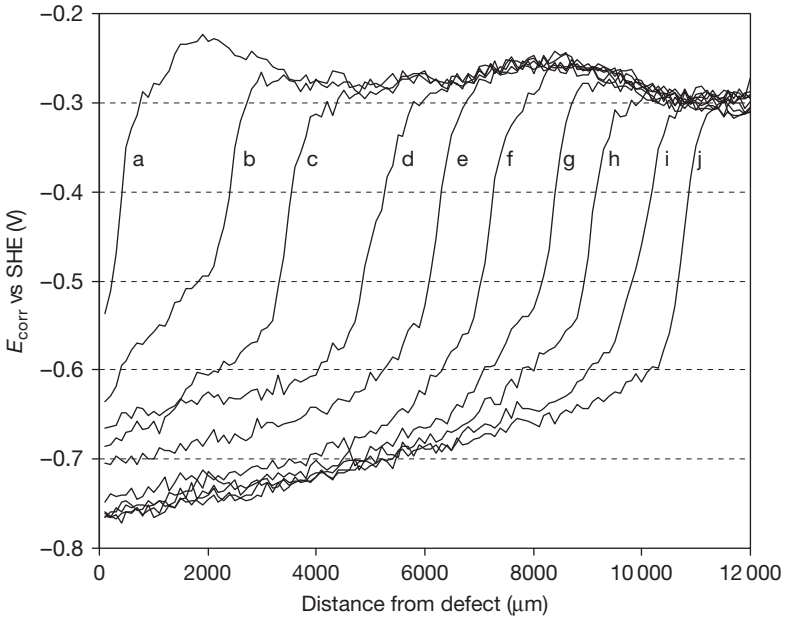


Figure 3 Time-dependent E_{corr} (versus standard hydrogen electrode, SHE) as a function of distance (x) from the defect edge profiles recorded for unpigmented 30 μm polyvinyl butyral (PVB) coatings on HDG zinc substrate. Electrolyte in contact with the defect is 0.86 mol dm^{-3} NaCl: curve (a) = 120 min, (b) = 240 min, (c–e) at 120 min intervals.

concentration is required. Once interfacial O_2 reduction begins, cations migrate towards the underfilm cathode, and anions migrate towards the defect anode, that is, anions from the experimental electrolyte are excluded from the electrolyte layer forming beneath the delaminated coating.[30]

In agreement with the earlier findings of Leidheiser[1,27] the value of k_d in eqn [1] was found to increase with the aqueous diffusion coefficient of the group I cation used in the experimental electrolyte, that is, in the order $Li^+ < Na^+ < K^+ < Cs^+$. The identity of the counter anion (Br^-, F^-, Cl^- ClO_4^-) was found to have no influence on k_d, and the absence of anions in the under-film region was confirmed by Auger analysis.[30] The electrolyte layer forming beneath the delaminated coating was found to be 2–3 μm thick. However, the diffusion coefficients calculated for cations migrating in this layer were calculated to be more than 10 times smaller than in aqueous solution. It was therefore proposed that the electrolyte layer takes the form of a gel comprising highly oxidized polymer in water.[30]

On the basis of the above, it would seem clear that the rate determining step in cathodic delamination is typically the transport of cations from the external electrolyte along the interface between the organic coating and the metal substrate. However, this notion should be regarded with caution because the coated metal systems used to study the intrinsic kinetics of cathodic delamination have largely been 'model' systems comprising unpigmented polymer films coated onto metals which are good electrocatalysts for cathodic O_2 reduction, and with no prior application of any chemical 'pretreatment' or 'conversion coating' to influence adhesion and/or rates of interfacial electron transfer. The rate determining step in any complex process is always the slowest step. In practical coated systems pigments capable of inhibiting cathodic oxygen reduction may be added to the organic coating, an inhibitory pretreatment applied to the metal substrate, or the elemental composition of the metal surface modified to reduce activity for oxygen reduction. Under these circumstances, the rates of cathodic delamination may be greatly diminished without any obvious influence on the facility of interfacial cation transport – as will be discussed further under *preventing cathodic delamination* given below.

2.14.2.1 Disbondment Mechanism

A question arises as to exactly how the presence of cathodic O_2 reduction and an alkaline environment produces disbondment of an organic coating from the underlying substrate. We should say immediately that it is very unlikely that any single, unique and universal disbondment mechanism exists. More

probably, there are a number of possible mechanisms that may occur alone or in combination depending on the nature of the coated metal system under consideration. In any particular system, the relevant mechanism will be that which proceeds at the fastest rate. Modifying the system to inhibit a particular mechanism may reduce the rate of cathodic delamination, but not prevent it completely if another mechanism is available.

The first possible mechanism is the alkaline dissolution of a preexisting (air-formed) oxide layer at the substrate–coating interface. In the case of metals such as zinc with strongly amphoteric oxides, which dissolve at moderately high pH, this would seem a reasonable hypothesis. Stratmann *et al.* have shown that the oxide layer on zinc actually thickens in the delaminated zone, but that this occurs via a dissolution-reprecipitation mechanism.[32] However, iron III oxides are only weakly amphoteric and dissolve only at very high pH (approaching 14). Consequently, the iron surface tends to remain passive in the delaminated zone.[29] Reduction of iron III to a more soluble iron II oxide is possible as potentials decrease.[35] However, Castle and Watts used X-ray photoelectron spectroscopy (XPS) to show that iron oxide within the delaminated region only became reduced in localized patches and that the delaminated region extended well beyond these.[36] They concluded that oxide dissolution is not a significant factor in cathodic delamination on iron.

The second possibility is that alkaline electrolyte physically displaces the coating at the substrate–coating interface.[14] Aqueous displacement was first suggested by Evans in terms of the ability of alkaline solution to creep over the metal surface.[37] The 'wettability' of an electrode surface increases with decreasing (electrolyte–electrode) interfacial energy and latter reaches a maximum at the potential of zero charge.[38,39] On this basis it has been pointed out that the affinity for water of an oxide covered metal surface will tend to increase as the pH becomes increasingly removed from the isoelectric point (pH point of zero charge) of that oxide.[40] If it is assumed that the isoelectric point lies near or below neutral pH, a competitive displacement of the coating by water will be promoted as pH, and therefore water affinity is increased. It has also been argued that competitive displacement may occur when oxide-polymer acid–base interactions make a significant contribution to coating adhesion.[41]

The third possibility is that alkaline electrolyte reacts directly with the coating polymer in such a way as to break covalent chemical bonds existing within the coating or between the substrate and the coating. Carboxylic esters are known to undergo alkaline hydrolysis (saponification), and this process has long been thought of as important in cathodic delamination.[36] Many polyesters, epoxy esters, alkyd resins and other polymers bearing ester functionality are susceptible to hydrolysis; and so are polymers bearing urethane and urea linkages. It has been shown using XPS that coatings based on melamine–formaldehyde and urea–formaldehyde coatings cross linked with epoxy resin oligomers cathodically delaminating from steel in 5% aqueous NaCl showed chemical change consistent with hydrolysis of ester, urethane, and/or urea linkages.[42,43] The same studies also showed that the delaminated steel surface was covered with an organic residue indicative of cohesive failure within the coating. Conversely, for cross linked polybutadiene coatings, which develop only very limited quantities of carboxylic and ester functionality during oxidative curing, evidence of hydrolysis damage is limited to a layer of more highly oxidized polymer immediately adjacent to the substrate surface.[44] Castle and Watts concluded that, whilst there is no universally applicable mechanism of cathodic disbondment, the more resistant the coating polymer is to alkaline hydrolysis the more likely is an interfacial separation rather than a cohesive failure due to coating degradation.[45]

Stratmann *et al.* have pointed out that OH^- is not the only cathodically produced species capable of destroying iron–oxygen–polymer bonds such as Fe–O–CO–R (were R = organic).[31] Many short lived intermediates of O_2 reduction, such as $O_2^{-\bullet}$, $HO_2^{\bullet}$, HO_2^-, H_2O_2, and $HO^{\bullet}$ may be more reactive than OH^- itself. The same authors go on to show that an epoxy coated iron exposed to 1 M aqueous NaOH through a coating defect delaminates six times more slowly than the same system exposed to 1 M aqueous NaCl and held in the presence of air.[31] The intermediates of cathodic O_2 reduction are capable of reacting with a wide range of organic functionalities (particularly in the case of $HO^{\bullet}$). Consequently, in their presence, carboxylic ester functionality and/or other functionalities susceptible to alkaline hydrolysis are not required for polymer degradation or disbondment to occur.

2.14.2.2 Preventing Cathodic Delamination

Under circumstances where disbondment occurs through physical displacement of the coating it might reasonably be assumed that disbondment could be discouraged by promoting coating adhesion. Funke has

described adhesion as the most important and decisive property of a coating.[46–48] Functionalized silanes have been widely used as reactive coupling agents for improving the adhesion of organic coatings to steel. Examples include vinyl silanes for improving wet adhesion of alkyd coatings,[49] amino and epoxy silanes for epoxy/polyamide and polyurethane coatings,[48] and various other silanes for improving adhesion of automotive electrocoats.[50] Marsh *et al.* treated steel substrates with prehydrolyzed allyltriethoxysilane prior to coating with unpigmented alkyd resin.[51] They found that silanization improved both wet and dry adhesion but subsequently found no relationship between adhesion and rates of cathodic delamination. Leidheiser did find that treating an iron surface with silane prior to coating with polybutadiene resulted in significantly reduced rates of cathodic delamination.[1] However, it is not clear that this effect had anything to do with adhesion because impedance measurements on silane treated iron also showed large increases in charge transfer resistance and, therefore, decreased electrochemical activity for cathodic oxygen reduction.[1] Overall, it would seem that improving coating adhesion alone may not improve resistance to cathodic delamination if an alternative disbondment mechanism, such as hydrolysis of the coating polymer, is available.[52]

Inorganic pretreatments or conversion coatings are frequently applied to metal substrates prior to coating with paint or lacquer as a means of both promoting adhesion and chemically inhibiting corrosion.[53] Pretreatments based on phosphoric or chromic acids react electrochemically with the metal and substantially replace the air-formed oxide film with a layer of metal phosphate or mixed chromium oxide.[53] Such a layer may disfavor cathodic delamination if it is more resistant to alkaline dissolution than the original oxide layer, and/or if it inhibits cathodic O_2 reduction. It has long been proposed that resistance to alkaline dissolution determines resistance to cathodic delamination in the case of metal phosphate pretreatment layers.[54,55] A recent study has similarly concluded that interfacial stability is strongly influenced by the chemical stability of conversion layers to hydroxide generated by cathodic O_2 reduction.[56] However, SKP studies of thin phosphate-based conversion coatings on zinc (galvanized steel) have shown that these also inhibit cathodic O_2 reduction and produce a corresponding cathodic (negative) shift in the potential of the polymer coated surface.[57] This effect is attributed to the replacement of n-semiconducting zinc oxide with an electrically insulating zinc phosphate layer that obviously reduces the driving force for cathodic delamination. Similarly, SKP studies of chromate conversion coatings on galvanized steel have shown a cathodic potential shift relative to the unmodified zinc surface.[58] This finding might seem paradoxical at first given the oxidizing nature of chromium VI species. However, cathodic inhibition by chromate has been reported by a number of authors.[59–61]

Corrosion inhibitor pigments based on sparingly soluble salts or ion-exchange solids are often dispersed in paint coatings with the aim of preventing or retarding corrosion-driven coating failure. The SKP technique has been used to study the influence of strontium chromate ($SrCrO_4$) pigment on the kinetics of cathodic delamination affecting polyvinylbutyral coatings adherent to the intact zinc surface of hot dip galvanized steel.[62] When the defective electrolyte was saturated with $SrCrO_4$, CrO_4^{2-} anions were excluded from the under-film electrolyte layer by the delamination cell electric field, and delamination rates were reduced by less than 25%. In contrast, dispersions of $SrCrO_4$ pigment in the PVB coating allowed CrO_4^{2-} diffusion directly into the under-film electrolyte layer. **Figure 4(a)** shows a series of x_{del} versus $(t_{del} - t_i)^{1/2}$ plots for various values of pigment volume fraction (ϕ_{SC}). The slopes of these plots decrease as ϕ_{SC} increases and deviations from linearity consistent with departure from pure cation migration control occur for $\phi_{SC} \geq 0.049$. The dependence of k_d values, obtained from initial x_{del} versus $(t_{del} - t_i)^{1/2}$ slopes, on ϕ_{SC} is given in **Figure 4(b)**, which shows that k_d tends to zero for $\phi_{SC} = 0.064$. On the basis of these findings, it was proposed that replacement of under-film O_2 reduction by a self-limiting reduction of CrO_4^{2-} to insoluble Cr(III) (hydr)oxide is the most significant factor in decreasing delamination rates.[62]

When the same SKP approach was used to study dispersions of silica and bentonite based rare earth (Ce^{3+}), cation-exchange pigments in a PVB coating on zinc (galvanized steel) $Ce^{3+}_{(aq)}$ was found to exchange directly into the under-film electrolyte and significantly retard coating delamination.[63] On a nonpretreated zinc surface, delamination was not halted completely because the under-film exchange of Na^+ by Ce^{3+} was itself incomplete. The resulting precipitation of $Ce(OH)_3$ increased the under-film electrolyte layer resistivity, but, unlike the Cr(III) (hydr)oxide described above, $Ce(OH)_3$ did not form a coherent surface film capable of stifling O_2 reduction at the cathodic delamination front.[27] However, when the same cation-exchange pigments were used in conjunction with a chromate based pretreatment this appeared to influence the surface nucleation of $Ce(OH)_3$ to

produce a coherent film, stifling O_2 reduction and producing a profound inhibition of cathodic delamination.[64] Lyon *et al.* have reported a similar synergistic activity between zinc–phosphate and commercial Ca^{2+} cation-exchanged silica (Shieldex™).[65]

There has been recent interest in the use of intrinsically conducting polymers (ICPs) for corrosion control and the topic has been subject to review.[66,67] Many studies have focused on polyaniline (PAni).[68–70] PAni exists in various states interrelated by the redox and acid–base equilibria[71,72] but it is the emeraldine salt (ES) which is electronically conductive and most widely used for corrosion control. The ES is produced by protonating or 'doping' the partially oxidized PAni emeraldine base (EB) using Bronsted acids.[73–76] PAni ES may be applied to metal surfaces directly by electrodeposition[77,78] or deposition of water-borne dispersions.[79] However, PAni ES particles may also be dispersed in a conventional polymer binder.[80,81] The protection of carbon steel by PAni ES has been variously ascribed to substrate ennoblement, the formation of a passive oxide film and anodic inhibition through precipitation of an insoluble metal salt of the dopant anion.[66,67] The ES/LB (leuco base) redox couple is generally regarded as being responsible for establishing substrate potentials in the passive regime.[67,69,72,82] Furthermore, it has been suggested that PAni ES becoming reduced to LB is cyclically reoxidized by the atmospheric O_2, thus extending the period of ennoblement.[67,69,71,72,83,84] It has also been proposed that the displacement of cathodic oxygen reduction away from the metal–coating interface acts to suppress cathodic delamination.[66,82,85]

This last notion has been investigated using coatings comprising various volume fractions of particulate polyaniline (PAni) dispersed in a polyvinylbutyral binder applied to iron substrates.[86,87] SKP was used to measure substrate potentials in air and follow corrosion-driven cathodic delamination when 5% (w/v) (0.86 M) aqueous NaCl was brought into contact with a coating defect. Pani EB was found to have no effect on either substrate potential or delamination kinetics. **Figure 5** shows delamination kinetics for various

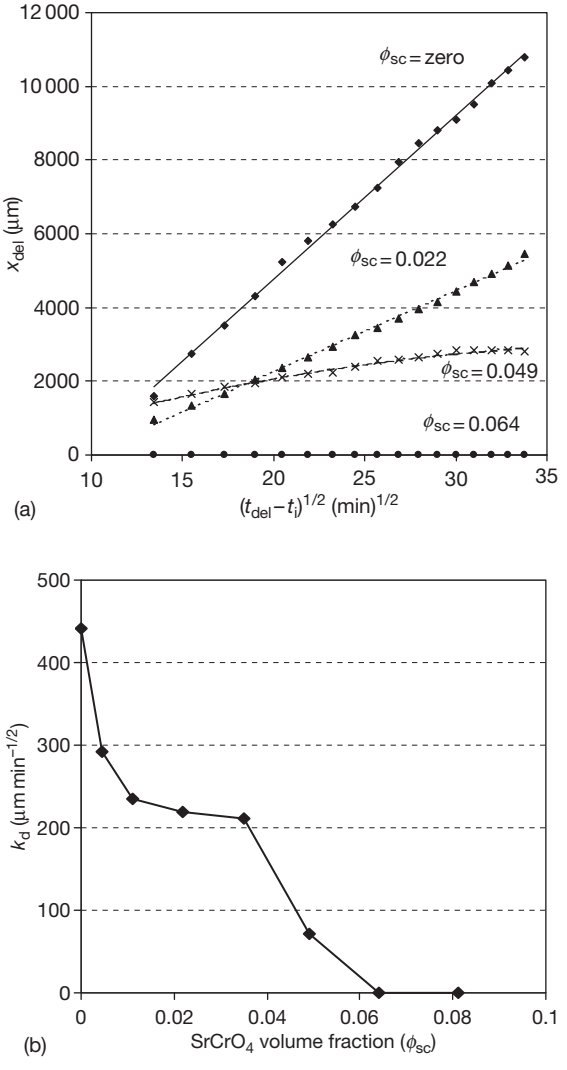


Figure 4 Delamination kinetics for $SrCrO_4$ pigmented 30 μm PVB coatings on HDG zinc substrate. Electrolyte in contact with the defect is uninhibited 0.86 mol dm^{-3} NaCl. (a) Plots of x_{del} versus $(t_{del} - t_i)^{1/2}$ for varying $SrCrO_4$ volume fraction (ϕ_{sc}). (b) Plot of the initial x_{del} versus $(t_{del} - t_i)^{1/2}$ slope (k_d) as a function of (ϕ_{sc}).

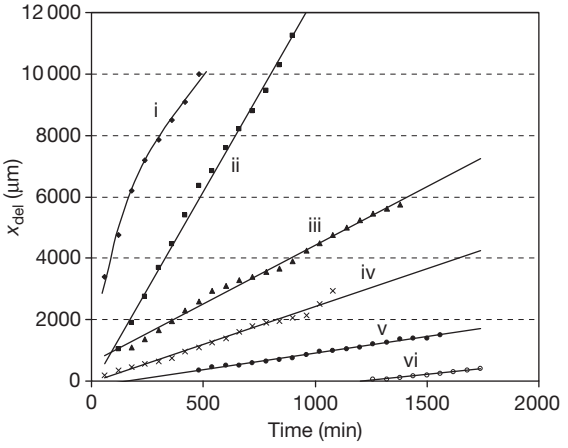


Figure 5 Plots of x_{del} versus time for 30 μm PVB/PAni–pTS coatings on iron with various PAni–pTS pigment volume fractions (ϕ_{pa}). In each case, the electrolyte in contact with the defect is 0.86 mol dm^{-3} NaCl (aq). ϕ_{pa} key: (i) 0, (ii) 0.05 (iii) 0.1 (iv) 0.15 (v) 0.2, and (vi) 0.25.

volume fractions (ϕ_{pa}) of *p*-toluene sulfonic acid doped Pani ES (Pani-pTS). It may be seen that when $\phi_{pa} \geq 0.05$ the x_{del} increases linearly with t_{del} and not parabolically according to eqn [1]. Delamination rate (dx_{del}/dt_{del}) decreases monotonically from 12.8 to 0.83 μm min^{-1} as ϕ_{pa} is increased from 0.05 to 0.25, a finding summarized in curve (ii) of **Figure 6**. Curve (i) of **Figure 6** shows the mean E_{intact} (potential of undelaminated surface) values plotted versus ϕ_{pa}. The critical volume fraction required for long-range electrical conduction in a percolating network is ~0.17[88] and E_{intact} appears to become constant and determined by the PAni ES/EB couple for $\phi_{pa} \geq 0.2$.

PAni ES doped using *p*-toluenesulfonic (H-pTS), camphorsulfonic (H-CS), phosphoric (H_3PO_4) and phenylphosphonic (H_2-PP) acids increased substrate potentials by up to 0.36 V and inhibited cathodic delamination, with the efficiency order H-pTS < H-CS < H_3PO_4 = H_2-PP.[87] Conversely, dopant sodium salts added to the corrosive electrolyte when the coating did not contain PAni were not found to inhibit delamination. On this basis, it is concluded that, on iron, the inhibition of cathodic delamination by polyaniline did arise primarily from cathodic O_2 reduction becoming relocated from the ennobled substrate onto the coating ES. However, $Fe_3(PO_4)_2$ and FePP salt films formed through redox reactions at the ES–substrate interface also contributed by hindering electron transfer at the substrate surface. Without these 'chemical' contributions to inhibition, ES could only slow up cathodic delamination on iron and not prevent it.[86,87] It should also be noted that the reoxidation of LB to ES is only likely to occur at significant rates when the coating is in contact with air, that is under nonimmersion conditions. PAni is therefore unlikely to be of significant benefit when a coated metal is immersed in aqueous electrolyte.

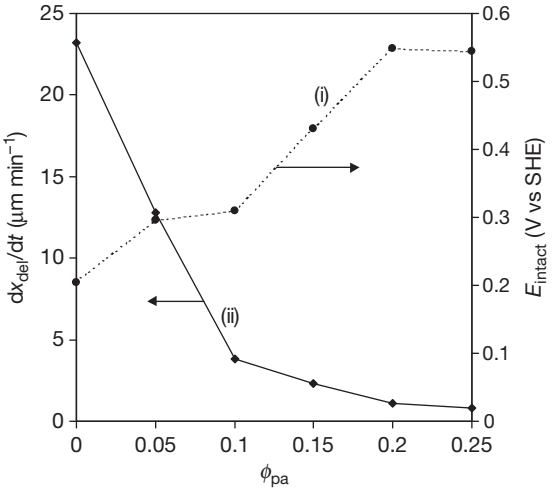


Figure 6 Summary plot (a) E_{intact} and (b) delamination rate (dx_{del}/dt) as a function of ϕ_{pa} for 30 μm PVB/PAni–pTS. Initial dx_{del}/dt_{del} is plotted in the case of the PVB ($\phi_{pa} = 0$) coating.

2.14.2.3 Modifying Metal Composition

Metals differ considerably in their electrocatalytic activity for cathodic oxygen reduction. Pure aluminum, for example, is virtually inactive as an oxygen cathode.[89] This inert behavior probably reflects the electronic properties of the surface aluminum oxide that, as an electrical insulator with a wide band gap, disfavors interfacial electron transfer. Ion implantation of aluminum or titanium into iron has been shown to reduce both surface activity and rates of cathodic delamination.[1,90] Hydrated oxide or hydroxide layers produced through the interaction of group II (alkaline earth) metal cations with hydroxide anions are similarly known to act as effective blockers of cathodic O_2 reduction.[91] On this basis, alloying with aluminum and/or group II elements might reasonably be expected to reduce the susceptibility of a given metal to cathodic delamination. In structural metals, the scope for alloy development is limited by the need to preserve the mechanical properties. In metallic coatings, there is greater freedom, and significant advances have been seen in the field of zinc-based (galvanized) coatings for steel.[92]

Souto and Scantlebury have shown that hot dip galvanized steels in which the zinc layer is alloyed with 5% aluminum (Galfan) and 55% aluminum (Aluzinc) exhibit a much increased resistance to cathodic delamination relative to a conventional (zinc only) system.[93] Paradoxically, when the same Zn and Zn–Al alloy coatings were studied using a rotating disc electrode all three were found to be highly active O_2 cathodes.[21] This suggests that the improved delamination resistance exhibited by the Zn–Al alloy coatings results from a fragmentation of cathodic sites rather than general surface deactivation. That is to say the heterogeneous microstructure of the Zn–Al alloy coatings, consisting as it does of interlocking zinc-rich and aluminum-rich phases, is reflected in local composition of the interfacial oxide layer. Blocking of electron transfer over aluminum-rich oxide regions would then disfavor the formation of a continuous delamination front and prevent the front from advancing from one zinc-rich region to the next.

Stratmann *et al.* used SKP to show that cathodic delamination occurred rapidly on zinc (electrogalvanized steel) but not on the zinc-rich intermetallic phase $MgZn_2$, which frequently forms in Zn–Mg alloy coatings.[94] The potential of the intact $MgZn_2$–oxide–polymer interface was found to be ~0.7 V more negative than the corresponding Zn–oxide–polymer interface. This increase in cathodic polarization was attributed to a large band gap in the $MgZn_2$ interfacial oxide, resembling that of MgO. Dissolution of magnesium oxide under cathodic conditions is unlikely because group II oxides are basic and become more insoluble as pH increases. However, $MgZn_2$ exposed to neutral aqueous NaCl at an organic coating defect was found to become rapidly dealloyed through Mg dissolution. In so doing, the defect potential rose to resemble freely corroding zinc (−0.7 V vs SHE). Potentials in the defect were actually higher than those in the intact coated area and no driving force existed for cathodic delamination.[94]

2.14.3 Filiform Corrosion

An underfilm corrosion phenomenon, producing 'hair-like' corrosion tracks on lacquer-coated steel, was first reported in the scientific literature by Sharman in 1944,[95] who used the phrase FFC to describe these features. FFC is regarded as a type of atmospheric corrosion, principally affecting organic-coated iron (steel) and aluminum surfaces, initiating at penetrative coating defects and producing characteristic thread-like under-film deposits of corrosion product. An example of the typical appearance of corrosion filaments on organic coated aerospace aluminum alloy is given in Figure 7. Although FFC attack is largely considered to be superficial in nature, it is aesthetically unpleasant and hence undesirable in modern organic-coated metal products.

Much of the research carried out in FFC prior to 1995 has been comprehensively reviewed elsewhere.[5,6]

Figure 7 Filiform corrosion observed on aluminum alloy AA2024-T3 coated with a clear film of PVB.

It is not our intention to recover the same ground, but rather to give a general overview of FFC and a subsequent outline of recent developments in the field.

2.14.3.1 General FFC Characteristics

It is generally accepted that FFC initiates when aggressive species, such as chloride ions, permeate through an organic coating, usually at a defect or weakness in the coating and accumulate at the metal–coating interface. FFC only occurs in the presence of atmospheric oxygen and high humidity, although the humidity range over which corrosion is observed is dependent upon the metal studied. Slabaugh *et al.*[96] observed FFC occurrence over a 30–95% relative humidity range for common aerospace aluminum alloys coated with a range of organic coating types, while FFC activity on iron (steel) samples is reported to cease at humidity levels below 60%.[5] The motion of individual corrosion filaments is largely directionally stable, propagating in relatively straight lines for appreciable distances. It has also been reported that filaments growing on aluminum alloy surfaces do so preferentially along the rolling or extrusion axis.[6,97] Such behavior is thought to arise from corrosion filaments following a 'path of least resistance' formed by lines of intermetallic precipitates present at the elongated grain boundaries of the rolled/extruded alloy microstructure. Corrosion filaments do not generally cross each other, nor do they propagate through a break in the surface coating.[5]

Corrosion filaments consist of a mobile, electrolyte-filled head, containing metal cations and aggressive anions (such as Cl^-), and a tail of dry, porous corrosion product. FFC head regions typically exhibit a low pH towards their leading edge (as low as pH 1) as a consequence of cation hydrolysis. The primary driving force for filament advancement is thought to be differential aeration, arising from facile mass transport of gaseous O_2 through the filament-tail. The oxygen concentration cell formed directs anodic metal dissolution towards the leading-edge of the filament-head, while the principal site of cathodic oxygen reduction lies towards the trailing-edge of the active head. Aggressive anions (Cl^-) and water are conserved in the filament head electrolyte and corrosion filaments are reported to continue propagating for long periods of time (months or years). Although these general principles are considered to hold true for FFC on both aluminum and iron surfaces, there are notable mechanistic differences,

and these will be discussed separately for each type of surface in the following sections.

It should also be noted that there are several anecdotal reports of FFC observed on coated magnesium surfaces.[5,6,98,99] However, magnesium underfilm corrosion has received relatively little in the way of scientific study and a description of FFC remains limited to observations of vigorous H_2 evolution and a low pH (2–3) within electrolyte filled head regions.[99] In fact, there seems to be a degree of confusion regarding magnesium FFC, as the term is commonly used to describe a type of localized corrosion observed when magnesium and its alloys are immersed in bulk chloride-containing electrolyte.[100,101]

2.14.3.2 FFC on Aluminum Surfaces

Filiform activity on aluminum has been shown to initiate at a range of different coating discontinuity sites such as cut edges,[5,96] scratches[6,96–98,102] and rivet heads.[6,99] At such defects, the presence of contaminating salts, such as chlorides will have resulted in production of $AlCl_3$ and Al(III) oxychloride salts ($Al(OH)_2Cl$ and $Al(OH)Cl_2$) at the exposed aluminum surface. The dissolution of these salts in the adsorbed humidity film present at high humidity subsequently produces an aqueous electrolyte capable of supporting corrosion. Chloride is known to disrupt the passivity of aluminum by reacting with the protective aluminum (hydr)oxide film to produce water soluble aluminum (III) oxychloride complexes. The passive layer breakdown is usually local in nature, resulting in the initiation of highly localized sites of anodic metal dissolution and producing pit-like excavations of the aluminum surface. The generation of the active heads of propagating FFC filaments probably results from the migration of such sites beneath coated regions adjacent to a penetrative defect. **Figure 8(a)** shows a possible mechanism by which corrosion filaments may become initiated through the presence of differential aeration. In this scheme, it is assumed that oxygen transport to the metal surface through the thin adsorbed electrolyte film at the exposed aluminum surface is significantly easier than through a relatively thick organic coating. Hence cathodic oxygen reduction is favored on the bare metal while anodic aluminum dissolution will tend to become concentrated at the O_2-deficient metal–electrolyte–coating interface. Correspondingly, the coating will become undercut while Cl^- ions will migrate beneath the delaminated coating to preserve

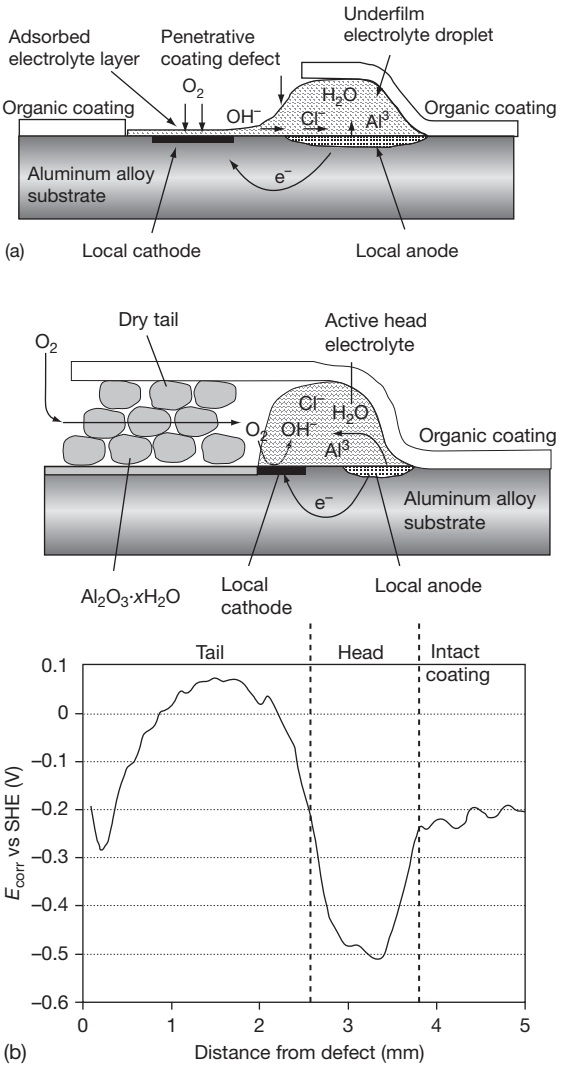


Figure 8 Schematic diagram showing the processes involved during (a) initiation and (b) propagation of a corrosion filament. A SKP-derived free corrosion potential distribution, measured along the length of a propagating filament on organic-coated AA2024-T3, showing typical E_{corr} values in each separate region is also given in (b).

electroneutrality and water will be drawn in by osmosis to produce an electrolyte droplet.

Once initiated, corrosion filaments propagate away from the coating defect, under the influence of an O_2 concentration cell of the type represented schematically in **Figure 8(b)**. Aluminum (III) cations produced by anodic dissolution at the leading edge of the filament head migrate towards a cathodic region at the trailing edge of the active head and combine with OH^- anions generated by cathodic oxygen reduction. Stepwise hydrolysis of Al^{3+} (or more

correctly $[Al(H_2O)_6]^{3+}$) cations eventually leads to the precipitation of the water insoluble corrosion product $Al(OH)_3$. **Figure 8(b)** also includes an empirically-derived free corrosion potential (E_{corr}) profile, obtained along the length of a propagating filament on aluminum by SKP potentiometry.[103] The profile shows that E_{corr} values tend to be lower towards the leading edge of the FFC active head and higher towards the trailing edge. This is consistent with the extent of anodic depolarization, and hence anodic dissolution, being greatest at the leading edge.

Constant migration of Cl^- anions toward the anodic leading edge results in the retention of all water-soluble ions and, through osmosis, all liquid water within the FFC active head. Several authors have used surface analysis techniques such as secondary ion mass spectrometry (SIMS)[103] and X-ray emission[104] to demonstrate that Cl^- ions applied at a penetrative defect are conserved within the filament head, with little evidence of entrapment in defect or tail regions of the FFC-affected surface. The high E_{corr} values observed in the tail region of the FFC filament represented in **Figure 2(b)** is due to the deposition of a thick, substantially dry, hydrated aluminum (III) oxide ($Al_2O_3{\cdot}xH_2O$) layer, formed as an $Al(OH)_3$ corrosion product left behind in the wake of the advancing filament head slowly losing water. Relatively easy O_2 diffusion through the porous $Al_2O_3{\cdot}xH_2O$ deposit allows the FFC differential aeration cell, and hence filament propagation, to be maintained over significant distances. The high E_{corr} values observed in the FFC tail imply increased passivity in these regions, which may help to explain empirical observations that individual corrosion filaments generally avoid intersecting others.[6] Such highly passive areas would tend to disfavor further anodic activity, and an encroaching FFC head region would be forced to redirect onto adjacent uncorroded areas.

Several authors have observed the formation of hydrogen gas bubbles within filament heads produced on aluminum surfaces[6,96,105] Hoch employed pH indicator dyes incorporated within organic coatings to determine pH values of 2–3 at leading edges of aluminum filament head.[99] Such levels of acidity will tend to maintain active aluminum dissolution in these regions. A combination of low pH, caused by proton release via partial hydrolysis of $[Al(H_2O)_6]^{3+}$ cations, negative E_{corr} values of ~ -0.5 V versus SHE,[103] along with oxygen deficiency at the leading edge of the filament head are thought to result in a secondary cathodic reaction involving hydrogen evolution.

Experimental support for an anodic disbondment mechanism of FFC on aluminum is principally derived from the observations of FFC behavior under different O_2 transport regimes. Ruggieri and Beck[5] showed that FFC advancement ceased when filament tails were sealed in such a way as to halt the ingress of air. They also demonstrated that stationary filaments became reactivated when organic coatings above tail regions were cut open. Lenderink[102] provided further support for the importance of through-tail O_2 diffusion by reporting that FFC propagation rates remained unaffected on samples where a metal foil had been deposited over the organic coating to obstruct through-coating O_2 transport.

The use of conventional electrochemical techniques to study FFC has routinely been limited by the absence of bulk electrolyte, the presence of highly resistive organic coatings and the small size of the local corrosion cell. However, in the last decade, significant developments in the SKP have greatly facilitated the quantitative electrochemical investigation of FFC. It has been shown that SKP is capable of visualizing the free corrosion potential (E_{corr}) distributions associated with individual corrosion filaments.[106–108] *In situ*, SKP scanning has also been used to generate a series of time-dependent E_{corr} distribution patterns that provide information on filament initiation, propagation, and termination rates within a filament population.[97,103] FFC propagation rates on AA2024-T3, quantified by SKP,[103] are shown to be strongly dependent on the quantity of the initiating chloride, confirming that the ability to control the quantity of chloride ion entering a filament-head population is critical to the reproducibility of any quantitative FFC test procedure. Other *in situ* techniques, such as Fourier transform infrared (FTIR) microspectroscopy, have also been used in conjunction with SKP scanning to give additional information regarding the constituent salts present within propagating electrolyte-filled FFC heads.[109]

The same *in situ* techniques have also been employed to study inhibition of FFC by in-coating pigments based on both conventional Cr(VI) chemistry[110] and novel anion-exchange hydotalcite minerals.[111–113] Previous *ex situ* studies, using a combination of electron microscopy and X-ray and SIMS techniques, showed that Cr(VI) species inhibit FFC on copper-containing AA2024 primarily by stifling S-phase particle dealloying and subsequent cathodic O_2 reduction. *In situ* SKP studies confirmed this finding, but in addition demonstrated that a significant depression in E_{corr} of the intact coated metal

surface resulted from a reversible, nonfaradaic adsorption of Cr(VI) species, which also contributed to FFC inhibition.[110]

2.14.3.3 Surface-Active FFC on Aluminum

Within the last decade it has been established that thermo-mechanical processing and/or shear deformation of aluminum alloys can produce a surface microstructure which differs considerably from the bulk.[114–123] Deformed aluminum alloy surface layers are typically up to a few micrometers thick, consisting of fine, submicron grain sizes where the presence of second phase particles stabilize grain boundaries and prevent recrystallization. Many researchers have so far demonstrated that a strong correlation exists between the presence of a deformed surface layer and susceptibility of painted alloy samples to FFC.[114–123] In several cases, the severity of FFC attack is increased by heat treatment, and attributed to thermally-promoted precipitation of intermetallic particles in the surface layer. The morphology of FFC on surface-activated samples is considered to differ from that discussed in the previous section and shown in **Figure 7**. FFC attack on samples comprising a deformed surface layer proceeds rapidly and is highly cosmetic and superficial in nature, limited only to the depth of the surface layer itself. Frequently FFC takes the form of numerous narrow filaments, propagating on a parallel course and separated laterally by only small distances, giving the impression of a single coalesced corrosion front. **Figure 9** shows this type of cosmetic FFC affecting automotive alloy AA6016-T4 coated with a clear organic film, where corrosion was initiated by controlled application of aqueous hydrochloric acid to a penetrative coating defect.

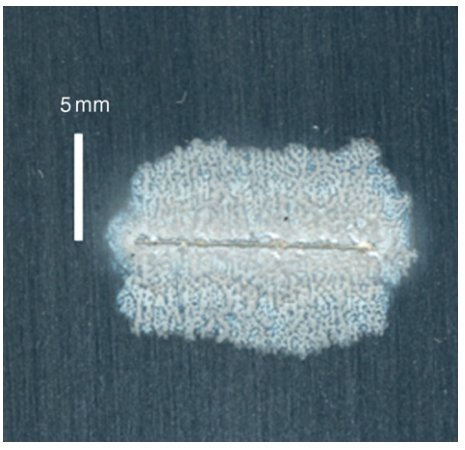


Figure 9 Filiform corrosion observed on aluminum alloy AA6016-T4 overcoated using PVB. Courtesy of A. Coleman.

Surface-active FFC has been reported on a wide range of alloys, ranging from model Al–Fe–Mg–Si[123] and binary Al–X (X = Fe, Cu, Mn and Mg)[120] systems through to automotive and architectural 5000[118,122] and 8000[114–116] series alloys. However, the majority of research in this area has concentrated on 3000 (Al–Mn)[114–119,121] and 6000 (Al–Mg–Si)[124–128] series alloys designed for automotive panel applications. The increased FFC susceptibility of AA3005 samples comprising damaged surface layers is thought to derive from preferential precipitation of noble Mn and Fe rich intermetallics,[119] leading to a cathodic activation of the surface layer with respect to the bulk alloy. Several authors have confirmed that when the surface layer is removed, the extent of FFC is considerably reduced and the type of attack is restricted to a much slower 'deep pitting' mode of corrosion,[117,122] more closely resembling the type of features shown in **Figure 7**.

In situ SKP potentiometry has recently been used to follow the kinetics of FFC propagation and measure underfilm corrosion potentials associated with head and tail regions of surface-active FFC affecting organic-coated AA6016.[125] The authors report that for this alloy, surface abrasion and subsequent heat-treatment produces a significant depression in surface potential compared with the bulk material and this forms an additional thermodynamic driving force for rapid 'surface-active' corrosion. They conclude that although differential aeration remains a major driving force for FFC propagation, the filament cell potential will also contain a contribution arising from the dissimilar electrochemical characteristics of the surface deformed layer present at the leading edge of the FFC head and bulk alloy exposed by layer dissolution. A schematic representation of the mechanism of 'surface-active' FFC, where the deformed layer reacts as a sacrificial anode with respect to bulk AA6016, is given in **Figure 10**. Also included is a representation of the expected relationship of corrosion potential with distance across the FFC cell, confirmed empirically by *in situ* SKP measurements.[125]

Several authors have shown that problems of increased FFC susceptibility, through the introduction of a damaged layer by thermomechanical alloy processing, can be controlled by etching the surface using acid or alkali cleaner prior to coating.[117,122,125] However, automotive manufacturers frequently rectify surface defects in formed panels by mechanical

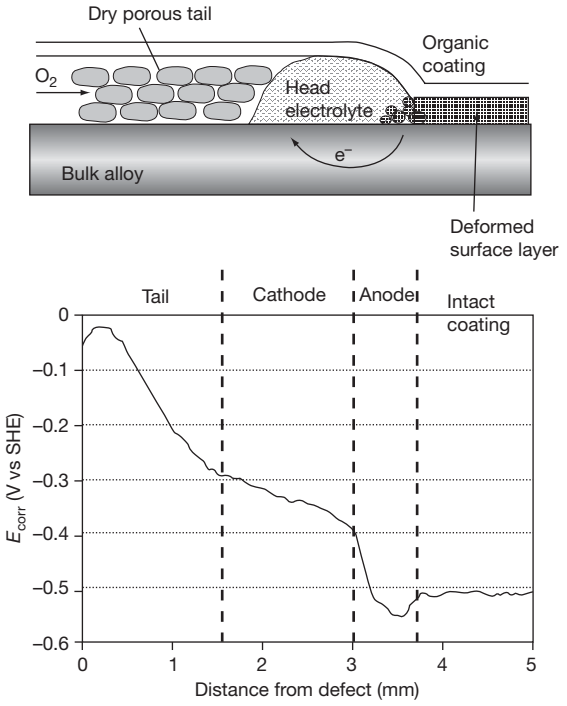


Figure 10 Schematic representation of the physical regions within a surface active FFC corrosion cell and proposed relationship with SKP-derived corrosion potential versus distance profiles.

grinding immediately prior to pretreatment and painting. This procedure reintroduces a deformed layer, subsequently reactivating the surface with respect to FFC and is liable to serious underfilm cosmetic failure. To avoid the introduction of a costly postgrinding etch or cleaning process, researchers are investigating new coating and/or pretreatment technologies which are capable of inhibiting FFC attack on surface-damaged alloys. To date, coatings based on phenylphosphonic acid dissolved in a polyvinylbutyral matrix are among the most promising candidates, demonstrating profound inhibition of FFC on AA6111-T4 automotive alloy samples comprising shear-deformed surface active layers.[126]

2.14.3.4 FFC on Iron (Steel) Surfaces

The general consensus regarding the mechanism of FFC on iron is that filament advance remains driven by differential aeration,[5,6] arising from facile mass transport of atmospheric O_2 through the filament-tail. As in the case of aluminum, cathodic oxygen reduction is considered to be localized at the droplet trailing-edge and anodic metal dissolution located at the leading-edge of the electrolyte filled head. However, although this anodic disbondment mechanism is largely accepted for FFC on aluminum, there persists some controversy regarding the location of the anodic and cathodic sites for corrosion filaments propagating on a coated iron surface. Since, as discussed in the preceding section, organic-coated iron surfaces are well known to undergo corrosion-driven delamination via a cathodic disbondment route, several authors have suggested that a similar process plays a major role in the mechanism of FFC on iron. In this alternative mechanism, cathodic activity is proposed to occur in a narrow area in the front part of the filament head,[6,102,129] although it has also been suggested that a site of cathodic activity exists a small distance in front of the filament leading edge.[130,131] Drawing on observations previously noted for blister formation on steel, Funke[130,131] proposes a mechanism whereby FFC is characterized by the development of an underfilm cathode, which precedes the head electrolyte droplet. This cathodic region appears at some distance from the front of the head and grows due to the generation of underfilm hydroxyl ions, up to a point where the trailing filiform head draws near and the two regions merge. At this point OH^- ions migrate towards the anode while metal cations present within the head electrolyte migrate forward in the direction of the preceding cathode. This results in a polarity reversal where the previously cathodic front region is transformed to become the anodic leading edge of the filiform head. The whole process is then repeated and is considered to be consistent with both the observed saltatory (jumping) movement of FFC on iron and the segmented appearance of corrosion tracks. A schematic representation of this theory is shown in **Figure 11**.

Ruggieri and Beck[5] dispute this hypothesis, pointing out that the model is inconsistent with asymmetric separation of the cathodic and anodic zones in the corrosion cell, and that previous studies had shown no experimental evidence of cathodic detachment in the vicinity of the propagating filaments. In addition, since the preceding cathode directs the FFC, it is implicit that ionic current must flow from the anode through the coating. The size of the cathode would thus be limited by the high resistance of an adherent metal–coating interface. It could however be argued that diffusion of ions from the preceding cathodic area to the head might occur through a layer of adsorbed water existing at the coating–metal interface. Ruggeri and Beck calculate that the thickness of the aqueous layer required would be 3.5×10^{-5} mm,[5]

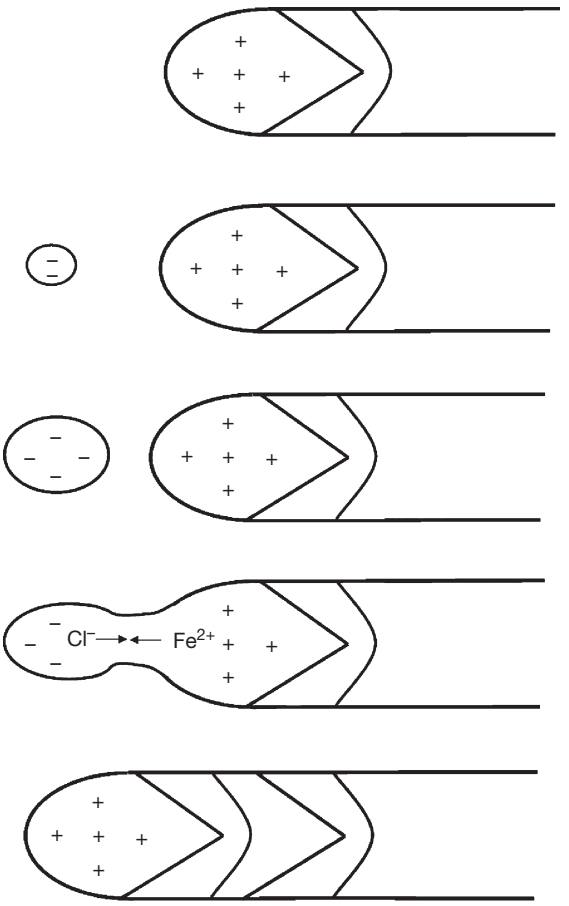


Figure 11 Filiform corrosion via a cathodic disbondment mechanism as proposed by Funke[130,131] for a protective organic coating on iron (steel).

which would tend to indicate a region of coating disbondment. As mentioned previously, the lack of empirical observation of the regions of coating disbondment preceding a propagating filiform head is taken as a further argument for rejecting the cathodic disbondment FFC mechanism.

Another point of debate concerns the role of the cation associated with the initiating anion (usually chloride). Several authors argue that although initiating anions are conserved in the filament head electrolyte, the corresponding cations present consist solely of Fe^{2+} and Fe^{3+} produced by the metallic corrosion process.[132,133] X-ray spectroscopic and SIMS evidence are presented showing that filament heads do not contain the cation of the initiating salt. Therefore, if corrosion is initiated by applying a salt such as a group (I) chloride to a penetrative coating defect, then a question mark arises as to the eventual fate of the group (I) cations.

To address this issue Williams and McMurray[133] used a combination of *in situ* SKP and *ex situ* imaging SIMS measurements to elucidate the role of group (I) cations in the initiation of FFC on iron by aqueous group (I) chloride salts and to additionally determine the contribution of cathodic disbondment processes in filament propagation. It was demonstrated that group (I) chloride – initiated FFC proceeded via a two stage mechanism. Phase 1 of the underfilm corrosion process comprised the radial growth of a depassivated, roughly circular area surrounding the penetrative defect onto which a controlled quantity of initiating group (I) chloride was originally applied, indicating localized anodic iron dissolution at the defect site coupled to cathodic delamination of the surrounding coating. Imaging SIMS confirmed that the area was abundant in the group (I) cation associated with the initiating salt, showing that phase 1 consisted of group (I) cation migration beneath the delaminating coating to form an alkaline under-film electrolyte layer. As phase 1 continued, group (I) cations became progressively replaced in the vicinity of the defect site by Fe^{2+} cations, which in turn were prevented from entering the alkaline under-film layer due to the insolubility of Fe(II) hydroxide. Once group (I) cations were exhausted in the defect region, phase 1 cathodic delamination ceased and phase II commenced, where a polarity reversal occurred in which anodic activity was driven beneath the delaminated coating. Underfilm anodic fragments formed corrosion filament heads which propagated across the well-defined circular zone of cathodically delaminated coating left by phase 1 at a constant speed of ~0.3 $\mu m\,min^{-1}$. Filaments were also shown to propagate beyond this zone with only a minimal decrease in velocity. For these filaments no evidence of cathodic activity preceding the head region was observed. The entire process is exemplified by the selection of E_{corr} distribution maps, obtained by *in situ* SKP scanning of a PVB-coated iron sample after initiation of underfilm corrosion using a controlled quantity of NaCl, given in **Figure 12**. When $FeCl_2$ was substituted for NaCl as the initiating salt, a cathodic delamination phase was not observed, although FFC occurred as before and propagation rates similar to those initiated with group (I) chlorides were measured. Because filament velocities were similar beneath cathodically delaminated and intact coatings, it was concluded that cathodic delamination cannot be a rate determining step in the propagation of FFC on organic-coated iron.[133]

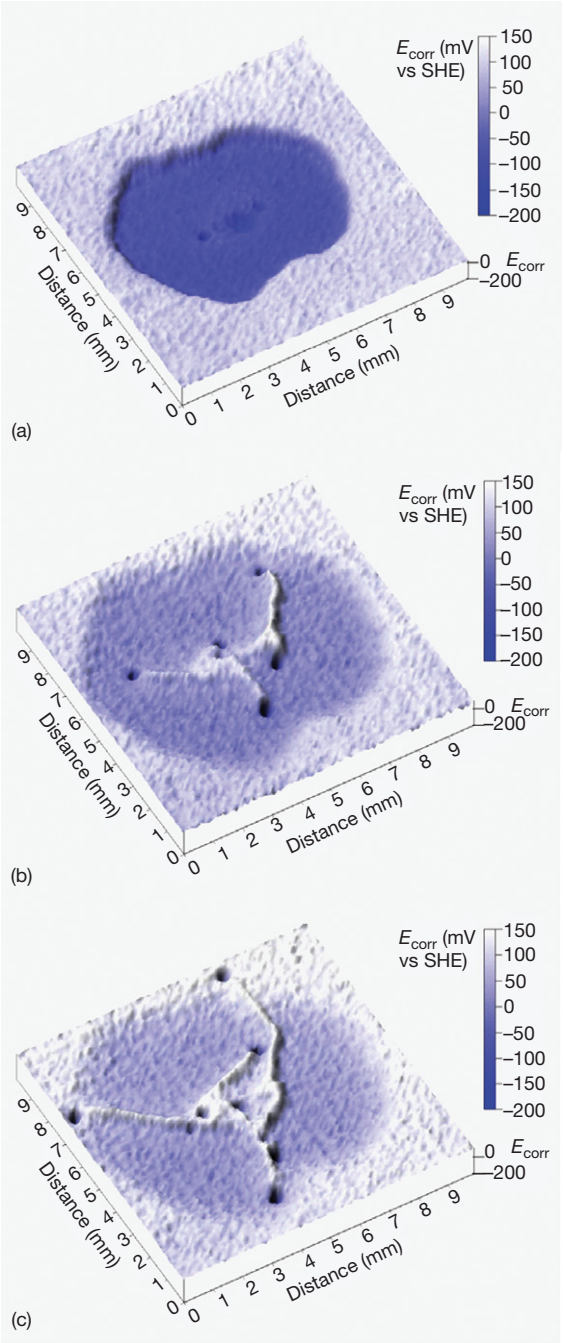


Figure 12 Surface maps showing E_{corr} distributions measured over a PVB coated iron surface at $t =$ (a) 16, (b) 130, and (c) 240 h, following initiation of underfilm corrosion at a penetrative coating defect using 2 µl of 10^{-3} mol dm^{-3} NaCl (aq). (a) shows an initial cathodic disbondment process (phase 1), while (b) and (c) respectively demonstrate the propagation of corrosion filaments both within and beyond the boundary of the well defined, cathodically delaminated zone (phase II).

The findings of this work imply that the hydrolytic instability of ferrous ions, generated by underfilm anodic activity, precludes Fe^{2+} cations from coupling with remote cathode sites in both phases of the corrosion process. However, the possibility of short range coupling over distances of <0.1 mm, that is, less than the lateral resolution of the SKP used in this study, cannot be ruled out. Indeed a subsequent higher resolution scanning Kelvin probe force microscope (SKPFM) investigation of iron FFC identified a possible cathodic area ~10–20 µm in front of the principally anodic leading edge of the head electrolyte. A recent time-lapse photographic investigation of individual corrosion filaments propagating on iron has shown that dark rings of a similar dimension form at the anterior of electrolyte filled heads,[134] and are associated with the saltatory motion of FFC. A photographic image showing this phenomenon, recorded using an optical microscope under reflective illumination, is given in **Figure 13**. It is considered that an increase in local pH at the head perimeter, caused by facile through-coating oxygen ingress at this point is responsible for the ring formation. When oxygen was excluded from the environment in contact with the filament head, but allowed to ingress through the porous tail, filaments continued to propagate at the same rate measured in air. However, under these conditions, neither ring formation nor filament saltation was observed and the filament advancement was steady in nature. The authors conclude that this apparent anterior cathodic activity is not a rate determining process in FFC on iron, but is an important factor in constraining the lateral spread of filament head electrolyte.[134]

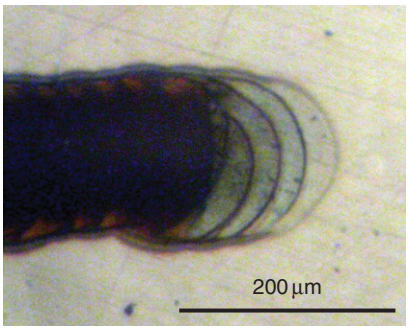


Figure 13 Magnified optical image of the head region of a corrosion filament propagating on a coated iron substrate in humid air, showing the formation of 'cathodic' corrosion product rings. Courtesy of T. Watson.

Acknowledgments

The authors would like to thank Drs. Trystan Watson and Andrew Coleman for supplying some of the diagrams shown in this work.

References

1. Leidheiser, H. *J. Adhes. Sci. Technol.* **1987**, *1*, 79.
2. Greenfield, D.; Scantlebury, D. J. *J. Corros. Sci. Eng.* **2000**, *2*(26). ISSN 1466–8858.
3. Grundmeyer, G.; Simoes, A. In *Encyclopedia of Electrochemistry*; Bard, A. J., Stratmann, M., Frankel, J., Eds.; Wiley-VCH: New York, 2003; Vol. 5, pp 500–566.
4. Ogle, K.; Simoes, A. In *Encyclopedia of Electrochemistry*; Bard, A. J., Stratmann, M., Frankel, J., Eds.; Wiley-VCH: New York, 2003; Vol. 4, pp 460–499.
5. Ruggieri, R. T.; Beck, T. R. *Corrosion* **1983**, *39*, 452.
6. Bauista, A. *Prog. Org. Coat.* **1996**, *28*, 49.
7. Mayne, J. E. O. *J. Oil Colour Chem. Assoc.* **1950**, *31*, 538.
8. Funke, W. In *Corrosion Control by Organic Coatings*; Leidheiser, H., Ed.; NACE: Houston, 1981; pp 97–102.
9. Gowers, K. R.; Scantlebury, J. D. *Corros. Sci.* **1983**, *23*, 935.
10. Nguyen, T.; Hubbard, J. B.; McFadden, G. B. *J. Coat. Technol.* **1991**, *63*, 43.
11. Mayne, J. E. O. *J. Oil Colour Chem. Assoc.* **1949**, *32*, 481.
12. Skoulikidis, T.; Ragoussis, A. *Corrosion* **1992**, *48*, 666.
13. Leidheiser, H.; Granata, R. D. *IBM J. Res. Dev.* **1988**, *32*, 582.
14. Koehler, E. L. *Corrosion* **1977**, *33*, 209.
15. Leidheiser, H.; Kendig, M. W. *Corrosion* **1979**, *32*, 69.
16. Leidheiser, H.; Wang, W. *J. Coat. Technol.* **1981**, *53*, 77; *Proc. Electrochem. Soc.* **1984**, *84*, 239.
17. Ritter, J. J. *J. Coat. Technol.* **1982**, *54*, 51.
18. Leidheiser, H.; Igetoft, L.; Wang, W. *Prog. Org. Coat.* **1983**, *11*, 19.
19. Koehler, E. L. *Corrosion* **1984**, *40*, 5.
20. Stratmann, M.; Streckel, H.; Feser, R. *Corros. Sci.* **1991**, *32*, 467.
21. Sharman, J. D. B.; Sykes, J. M.; Handyside, T. *Corros. Sci.* **1993**, *35*, 1375.
22. Stratmann, M.; Feser, R.; Leng, A. *Electrochim. Acta* **1994**, *39*, 1207–1214.
23. Furbeth, W.; Stratmann, M. *Fresenius J. Anal. Chem.* **1995**, *353*, 337.
24. Furbeth, W.; Stratmann, M. *Prog. Org. Coat.* **2000**, *39*, 23.
25. Lorimer, J. W., Ed. *Solubility Data Series. Alkaline Earth Hydroxides in Water and Aqueous Solutions;* IUPAC, Pergamon Press: Oxford, 1992; Vol. 52.
26. Ritter, J. J.; Kruger, J. In *Corrosion Control by Organic Coatings*; Leidheiser, H., Ed.; NACE: Houston, 1981; p 28.
27. Leidheiser, H.; Wang, W. *J. Coat. Technol.* **1981**, *53*, 77.
28. Watts, J. F.; Castle, J. E. *J. Mater. Sci.* **1984**, *19*, 2259.
29. Leng, A.; Streckel, H.; Stratmann, M. *Corros. Sci.* **1999**, *41*, 547.
30. Leng, A.; Streckel, H.; Stratmann, M. *Corros. Sci.* **1999**, *41*, 579.
31. Leng, A.; Streckel, H.; Stratmann, M. *Corros. Sci.* **1999**, *41*, 599.
32. Furbeth, W.; Stratmann, M. *Corros. Sci.* **1999**, *43*, 207.
33. Furbeth, W.; Stratmann, M. *Corros. Sci.* **1999**, *43*, 228.
34. Furbeth, W.; Stratmann, M. *Corros. Sci.* **1999**, *43*, 243.
35. Koehler, E. L. *J. Electrochem. Soc.* **1985**, *132*, 1005.
36. Castle, J. E.; Watts, J. F. In *Corrosion Control by Organic Coatings*; Leidheiser, H., Ed.; NACE: Houston, 1981; p 78.
37. Evans, U. R. *The Corrosion and Oxidation of Metals*; St Martins Press: NY, 1960.
38. Bard, A. J.; Faulkner, L. R. *Electrochemical Methods – Fundamentals and Applications*; Wiley: NY, 1980; pp 494–496.
39. Delahay, P. *Double Layer and Electrode Kinetics*; Interscience, 1966; pp 17–32.
40. Bolger, J. C.; Michaels, A. S. In *Interface Corrosion for Polymer Coatings*; Weiss, P., Cheever, G. D., Eds.; Elsevier: NY, 1968; p 3.
41. Fowkes, F. M. *Ind. Eng. Chem. Prod. Res. Dev.* **1978**, *17*, 3.
42. Hammond, J. S.; Holubka, J. W.; Dickie, R. A. *J. Coat. Technol.* **1979**, *51*, 45.
43. Hammond, J. S.; Holubka, J. W.; Dickie, R. A. *Corros. Sci.* **1981**, *21*, 239.
44. Dickie, R. A.; Hammond, J. S.; Holubka, J. W. *Ind. Eng. Chem. Prod. Res. Dev.* **1981**, *20*, 339.
45. Castle, J. E.; Watts, J. F. *Ind. Eng. Chem. Prod. Res. Dev.* **1985**, *24*, 361.
46. Funke, W. *J. Oil Colour Chem. Assoc.* **1979**, *62*, 63.
47. Funke, W. *J. Oil Colour Chem. Assoc.* **1985**, *68*, 229.
48. Funke, W.; Leidheiser, H. *J. Oil Colour Chem. Assoc.* **1987**, *70*, 121.
49. Marsh, J.; Scantlebury, J. D.; Lyon, S. B. In *Surfaced Engineering: Fundamentals of Coatings*; Datta, P. K., Gray, J. S., Eds.; Royal Society of Chemistry: Cambridge, 1993; Vol. 1, p 568. Publication no. 126.
50. Sabata, A.; Van Ooij, W. J.; Koch, R. J. *J. Adhes. Sci. Technol.* **1993**, *7*, 1153.
51. Marsh, J.; Scantlebury, J. D.; Lyon, S. B. *Corros. Sci.* **2001**, *43*, 829.
52. Harun, M. K.; Marsh, J.; Lyon, S. B. *Prog. Org. Coat.* **2005**, *54*, 317.
53. Suzuki, I. *Corrosion Resistant Coatings Technology*; Marcel Dekker: New York, 1989; pp 167–205.
54. Wiggle, R. R.; Smith, A. G.; Petrocelli, J. V. *J. Paint Technol.* **1968**, *40*, 174.
55. Sommer, A. J.; Leidheiser, H. *Corrosion* **1987**, *43*, 661.
56. Ogle, K.; Morel, S.; Meddahi, N. *Corros. Sci.* **2005**, *47*, 2034.
57. Klimow, G.; Fink, N.; Grundmeier, G. *Electrochim. Acta* **2007**, *53*, 1290.
58. Forget, L.; Delhalle, J.; Mekalif, Z. *Mater. Corros.* **2001**, *52*, 181.
59. Kuznetsov, Yu. I.; Rosenfeld, I. L.; Agalarova, T. A. *Prot. Met.* **1982**, *18*, 438.
60. Rosenfeld, I. L. *Corrosion Inhibitors*; McGraw Hill: New York, 1981; p 145.
61. Zhao, J.; Frankel, G.; McCreery, R. L. *J. Electrochem. Soc.* **1998**, *145*, 2258.
62. Williams, G.; McMurray, H. N. *J. Electrochem. Soc.* **2001**, *148*, B377.
63. McMurray, H. N.; Williams, G.; Worsley, D. A. *J. Electrochem. Soc.* **2002**, *149*, B154.
64. McMurray, H. N.; Williams, G.; Worsley, D. A. *Corros. Eng. Sci. Technol.* **2003**, *38*, 112.
65. Zin, I. M.; Pokhmurski, V. I.; Scantlebury, J. D.; Lyon, S. B. *J. Electrochem. Soc.* **2001**, *148*, B293.
66. Tallman, D. E.; Spinks, G. M.; Dominis, A. J.; Wallace, G. G. *J. Solid State Electrochem.* **2002**, *6*, 73.
67. Spinks, G. M.; Dominis, A. J.; Wallace, G. G.; Tallman, D. E. *J. Solid State Electrochem.* **2002**, *6*, 85.
68. Wessling, B. *Adv. Mater.* **1994**, *6*, 226.
69. Wessling, B. *Synth. Met.* **1997**, *85*, 1313.

70. Kinlen, P. J.; Ding, Y.; Silverman, D. C. *Synth. Met.* **1997**, *85*, 1327.
71. Bernard, M. C.; Joiret, S.; Goff, A. H.-Le.; Phong, P. V. *J. Electrochem. Soc.* **2001**, *148*, B12.
72. Kinlen, P. J.; Menon, V.; Ding, Y. *J. Electrochem. Soc.* **1999**, *146*, 3690.
73. Lu, X.; Ng, H. Y.; Xu, J.; He, C. *Synth. Met.* **2002**, *128*, 167.
74. Lu, W.; Eisenbaumer, R. L.; Wessling, B. *Synth. Met.* **1995**, *71*, 2163.
75. Wessling, B. *Synth. Met.* **1999**, *102*, 1396.
76. Tawde, S.; Mukesh, D.; Yakhmi, J. V. *Synth. Met.* **2002**, *125*, 401.
77. Bernard, M. C.; Hugot-Le Goff, A.; Joiret, S.; Ginh, N. N.; Toan, N. N. *J. Electrochem. Soc.* **1999**, *146*, 995.
78. Galkowski, M.; Malik, M. A.; Kulesza, P. J.; Bala, H.; Miecznikowski, K.; Wlodarczyk, R.; Adamczyk, L.; Chojak, M. *J. Electrochem. Soc.* **2003**, *150*, B249.
79. Posdorfer, J.; Wessling, B. *Synth. Met.* **2001**, *119*, 363.
80. Talo, A.; Forsén, O.; Yläsaari, S. *Synth. Met.* **1999**, *102*, 1394.
81. Kaiser, A. B.; Liu, C.-J.; Gilberd, P. W.; Chapman, B.; Kemp, N. T.; Wessling, B.; Partidge, A. C.; Smith, W. T.; Shapiro, J. S. *Synth. Met.* **1997**, *84*, 699.
82. Schauer, T.; Joos, A.; Dulog, L.; Eisenbach, C. D. *Prog. Org. Coat.* **1998**, *33*, 20.
83. Tallman, D. E.; Pae, Y.; Biervagan, G. P. *Corrosion* **1999**, *55*, 779.
84. Gašparac, R.; Martin, R. C. *J. Electrochem. Soc.* **2001**, *148*, B138.
85. Wessling, B.; Posdorfer, J. *Synth. Met.* **1999**, *102*, 1400.
86. Holness, R. J.; Williams, G.; Worsley, D. A.; McMurray, H. N. *J. Electrochem. Soc.* **2005**, *152*, B73.
87. Williams, G.; Gabriel, A.; Cook, A.; McMurray, H. N. *J. Electrochem. Soc* **2006**, *153*, B425.
88. Sahimi, M. *Applications of Percolation Theory*; Taylor and Francis Ltd: London, 1994; pp 201–207.
89. Dafydd, H.; Worsley, D. A.; McMurray, H. N. *Corros. Sci.* **2005**, *47*, 3006.
90. Granata, R. D.; De Crosta, M. A.; Leidheiser, H. *Proc. Int. Cong. Met. Corros.* **1984**, 264–268.
91. Davies, J. R. *Corrosion: Understanding the Basics*; ASTM International, 2000; p 403.
92. Suzuki, I. *Corrosion Resistant Coatings Technology*; Marcel Dekker: New York, 1989; pp 21–108.
93. Souto, R. M.; Scantlebury, D. J. *Prog. Org. Coat.* **2005**, *53*, 63.
94. Hausbrand, R.; Stratman, M.; Rohwerder, M. *Steel Res.* **2003**, *74*, 253.
95. Sharman, C. F. *Nature* **1944**, *153*, 621.
96. Slabaugh, W. H.; Dejager, W.; Hoover, S. E.; Hutchinson, L. L. *J. Paint Technol.* **1972**, *44*, 76.
97. Williams, G.; McMurray, H. N.; Hayman, D.; Morgan, P. C. *Phys. Chem. Comm.* **2001**, *6*, 26–31.
98. Van Loo, M.; Laiderman, D. D.; Bruhn, R. R. *Corrosion* **1953**, *9*, 277.
99. Hoch, G. M. A Review of Filiform Corrosion. In *Localised Corrosion*; Staehle, R. S., Brown, B. F., Kruger, J., Agrawal, A., Eds.; NACE: Houston, 1974; p 134.
100. Lunder, O.; Lein, J. E.; Hesjevik, S. M.; Aune, T. K.; Nisancioglu, K. *Werkst. Korros.* **1994**, *45*, 331.
101. Schmutz, P.; Guillaumin, V.; Lillard, R. S.; Lillard, J. A.; Frankel, G. S. *J. Electrochem. Soc.* **2003**, *150*, B99.
102. Lenderink, H.J.W. PhD thesis, Technical University of Delft, 1995.
103. Williams, G.; McMurray, H. N. *J. Electrochem. Soc.* **2003**, *150*, B380.
104. Szymanski, R.; Jamieson, D. N.; Hughes, A. E.; Mol, A.; Van der Zwaag, S.; Ryan, C. G. *Nucl. Instrum. Methods. Phys. Res., Sect. B* **2002**, *190*, 365.
105. Mol, J. M. C.; Hinton, B. R. W.; Van der Weijde, D. H.; De Wit, J. H. W.; Van der Zwaag, S. *J. Mater. Sci.* **2000**, *35*, 1629.
106. Schmidt, W.; Stratmann, M. *Corros. Sci.* **1998**, *40*, 1441.
107. Grundmeier, G.; Schmidt, W.; Stratmann, M. *Electrochim. Acta* **2000**, *45*, 2515.
108. Le Bozec, N.; Persson, D.; Nazarov, A.; Thierry, D. *J. Electrochem. Soc.* **2002**, *149*, B403.
109. LeBozec, N.; Persson, D.; Thierry, D. *J. Electrochem. Soc.* **2004**, *151*, B440.
110. McMurray, H. N.; Williams, G.; O'Driscoll, S. *J. Electrochem. Soc.* **2004**, *151*, B406.
111. Williams, G.; McMurray, H. N. *Electrochem. Solid-State Lett.* **2003**, *6*, B9.
112. Williams, G.; McMurray, H. N. *Electrochem. Solid State Lett.* **2004**, *7*, B13.
113. McMurray, H. N.; Williams, G. *Corrosion* **2004**, *60*, 219.
114. Leth-Olsen, H.; Nisancioglu, K. *Corros. Sci.* **1998**, *40*, 1179.
115. Leth-Olsen, H.; Afseth, A.; Nisancioglu, K. *Corros. Sci.* **1998**, *40*, 1195.
116. Leth-Olsen, H.; Nordlien, J. H.; Nisancioglu, K. *Corros. Sci.* **1998**, *40*, 2051.
117. Afseth, A.; Nordlien, J. H.; Scamans, G. M.; Nisancioglu, K. *Corros. Sci.* **2001**, *43*, 2093.
118. Afseth, A.; Nordlien, J. H.; Scamans, G. M.; Nisancioglu, K. *Corros. Sci.* **2001**, *43*, 2359.
119. Afseth, A.; Nordlien, J. H.; Scamans, G. M.; Nisancioglu, K. *Corros. Sci.* **2002**, *44*, 2491.
120. Afseth, A.; Nordlien, J. H.; Scamans, G. M.; Nisancioglu, K. *Corros. Sci.* **2002**, *44*, 2529.
121. Afseth, A.; Nordlien, J. H.; Scamans, G. M.; Nisancioglu, K. *Corros. Sci.* **2002**, *44*, 2543.
122. Zhou, X.; Thompson, G. E.; Scamans, G. M. *Corros. Sci.* **2003**, *45*, 1767.
123. Ambat, R.; Davenport, A. J.; Afseth, A.; Scamans, G. *J. Electrochem. Soc.* **2004**, *151*, B53.
124. Liu, Y.; Zhou, X.; Thompson, G. E.; Hashimoto, T.; Scamans, G. M.; Afseth, A. *Acta Mater.* **2007**, *55*, 353.
125. McMurray, H. N.; Coleman, A. J.; Williams, G.; Afseth, A.; Scamans, G. M. *J. Electrochem. Soc.* **2007**, *154*, C339.
126. Coleman, A. J.; McMurray, H. N.; Williams, G.; Afseth, A.; Scamans, G. M. *Electrochem. Solid-State Lett.* **2007**, *10*, C35.
127. Scamans, G. M.; Afseth, A.; Thompson, G. E.; Liu, Y.; Zhou, X. *Mater. Sci. Forum* **2006**, *519–521*, 647.
128. Svenningsen, G.; Larsen, M. H.; Lein, J. E.; Nordlien, J. H.; Nisancioglu, K. In Proceedings of the 9th International Conference on Aluminium Alloys (ICAA9); Nie, J. F., Morton, A. J., Muddle, B. C., Eds.; Institute of Material Engineering Australasia Ltd.: Melbourne, Australia, 2004; p 818.
129. Leblanc, P. P.; Frankel, G. S. *J. Electrochem. Soc.* **2004**, *151*, B105.
130. Funke, W. *Prog. Org. Coat.* **1981**, *9*, 29.
131. Funke, W. *Ind. Eng. Chem. Prod. Res. Dev.* **1985**, *24*, 343.
132. Koehler, E. L. *Corrosion* **1977**, *33*, 209.
133. Williams, G.; McMurray, H. N. *Electrochem. Commun.* **2003**, *5*, 871.
134. Watson, T. M.; McMurray, H. N.; Williams, G.; Coleman, A. C. In *Corrosion in Marine and Saltwater Environments II*; Proceedings of the 206th Electrochemical Society Meeting., Honolulu, Hawaii, USA, October 2004; Shifler, D. A., Tsuru, T., Natishan, P. M., Ito, Y., Eds.; ECS, 2005; p 321.

2.13 Flow Assisted Corrosion

H. G. Schmitt and M. Bakalli
Iserlohn University of Applied Sciences, Frauenstuhlweg 31, D-58644 Iserlohn, Germany

Abbreviations

DRAs Drag-reducing agents
FEC Flow enhanced corrosion
FED Freak energy density
FILC Flow-induced localized corrosion

Symbols

a Exponent (constant)
A Amplitude of the noise current density from the polynomial fitting ($A\,m^2$)
b Microelectrode width (m)
c_0, ***c**** Solution concentration at the electrode surface ($mol\,m^{-3}$)
c_{bulk} Bulk solution concentration ($mol\,m^{-3}$)
c_S Concentration at the entrance of the pores ($mol\,m^{-3}$)
d Diameter of the pipe/characteristic diameter/ thickness layer (m)
d_N Nernst diffusion layer thickness (m)
D Diffusion coefficient ($m^2\,s^{-1}$)
D_{Pr} Diffusion coefficient in the Prandtl layer ($m^2\,s^{-1}$)

D_S Diffusion coefficient in the pores of the scale ($m^2\,s^{-1}$)
E_D Hydrodynamic energy density (Pa)
E_D^F Freak energy density (Pa)
F Faraday constant (96 485) ($C\,mol^{-1}$)
Fr Froude number
g Acceleration due to gravity ($m\,s^{-2}$)
h_{eff} Effective height of the liquid film (m)
j_c Corrosion current density ($A\,m^{-2}$)
j_D Diffusion controlled current density ($A\,m^{-2}$)
k Mass transport coefficient ($m\,s^{-1}$)
k Constant on the Léveque equation ($N\,m^4\,A^{-3}$)
k_1, k_2, k_3 Constants
K Average height of all roughness peaks
l Tube or channel length/microelectrode length (m)
L Characteristic length (m)
m_{Pr} Mass transport rate in the Prandtl's boundary layer ($mol\,m^{-2}\,s^{-1}$)
m_S Mass transport rate in the pores of the scale ($mol\,m^{-2}\,s^{-1}$)
n Exponent (depends on Re)
Pe Dimensionless Peclet number
R Inner radius of the pipe (m)
r_{RC} Radius of the rotated cage (m)
Re Dimensionless Reynolds number
Re_{RC} Dimensionless Reynolds number in rotated cage
Sc Dimensionless Schmidt number
Sh Dimensionless Sherwood number
Ta Dimensionless Taylor number
u^+ Normalized (dimensionless) flow velocity
u_b Bulk flow velocity ($m\,s^{-1}$)
u_c Velocity at the pipe centreline ($m\,s^{-1}$)
u_{crit} Critical flow velocity ($m\,s^{-1}$)
u_p Particle velocity ($m\,s^{-1}$)
u_r Axial velocity ($m\,s^{-1}$)
u_{SL} Superficial liquid velocity ($m\,s^{-1}$)
u_{SG} Superficial gas velocity ($m\,s^{-1}$)
u_t Translational velocity of the slug front ($m\,s^{-1}$)
u_f Average velocity of the liquid film ($m\,s^{-1}$)
u_x Shear velocity ($m\,s^{-1}$)
U Average fluid velocity in the pipe ($m\,s^{-1}$)
U_0 Flow velocity at the boundary layer ($m\,s^{-1}$)
v Corrosion rate ($g\,m^{-2}\,h$)
w Energy density/mass transport rate/corrosion rate ($Pa/mol\,m^{-2}\,s^{-1}/g\,m^{-2}\,h$)
w_{er} Mass loss rate in erosion corrosion ($g\,m^{-2}\,h$)
Y Axial position/distance from the wall (m)
y^+ Normalized (dimensionless) wall distance
Z Number of charges
δ_{Pr} Prandtl's boundary layer thickness
Δc Concentration gradient ($mol\,m^{-3}$)
Δp, ΔP Pressure loss (Pa)
ϵ Relative roughness
λ Tube friction factor
μ Hydrodynamic viscosity (Pa s)
ν Kinematic viscosity of flowing fluid ($m^2\,s^{-1}$)
ω Wave frequency/rotational speed (s^{-1})
$\Phi_{destruct}$ Rate of scale destruction
Φ_{scale} Rate of scale formation
ρ Density ($kg\,m^{-3}$)
σ_B Fracture stress of the scales (Pa)
τ_{RC} Wall shear stress in rotated cage (Pa, $N\,m^{-2}$)
τ_w Wall shear stress (Pa, $N\,m^{-2}$)

2.13.1 Introduction

Flowing media are encountered in nearly all technical systems, for example, in heat exchangers, pumps, hydraulic machines, steam generators, different components of power stations, chemical plants, and the oil and gas industry, and in many cases, corrosion failures occurred due to flow effects. Understanding the mechanisms involved needs a detailed knowledge of both corrosion (specifically electrochemical corrosion) and fluid dynamics.

On the corrosion side, the type of material and the status of its surface (scale-free or scale-covered) are decisive. On the fluid dynamics side, the type of flow regime must be considered (developed laminar and turbulent flow; disturbed flow; one-phase, two-phase, multiphase flow without and with solid particles; and in the case of multiphase flow, horizontal, inclined, and vertical flow).[1–11]

2.13.2 Flow Dynamic Basics

2.13.2.1 Laminar and Turbulent Flow

The intensity of flow can be characterized as either laminar or turbulent. In laminar pipe flow, the profile of the local flow velocity, u_r, is parabolic (**Figure 1(a)**) and can be calculated according to eqn [1].[12]

$$u_r = u_c[1 - (y/R)^2] \qquad [1]$$

where u_r is the axial velocity, u_c is the velocity at the pipe centerline, y is the axial position, and R is the inner radius of the pipe. Equation [1] is an analytical solution of the Navier–Stokes equation and the continuity equations of the fluid flow.

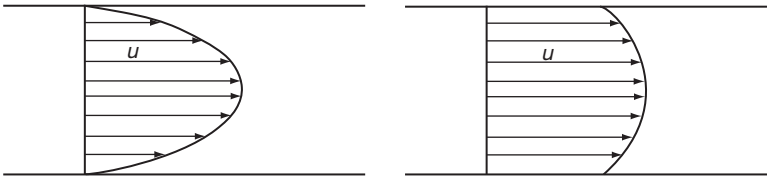


Figure 1 Velocity profile in (a) laminar flow and (b) turbulent flow.

In turbulent pipe flow, the profile of the local flow velocity, u_r, is logarithmic (**Figure 1(b)**) and can be calculated according to eqn [2].[12]

$$u_r = u_c[1 - (y/R)]^{1/n} \qquad [2]$$

where n depends on the Reynolds number (Re). Equation [2] is based on extensive experimental analysis, because no exact solution of the Navier–Stokes equation can be given for turbulent pipe flow.

The dimensionless Re is commonly used to characterize the nature of fluid flow. For liquid flow through pipes, Re is defined according to eqn [3].[12]

$$Re = \frac{uL}{\nu} = \frac{u\rho L}{\mu} \qquad [3]$$

where u is the average fluid velocity in the pipe ($m\,s^{-1}$), L, the characteristic length (for pipe flow the diameter, d, of the pipe) (m), and ν, the kinematic viscosity of flowing fluid ($m^2\,s^{-1}$), which can be expressed as the ratio of the hydrodynamic viscosity μ (Pa s) and the density ρ ($kg\,m^{-3}$).

From the Re, the flow regime can be derived. For Newtonian liquids, pipe flow is in the:

- developed laminar flow regime, if $Re < 2300$;
- transition (buffer) flow regime, if $2300 < Re < 4000$;
- turbulent flow regime, if $Re > 4000$.

In flow systems with real liquids, energy losses occur as a result of friction at the solid wall and due to eddies (vortices) in the near-wall region. In the laminar flow regime, the pressure loss in a pipe or channel is proportional to the average flow velocity, u, and can be quantified according to the Hagen–Poiseuille equation (eqn [4]).[13]

$$\frac{\Delta p}{l} = 32\mu u \frac{1}{d^2} \qquad [4]$$

where Δp is the pressure loss, l is the tube or channel length, and d is the characteristic diameter. However, in technical systems, the surface roughness has to be considered with a tube friction factor λ (eqn [5]), which depends on both the surface roughness and the flow regime (Re).[13]

$$\Delta p = \lambda \frac{l}{d} \frac{u^2}{2} \rho \qquad [5]$$

For laminar flow ($Re < 2300$), the friction factor is a direct function of the Re (eqn [6]):

$$\lambda = \frac{64}{Re} \qquad [6]$$

In the case of turbulent flow, the friction factor is strongly dependent on the Re, and Δp will increase proportional to u^2. For $2300 \leq Re \leq 10^5$, the following relationship was found by Blasius for smooth surfaces (eqn [7]):

$$\lambda = \frac{0.3164}{Re^{1/4}} \qquad [7]$$

In the case of small Re, the thickness of the boundary layer may be sufficient to cover all roughness peaks. The pressure loss will be slightly greater than in the case of smooth walls. However, it occurs according to the same physical laws relevant to the laminar region. With increasing Re, the boundary layer becomes thinner and thinner, and the roughness elements penetrate through the boundary layer into the main flow stream resulting in considerable drag and pressure loss. Thus, with increasing flow velocity, the pressure loss becomes more dependent on the surface roughness. At high roughness, the viscous forces in the near-wall boundary layer become negligible.

Regardless of the multitude of roughness forms and their distribution, a relative roughness, ε, is defined according to eqn [8]:

$$\varepsilon = \frac{K}{d} \qquad [8]$$

where K is the average height of all roughness peaks. Equation [8] also indicates that wall effects decrease with increasing pipe diameter. In the case of fully developed flow, the friction factor λ becomes

independent of the *Re* and can be calculated according to eqn [9] (Prandtl, Karman).[14]

$$\lambda^{-1/2} = 2 \log \frac{d}{K} + 1.14 \qquad [9]$$

These equations are empirical in nature. However, they clearly indicate that pressure losses in pipes strongly depend on the flow velocity (more general: the *Re*) and on fluid–wall interactions, which depend on the surface roughness on the solid side and on the nature of the boundary layer on the fluid side. In the following section, the nature of the boundary layer is looked at in more detail.

2.13.2.2 The Boundary Layer

The boundary layer is a near-wall region with non-uniform flow in a small layer between the bulk flow and the solid wall. This small layer is not constant over a greater length, but develops gradually in the flow direction. All friction effects occur within this layer, while in the bulk region, the flow is friction-free in the ideal case.

The development of the boundary layer is shown schematically in **Figure 2**. A thin plate is hit by a free flow with the velocity U_0. A boundary layer starts forming at the plate surface. The gradient of flow velocity, $u(y)$ – with y giving the distance from the plate surface – changes over the length of the plate. The thickness of the boundary layer increases with the distance, x, from the front of the plate. The Prandtl's boundary layer, δ_{Pr}, is defined as the distance y where the local flow velocity is only 1% smaller than the bulk velocity (eqn [10]).[15]

$$\delta_{Pr} = y_{(Uy=0.99U_0)} \qquad [10]$$

The flow in the boundary layer can be laminar or turbulent. In a small region after the leading edge of the plate where the boundary layer thickness is still small, the flow regime is laminar. With increasing thickness of the boundary layer along the plate, the boundary layer becomes increasingly unstable and turbulences are formed. The region where laminar flow changes to turbulent flow is called the transition region. The thickness of the boundary layer increases step-like along the plate. The time-averaged velocity profiles are shown in **Figure 2**. In the laminar region, the wall shear stresses, τ_w, decrease up to the transition region. There wall shear stresses τ_w increase spontaneously because the ratio between the change of flow velocity, du, and the distance, dy, from the plate increases spontaneously. However, τ_w decreases gradually along the plate length, x.

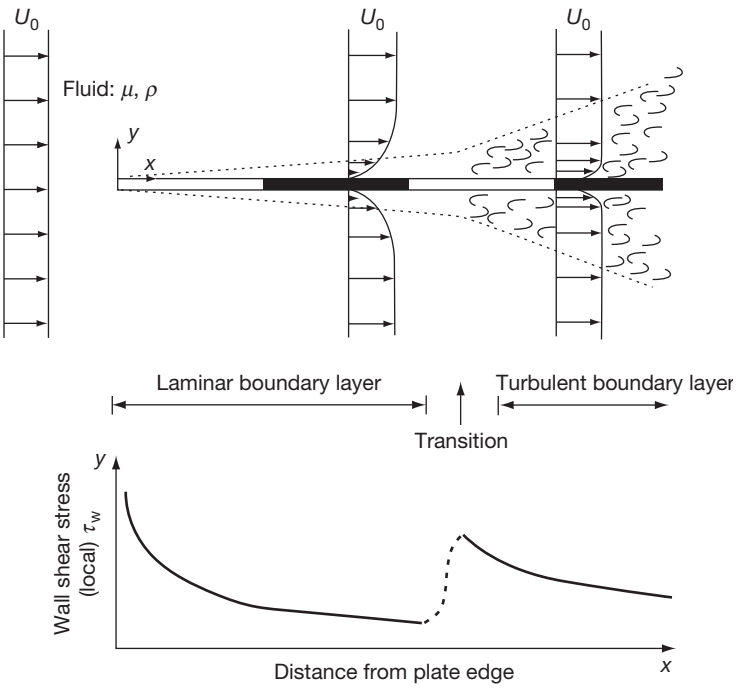


Figure 2 Boundary layer along a thin plate.

The properties of the boundary layer can be characterized by local *Re*, which are related either to the layer thickness, δ_{Pr}, or to the distance, x, from the front of the plate (eqns [11] and [12]).[15]

$$Re_{\delta_{Pr}} = \frac{U_0 \delta_{Pr}}{\mu} \quad [11]$$

$$Re_x = \frac{U_0 x}{\mu} \quad [12]$$

Laminar boundary layer: The layer thickness, δ_{Pr}, increases with the distance from the leading edge of the plate and decreases with the flow velocity, U_0. This can be expressed by the Blasius equation (eqn [13]).[15]

$$\delta_{Pr} = 5.0\sqrt{\frac{\upsilon x}{U_0}} = \frac{5.0x}{\sqrt{Re_x}} \quad [13]$$

The wall shear stress, τ_w, is directly related to the gradient of the flow velocity (eqn [14]).

$$\tau_W = \mu\left(\frac{dU}{dy}\right)_{y=0} = 0.332\ \mu\ \frac{U_0}{x}\sqrt{Re_y} \quad [14]$$

where μ is the hydrodynamic viscosity (Pa s).

Turbulent boundary layer: The turbulent boundary layer exhibits a much more complex structure than the laminar boundary layer. The major part of the boundary layer consists of the turbulent zone (**Figure 3**) with a chaotic turbulence activity and permanently changing flow pattern. The turbulent zone is again divided into an outer zone and an inner zone (often called buffer zone). Only in the inner zone are viscous effects encountered.[16,17]

Close to the wall, in the inner zone (**Figure 3**), the fluctuations are damped, resulting in a viscous sublayer with laminar flow profile. In general, the flow pattern in the viscous sublayer is taken as quasi-static. However, it has been shown experimentally that turbulences in the outer and, specifically, in the buffer zone can affect the flow situation in the viscous sublayer, causing fluctuations of the thickness and the flow pattern in the viscous sublayer.[18–20]

The three regions of the turbulent boundary layer can be mathematically resolved using the normalized (dimensionless) flow velocity, u^+ (eqn [15]), and the normalized (dimensionless) wall distance, y^+ (eqn [16]).

$$u^+ = \frac{u}{u_x} \quad [15]$$

u is mean velocity in turbulent boundary layer (m s^{-1})

$$y^+ = \frac{y u_x}{\nu} \quad [16]$$

where y is distance from the wall (m), $u_x = \sqrt{(\tau_w/\rho)}$ is shear velocity (m s^{-1}), ν, the kinematic viscosity (m^2 s^{-1}), τ_w, the wall shear stress (N m^{-2} or Pa), and ρ, the fluid density (kg m^{-3}).

The velocity distribution in the viscous sublayer ($y^+ < 5$), in the buffer region ($5 < y^+ < 30$–40), and the outer turbulent zone (30–$40 < y^+ < 100$) can be calculated according to the equations given in the semilogarithmic plot in **Figure 4**.[18]

The local flow velocity in the viscous sublayer can be calculated according to eqn [17]:

$$u = \frac{\tau_w}{\mu} y \quad [17]$$

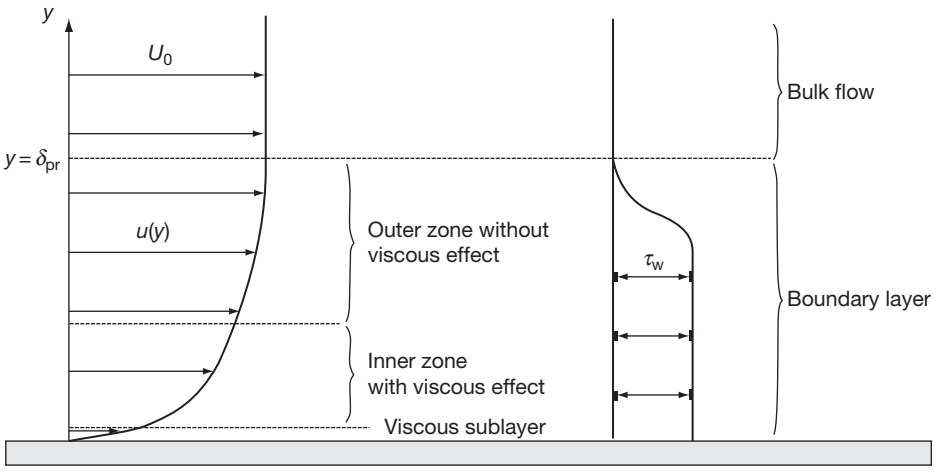


Figure 3 Turbulent boundary layer.

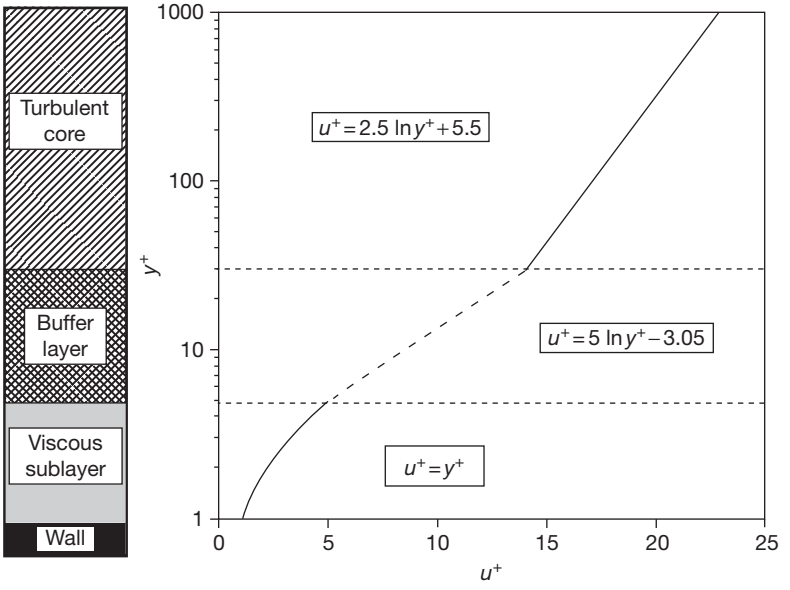


Figure 4 Structure of the turbulent boundary layer and distance-related velocity distribution. Reproduced from Davies, J. T. In *Turbulence Phenomena*; Academic Press: New York, 1972; Chapter 2, pp 121–143.

2.13.2.3 Flow Patterns

The interaction intensity between flowing media and solid walls depends on the number of phases and on the magnitude and relative ratio of the superficial velocities of each phase involved. (The superficial velocity is calculated separately for each phase assuming that no other phase is present). The easiest-to-describe flow pattern is the undisturbed, developed single phase flow. Much more complicated and more dangerous, with respect to flow-induced localized corrosion, (FILC) is the disturbed single phase flow as encountered downstream of weld beads, pipe fittings, valves, orifice plates, at pipe joints, upsets, bends, tees, elbows, or tube inlets of heat exchangers. The hydrodynamic description becomes even more challenging in the case of two-phase flow (liquid–liquid, liquid–gas, liquid–solid) and multiphase flow (liquid–liquid–gas, liquid–liquid–solid, liquid–gas–solid, liquid–liquid–gas–solid) in horizontal, inclined, or vertical flow systems, not to mention additional conditions of surface or flow directional disturbances. **Figures 5–7** give examples of mixed phase flow patterns.[9,21,22]

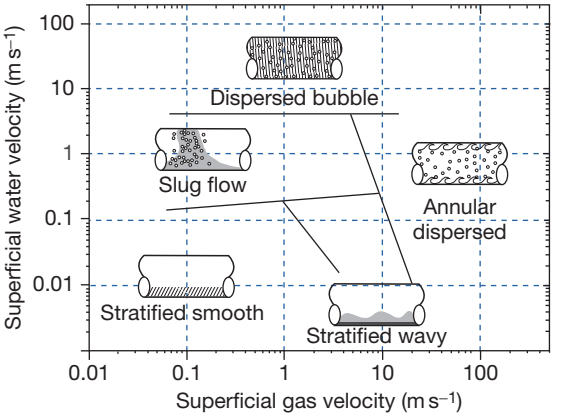


Figure 5 Flow patterns in horizontal liquid–gas pipe flow. Adapted from Lotz, U. *Velocity Effects in Flow Induced Corrosion*, Corrosion'90; NACE International: Houston, TX, 1990; Paper no. 27; Weisman, J. In *Handbook of Fluids in Motion*; Cheremisinoff, N. P., Gupta, R., Eds.; Ann Arbor Science Publishers: Ann Arbor, MI, 1983; pp 409–424; Govier, G. W.; Aziz, K. *The Flow of Complex Mixtures in Pipes*; Van Nostrand Reinhold: New York, 1972; pp 324, 326, 506.

What makes exact hydrodynamic descriptions so difficult is the fact that a multitude of fluid–wall interactions can take place, including mass transfer, heat transfer, fluctuating shear stresses parallel to the surface, fluctuating energy densities (fluctuating pressures) perpendicular to the surface, particle impact (erosion), and near-wall gas bubble collapse (cavitation) (**Figure 8**). Besides these kinds of interactions, the nature of the wall (surface roughness, hardness, ductility, and fracture toughness, presence of protective layers like films, coatings, or scales) plays a significant role. This is the reason why predictive mathematical models in flow-assisted corrosion are available only for relatively simple flow systems.

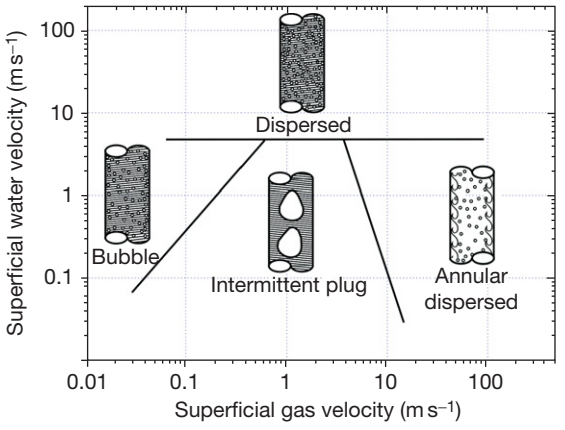


Figure 6 Flow patterns in vertical liquid–gas pipe flow. Adapted from Lotz, U. *Velocity Effects in Flow Induced Corrosion*, Corrosion'90; NACE International: Houston, TX, 1990; Paper no. 27; Weisman, J. In *Handbook of Fluids in Motion*; Cheremisinoff, N. P., Gupta, R., Eds.; Ann Arbor Science Publishers: Ann Arbor, MI, 1983; pp 409–424; Govier, G. W.; Aziz, K. *The Flow of Complex Mixtures in Pipes*; Van Nostrand Reinhold: New York, 1972; pp 324, 326, 506.

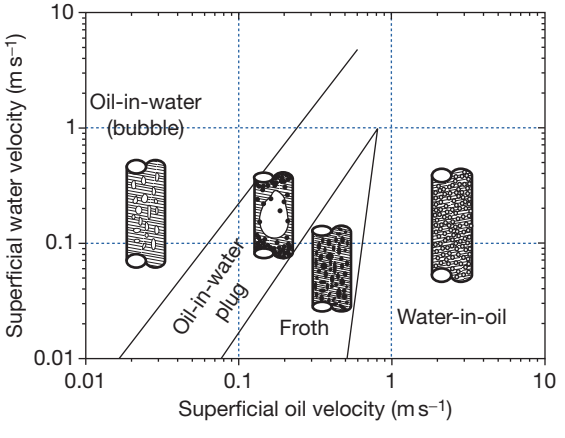


Figure 7 Flow patterns in vertical liquid–liquid (oil–water) pipe flow. Adapted from Lotz, U. *Velocity Effects in Flow Induced Corrosion*, Corrosion'90; NACE International: Houston, TX, 1990; Paper no. 27; Weisman, J. In *Handbook of Fluids in Motion*; Cheremisinoff, N. P., Gupta, R., Eds.; Ann Arbor Science Publishers: Ann Arbor, MI, 1983; pp 409–424; Govier, G. W.; Aziz, K. *The Flow of Complex Mixtures in Pipes*; Van Nostrand Reinhold: New York, 1972; pp 324, 326, 506.

2.13.3 Corrosion Aspects

Flow can enhance uniform corrosion or initiate and propagate localized corrosion. The first case is mostly encountered when the material's surface is bare (i.e., scale free) or is covered with a very porous and thus nonprotective scale. This case will be described as flow-enhanced corrosion (FEC). In contrast the acronym FILC will be used to describe flow-induced localized corrosion generally found under corrosion conditions when more or less protective scales cover the material's surface and are destructed by flow effects.[23–26]

2.13.3.1 Scale-free Surfaces

Under free corrosion conditions, the corrosion rate is flow dependent when the rest potential, $U_{\mathbf{R}}$, is in the potential range of the mass transport (diffusion)-controlled region of the cathodic partial current density–potential curve j^- (**Figure 9**). $U^*_{(\mathrm{Me})}$ and $U^*_{(\mathrm{CA})}$ are the equilibrium potential of the metal–metal ion redox reaction and the cathodic redox reaction, respectively. The corrosion current density, $j_{\mathbf{c}}$, is then equal to the diffusion current density, $j_{\mathbf{D}}$, of the cathodic partial current density–potential curve j^-, and $j_{\mathbf{D}}$ depends on the flow, according to the Fick's law (eqn [18]).[27]

$$j_{\mathrm{D}} = \frac{zFD\Delta c}{d_{\mathrm{N}}} \qquad [18]$$

where D is the diffusion coefficient of the cathodically active species ($\mathrm{m^2\,s^{-1}}$), z, the number of charges exchanged in the cathodic reaction, F, Faraday's constant (96 485 C $\mathrm{mol^{-1}}$), $d_{\mathbf{N}}$, the Nernst diffusion layer thickness (m), and Δc, the concentration gradient of the cathodically active species between the bulk solution (c_{bulk}) and the electrode surface (c^*). Under diffusion controlled conditions, c^* is equal to zero (**Figure 10**).

In laminar flow regimes, $d_{\mathbf{N}}$ ranges between 1 and 100 µm and depends on the bulk flow velocity, u_0, according to eqn [19], where n can take values between 0.5 and 1.[17]

$$d_{\mathrm{N}} = \frac{1}{u_0^n} \qquad [19]$$

The hydrodynamic layers next to the metal surface will exhibit a transition from pure diffusion and migration immediately next to the surface to a convection-dominated region further away from the surface. It is mathematically convenient to consider a simplified model of the interface that consists of a layer in which only diffusion and migration occur, known as the Nernst diffusion layer, in direct contact with bulk solution, which has a constant composition. The thickness, $d_{\mathbf{N}}$, of the Nernst diffusion layer is

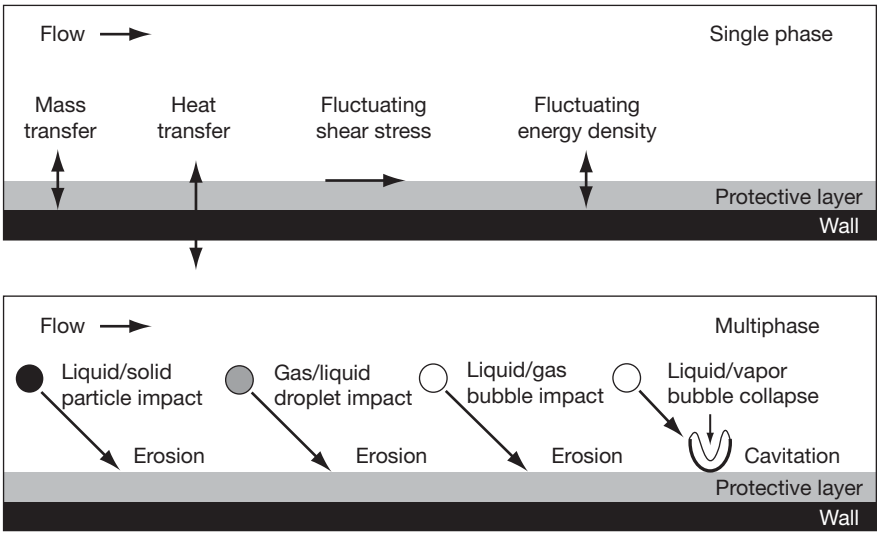


Figure 8 Fluid–wall interactions in flow systems. Reproduced from Heitz, E. In *Flow-Induced Corrosion: Fundamental Studies and Industry Experience*; Kennelley, K. J., Hausler, R. H., Silverman, D. C., Eds.; NACE International: Houston, TX, 1991; pp 1:1–1:29.

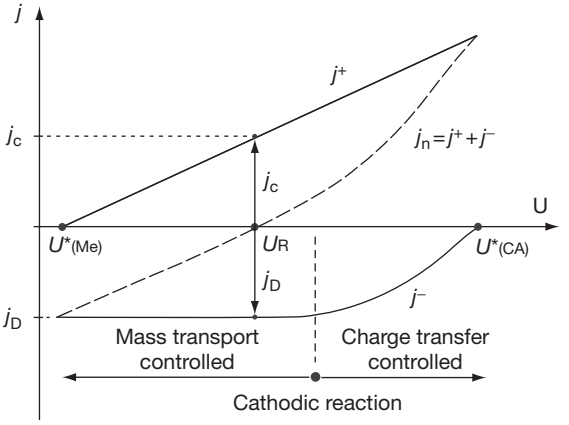


Figure 9 Current density–potential plot for diffusion controlled metal corrosion.

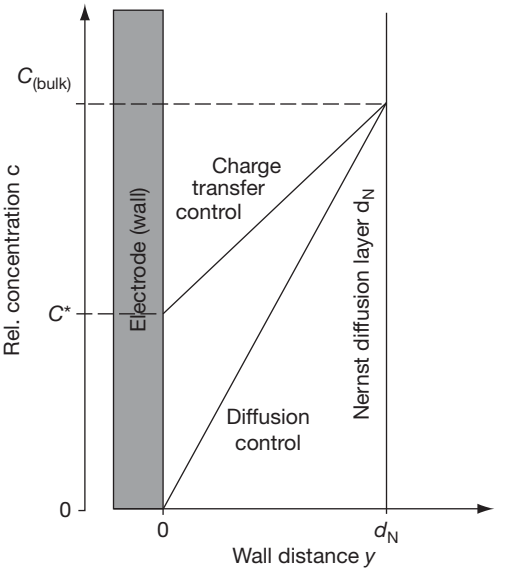


Figure 10 Concentration gradients in Nernst diffusion layer.

proportional to the thickness of the Prandtl boundary layer, δ_{Pr} (eqn [20]).

$$d_N \sim \left(\frac{D}{\nu}\right)^{1/3} \delta_{Pr} \qquad [20]$$

For aqueous solutions with kinematic viscosities, ν, in the order of $10^{-2}\,cm^2\,s^{-1}$ and diffusion coefficients in the order of $10^{-5}\,cm^2\,s^{-1}$, d_N ranges in the order of 10% of the Prandtl boundary layer.

The ratio between the kinematic viscosity and the diffusion coefficient is also called Schmidt number, *Sc* (eqn [21]).[17]

$$Sc = \frac{\nu}{D} = \left(\frac{u_0 l}{D}\right) \Big/ \left(\frac{u_0 l}{\nu}\right) = \frac{Pe}{Re} \qquad [21]$$

Pe is the Peclet number ($=u_0 l/D$). When the flow becomes turbulent (high *Re*, rough surfaces), the transport conditions become more complicated and can be expressed only semiquantitatively. For *Sc* in the order of 10^3, the diffusion boundary layer thickness is of the order of 1/6 of the thickness of the viscous sublayer. Thus, the diffusion boundary layer reacts directly to changes in the thickness of the viscous sublayer.

For turbulent flow, a general equation exists for mass transport correlations at different flow patterns, based on the dimensionless parameters *Re*, *Sc*, and *Sh* (eqn [22]).[28]

$$Sh = aSc^{b}\mathrm{Re}^{c} \quad [22]$$

The Sherwood number, *Sh*, describes the ratio of total mass transport to mass transport by molecular diffusion (eqn [23]).[29]

$$Sh = \frac{kL}{D} \quad [23]$$

where $k = D/d_{\mathrm{N}}$ is the mass transport coefficient ($\mathrm{m\,s^{-1}}$), and *L*, the characteristic length (m). The constants *a*, *b* and *c* in eqn [22] are given in **Table 1** for different flow devices.

2.13.3.2 Scale-Covered Surfaces

At scale-covered metal surfaces, molecular pore diffusion becomes the rate-determining step (**Figure 11**). This depends, of course, on the pore size and density. In general, the corrosion rate becomes nearly independent of flow under the assumption that the thickness or the existence of the scale is not influenced by the flowing medium, for example, due to enhanced dissolution or cracking and subsequent spalling. The major concentration gradient of electro-active species (transported to or from the metal surface at the bottom of the pores) occurs in the pores.[9,30] The flow intensity at the scale-covered surface will influence only the convective diffusion of electro-active species in the Prandtl boundary layer to the entrance of the pores, which generally results in only a small concentration gradient (**Figure 11**).

The mass transport in the laminar Prandtl boundary layer, d_{Pr}, and in the pores of the scale can be described by Fick's law (eqns [24] and [25]):

$$m_{\mathrm{Pr}} = D_{\mathrm{Pr}}\frac{c_{\mathrm{bulk}} - c_{\mathrm{s}}}{d_{\mathrm{Pr}}}\,(\text{Diffusion in Prandtl layer}) \quad [24]$$

$$m_{\mathrm{s}} = D_{\mathrm{s}}\frac{c_{\mathrm{s}} - c_{0}}{d_{\mathrm{s}}}\,(\text{Pore diffusion}) \quad [25]$$

Table 1 Mass transport relationships for important flow systems

Flow system	*Power law* $Sh = aSc^{b}Re^{c}$	*Range of validity*	*Characteristic length and Re*
Free rotated disc			
Laminar flow	$0.60Sc^{1/3}Re^{1/2}$	$10^2 < Re < 10^5$	$L = r, Re = \frac{\omega r^2}{\nu}$
Turbulent flow	$0.011Sc^{1/3}Re^{0.87}$	$Re > 10^6$	
Rotated disc in front of wall			
Laminar flow	$17.3Sc^{1/3}Re^{0.34}$	$9 \times 10^4 < Re < 6 \times 10^5$	$L = r, Re = \frac{\omega r^2}{\nu}$
Turbulent flow	$0.05Sc^{1/3}Re^{0.78}$	$Re > 6 \times 10^5$	
Free rotated cylinder			
Turbulent flow	$0.079Sc^{0.35}Re^{0.7}$	$10^2 << Re < 4 \times 10^5$	$L = r, Re = \frac{\omega r^2}{\nu}$
Coaxial cylinder	$0.0027Sc^{1/3}Re$	$Re < 2.7 \times 10^5$	$L = r, Re = \frac{\omega r^2}{\nu}$
Flat plate in channel and tube flow			
Laminar flow	$2.54Sc^{1/3}Re(d_e/l)^{1/3}$	$Re < 2300$	$L = d_{\mathrm{e}}, Re = \frac{d_e u}{\nu}$
Turbulent flow	$0.079Sc^{0.35}Re^{0.7}$	Re <2300	$d_e = \frac{4 \times \text{cross section are}}{\text{perimeter}}$
Flat plate in free area			
Laminar flow	$0.34Sc^{1/3}Re^{1/2}$	$Re < 5 \times 10^5$	$L = l, Re = \frac{lu}{\nu}$
Rotated cage			
Turbulent	NA	NA	$Re_{\mathrm{RC}} = \frac{\omega r_{RC}^2}{\nu}$
Impinging Jet			
Region A (laminar)	$1.51Re^{0.5}Sc^{0.33}\left(\frac{H}{d}\right)^{-0.054}$	$Re < 2000$	$L = d, Re = \frac{du}{\nu}$
Region B (transition)	$1.12Re^{0.5}Sc^{0.33}\left(\frac{H}{d}\right)^{-0.054}$	$4000 < Re < 16\,000$	*d*: nozzle diameter *u*: flow velocity in nozzle

Source: DIN 50920, Corrosion of Metals, Part I. *Corrosion Testing in Flowing Liquids; General*, (in German); Beuth Verlag: Berlin, Germany,1985; Papavinasam, S.; Revie, R.W.; Attard, M.; Demoz, A.; Sun, H.; Donini, J.C.; Michaelian, K. In *Inhibitor Selection for Internal Corrosion Control of Pipelines: 1. Laboratory Methodologies*, Corrosion'99; NACE International: Houston, TX, 1999; Paper no. 1; Poulson, B. *Corros. Sci*. **1983**, *23*(4), 391–430; Poulson, B.; Robinson, R. *Corros. Sci*. **1986**, *26*(4), 265–280.

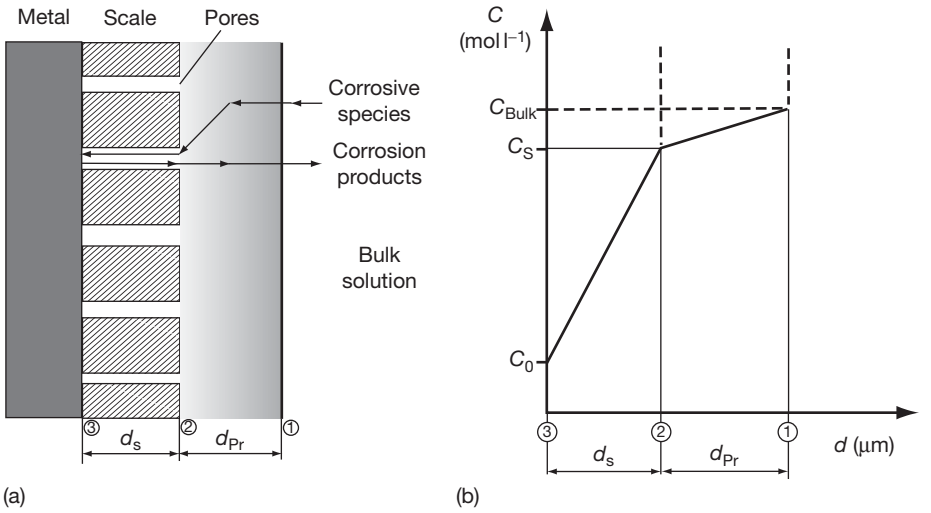


Figure 11 (a) Corrosion process on scale-covered metal surfaces and (b) concentration gradient of electro-active species from the bulk solution to the metal surface.

where m_{Pr} is the mass transport rate in the Prandtl boundary layer (mol m^{-2} s^{-1}); m_S is the mass transport rate in the pores (mol m^{-2} s^{-1}); D is the diffusion coefficient of the electro-active species (i) under convective diffusion in the Prandtl layer (D_{Pr}) and (ii) in the pores of the scale (D_S); c, the concentration (mol m^{-3}) of the electro-active species (i) in the bulk (c_{bulk}), (ii) at the entrance of the pores (c_S), and (iii) at the metal surface (c_0); and d_S is the thickness of the layer (m).

The kinetics of corrosion processes at scale-covered materials results from a combination of the following simplified partial processes:

1. Transport of corrosive species to the scale (eqn [26])[9]:

$$v_1 = k_1(c_{bulk} - c_s) \qquad [26]$$

This process is controlled by the combination of convection and diffusion. The coefficient k_1 is flow-dependent.

2. Transport of electro-active species through the pores of the scale (eqn [27]):

$$v_2 = k_2(c_{bulk} - c_s) \qquad [27]$$

This process is mainly controlled by diffusion. The coefficient k_2 is strongly dependent on the pore structure of the scale but is not flow-dependent.

3. Interface reaction (eqn [28]):

$$v_3 = k_3 c_0 \qquad [28]$$

The coefficient k_3 is potential-dependent, but not flow-dependent. Under conditions of stationary mass transport rate, the following equations are valid (eqns [29] and [30]):

$$c_{bulk} = (c_{bulk} - c_s) + (c_s - c_0) + c_0 \qquad [29]$$

or

$$v = v_1 + v_2 + v_3 \qquad [30]$$

Regarding eqns [26]–[28] and the assumptions in eqns [29] and [30], the following relationships will result (eqns [31] and [32]):

$$c_{bulk} = \frac{v}{k_1} + \frac{v}{k_2} + \frac{v}{k_3} \qquad [31]$$

with

$$v = v_1 + v_2 + v_3 = \frac{c_{bulk}}{(1/k_1) + (1/k_2) + (1/k_3)} \qquad [32]$$

where v is the corrosion rate (g m^{-2} h); k_1, k_2, k_3 are coefficients, where only k_1 is flow-dependent. k_2 and k_3 depend on the pore structure (size, density). As the scale is growing and the porosity is decreasing, that is, $k_2 \rightarrow 0$, the flow velocity effect will decrease.

The rate of scale formation, Φ_{scale}, is proportional to the mass conversion rate (corrosion rate) according to eqn [33]:

$$\Phi_{scale} = K_1 v \qquad [33]$$

The rate of scale destruction is influenced by the flow intensity, which is expressed mostly in terms of wall shear stress, τ_w (eqn [34]):

$$\Phi_{\text{destruct}} = K_2 \tau_w^a \quad [34]$$

The proportionality factor, K_2, is mainly influenced by the binding forces between the particles (e.g., crystallites) forming the scale and the adhesion between scale and base material. The exponent, a, relates to the flow pattern and the contribution of erosion and cavitation. For a dynamic equilibrium ($\Phi_{\text{scale}} = \Phi_{\text{destruct}}$), it follows that

$$K_1 v = K_2 \tau_w^a \quad [35]$$

$$v = \frac{K_2}{K_1} \tau_w^a = K_3 \tau_w^a \quad [36]$$

These equations (eqns [35] and [36]) describe the relation between the mass conversion rate (corrosion rate) and the flow intensity, expressed in terms of wall shear stress, τ_w. The proportionality factor, K_3, describes the density of the scale, its porosity, the binding forces between the particles forming the scale, and finally its fracture stress.[30–32]

2.13.3.2.1 Scale destruction

For each corrosion system, a critical flow intensity I_{crit} can be identified above which the scale is destroyed and its protectiveness is lost (**Figure 12**). The corrosion rate increases abruptly and finally becomes purely mass transport controlled according to the boundary conditions of a scale-free system.[25,30]

The classic mechanism of scale destruction above the critical flow intensity is sketched in **Figure 13**. It is assumed that near-wall microturbulences cause local thinning of the scale, which locally becomes porous and finally is completely carried away by the flow. Once a scale-free, unprotected local surface area has formed, the local flow intensity prevents the reformation of the protective scale, and local penetrations start producing the characteristic pattern of FILC. While it is possible that the scale is destroyed particle by particle and crystal by crystal during interaction with the flowing media, more recent observations revealed that scales of oxides, carbonates, phosphates, and sulfides rather crack and spall.[33,34] **Figures 14 and 15** give examples of iron carbonate and iron sulfide scales from which parts of the scales experienced cracking and spalling. The fracture stresses of such scales have been measured[33–35] and range in the order of 75–720 MPa for iron carbonate scale from CO_2 corrosion of steel and 50–150 MPa for iron sulfides from H_2S corrosion of steel. The fracture stresses are influenced by the porosity and thickness of the scale and increase dramatically when formed in the presence of effective corrosion inhibitors.[36,37] **Figure 16** shows the failure mode diagram of corrosion-produced iron carbonate scales in the absence and presence of corrosion inhibitor. The corrosion inhibitor decreases the scale thickness, and hence, increases the critical fracture strain. The data were used to evaluate the surface fracture stress of iron carbonates (**Figure 17**).[34,38]

In view of the magnitude of the fracture stresses and adhesion strengths measured for corrosion product scales,[34,35,38] which number in the MPa range (**Table 2**), it appears difficult to understand that wall shear stresses in the range of 10–100 Pa (technical systems) or at maximum up to 10^3–10^4 Pa (extreme local flow intensities) can cause cracking and spalling of scales.

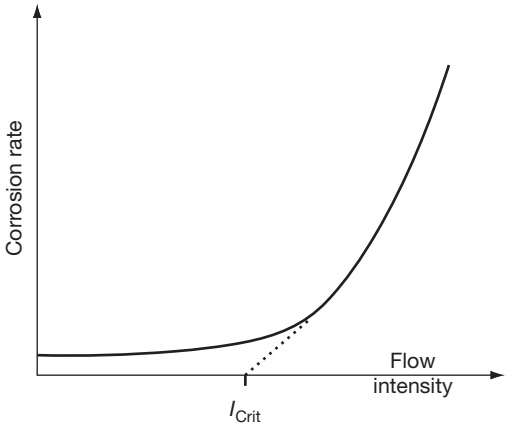


Figure 12 Dependence of corrosion rate on flow intensity.

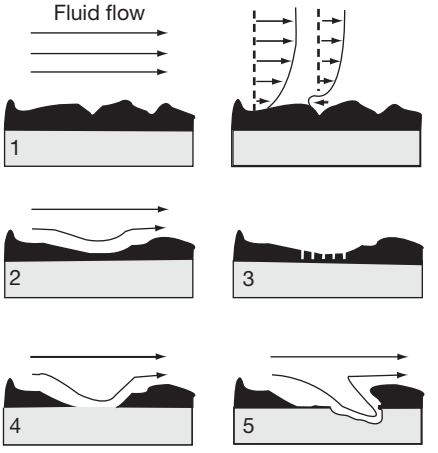


Figure 13 Classic mechanism for initiation of FILC. Reproduced from Pini, G.; Weber, J. *Technische Rundschau Sulzer* (in German) **1979**, *2*, 69.

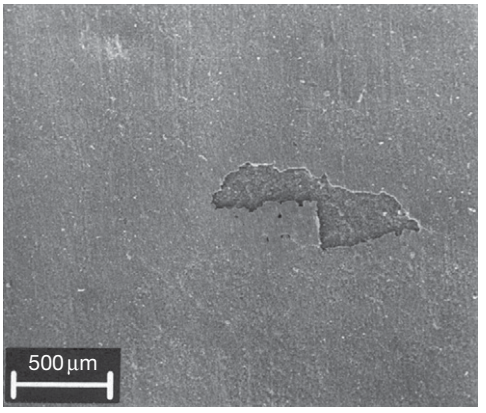


Figure 14 Cracking and spalling of iron carbonate scale during flow experiments. Reproduced from Schmitt, G.; Mueller, M.; Papenfuss, M.; Strobel-Effertz, E. In *Understanding Localized CO_2 Corrosion of Carbon Steel from Physical Properties of Iron Carbonate Scales*, Corrosion'99; NACE International: Houston, TX, 1999; Paper no. 38.

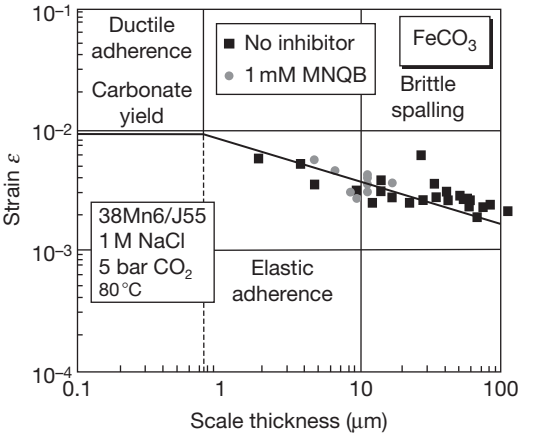


Figure 16 Failure mode diagram of $FeCO_3$ (data from 4-point-bent-tests). Reproduced from Schmitt, G.; Mueller, M.; Papenfuss, M.; Strobel-Effertz, E. In *Understanding Localized CO_2 Corrosion of Carbon Steel from Physical Properties of Iron Carbonate Scales*, Corrosion'99; NACE International: Houston, TX, 1999; Paper no. 38.

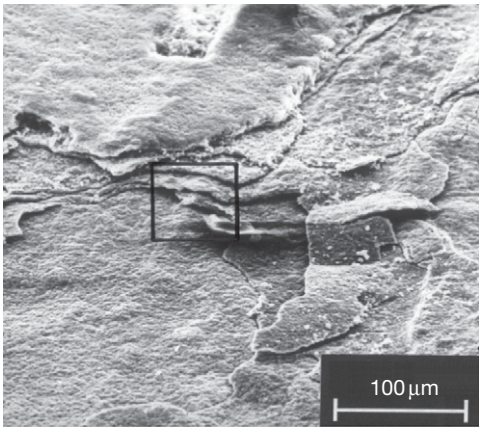


Figure 15 Cracking and spalling of iron sulfide scale during flow experiments. Reproduced from Schmitt, G.; Bosch, C.; Pankoke, U.; Bruckhoff, W.; Siegmund, G. In *Evaluation of Critical Flow Intensities for FILC in Sour Gas Production*, Corrosion'98, NACE International: Houston, TX, 1998; Paper no. 46.

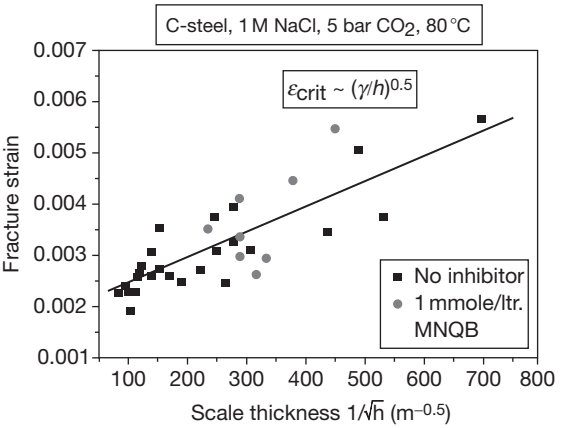


Figure 17 Evaluation of surface fracture energy γ of iron carbonate scales from fracture strain measurements. Reproduced from Schmitt, G.; Bosch, C.; Mueller, M.; Siegmund, G. In *A Probabilistic Model for Flow Induced Localized Corrosion*, Corrosion'2000; NACE International: Houston, TX, 2000; Paper no. 49.

As it is clearly proven that flow-induced fluid dynamic forces can destroy protective layers, it is necessary to define such forces, which obviously act not parallel to the surface (as the wall shear stress) but rather perpendicular to the surface (**Figure 18**). This hydrodynamic force perpendicular to the wall shall be called hydrodynamic energy density, E_D (Pa). For the initiation of FILC, the hydrodynamic energy density must be, at least locally, equal or greater than the fracture stress of the scales, σ_B (Pa) (eqn [37]).

$$E_D \geq \sigma_B \qquad [37]$$

While the wall shear stress, τ_w (Pa), has been clearly defined (eqn [14]) and measured for a long time, the hydrodynamic energy density, E_D (Pa), has been defined only recently[39,40] during investigations,

aiming at identifying the interaction energy of singly near-wall turbulence elements via microelectrochemical flow measurements.

This approach used the Léveque equation[41](eqn [38]), which correlates the wall shear stress with the diffusion current density of a tracer redox system at a microelectrode (diameter 15–200 µm) under defined flow conditions.

$$\tau_w = k[j_D]^3 \qquad [38]$$

Table 2 Comparison of system forces

Forces	*Magnitude (Pa)*
Wall shear stress	
Technical systems	10^0–10^2
Extreme flow intensities	10^3–10^4
Adherence of scales	10^6–10^7
Fracture stress of scales	
High porosity	10^7–10^8
Low porosity	10^8–10^9

Source: Schmitt, G.; Mueller, M.; Papenfuss, M.; Strobel-Effertz, E. In *Understanding Localized CO_2 Corrosion of Carbon Steel from Physical Properties of Iron Carbonate Scales*, Corrosion'99; NACE International: Houston, TX, 1999; Paper no. 38; Schmitt, G.; Bosch, C.; Pankoke, U.; Bruckhoff, W.; Siegmund, G. In *Evaluation of Critical Flow Intensities for FILC in Sour Gas Production*, Corrosion'98; NACE International: Houston, TX, 1998; Paper no. 46; Schmitt, G.; Bosch, C.; Mueller, M.; Siegmund, G. In *A Probabilistic Model for Flow Induced Localized Corrosion*, Corrosion'2000; NACE International: Houston, TX, 2000; Paper no. 49; Schmitt, G.; Mueller, M. In *Critical Wall Shear Stresses in CO_2 Corrosion of Carbon Steel*, Corrosion'99; NACE International: Houston, TX, 1999; Paper no. 44.

with

$$k = \frac{\mu}{0.81^3 z^3 c_0 l^2 b^3 D^2} \qquad [39]$$

where k is a constant depending on the dynamic viscosity, μ (Pa s), of the electro-active species, the number z of electrons exchanged in the tracer redox system, the concentration, c_0 (mol m^{-3}), and diffusion coefficient, D (m^2 s^{-1}), of the redox system, and the geometry of the microelectrode: the length, l (m), and the width, b (m).

The measured flow dependent parameter is j_D (A m^{-2}), the diffusion current density of the mass transport-controlled redox reaction. Most commonly, the hexacyano-ferrate(II/III) redox system (eqn [40]) is used as tracer redox system in such measurements, because of its simple one-electron transfer and known system parameters.

$$[Fe(CN)_6]^{3-} + e^- \rightarrow [Fe(CN)_6]^{4-} \qquad [40]$$

Under turbulent flow conditions, the diffusion-controlled current densities, $j_D(t)$, measured at microelectrodes are fluctuating with time (**Figure 19**) due to fluctuations in the Nernst diffusion layer, which is caused by turbulence elements disturbing the viscous sublayer of the turbulent boundary layer. In order to evaluate the information contained in the noise of the diffusion current densities, the noise data (after subtraction of the DC part) have been subjected to a Wavelet transform, which yields the signal amplitudes in the frequency and time domain (**Figure 20**).

For each time resolution, an amplitude/frequency spectrum is obtained. It appears that at different times

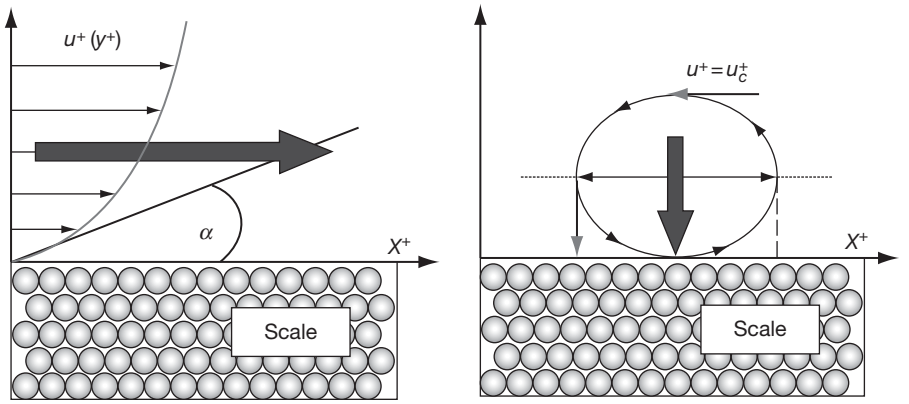


Figure 18 Interaction of hydrodynamic forces with scales.

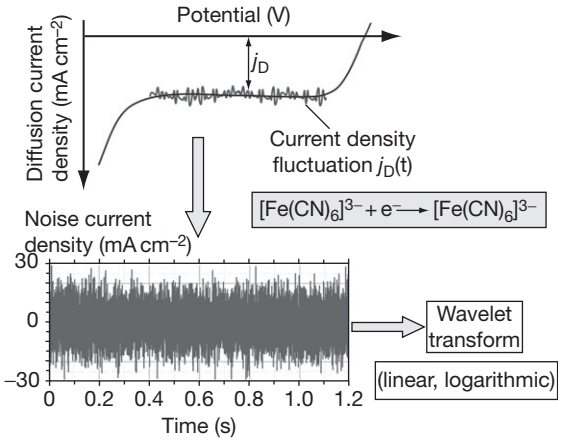


Figure 19 Evaluation of diffusion controlled current density noises.

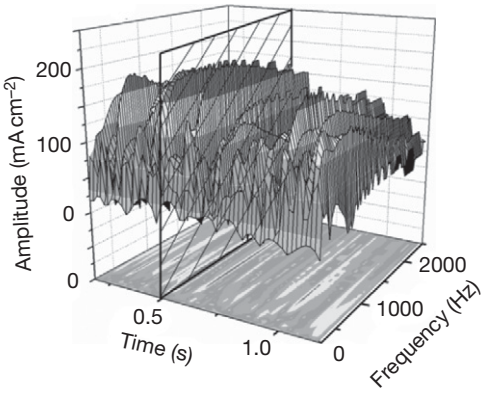


Figure 20 Linear wavelet transform of current noise signal.

different amplitude fluctuations are encountered. Furthermore, no discrete frequency spectra are observed even at very high time resolution. Based on this, broad frequency range simulations have been performed with the aim of investigating whether the frequency spectra obtained can provoke high-intensity singularities, which significantly exceed the general noise level. The simulation was performed according to the algorithm expressed in eqn [41].[39,40]

$$j_D(t) = \sum_{\omega} A(\omega)\sin(\omega t)_{t_0} \qquad [41]$$

At an arbitrarily chosen time (e.g., at the time 0.5 s, Figure 20), the corresponding amplitude/frequency spectrum is obtained (Figure 21). If all amplitudes at all frequencies in the selected frequency range are added for this specific time according to eqn [41], a singularity arises that represents the same singular

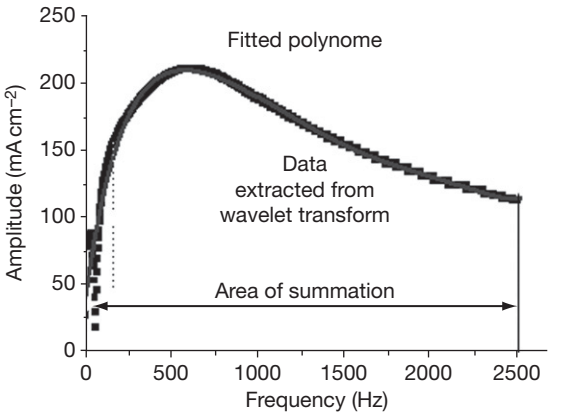


Figure 21 Polynomial fitting (fourth grade) for time cut in Wavelet data.

situation when oceanic 'freak waves' (Rogue Waves)[42,43] are formed. In this case, a group of surface waves with different wave heights (amplitudes) travels in the same direction, but with different velocities. When all waves of this group are phased in for a single moment, that is, when they meet at the same time, a single wave is formed that piles up orders of magnitude higher than the average height of the group of waves. This singularity contains the power of all waves included, and the volume elements of this singular wave contain the highest energy density possible under these conditions.

In order to apply the same algorithm on the amplitude/frequency profile chosen in the example according to Figure 21, the amplitude/frequency function is fitted with a polynomial equation, which is inserted in eqn [41] for the amplitude A. The summing-up range shall include the relevant frequency range (e.g., 1–2500 Hz; see Figure 21). Within a unit time of 1 s, the summing-up procedure will be repeated in very small time steps, for example, 1/250 000 s.[39,40]

The results of this calculation are presented in Figure 22. A singularity is found in the time range close to $t = 1$ s, the amplitude of which exceeds the general noise level by orders of magnitude. Taking into account that the amplitudes represent diffusion current densities, the maximum in Figure 22 represents a fictive maximum diffusion current density in the order of 500.000 mA cm^{-2}. This phenomenon repeats at all unit time intervals.

The physical meaning of this maximum current density (called 'freak current density') is that, at the time of this 'freak' event, volume elements of the 'freak wave' in a near-wall turbulence element close

to the viscous sublayer is accelerated perpendicular toward the solid surface (**Figure 23**), reaching through the viscous sublayer down to the solid. If a microelectrode is present on the surface, the viscous sublayer, and hence, the Nernst diffusion layer is deformed for a very short time (in the range of milliseconds) to thicknesses that allow extreme high mass transport rates.

The energy density, E_D, of a freak wave volume element accelerated toward the surface can be expressed according to classic wave dynamics with eqn [42].[44]

$$w = \frac{dE}{dV} = \frac{1}{2}\rho A^2 \omega^2 \qquad [42]$$

where w is energy density (Pa), ρ is density (kg m^{-3}), A is wave amplitude (m), and ω is wave frequency (s^{-1}). Both energy density and wall shear stress have the unit Pascal. It is, therefore, assumed that the same relation used in the Léveque equation can be used to calculate the 'freak' energy density, E_D^F, from the freak current density, j_D^F (eqn [43]):

$$w_F = \tau_F = E_D^F = k_{red/ox}[j_D^F(t)]^3 \qquad [43]$$

or by using eqn [41]:

$$\tau_F = E_D^F = k_{red/ox}\left(\sum_{\omega} A(\omega)\sin(\omega t_F)\right)^3 \qquad [44]$$

$k_{red/ox}$ is a calibration factor which relates to the redox system used for measuring the diffusion-controlled current noise fluctuations. For the hexacyanoferrate redox system, this calibration factor has been evaluated to[45]:

$$k_{red/ox} = 2.14 \times 10^{-8} \quad (\mathrm{Nm^4A^{-3}}) \qquad [45]$$

A freak amplitude of 500 000 mA cm^{-2} (**Figure 22**) would then amount to a freak energy density (FED) of ~3 GPa, which is by far enough to crack any protective scale. For comparison, the fracture stress of iron carbonate single crystals have been estimated to range in the order of 1.4 GPa, while fracture stresses of iron carbonate layers from CO_2 corrosion of steel have been measured to range in the order of 0.075–0.72 GPa.[34]

Thus, the critical flow intensity, which can cause cracking and subsequent spalling of protective scales, and therefore, will initiate FILC that can be determined with eqn [43] using mass transfer-controlled current density noise data from flush-mounted microelectrodes with subsequent Wavelet transform treatment.

Based on this FED model, the effect of additives (inhibitors, drag reducers/flow improvers) on the prevention of FILC can easily be understood. Assuming that the critical fracture stress of an iron carbonate scale is 200 MPa (**Figure 24**), a FED of about 3000 MPa encountered under given flow conditions in a CO_2 corrosion system is by far enough to crack the scale. Addition of 1 mM C12-quat(PEGmod) (polyethylenglycol-modified C12-quaternary alkylammonium halide) reduces the FED significantly

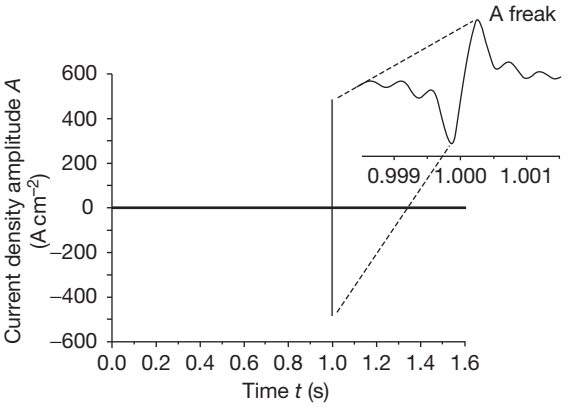


Figure 22 Simulation of the noise current density based on Wavelet transform data. Phasing-in of amplitudes in the unit time region. Reproduced from Schmitt, G.; Werner, C.; Bakalli, M. In *Evaluation of Local Energy Densities in Disturbed Flow: A New Approach to Characterize Inhibitor Efficiencies to Mitigate Erosion Corrosion*, Proceedings of European Conference, EUROCORR'2004, Nice, France, Sept 12–16, 2004; European Federation of Corrosion, The Institute of Materials, London, 2004.

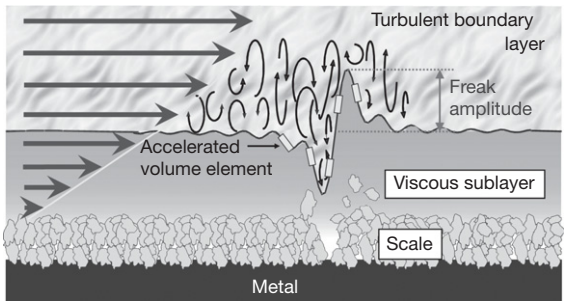


Figure 23 Model of freak wave impact on protective scales. Reproduced from Schmitt, G.; Werner, C.; Bakalli, M. In *Evaluation of Local Energy Densities in Disturbed Flow: A New Approach to Characterize Inhibitor Efficiencies to Mitigate Erosion Corrosion*, Proceedings of European Conference, EUROCORR'2004, Nice, France, Sept 12–16, 2004; European Federation of Corrosion, The Institute of Materials, London, 2004.

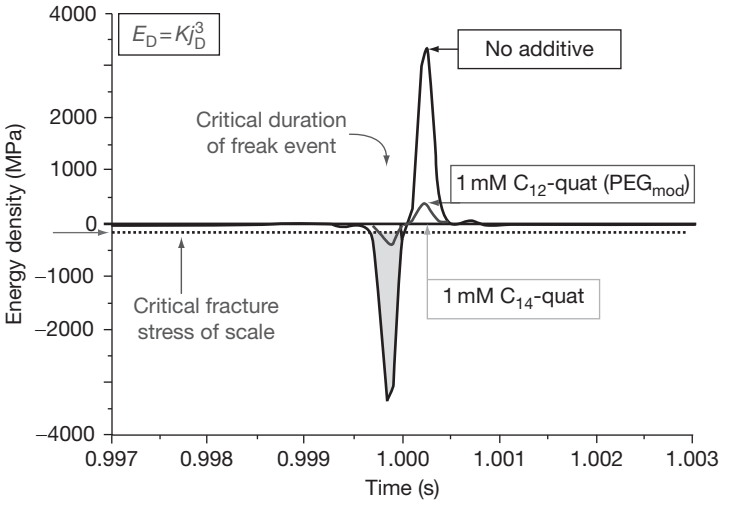


Figure 24 Effect of additives on freak energy densities. Reproduced from Schmitt, G.; Werner, C.; Bakalli, M. In *Fluid Mechanical Interactions of Turbulent Flowing Liquids with the Wall – Revisited with a New Electrochemical Tool*, Corrosion'2005; NACE International: Houston, TX, 2005; Paper no. 344.

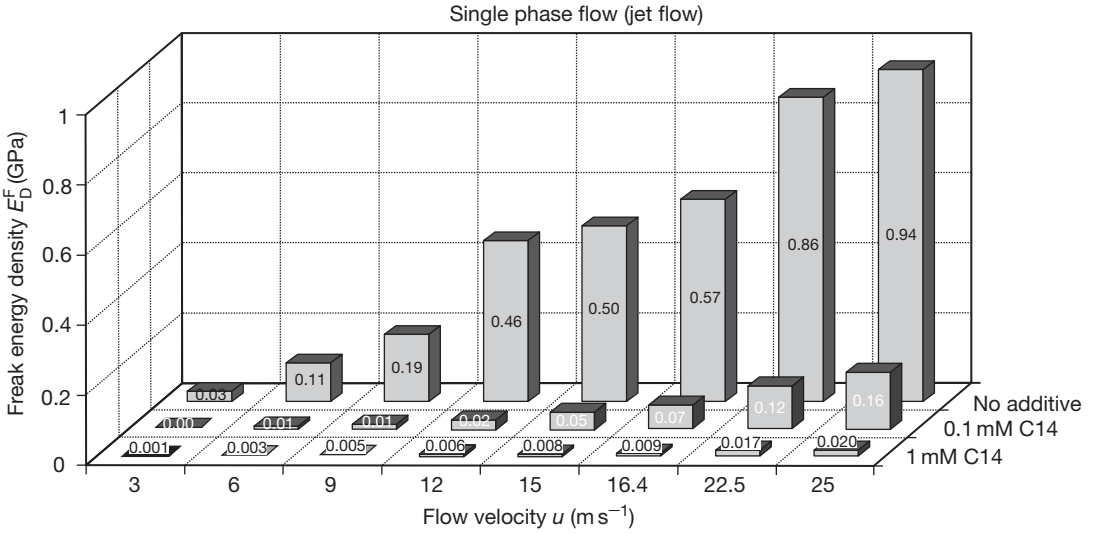


Figure 25 Effect of flow velocity and C14-Quat concentration on maximum (freak) energy density in single phase flow. Reproduced from Schmitt, G.; Bakalli, M. In *Quantification of Maximum Energy Densities Between Flowing Liquids and Solid Walls – Effect of Flow Improving Additives*, Proceedings of European Conference, EUROCORR'2006, Maastricht, The Netherlands, Sept. 24–28, 2006; European Federation of Corrosion, The Institute of Materials: London, 2006.

under otherwise constant conditions, and cannot, however, reduce it below the fracture stress of the scale. Only with 1 mM C14-quat (tetradecyltrimethylammonium bromide) could the freak energy density be reduced significantly below the fracture stresses of the scale, and thus, the onset of FILC be prevented.[39,40]

The effect of C14-Quat concentrations at different jet nozzle velocities on the maximum (freak) energy densities encountered in aqueous liquids with kinematic viscosities in the order of 1 mPa s is very significant and consistent and supports the FED approach.

2.13.3.2.2 Corrosion intensity after scale destruction

After flow-induced local destruction of protective scales, layers, or films, the surface of the base material

is attacked like a bare metal surface, depending on the environmental conditions, that is, depending on the relative contributions of the corrosion and the erosion part to the rate of materials loss. These contributions can be estimated from the exponent a in eqn [46], which aims to relate the mass loss rate, w_{er}, in erosion corrosion to the bulk flow velocity, u_b.[4]

$$w_{er} \propto u_b^a \qquad [46]$$

For mass control of the corrosion reaction, the exponent a is 0.33 for laminar flow and 0.8–1.0 for developed turbulent flow. It is zero for charge transfer (activation) control and somewhere between 0 and 1 under mixed control. The exponent a is 1 also for materials with passive films under activation/repassivation control. If erosive effects are contributing, a is >1, for example, 2–3 in the case of solid particle impingement or even 5–8 in the case of cavitation attack or liquid droplet impingement in high-speed gas flow.

While the bulk flow velocity in full bore flow or superficial flow velocities of the phases in question are readily measured and controlled, it is appreciated that, specifically, in disturbed flow, the flow characteristics in the direct vicinity of the wall are much more important than the bulk flow situation. Therefore, the trend is increasing to evaluate at least the existing wall shear stresses. However, bulk parameters are still used to characterize a given flow system.

2.13.3.3 Tools for Investigating Flow-Assisted Corrosion

A great variety of laboratory tools is available for investigating flow-assisted corrosion.[46] Lab tools used include (**Figures 26–28**)[28,47–50]:

- rotated probes;
- pipes and channels; and
- jet impingement.

Rotated probes with axial symmetry include the free-rotated disc, the disc in front of a wall, the free-rotated cylinder, and the coaxial cylinder (**Figure 26**).[28] Another set-up of rotated probes is the rotated cage (**Figure 27**), which contains coupons arranged between holders in such a way that the outer surface of the coupons forms parts of the surfaces of a rotating cylinder.[48]

The industrially prevalent type of flow exists in pipes. Therefore, flow studies with tubes/pipes and channels (**Figure 26**) are performed in those cases where a direct correlation with real flow situations in pipes is needed.

A very powerful tool is the impinging jet (**Figure 28**), which is specifically suitable to investigate susceptibilities of materials to FILC under severe flow intensities and to study the efficiency of corrosion mitigation, for example, the application of corrosion inhibitors under such conditions. A newly

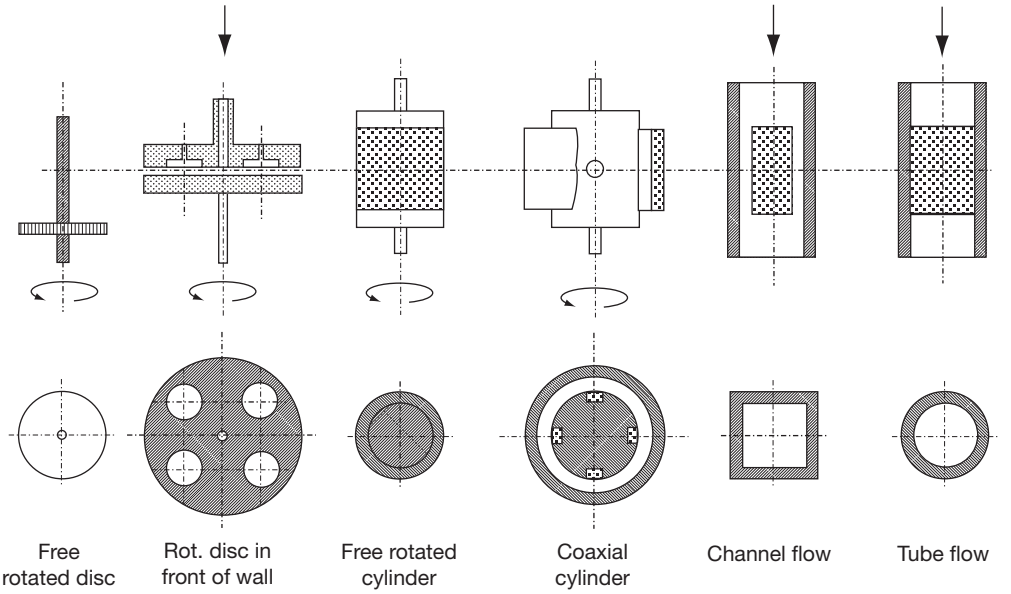


Figure 26 Standardized tools to investigate flow-assisted corrosion. Reproduced from DIN 50920, Corrosion of Metals, Part I; In *Corrosion Testing in Flowing Liquids; General* (in German); Beuth Verlag: Berlin, Germany, 1985.

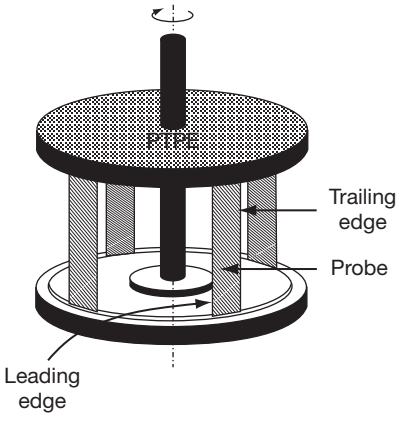


Figure 27 Rotated cage arrangement. Reproduced from Schmitt, G.; Bruckhoff, W. In *Relevance of Laboratory Experiments for Investigation and Mitigation of Flow Induced Corrosion in Gas Production*, Corrosion'88; NACE International: Houston, TX, 1988; Paper no. 357.

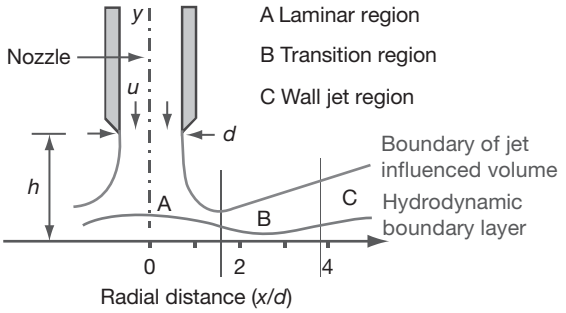


Figure 28 Hydrodynamic flow pattern at the submerged impinging jet. Reproduced from Dawson, J. L.; Shih, C. C. In Corrosion'90; NACE International: Houston, TX, 1990; Paper no. 21.

developed modification is the gas-pulsed impinging jet (**Figure 29**), which can simulate the damaging effect of slug flow without using a slug flow loop.[46]

Mass transport relationships available for the different experimental tools are listed in **Table 1**. The merits and limitations of each tool are discussed in the literature.[46] Here, the authors clearly differentiate between corrosion experiments (where corrosion rates are directly calculated from time and surface related mass loss measurements and where localized corrosion can be observed by visual inspection of the corroded specimens) and electrochemical experiments, which may give information on corrosion rates as well, but should rather be used for mechanistic investigations, specifically to determine the nature of rate-determining steps in corrosion mechanisms.

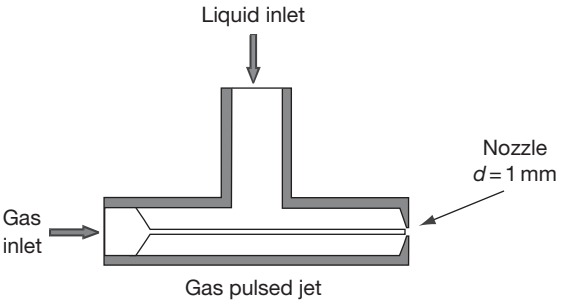


Figure 29 Gas-pulsed impinging jet. Reproduced from Schmitt, G.; Bakalli, M. In *A Critical Review of Measuring Techniques for Corrosion Rates under Flow Conditions*, Corrosion'2006; NACE International: Houston, TX, 2006; Paper no. 593.

2.13.3.3.1 Rotating disc electrode

The rotating disc electrode (RDE) is preferentially used to investigate mass transport phenomena in laminar flow conditions. According to the Levich equation (eqn [47]),[51] the mass transport rate, w, in a mass transport-controlled system is constant over the disk surface and proportional to the square root of the angular velocity, ω, of the disc:

$$w = k\sqrt{\omega} \qquad [47]$$

where w is the mass transport rate, for example, expressed in terms of mass loss rates ($g\,m^{-2}\,h^{-1}$), ω is the angular velocity in radian per second, and k is the constant, containing the diffusion coefficient, D, of the rate-determining diffusing species ($m^2\,s^{-1}$), its concentration, C_{bulk}, in the bulk solution ($mol\,m^{-3}$), and the kinematic viscosity, ν, of the solution investigated ($m^2\,s^{-1}$) (eqn [48]).

$$k = 0.6205 D^{2/3} \nu^{-1/6} C_{bulk} \qquad [48]$$

The mass transport rate can also be studied electrochemically. In the case of mass transport control, the diffusion-limited current density measured at a rotated disc electrode is proportional to the square root of the angular velocity, ω (eqn [49]):

$$j_D = 0.6205 nFD^{0.66} \nu^{-0.167} \omega^{0.5} C_{bulk} \qquad [49]$$

with the number of electrons n exchanged in the specific electrode reaction, F ($96\,485\,C\,mol^{-1}$), and the other parameters as defined earlier.

A graph of j_D versus $\sqrt{\omega}$ or w versus $\sqrt{\omega}$ (**Figure 30**) should give a straight line going through the origin if the electrode reaction in question is purely

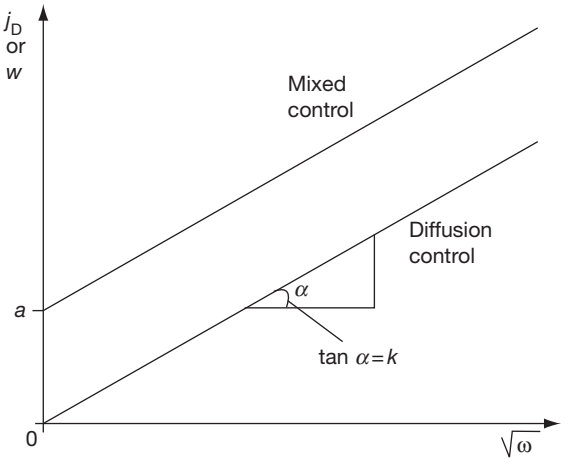


Figure 30 Levich plot for diffusion controlled and mixed controlled corrosion rate at rotated disc.

diffusion controlled. If the straight line cuts the ordinate according to the equation $w = a + k\sqrt{\omega}$, there are either experimental problems (e.g., shaft wobbling, formation of vortices, entrainment of gas bubbles, etc.) or there is an additional electrode reaction involved that is not diffusion limited. From the slope of the straight line in the j_D versus $\sqrt{\omega}$ diagram, the diffusion coefficient of the species determining the electrode reaction can be measured very accurately (eqn [49]). The flow pattern at the surface of the rotated disc can be described in terms of dimensionless parameters (Table 1).

2.13.3.3.2 Rotating cylinder electrode

The rotating cylinder electrode (RCE) has the advantage that the complete flow regime from laminar to turbulent can be tested in the same arrangement. Three flow regimes can be differentiated[52,53]:

1. Laminar flow at low rotational speeds of the cylinder (the velocity profile is tangential).
2. Laminar flow superimposed with some turbulent vortices in radial and axial motion above a critical Taylor number ($Ta = Re[(r_o - r_i)/r_i]$, where Ta is dimensionless Taylor number, Re is Reynolds number, r_o is outer cylinder radius (static), r_i is inner cylinder radius (rotating)).
3. Fully turbulent flow above critical Re.

The transition from laminar to turbulent flow is ~200. Below this critical Re, the Sh is constant ($Sh = 37$) and independent of Re.[54] At $Re > 200$, that is, in the turbulent region, the following dimensionless equation can be used for the system rotating inner cylinder/static outer cylinder (eqn [50])[55]:

$$Sh = 0.00791 Re^{0.7} Sc^{0.356} \quad [50]$$

This equation correlates with data from mass transfer, heat transfer, and friction measurements at cylinders rotating in liquids. The flow intensity at the cylinder wall can be calculated from eqn [51][56]:

$$\tau_w = 0.0791 Re^{-0.3} \rho r^2 \omega^2 \quad [51]$$

where τ_w is the wall shear stress (Pa), ρ is the solution density (kg m^{-3}), r is the cylinder radius (m), and ω is the rotational speed (s^{-1}).

Corrosion rate at a rotating cylinder is correlated with wall shear stresses using eqn [52].

$$w = a\tau^b \quad [52]$$

where w is the corrosion rate (mm a^{-1}), τ is the wall shear stress (Pa), and a and b are constants depending on solution parameters, chemistry, and corrosion product film and require definition.

2.13.3.3.3 Rotating cage

The rotating cage (Figure 27) is commonly used to screen the susceptibility of materials to FILC and to evaluate corrosion inhibitors for their efficiency to prevent initiation of FILC. Due to the many experimental advantages and the successful use of rotated cage data in service, attempts have been made to quantify the flow intensities encountered at coupons rotated in the rotated cage. Equation [53] was developed as an approximation to estimate the wall shear stress at coupons in the rotated cage[57,58]:

$$\tau_{RC} = 0.0791 Re_{RC}^{-0.3} \rho r_{RC}^2 \omega^{2.3} \quad [53]$$

where τ_{RC} is the wall shear stress in rotated cage (Pa), Re_{RC} is the Reynolds number of rotated cage $\left(Re_{RC} = \omega r_{RC}^2/\nu\right)$, r_{RC} is the radius of the rotated cage (m), ρ is the solution density (kg m^{-3}), and ω is the rotational speed (s^{-1}).

This equation is valid only for a certain geometry of the rotated cage with a certain number of coupons in a certain position. Comparison of critical wall shear stresses obtained from both rotated cage and impinging jet experiments for initiation of FILC in different corrosion systems revealed that the critical wall shear stresses calculated according to eqn [53] are about a factor of 2 higher than the values received via eqn [56] for the impinging jet.[59]

2.13.3.3.4 Pipe and channel flow

Flow experiments in tubes (pipes) or channels are rated as most service-related and most predictive, because the flow pattern in pipes and channels is encountered in technical plants and is well defined by equations for mass, heat, and momentum transfer. In the lab, pipe and channel flow experiments are generally performed in flow loops. The corrosion-relevant flow intensity at pipe or channel walls is best defined by the wall shear stress, τ_w, which can be measured as a pressure drop ΔP along a given length ΔL of the pipe with a diameter d (eqn [54])[13,15]:

$$\tau_w = \left(\frac{\Delta P}{\Delta L}\right)\left(\frac{d}{4}\right) \qquad [54]$$

The wall shear stress can also be measured electrochemically at microelectrodes flush mounted in pipe walls using the Léveque Equation (eqn [38]).[41]

For a microelectrode with disc-shaped electrode surface (diameter 15–60 μm) located in the region of highest turbulence of an impinging jet with a hexacyanoferrate (II/III) solution, a factor $k = 3 \times 10^{-8}$ N m^4 A^{-3} (eqn [38]) was obtained.[39,40] This value is given here for orientation on its range of magnitude. The microelectrode approach offers the advantage of measuring local wall shear stresses also at sites of disturbed flow (at obstacles, steps, grooves, etc.). Microelectrodes with diameters in the range of the size of microturbulence elements (e.g., ≪100 μm) allow the mapping of local wall shear stresses at surfaces with high lateral resolution.

The microelectrode approach was successfully used to quantify local and maximum fluid–wall interactions under slug flow intensities[40,60] expressed in terms of Froude numbers. The Froude number (Fr) is defined according to eqn [55].[61]

$$Fr = \frac{u_t - u_f}{\sqrt{g \cdot h_{\text{eff.}}}} \qquad [55]$$

where u_t is the translational velocity of the slug front (m s^{-1}), u_f is the average velocity of the liquid film (m s^{-1}), g is the acceleration due to gravity (m s^{-2}), and h_{eff} is the effective height of the liquid film (m). The higher the Fr, the higher the slug flow intensity.

2.13.3.3.5 Impinging jet

The impinging jet, specifically, the submerged impinging single phase jet (Figure 28), has proved its usefulness to investigate flow effects in corrosion systems, specifically, the initiation of FILC. The flow pattern at the impinged surface has been extensively studied and several detailed mathematical expressions of the hydrodynamics are available.[62,63] On the impinged surface, three hydrodynamic regions can be differentiated (Figure 28): the stagnation region (region A), a region with high turbulence intensities (region B), and the wall jet region (region C). For region B, the wall shear stress can be calculated according to eqn [56][47,63]:

$$\tau_w = 0.0447 \cdot \rho \cdot u_0^2 \cdot Re^{-0.182}\left(\frac{x}{d}\right)^{-2} \qquad [56]$$

where τ_w is the wall shear stress (Pa), ρ is the fluid density (kg m^{-3}), d is the inner diameter of the jet nozzle (m), u_0 is the fluid velocity at the front of the jet nozzle (m s^{-1}), x is the radial distance from the jet nozzle center line (m), Re is based on the dimensions of the jet ($Re = u_0 d/\nu$), and ν is the kinematic viscosity (m^2 s^{-1}). With microelectrodes flush mounted into the impinged surface, local wall shear stresses can be measured using the Léveque equation (eqn [38]). The impinging jet is commonly used to test the efficiencies of additives (inhibitors, flow improvers) for preventing FILC.

A modification of the impinging single phase jet is the gas-pulsed jet, which has been developed to test under flow intensities exceeding slug flow conditions. Figure 31 shows the wavelet transform plots of the time-related frequency/amplitude distributions obtained at microelectrodes impinged with a gas pulsed jet (Figure 29)[50,64] at a constant superficial liquid velocity, v_{SL}, of 12 m s^{-1} (at the jet nozzle) and two different superficial gas velocities, v_{SG} (Figure 31(a): 16 m s^{-1} and Figure 31(b): 22.3 m s^{-1}) at 1 gas pulse per second with a duration of 50 ms. The gas-pulsed impingement under conditions according to Figure 31(b) yielded responses with considerable intensities in a frequency range (>1000 Hz) clearly exceeding the typical range encountered in a moving slug flow at $Fr = 13$ (Figure 31(c)). The gas-pulsed impinging jet is a new versatile and small-scale experimental device to test flow-improving additives and the efficiency of corrosion inhibitors under slug flow conditions.

Equations for determining wall shear stresses and limiting current densities in different experimental set-ups are collected in Table 3.

2.13.3.3.6 Maximum local flow intensities

In most technical corrosion systems, the materials applied are exposed to the flowing media with a scale-covered surface. This makes scale cracking

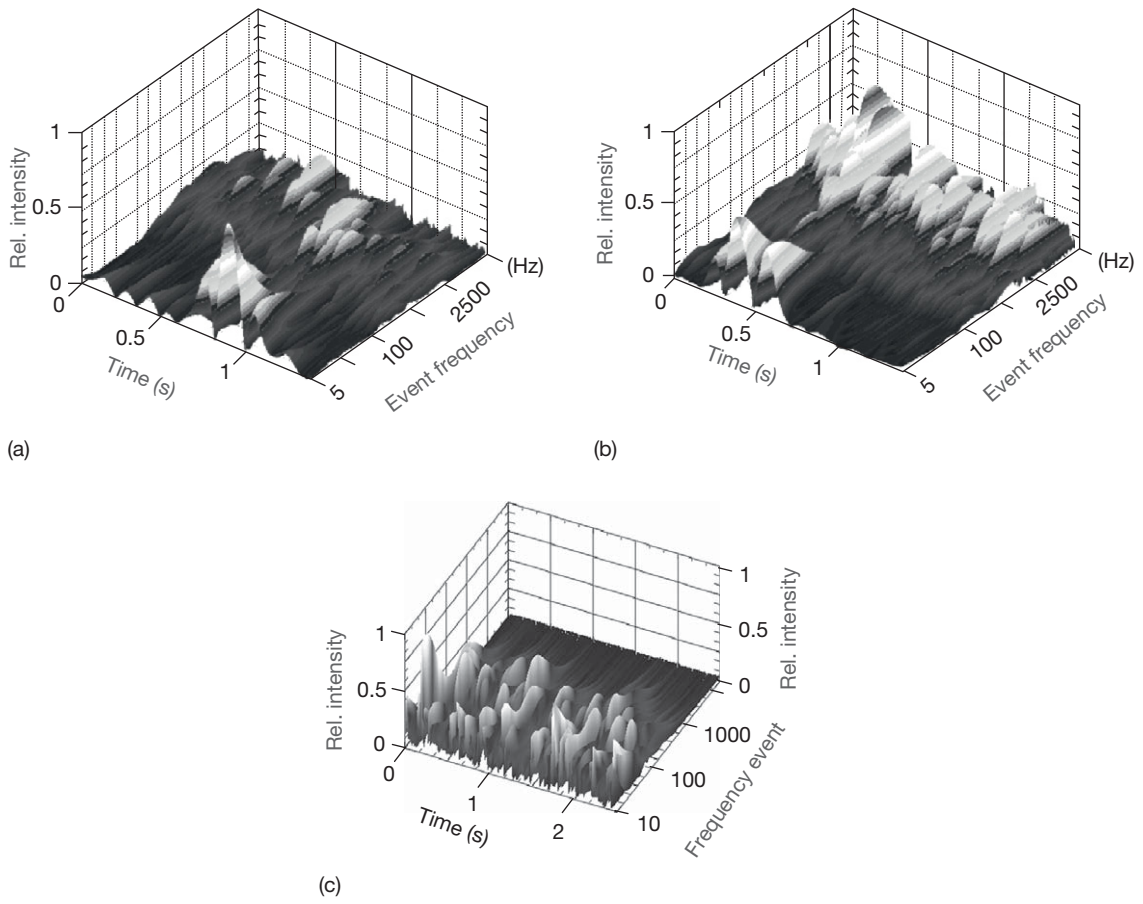


Figure 31 Flow intensities at surface impinged with the gas-pulsed jet (a, b) and treated with a moving horizontal slug (c). Reproduced from Schmitt, G.; Al-Janabi, Y. T.; Plagemann, P.; Bakalli, M. *Can the Gas-Pulsed Impinging Jet Simulate Slug Flow?*, Proceedings of European Conference, EUROCORR'2003, Budapest, Hungary, Sept 28–Oct 2, 2003; European Federation of Corrosion, The Institute of Materials: London, 2003.

and spalling the prevailing mechanism of FILC initiation. The evaluation of critical flow intensities, therefore, needs the selection of the appropriate experimental tools, which at acceptable experimental flow conditions can produce maximum (local) flow energy densities high enough to overcome the fracture stresses of the scales involved. (Editor's note: The theory presented here is based on a hypothesis to explain the creation of mechanical forces that are sufficient to fracture and remove protective layers and scales. While the hypothesis is plausible and provides an explanation for a number of experimental flow results, further work is needed to confirm this hypothesis and to understand fully the mechanism of scale removal.)

The maximum (freak) energy densities encountered in different flow systems have been quantified by microelectrochemical current noise measurements and subsequent Wavelet transform analysis using the algorithm outlined elsewhere in this volume. The maximum local flow intensities achieved with different experimental flow devices under single-phase and two-phase flow conditions are compared in **Figure 32** with a logarithmic scale for liquids with viscosities in the range of 0.8–1.1 m Pa s (e.g., aqueous solutions with moderate concentrations of dissolved solids).[39,40,45,67]

The maximum energy densities encountered under slug flow conditions range from 3 MPa (Froude number 5: $u_{SL} = 0.5\ m\ s^{-1}$; $u_{SG} = 2.0\ m\ s^{-1}$) to 30 MPa (Froude number 13: $u_{SL} = 1.5\ m\ s^{-1}$; $u_{SG} = 6.0\ m\ s^{-1}$).[39,40] With the gas-pulsed impinging jet energy densities can easily be achieved, which are 60–110 times higher than under severe moving slug flow conditions ($Fr = 13$).[67] It appears that the superficial liquid velocity (u_{SL}) is more important for achieving high maximum (freak) energy densities than the superficial gas

Table 3 Equations for determining wall shear stresses and limiting current densities with different lab tools for investigation of flow-assisted corrosion

Exp. tool	*Method*		*Wall shear stress*	*Reference*	*Limiting current density*	*Reference*
	Mass loss	*Electrochemical*				
Rotating disk	+	+	$\tau = 6.302\ \nu\rho\omega\sqrt{Re}$	66	$j_D = 0.6205\ nFD^{0.66}\nu^{-0.167}\ \omega^{0.5}C_{bulk}$	55
Rotating cylinder	+	+	$\tau_{RCE} = 0.0791\ Re_c^{-0.3}\ \rho r_c^2\omega^2$	58	$j_D = 0.0791\ nFC_{bulk}(\omega r)^{0.7}(r/\nu)^{-0.3}(\nu/D)^{-0.644}$	17
Rotating cage	+	–	$\tau_{RC} = 0.0791\ Re_{RC}^{-0.3}\rho r_{RC}^2\omega^{2.3}$	30	$j_D^3 - k\tau$, Léveque equation at	
			$\tau = \eta k(j^3 + 2k_c^2\ dj/dt)$ (electrochemically)	67	microelectrodes under stationary condition	45, 70
Flow loop	+	+	$\tau_w = (\Delta P/\Delta L)(d/4)$	13, 15	(a) Tube: $j_D = 0.320\ nFD^{2/3}\nu^{-1/6}l^{-1/3}r^{-1/3}u^{0.33}C_{bulk}$	17, 71
					l, r – tube electrode length and diameter, respectively	17
					(b) Channel: $j_D = 0.81\ nFD^{2/3}\ \nu^{-1/6}u^{1/2}C_{bulk}$	
Impinging jet	+	+	$\tau_w = 0.0447\rho u_0^2\ Re^{-0/182}x/d^{-2}$	65, 66	$j_D = knFC_{bulk}\ D^{2/3}\nu^{-1/6}(0/637u/r)^{1/2}$ r: electrode radius	71
					$j_D^3 = k\tau$, Levêque equation, stationary condition	70

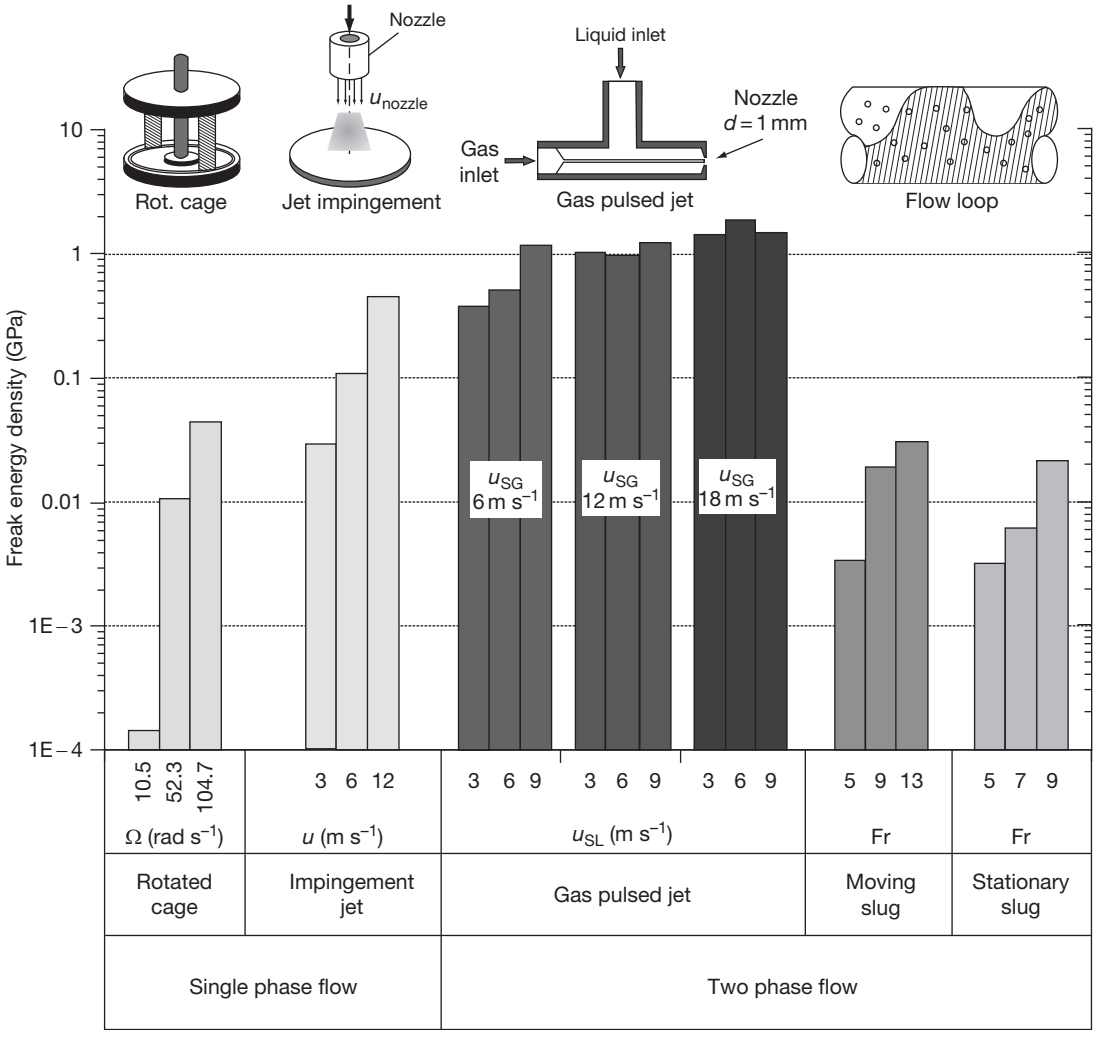


Figure 32 Maximum (freak) flow intensities in different flow systems. Adapted from Schmitt, G.; Werner, C.; Bakalli, M. In *Evaluation of Local Energy Densities in Disturbed Flow: A New Approach to Characterize Inhibitor Efficiencies to Mitigate Erosion Corrosion*, Proceedings of European Conference, EUROCORR'2004, Nice, France, Sept 12–16, 2004, European Federation of Corrosion, The Institute of Materials: London, 2004; Schmitt, G.; Werner, C.; Bakalli, M. In *Fluid Mechanical Interactions of Turbulent Flowing Liquids with the Wall – Revisited with a New Electrochemical Tool*, Corrosion'2005; NACE International: Houston, TX, 2005; Paper no. 344; Schmitt, G.; Bakalli, M. In *Quantification of Maximum Energy Densities Between Flowing Liquids and Solid Walls – Effect of Flow Improving Additives*, Proceedings of European Conference, EUROCORR'2006, Maastricht, The Netherlands, Sept. 24–28, 2006, European Federation of Corrosion, The Institute of Materials: London, 2006; Nakoryakov, V. E.; Kashinsky, O. N.; Kozmenko, B. K. In *Measuring Techniques in Gas–Liquid Flows*; Delhaye, J. M., Cognet, C., Eds.; Springer: Heidelberg, 1984; pp 695–721.

velocity. In the gas-pulsed jet approach, the gas pulse was injected once per second with duration of 50 ms. This indicates that the gas pulsed impinging jet can be used as an alternative experimental tool to the space-, time-, and cost-consuming slug flow loop, at least for screening purposes.

In single phase flow, the jet impingement allows the coverage of a wide range of maximum energy densities under easily achieved flow conditions. Thus, with the usual jet cell geometries, jet nozzle velocities of 3 or 12 m s^{-1} create maximum energy densities of about 28 and 940 MPa, respectively.[45] The flow intensities produced in the rotating cage are comparatively smaller. At angular velocities as low as 10.5 rad s^{-1} (equivalent to 100 rpm), the maximum energy density reaches only 0.14 MPa. However, even at 100.5 rad s^{-1}

(equivalent to 1000 rpm), the maximum local flow intensities amount only to 44 MPa. In view of the fracture stresses and adherence forces generally encountered for scales formed under corrosion conditions (Figure 33), the rotating cage finds its application for the quantification of critical flow conditions in the moderate range of flow intensities. The impinging jet and, even more, the gas-pulsed impinging jet can be used up to extreme fracture stresses of scales. Figure 33 demonstrates once more the importance of the FED approach for the understanding of the initiation step of FILC.

2.13.3.4 Erosion Corrosion and Cavitation Corrosion

Although FILC can occur already in single phase liquid flow, localized material attack very often is encountered in two-phase and multiphase flow. Thus, solid particle impact in a liquid or gas phase, liquid droplet impact in a gas phase, gas bubble impact in a liquid phase, and near-wall vapor bubble collapse in a liquid phase (Figure 8) can damage material surfaces with a mechanical and a corrosive contribution. While the material attack

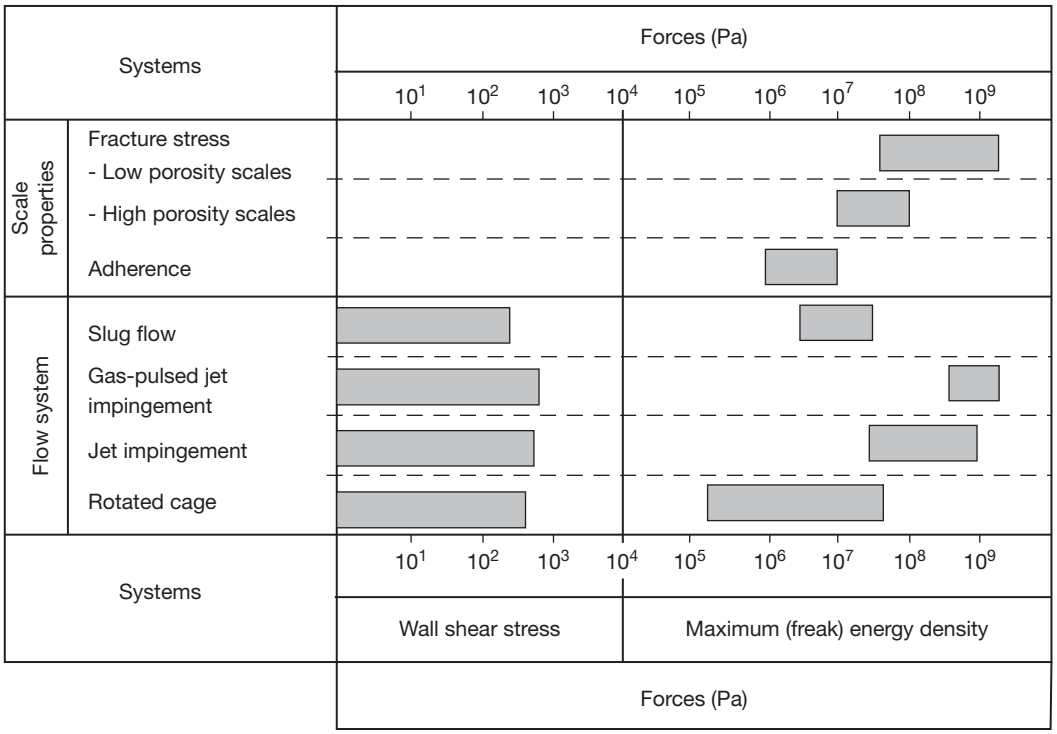


Figure 33 Comparison of mechanical and hydrodynamic forces. Adapted from Schmitt, G.; Mueller, M.; Papenfuss, M.; Strobel-Effertz, E. In *Understanding Localized CO_2 Corrosion of Carbon Steel from Physical Properties of Iron Carbonate Scales*, Corrosion'99; NACE International: Houston, TX, 1999; Paper no. 38; Schmitt, G.; Bosch, C.; Pankoke, U.; Bruckhoff, W.; Siegmund, G. In *Evaluation of Critical Flow Intensities for FILC in Sour Gas Production*, Corrosion'98; NACE International: Houston, TX, 1998; Paper no. 46; Schmitt, G.; Bosch, C.; Mueller, M.; Siegmund, G. In *A Probabilistic Model for Flow Induced Localized Corrosion*, Corrosion'2000; NACE International: Houston, TX, 2000; Paper no. 49; Schmitt, G.; Mueller, M. In *Critical Wall Shear Stresses in CO_2 Corrosion of Carbon Steel*, Corrosion'99; NACE International: Houston, TX, 1999; Paper no. 44; Schmitt, G.; Werner, C.; Bakalli, M. In *Evaluation of Local Energy Densities in Disturbed Flow: A New Approach to Characterize Inhibitor Efficiencies to Mitigate Erosion Corrosion*, Proceedings of European Conference, EUROCORR'2004, Nice, France, Sept 12–16, 2004, European Federation of Corrosion, The Institute of Materials: London, 2004; Schmitt, G.; Werner, C.; Bakalli, M. In *Fluid Mechanical Interactions of Turbulent Flowing Liquids with the Wall – Revisited with a New Electrochemical Tool*, Corrosion'2005; NACE International: Houston, TX, 2005; Paper no. 344; Schmitt, G.; Bakalli, M. In *Quantification of Maximum Energy Densities Between Flowing Liquids and Solid Walls – Effect of Flow Improving Additives*, Proceedings of European Conference, EUROCORR'2006, Maastricht, The Netherlands, Sept. 24–28, 2006, European Federation of Corrosion, The Institute of Materials: London, 2006; Nakoryakov, V. E.; Kashinsky, O. N.; Kozmenko, B. K. In *Measuring Techniques in Gas–Liquid Flows*; Delhaye, J. M., Cognet, C., Eds.; Springer: Heidelberg, 1984; pp 695–721.

produced by the first three cases of impingement is generally termed erosion or – in the case of corrosive contributions – erosion corrosion, the material damage by near-wall vapor bubble collapse is called cavitation or – in the case of corrosive contributions – cavitation corrosion. Typical examples of erosion corrosion and cavitation corrosion are given in Figures 34–36.

2.13.3.4.1 Solid particle impingement

Solid particles entrained in a flowing liquid can damage a material surface if they impact with a kinetic energy above critical values, which are material specific and relate to the nature and properties of scales, layers, or films, and the base material. The effective kinetic impact energy of the particle depends on the impact velocity, u_p, the impact angle, and the particle density, size, shape, relative hardness, and microroughness. At low effective impact energies, the particle impact destroys only the scales, layers, or films and activates the surface for corrosion. In this case, the kinetics of reformation of the scales, layers, or films ('repassivation') play an essential role for the final corrosion rate. At higher impact energies, the particles may penetrate into the surface of the material and can even cut out solid particles.

The rate of erosion corrosion, w_{er}, is, therefore, proportional to the kinetic energy of the impinging particle, that is, proportional to u_p^2, and the frequency of impacts, that is, proportional to u_p. Thus, in a first approximation, the rate of erosion corrosion correlates with the particle velocity in a cubic relationship (eqn [57])[68]:

$$W_{er} \propto u_p^3 \qquad [57]$$

Under conditions where the effective particle impact energy is above the critical value for scale, layer, or film destruction, but below the critical value for eroding the base material, the rate of erosion corrosion, is generally, directly proportional to the particle velocity and its concentration.[68] Velocity exponents up to 3 are encountered under conditions of base material erosion.

Impact angle of particles in nondisturbed turbulent pipe flow are generally below 5°. However, under

Figure 34 Erosion corrosion and cavitation corrosion at pump impeller (cast duplex stainless steel).

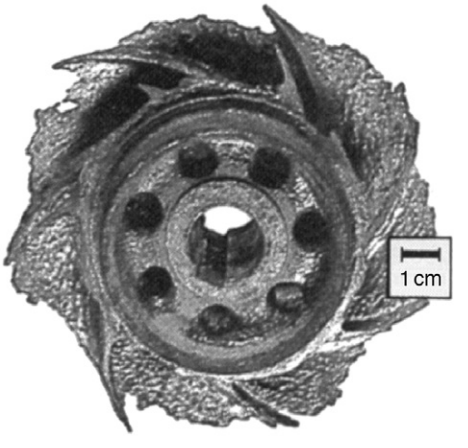


Figure 35 Erosion corrosion and cavitation corrosion at pump impeller (cast carbon steel).

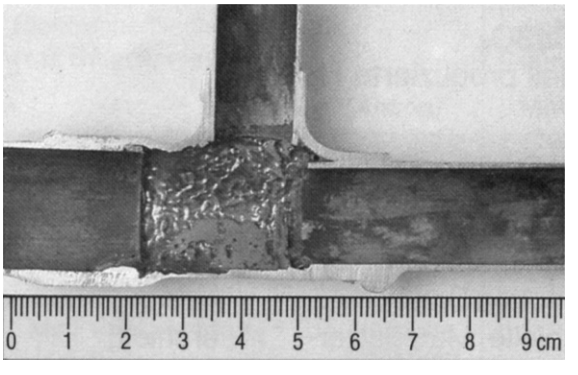


Figure 36 Erosion corrosion at flow disturbance in copper pipe for cold tap water.

conditions of disturbed flow, a wide range of impact angle will be encountered. Erosion corrosion of brittle materials by slurries increases, with the particle impact angle reaching a maximum at an angle of 90°. Ductile materials develop erosion corrosion maxima at angles in the range of 15°–50°.[23] Erosive attack drops dramatically when the impacting particles are softer than the base material.

2.13.3.4.2 Liquid droplet impingement

Liquid droplets entrained with fast-flowing gases or vapors can impinge on the surface with energy densities far higher than single phase or multiphase liquid flows. The result is the destruction of scales, layers, or films and, in severe cases, also the deformation of the surface to form discrete sharp-edged pits, which are gradually deepened by continued droplet impacts. Under such conditions, the mechanical part of erosion corrosion dominates, as in cavitation corrosion. However, the overall mass loss is still influenced by the kinetics of repassivation processes, which can be enhanced by appropriate alloying (e.g., increase of Mo content in stainless steels) or by the addition of powerful corrosion inhibitors. Liquid droplet impingement is encountered in low-pressure steam turbine blades operated with wet steam or in aircraft and helicopter rotors frequently operated in rain.

2.13.3.4.3 Gas bubble impingement

Gas bubbles entrained with fast-flowing fluids can damage protective scales, layers, or films locally when the impact energy is high enough. This may happen specifically at sites of disturbed flow, for example, at the inlets of heat exchanger tubes where air bubbles can develop a substantial radial velocity component. The morphology of the damage starts with isolated 'horseshoes,' which grow together and finally cause considerable roughening of the surface.

2.13.3.4.4 Cavitation

When gas bubbles implode in very close vicinity to the surface, they form very small liquid 'microjets,' which impact on the surface with flow velocities of up to 100 m s^{-1} (**Figure 37**). This is enough to destroy protective layers and deform the material in a way that particles of the base material are loosened and carried away. Cavitation bubbles with diameters <1 mm can form at sites where, due to local turbulence conditions (e.g., at sharp diameter changes in pipes, edges of propellers, turbine blades or pump impellers), the hydrostatic pressure decreases for a very short time below the saturation vapor pressure of the liquid. The collapse occurs downstream in higher pressure regions. Vibrating solid surfaces, for example, the cooling water side of diesel engine cylinder liners or the shell-side of heat exchanger tubing, may also experience cavitation attack.

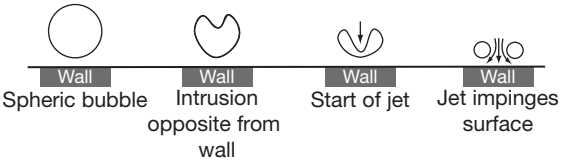


Figure 37 Collapse (implosion) of cavitation bubble with jet formation.

In cavitation corrosion, the mechanical contribution is by far dominating. However, the corrosion contribution should not be neglected, because it can be influenced significantly by enhancing the repassivation kinetics and by using appropriate corrosion inhibitors. Even the mechanical impact can be reduced by flow improving additives.

2.13.3.5 Corrosion Control

The basic methods for controlling flow-assisted corrosion and/or FILC include design modifications, changing of flow regime, and the improvement of the fluid mechanical properties of the flowing medium.

2.13.3.5.1 Fluid mechanics and design

The majority of failures by flow-assisted corrosion or FILC are based on poor design and ignorance of fluid mechanical principles. Therefore, flow related design is the most important corrosion control method. The design should primarily aim at reducing the flow velocity. The importance of this measure can be derived from eqn [46], which relates the corrosion rate to the flow velocity. For mass-controlled corrosion reactions, the exponent of the velocity term is 0.33 for laminar, but up to 1 for turbulent flow. For flow conditions causing destruction of protective scales, the exponent can reach values up to 3 and amount to 8 for cavitation conditions.

Rules of thumb have been empirically developed in different industries to estimate critical flow velocities, which should not be exceeded. One rule that was originally established in power generation for preventing FILC in pipes for wet steam has been adapted by the oil and gas industry as API-RP 14E,[69,70] where the critical flow velocity is inversly

proportional to the square root of the medium density (eqn [58]):

$$u_{crit} = C(\text{Density})^{-1/2} \quad [58]$$

With a proportionality factor C of 7.6 and the density given in ($kg\,m^{-3}$), the critical flow velocity is obtained in ($m\,s^{-1}$). The proportionality factor is 100 (for inhibited systems 150–200) if the density is given in lbs ft^{-3}. The critical velocity is then obtained in $ft\,s^{-1}$. For liquid flows, the proportionality factor of 100 is very conservative; for gas flows, it underestimates the potential for erosion. This holds also for information contained in BS EN ISO 13703:2001 developed for offshore applications.[71]

For plumbing systems in tap water distribution, critical flow velocities are given, for example, for steel ($4\,m\,s^{-1}$) and copper ($2\,m\,s^{-1}$). Many critical parameters have been worked out for materials in seawater application. Maximum design velocities for CuNi10 (UNS C70600) piping for seawater have been related to pipe diameters based on measurements of critical wall shear stresses (**Figure 38**).[72] For CuNi30 (UNS C71500) 4.5 $m\,s^{-1}$ and CuNi10 (UNS C70600) 3.6 $m\,s^{-1}$ have been accepted industrially as critical flow velocities in large piping for unpolluted seawater.[72] Critical flow parameters for copper alloy tubing are collected in **Table 4**.[7,10] In the chemical industry, the general (conservative) design rule is to prevent flow velocities above 1.5 $m\,s^{-1}$.

Proper design should also avoid abrupt changes in the flow system geometry, for example, steps, grooves, solder beads, rectangular elbows, tees, etc. Experience shows that flow disturbances as small as corrosion pits (even smaller than 1 mm) or burrings on cut ends of tubing may create sufficient turbulence to destroy protective layers and initiate FILC. Erosion corrosion at heat exchanger tube-inlets can be solved by plastic inserts.

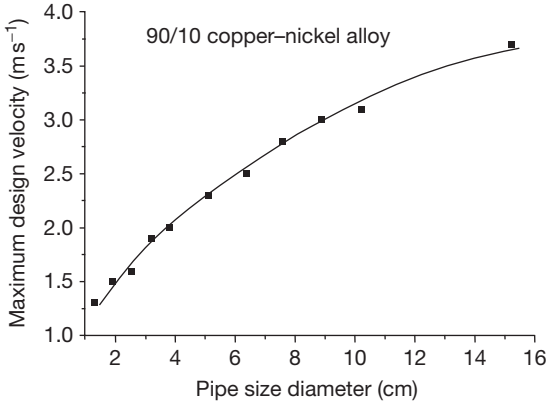


Figure 38 Maximum design velocities of piping for seawater. Reproduced from Efird, K. D. *Corrosion* **1973**, *33*(1), 3–8.

2.13.3.5.2 Materials selection

Selection of materials with higher resistance to flow-assisted corrosion is another option for corrosion control by design. Flow resistance is often related to the hardness of the material. **Figure 39** gives an example for copper and copper alloys in tap water.[73] The erosion corrosion performance of valve materials in seawater is compared in **Figure 40**.[74] Material selection is also helpful for mitigating cavitation corrosion. **Figure 41** presents different materials with respect to the relative material loss by cavitation.[75]

The variety of materials is great for application in fast-flowing media, for example, ferritic, ferritic-austenitic (duplex) and austenitic steels, NiCr alloys and titanium and its alloys (**Figure 42**).[6] To combat erosion corrosion by fluids containing solid particles, successful developments have been made with cast alloys and duplex steels. However, materials should not be selected only by their flow resistance, but also with consideration of susceptibility to localized corrosion such as pitting, which is a common problem with stainless steels and also nickel-base alloys, specifically in slow-flowing and stagnant media and under deposits. In these cases, decreasing the flow rate may not be the appropriate remedy for corrosion control.

In special applications, surface protection by metallic, organic, or ceramic coatings can be a solution to prevent FILC. The adhesion and mechanical stability of such protective layers is important.

2.13.3.5.3 Environment control

Environment control includes influencing the fluid chemistry and the fluid hydrodynamics (flow improvement).

Reducing the corrosivity of the flow system by removal or exclusion of corrosive components is surely an efficient step to control flow-assisted corrosion. An example would be the removal of oxygen from seawater or adjustment of pH. However, this is often not possible or not feasible. In these cases, the selection of appropriate additives can be very effective. Corrosion inhibitors are known to foster the formation of more flow-resistant protective scales with higher fracture stress, and hence, improve the resistance to FILC.

Table 4 Critical flow parameters for copper alloy tubing in seawater

Alloy	*Critical shear stress ($N m^{-2}$)*	*Critical velocity in 25-mm tube ($m s^{-1}$)*	*Design velocity ($m \cdot s^{-1}$)*[a]	*Maximum tubular design velocity ($m s^{-1}$)*
Cupronickel with Cr	297	12.6	8.6	9
CuNi30	48	4.6	3.1	4.5–4.6
CuNi10	43	4.3	2.9	3–3.6
Al Bronze			2.7	
Arsenical A1 Brass	19	2.7	1.9	2.4
Inhibitory Admirality				1.2–1.8
Low Si Bronze				0.9
P-deoxidized Copper	9.6	1.9	1.3	0.6–0.9

[a]Based on 50% critical shear stress.

Source: Roberge, P. In *Corrosion Testing Made Easy*; Syrett, B. C., Series, Eds.; NACE International: Houston, TX, 2004; Postlethwaite, J.; Nesic, S. In *Uhlig's Corrosion Handbook*, 2nd ed.; Winston Revie, R., Ed.; Wiley: New York, 2000; Chapter 15, pp 249–272; Smart, J.S., III *A Review of Erosion Corrosion in Oil and Gas Production*, Corrosion'90; NACE International: Houston, TX, 1990; Paper no. 10.

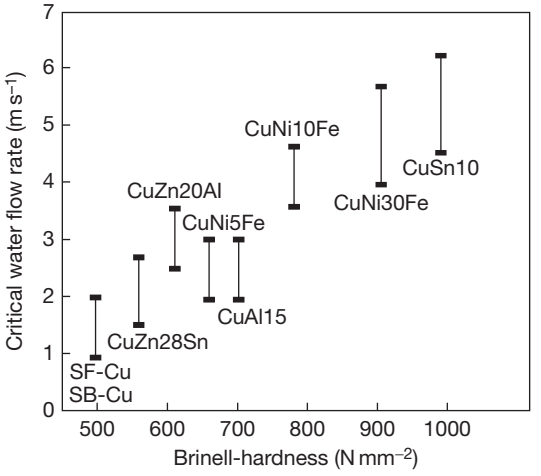


Figure 39 Maximum allowable flow velocity for water in copper and copper alloy piping. Reproduced from Sick, H. *Werks. Korros.* **1972**, *23*, 12.

The efficiency of corrosion inhibitors in flowing media depends on local flow intensities (**Figure 43 (a)**). Above critical flow intensities, inhibitors lose their efficiency. Under these conditions, the material surface is subject to FILC. Critical flow intensities for FILC initiation can also depend on the inhibitor concentration (**Figure 43(b)**). This depends on the chemical structure of the active inhibitor compounds and the formulation of inhibitor packages.[76] Economical inhibition of corrosion under turbulent flow conditions needs functional additives that inhibit under high critical flow intensities at low concentrations. In any case, the inhibitor concentration should be adjusted to the flow intensity and to a given target corrosion rate. With increasing flow intensity, an increase of the inhibitor concentration is likely to be necessary to reduce the corrosion rate below the target value. Failure prevention must be the primary goal for the selection of the type and the concentration of a functional additive.

However, reduction of the corrosion rate is not the only effect of corrosion inhibitors. They may also influence the flow dynamic property of the fluid, that is, by acting as a drag reducer, or in other words, as a flow improver.[37] There is literally no flow process that cannot be influenced by adding such 'drag-reducing agents' (DRA): frictional resistance of turbulent pipe and boundary layer flow is diminished; local energy losses at flow obstacles are reduced; the periodic formation of vortices in the flow field behind a cylinder experiences a shift to lower frequencies; resistance to laminar flow in porous media is increased; sedimentation and flocculation processes become accelerated; the efficiency of hydraulic transport of solids rises; the onset of cavitation starts at higher flow velocities and causes less damage to materials; flow noise becomes less intense. DRA have been suggested in various technical fields, for example, in industrial water distribution systems to reduce the pumping energy, in oil and gas transport as well as in hydraulic transport of solids (e.g., coal), to increase (e.g., double) the capacity of pipelines, in waste water pipe networks to reduce the residence time of waste water in sewers at peak loads, in ship technology to reduce the flow friction in order to increase the speed of ships (e.g., cut in half the

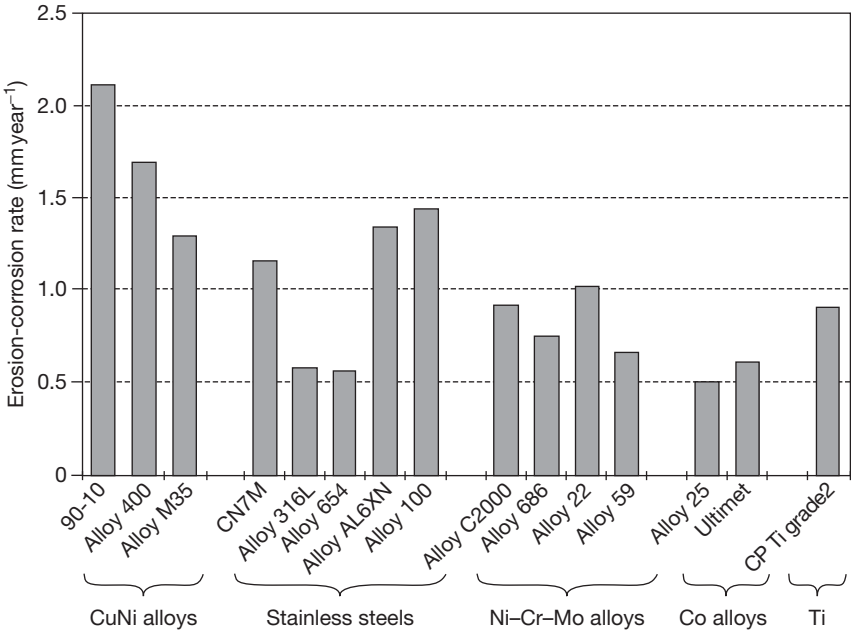


Figure 40 Valve materials in seawater. Reproduced from Ruedisueli, R. L.; Aylor, D. M. In *Erosion Resistance Testing for Candidate Seawater Valve Materials*, Corrosion'2001; NACE International: Houston, TX, 2001; Paper no. 468.

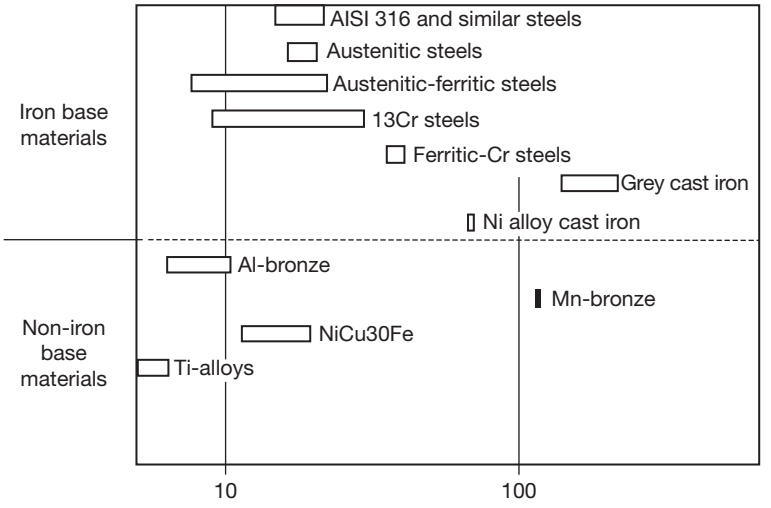


Figure 41 Cavitation resistance of different pump materials. Reproduced from Weber, J. VDI-Report No. 365, 1980, pp 73.

drag of a super-tanker), in fire fighting technology to reduce the jet spray and increase the throwing distance of the water jet from the fire-hose nozzle (e.g., by a factor of 5).

DRAs include polymers like polyethylenoxides, polyacrylamides, guar gum, sodium carboxymethylcelluloses, and other polymers, which are, however, only effective at mole masses $>10^6$ Da. Besides polymers, complex detergent mixtures, for example, an equimolar mixture of cetyltrimethylammonium bromide (CTAB) and 1-naphthole, and other quaternary ammonium salts can exhibit drag-reducing activity.[8,37,77] However, these compounds become DRAs only above the critical micelle concentration (cmc). It is assumed that near-wall turbulence elements interact with the DRA under energy dissipation into internal energy by viscous actions. For polymers, this implies shear-induced

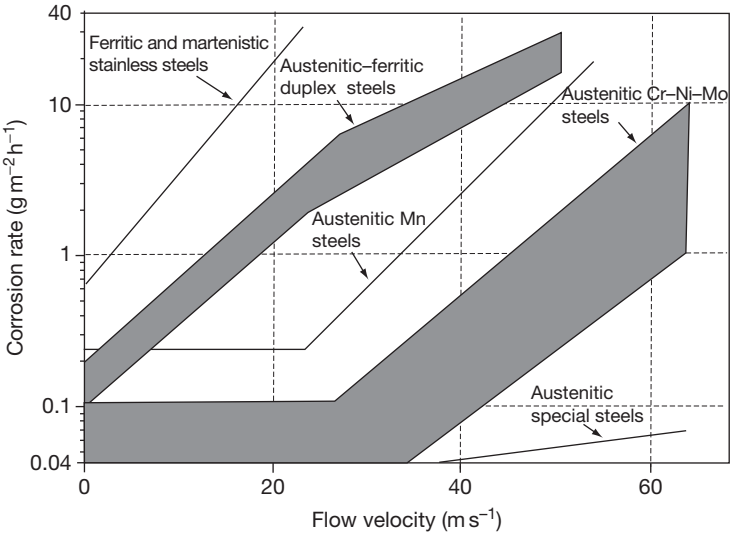


Figure 42 Materials performance in media with high levels of hydrogen sulfide and low pH. Reproduced from Weber, J. *Br. Corros. J.* **1992**, *27*(3), 193–199.

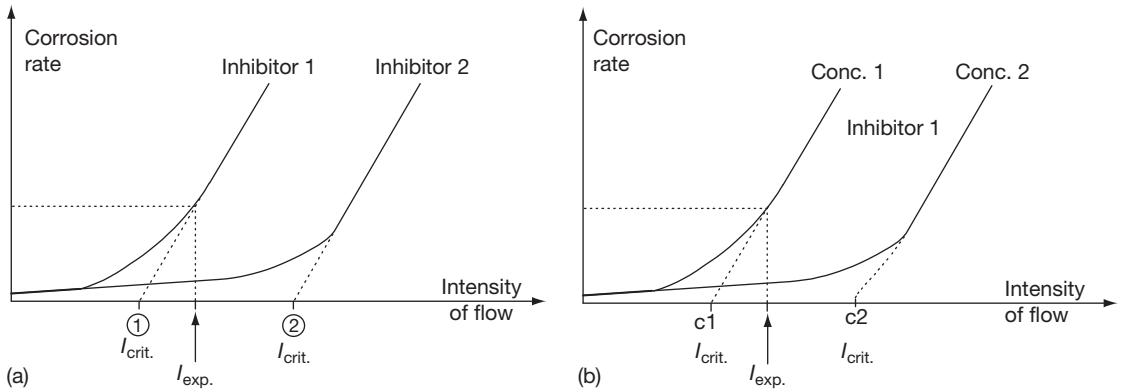


Figure 43 Effect of flow intensity on corrosion inhibitor performance.

coil-stretch actions, for surface active compounds, shear-induced deformation of supramolecular micelle structures.

The effect of additives in two-phase flow systems was observed, for example, in brine/hydrocarbon mixtures under annular mist flow conditions with a superficial gas velocity of 18.4 m s^{-1} and a superficial liquid velocity of 0.015 m s^{-1} at 80°C, and CO_2 partial pressures of up to 25 bar (**Figure 44**).[8] In the presence of 1 mmol l^{-1} additive, the relationship between wall shear stress and gas density (CO_2 gas pressure) significantly deviated from the linear relationship measured in the absence of additives. The highest effect was measured at higher CO_2 pressures in the presence of higher molecular alkylamines. The wall shear stress could be reduced by a factor of 5, thus reducing the risk of FILC.[8]

The drag-reducing efficiency of C14-quat (tetradecyl-trimethylammonium bromide) in single phase flow and under slug flow conditions is exemplified in **Figure 25** and **Figure 45**, respectively. The interaction intensity between aqueous liquids and the solid wall is, here, characterized by the maximum (freak) energy density of the flow system.

2.13.3.5.4 Cathodic protection

Cathodic protection could also be used for corrosion control in flow systems. The application is, however,

limited because it is technically often difficult to provide the appropriate protection current inside pipes, valves, or pumps, by auxiliary anodes due to difficult geometries. Furthermore, evolution of hydrogen might be a problem, for example, in the case of titanium and other materials that could experience hydrogen embrittlement.

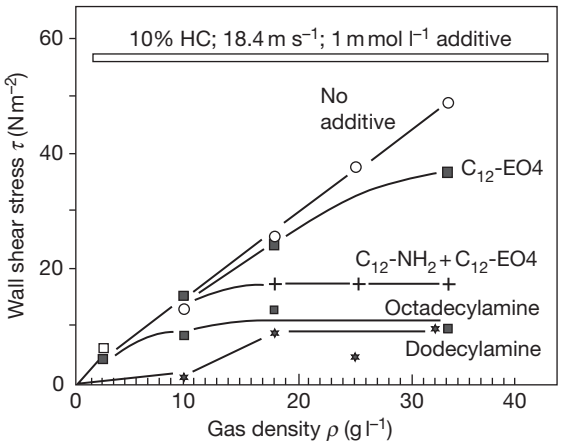


Figure 44 Flow improving effect of additives in annular mist flow loop under CO_2 corrosion conditions. Reproduced from Schmitt, G.; Bücken, W. In *Effect of Corrosion Inhibitors and Surfactants on the Fluid Dynamic Properties of Multiphase Annular Mist Flow in CO_2 Corrosion Systems*, Corrosion'92; NACE International: Houston, TX, 1992; Paper no. 6.

2.13.3.5.5 Control of solid particle impingement attack

Control measures can include appropriate material selection, application of functional additives, and flow adjustment. Material selection often ends up with carbon steel because of its relatively low price. Erosive metal loss is then generally compensated by higher wall thicknesses. Stainless iron-base alloys exhibit satisfactory erosion corrosion resistance due to fast healing passive layers, but are more costly. The materials listed in **Figure 40** in terms of corrosion resistance for valve applications also exhibit good resistance to erosion corrosion. However, material selection has to account also for the concentration and the nature (hardness) of the suspended solids. Criteria are different for flow systems with low concentrations of suspended particles and for slurries with high concentration of solids. High-Cr–Mo alloys with high carbon content (e.g., 20–28% Cr, 2% Mo, 2–2.5% C) have been successfully used for slurries with pH values down to 4. For more corrosive environments, the carbon content should be reduced. For sand and gravel pumps, Ni-hard alloys (4% Ni, 2% Cr) found successful application in the case of mildly corrosive conditions. Also rubber-lined pumps are an option for abrasive slurries (particle size >3 mm).

The literature reports on the excellent performance of Stellite (cast Co alloy) and silicon carbide during the handling of slurries with pumps. **Figure 46**

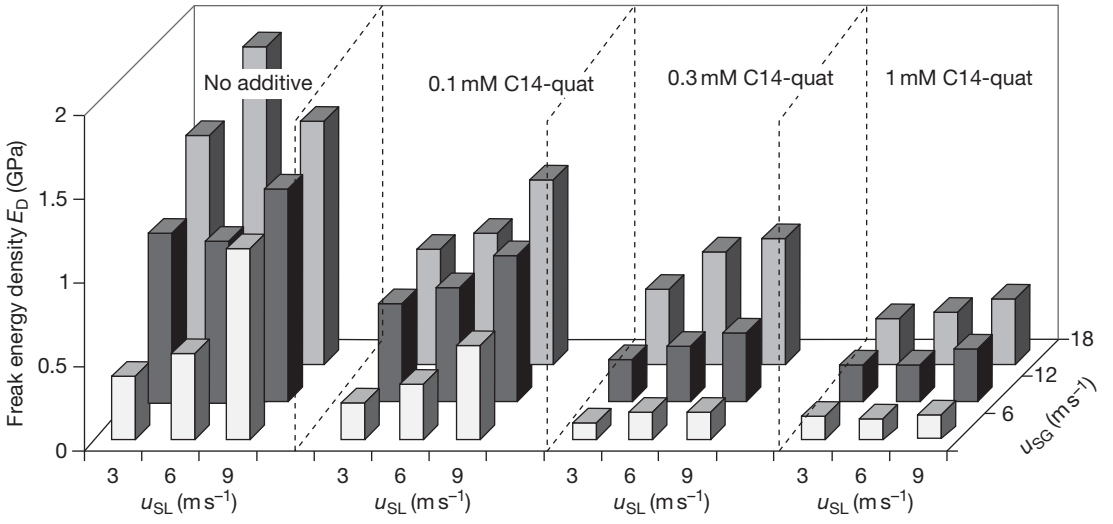


Figure 45 Effect of flow velocity and C14-Quat concentration on the maximum (freak) energy density under slug flow conditions simulated via gas-pulsed impinging jet. Reproduced from Schmitt, G.; Bakalli, M.; Hoerstemeier, M. In *Contribution of Drag Reduction to the Performance of Corrosion Inhibitors in One- and Two-Phase Flow*, Corrosion'2007; NACE International: Houston, TX, 2007; Paper no. 615.

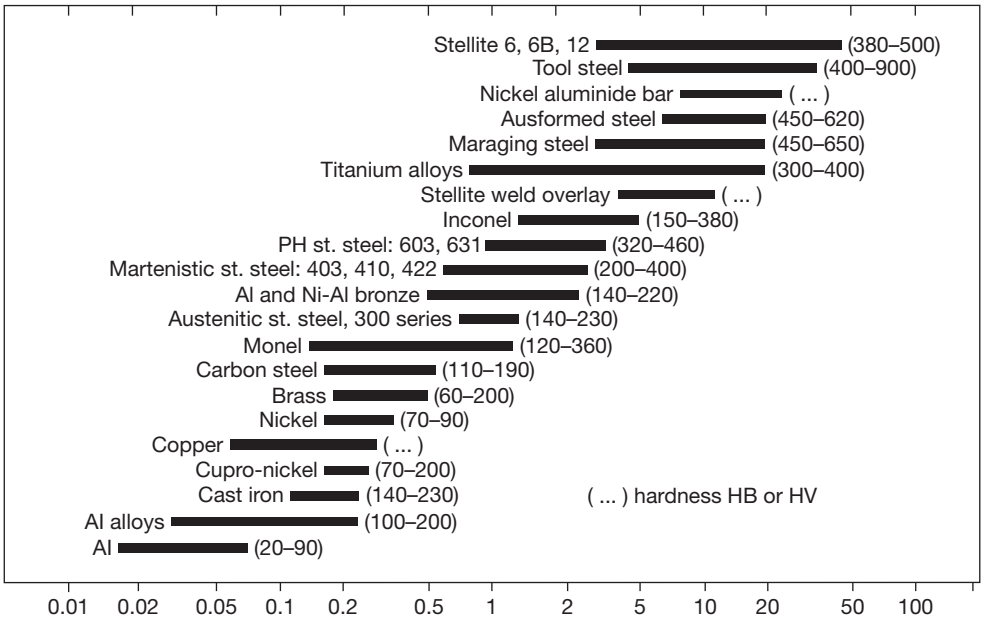


Figure 46 Normalized erosion resistance of materials relative to 316 SS (HV 170). Reproduced from Heymann, F. J. In *ASM Metals Handbook*: *Friction, Wear and Lubrication Technology*, Vol. 18; ASM: Metals Park, OH, 1992; pp 221–232.

gives an overview of the erosion resistance of various materials, normalized relative to 316 austenitic stainless steel with a hardness of HV 170.[78] These data relate mainly to impingement and cavitation test results and fit together, because the mechanical part dominates in both failure mechanisms. However, these data should be taken only as a first orientation.

2.13.3.5.6 Control of liquid droplet impingement attack and cavitation

Due to similar effects, failure prevention under conditions of liquid droplet impingement and cavitation can use similar options in material selection and flow control by design. **Figures 41** and **46** show a good orientation for material selection. However, polymeric materials, for example, high-density polyethylene (HDPE) or rubber, could also provide a solution. Flow control of droplet impingement should consider the droplet density (amount of impacting liquid), droplet size, and angle of droplet impact. Design options depend on the technical situation. If feasible, droplets should be removed, for example, by cyclones, or deflected with impingement plates. Guidelines for avoiding cavitation by proper design of pumps, orifices, valves, and elbows, are given in the literature.[79] They primarily aim at changing the hydraulic conditions. For centrifugal pumps, the major design parameter is the pressure difference between the total absolute pressure and the vapor pressure at the pump suction. Corrosion inhibitors have been successful in reducing cavitation attack in cooling systems of diesel engines. Another option is to use flow-improving additives (superhigh molecular polymers, micellar aggregates).

2.13.3.5.7 Predictive modelling

Computational fluid dynamics (CFD) methods have been successfully applied to predict flow fields and mass transfer rates for various system geometries in single phase and multiphase flow.[80,81] For fluids with entrained solid particles, the CFD procedure includes flow simulation, calculation of particle tracking, and assessment of erosion rates.[81–87] A comparison of results from different CFD approaches with experimental data indicates that CFD is a powerful tool in predicting erosion and erosion corrosion.[86]

References

1. Heitz, E.; Weber, S.; Liebe, R. In *Erosion Corrosion and Erosion of Various Materials in High Velocity Flows Containing Particles*, Proceedings of Symposium on Flow-Induced Corrosion; Fundamental Studies and Industry Experience; Kennelley, K. J., Hausler, R. H., Silverman, D. C., Eds.; NACE: Houston, TX, 1991; pp 5:1–5:15.
2. Hausler, R. H.; Schmitt, G. *Hydrodynamic Flow Effects on Corrosion Inhibition,* Corrosion'2004; NACE International: Houston, TX, 2004; Paper no. 402.

3. Schmitt, G.; Gudde, T. *Local Mass Transport Coefficients and Local Wall Shear Stresses at Flow Disturbances,* Corrosion'95; NACE International: Huston, TX, 1995; Paper no. 102.
4. Schmitt, G. In *Flow Induced Localized Corrosion – A Challenge for Corrosion Inhibition in Oil and Gas Production,* Proceedings of the 15th Annual Conference Corrosion Problems in Industry, Alexandria, Egypt, Oct. 9–10, 1996; pp 1–15.
5. Schmitt, G.; Pankoke, U.; Bosch, C.; Gudde, T.; Strobel-Effertz, E.; Papenfuß, V.; Bruckhoff, W. In *Initiation of Flow Induced Localized Corrosion in Oil and Gas Production. Hydrodynamic Forces vs. Mechanical Properties of Corrosion Product Scales,* Proceedings of the 13th International Corrosion Congress, Melbourne, Australien, Nov. 25–29, 1996; Vol. 2, Paper no. 150.
6. Weber, J. *Br. Corros. J.* **1992**, *27*(3), 193–199.
7. Roberge, P. In *Corrosion Testing Made Easy*; Syrett, B. C., Series, Eds.; NACE International: Houston, TX, 2004.
8. Schmitt, G.; Bücken, W. In *Effect of Corrosion Inhibitors and Surfactants on the Fluid Dynamic Properties of Multiphase Annular Mist Flow in CO_2 Corrosion Systems,* Corrosion'92; NACE International: Houston, TX, 1992; Paper no. 6.
9. Lotz, U. *Velocity Effects in Flow Induced Corrosion,* Corrosion'90; NACE International: Houston, TX, 1990; Paper no. 27.
10. Postlethwaite, J.; Nesic, S. In *Uhlig's Corrosion Handbook*, 2nd ed.; Winston Revie, R., Ed.; Wiley: New York, 2000; Chapter 15, pp 249–272.
11. Efird, K. D. In *Uhlig's Corrosion Handbook*, 2nd ed.; Winston Revie, R., Ed.; Wiley: New York, 2000; Chapter 14, pp 233–248.
12. Munson, B.; Young, D. F.; Okiishi, T. H. *Fundamentals of Fluid Mechanics*, 3rd ed.; Wiley: New York, 1988; Chapter 8, pp 459–549.
13. Oertel, H. *Prandtl–Führer durch die Strömungslehre,* (in German), 10th ed.; Carl Hanser: München Wien, 1990.
14. VDI-Wärmeatlas. Druckverlust bei der Strömung durch Rohre (in German) 8th ed.; VDI: Dusseldorf, Germany, 1997.
15. Schlichting, H.; Gersten, K. *Boundary Layer Theory,* 8th ed.; Springer-Verlag: Berlin, 2003.
16. Rathnasingham, R.; Breuer, K. S. *Phys. Fluids* **1997**, *9*, 1867.
17. Štulik, K.; Pacakova, V. *Electroanalytical Measurements in Flowing Liquids*; Ellis Horwood: Chichester, 1986.
18. Davies, J. T. *Turbulence Phenomena*; Academic Press: New York, 1972; Chapter 2, pp 121–143.
19. Durst, F.; Jovanovic, J.; Sender, J. *J. Fluid Mech.* **1995**, *295*, 305–335.
20. Nepomuceno, H. G.; Lueptow, R. M. *Phys. Fluids* **1997**, *9*, 2732–2739.
21. Weisman, J. In *Handbook of Fluids in Motion*; Cheremisinoff, N. P., Gupta, R., Eds.; Ann Arbor Science Publishers: Ann Arbor, MI, 1983; pp 409–424.
22. Govier, G. W.; Aziz, K. *The Flow of Complex Mixtures in Pipes*; Van Nostrand Reinhold: New York, 1972; pp 324, 326, 506.
23. Heitz, E. In *Flow-Induced Corrosion: Fundamental Studies and Industry Experience*; Kennelley, K. J., Hausler, R. H., Silverman, D. C., Eds.; NACE International: Houston, TX, 1991; pp 1:1–1:29.
24. Schmitt, G.; Bruckhoff, W.; Faessler, K.; Blummel, G. In *Flow Loop versus Rotating Probes – Correlations Between Experimental Results and Service Applications,* Corrosion'90; NACE International: Houston, TX, 1990; Paper no. 23.
25. Lotz, U.; Heitz, E. *Werkst. Korros.* **1983**, *34*, 454–461.
26. Deslouis, C. *Elec. Chim. Acta* **2003**, *48*, 3279–3288.
27. Heinze, J. *Ber. Bunsengesellsch. Phys. Chem.* **1980**, *84*, 785.
28. DIN 50920, Corrosion of Metals, Part I. *Corrosion Testing in Flowing Liquids; General,* (in German); Beuth Verlag: Berlin, Germany, 1985.
29. Newman, J. S. *Electrochemical Systems*; Prentice Hall: Englewood Cliffs, NJ, 1971.
30. Pini, G.; Weber, J. *Technische Rundschau Sulzer,* (in German) **1979**, *2*, 69.
31. Epstein, N. In *Fouling in Heat Exchangers*; Proceedings of the 6th International Heat Transfer Conference, Toronto: Canada, 1978; p 235.
32. Schmitt, G.; Buecken, W.; Fanebust, R. *Corrosion* **1992**, *48*(5), 431–440.
33. Schmitt, G.; Gudde, T.; Strobel-Effertz, E. In *Fracture Mechanical Properties of CO_2 Corrosion Product Scales and Their Relation to Localized Corrosion,* Corrosion'96; NACE International: Houston, TX, 1996; Paper no. 9.
34. Schmitt, G.; Mueller, M.; Papenfuss, M.; Strobel-Effertz, E. In *Understanding Localized CO_2 Corrosion of Carbon Steel from Physical Properties of Iron Carbonate Scales,* Corrosion'99; NACE International: Houston, TX, 1999; Paper no. 38.
35. Schmitt, G.; Bosch, C.; Pankoke, U.; Bruckhoff, W.; Siegmund, G. In *Evaluation of Critical Flow Intensities for FILC in Sour Gas Production,* Corrosion'98; NACE International: Houston, TX, 1998; Paper no. 46.
36. Schmitt, G.; Bosch, C.; Bauer, H.; Mueller, M. In *Modelling the Drag Reduction Effect of CO_2 Corrosion Inhibitors,* Corrosion'2000; NACE International: Houston, TX, 2000; Paper no. 2.
37. Schmitt, G. In *Drag Reduction – A Widely Neglected Property of Corrosion Inhibitors,* Proceedings of the 9th European Symposium Corrosion Inhibitors (9 SEIC), Ferrara/Italy, Ann. Univ. Ferrara, N.S., Sez. V, Suppl. no. 11, 2000; pp 1089–1122.
38. Schmitt, G.; Bosch, C.; Mueller, M.; Siegmund, G. In *A Probabilistic Model for Flow Induced Localized Corrosion,* Corrosion'2000; NACE International: Houston, TX, 2000; Paper no. 49.
39. Schmitt, G.; Werner, C.; Bakalli, M. In *Evaluation of Local Energy Densities in Disturbed Flow: A New Approach to Characterize Inhibitor Efficiencies to Mitigate Erosion Corrosion*, Proceedings of European Conference, EUROCORR'2004, Nice, France, Sept 12–16, 2004, European Federation of Corrosion, The Institute of Materials: London, 2004.
40. Schmitt, G.; Werner, C.; Bakalli, M. In *Fluid Mechanical Interactions of Turbulent Flowing Liquids with the Wall – Revisited with a New Electrochemical Tool,* Corrosion'2005; NACE International: Houston, TX, 2005; Paper no. 344.
41. Deslouis, C.; Tribollet, B.; Cromolet, R.; Vlachos, G. *La Houile Blanche,* (in French), **1987**, *6*, 459–467.
42. Onorato, M.; Osborne, A. R.; Serio, M.; Bertone, S. *Phys. Rev. Lett.* **2001**, *25*, 5831–5834.
43. Osborne, A. R.; Onorato, M.; Serio, M. *Phys. Lett. A* **2000**, *275*, 386–393.
44. Hering, E.; Martin, R.; Stohrer, M. *Physic for Engineers*, 8th ed.; Springer-Verlag: Berlin, 2002; Chapter 5, pp 386–390.
45. Schmitt, G.; Bakalli, M. In *Quantification of Maximum Energy Densities Between Flowing Liquids and Solid Walls – Effect of Flow Improving Additives*, Proceedings of European Conference, EUROCORR'2006, Maastricht, The Netherlands, Sept. 24–28, 2006, European Federation of Corrosion, The Institute of Materials: London, 2006.
46. Schmitt, G.; Bakalli, M. In *A Critical Review of Measuring Techniques for Corrosion Rates under Flow Conditions,* Corrosion'2006; NACE International: Houston, TX, 2006; Paper no. 593.

47. Dawson, J. L.; Shih, C. C. *Corrosion'90*; NACE International: Houston, TX, 1990; Paper no. 21.
48. Schmitt, G.; Bruckhoff, W. In *Relevance of Laboratory Experiments for Investigation and Mitigation of Flow Induced Corrosion in Gas Production,* Corrosion'88; NACE International: Houston, TX, 1988; Paper no. 357.
49. Matsumura, M.; Oka, Y.; Okumoto, S.; Furaya, H. In *Jet-in-slit Test for Studing Erosion Corrosion, Laboratory Corrosion Tests and Standards, ASTM STP 866*; American Society for Testing and Materials: Philadelphia, PA, 1985; pp 358–372.
50. Schmitt, G.; Al-Janabi, Y. T.; Plagemann, P.; Bakalli, M. *Can the Gas-Pulsed Impinging Jet Simulate Slug Flow?* Proceedings of European Conference, EUROCORR'2003, Budapest, Hungary, Sept 28–Oct 2, 2003, European Federation of Corrosion, The Institute of Materials: London, 2003.
51. Levich, V. G. *Physicochemical Hydrodynamics*, Prentice-Hall: Englewood Cliffs, NJ, 1962.
52. Newman, J. S. *Electrochemical Systems*, Prentice Hall: Englewood Cliffs, NJ, 1973.
53. Mizushina, T. In *Advances in Heat Transfer*, Vol. 7; Irvine, T. F., Hartnett, J. P., Eds.; Academic Press: New York, 1971, pp 87–161.
54. Gabe, D. R. *J. Appl. Electrochem.* **1974**, *4*, 91–108.
55. Eisenberg, E.; Tobias, C. W.; Wilke, C. R. *J. Electrochem. Soc.* **1954**, *101*(6), 306–319.
56. Silverman, D. C. *Corrosion* **1984**, *40*(5), 220–226.
57. Papavinasam, S.; Revie, R. W.; Attard, M.; Demoz, A.; Sun, H.; Donini, J. C.; Michaelian, K. In *Inhibitor Selection for Internal Corrosion Control of Pipelines: 1. Laboratory Methodologies,* Corrosion'99; NACE International: Houston, TX, 1999; Paper no. 1.
58. ASTM G184, *Standard Practice for Evaluating and Qualifying Oilfield and Refinery Corrosion Inhibitors using the Rotating Cage*; ASTM: West Conshohocken, PA, 2005.
59. Schleich, W.; Feser, R.; Schmitt, G.; Haarmann, S.; Schnier, K. In *Effect of Seawater Chlorination on the Erosion Behaviour of CuNi 90/10*, Proceedings of European Conference, EUROCORR'2007, Freiburg, Germany, Sept 9–13, 2007, European Federation of Corrosion, The Institute of Materials: London, 2007.
60. Schmitt, G.; Bosch, C.; Plagemann, P.; Moeller, K. In *Local Wall Shear Stress Gradients in the Slug Flow regime – Effect of Hydrocarbon and Corrosion Inhibitor,* Corrosion'2002; NACE International: Houston, TX, 2002; Paper no. 224.
61. Jepson, W. P.; Zhou, S. In *Experimental Study of Multiphase Slug Flow Characteristics and their Effect on Corrosion in Pipelines,* Corrosion'92; NACE International: Houston, TX, 1992; Paper no. 81.
62. Giralt, F.; Trass, O. *Can. J. Chem. Eng.* **1975**, *53*, 505–511.
63. Giralt, F.; Trass, O. *Can. J. Chem. Eng.* **1976**, *54*, 148–155.
64. Schmitt, G.; Möller, K.; Werner, C.; Plagemann, P.; Deslouis, C.; Bosch, C.; Schöning, M. J.; Janabi, Y. T.; Belghazi, A. In *Microturbulence Intensities in Disturbed Flow Regimes Quantified by Micro-electrochemical Measurements,* Proceedings of the 53th Annual Meeting on International Society of Electrochemistry (ISE), Electrochemistry in Molecular and Microscopy Dimensions; Düsseldorf, Germany, Sept. 15–20, 2002.
65. Nakoryakov, V. E.; Kashinsky, O. N.; Kozmenko, B. K. In *Measuring Techniques in Gas–Liquid Flows*; Delhaye, J. M., Cognet, C., Eds.; Springer: Heidelberg, 1984; pp 695–721.
66. Toth, K.; Štulik, K.; Kutner, W.; Feher, Z.; Lindner, A. *Pure Appl. Chem.* **2004**, *76*(6), 1119–1138.
67. Schmitt, G.; Bakalli, M.; Hoerstemeier, M. In *Contribution of Drag Reduction to the Performance of Corrosion Inhibitors in One- and Two-Phase Flow,* Corrosion'2007; NACE International: Houston, TX, 2007; Paper no. 615.
68. Lotz, U.; Postlethwaite, J. *Corros. Sci.* **1990**, *30*, 95.
69. API-RP 14E. *Recommended Practice for Design and Installation of Offshore Production Platform Piping Systems*, 5th revised ed.; American Petroleum Institute: Washington, DC, 2000.
70. Smart, J. S., III *A Review of Erosion Corrosion in Oil and Gas Production,* Corrosion'90; NACE International: Houston, TX, 1990; Paper no. 10.
71. BS EN ISO 13703. *Petroleum and Natural Gas Industries – Design and Installation of Piping Systems on Offshore Production Platforms*; 2001.
72. Efird, K. D. *Corrosion* **1973**, *33*(1), 3–8.
73. Sick, H. *Werks. Korros.* **1972**, *23*, 12.
74. Ruedisueli, R. L.; Aylor, D. M. *Erosion Resistance Testing for Candidate Seawater Valve Materials,* Corrosion'2001; NACE International: Houston, TX, 2001; Paper no. 468.
75. Weber, J. VDI-Report No. 365.1980, 73.
76. Schmitt, G. In *Hydrodynamic Limitations of Corrosion Inhibitor Performance,* Proceedings of the 8th European Symposium on Corrosion Inhibitors (8 SEIC); Ferrara/Italy; Ann. Univ. Ferrara, N.S., Sez. V, Suppl. N.10,, 1995; pp 1075–1099.
77. Schmitt, G. In *Corrosion Inhibitors in the Mirror of the Ferrara Conferences,* Proceedings of the 10th European Symposium on Corrosion Inhibitors (10 SEIC); Ferrara/Italy; Ann. Univ. Ferrara, N.S., Sez. V, Suppl. N.12, 2005.
78. Heymann, F. J. In *ASM Metals Handbook: Friction, Wear and Lubrication Technology,* Vol. 18; ASM: Metals Park, OH, 1992, pp 221–232.
79. Tullis, J. *Hydraulics of Pipelines: Pumps, Valves, Cavitation, Transients*; Wiley: New York, 1989; pp 59, 78.
80. Postlethwaite, J.; Nesic, S.; Adamopoulos, G.; Bergstrom, D. J *Corros. Sci.* **1993**, *35*, 627.
81. Zeisel, H.; Durst, F. In *Computation of Erosion-Corrosion Processes in Separated Two-Phase Flows*, Proceedings of Symposium on Flow-induced Corrosion, Fundamental Studies and Industry Experience; Kennelley, K. J., Hausler, R. H., Silverman, D. C., Eds.; NACE International: Houston, TX, 1991; pp 9:1–9:21.
82. Wallace, M. S.; Peters, J. S.; Scanlon, T. J.; Dempster, W. M.; McCulloch, S.; Ogilvie, J. B. In *CFD-based Erosion Modelling of Multi-orifice Choke Valves,* Proceedings of 2000 ASME Fluids Engineering Summer Meeting; Boston, MA, 11–15 June, 2000; Paper No. FEDSM2000–11244.
83. Wallace, M. S. Ph.D. Thesis, University of Strathclyde, Glasgow, 2001.
84. Edwards, J. K. Ph.D. Thesis, The University of Tulsa, 2000.
85. Chen, X.; McLaury, B. S.; Shirazi, S. A. *Comput. Fluids* **2004**, *33*(10), 1251–1272.
86. Barton, N. A. Erosion in elbows in hydrocarbon production systems: Review document, HSE Report 15; TÜV NEL Ltd, Scotisch Enterprise Technology Park, East Kilbride, Glasgow, 2003.
87. Mazumder, Q. H.; Shirazi, S. A.; McLaury, B. S. In *A Mechanistic Model to Predict Sand Erosion in Multiphase Flow in Elbows Downstream of Vertical Pipes,* Corrosion'2004; NACE International: Houston, TX, 2004; Paper no. 662.

2.15 Tribocorrosion

R. J. K. Wood
School of Engineering Sciences, Lanchester Building, University of Southampton, Southampton SO17 1BJ, UK

Abbreviations

AISI American Iron and Steel Institute
ASTM American Society for Testing and Materials
BS Bovine serum
CMP Chemical mechanical polishing
CRA Corrosion resistant alloy
CS Carbon steel
DLC Diamond-like carbon
ECN Electrochemical noise
HVOF High velocity oxy-fuel
KE Kinetic energy
NAB Nickel aluminum bronze
OCP Open circuit potential
PBS Phosphate buffer solution
RCE Rotating cylinder electrode
RDE Rotating disc electrode
SEM Scanning electron microscope
SF Synovial fluid
SWR Specific wear rate
UNS Unified Numbering System
WJE Wall jet electrode

Symbols

a Scaling constant
a_j Jump distance (m)
a_r Asperity radius (m)
A Apparent area of contact (m^2)
A_a Affected area (m^2)
A_{crater} Crater area (m^2)
$\dot{A}_e$ Normalized rate of active surface area (m^2)
A_r Real contact area (m^2)
A_s Surface area of electrode (m^2)
A_{atom} Atomic weight of the dissolving metal during electrochemical corrosion
b_a Anodic Tafel slope (m V decade^{-1})
B Migration kinetic constant ($\alpha zFa_j/RT$)
C Pure static corrosion (mg s^{-1})
C_b Bulk concentration of species (M, mM, mol cm^{-3})
C_{eq} Concentration of soluble ferrous ions at equilibrium (M, mM, mol cm^{-3})
C_k Cutting characteristic velocity (m s^{-1})
C_p Weight of solids per unit volume (kg m^{-1})
C_s System constant
C_v Solids volume fraction
C' Solid free corrosion rate (mg s^{-1})
C'' Corrosion rate under wear-corrosion (mg s^{-1})
C_0 Concentration of species at electrode surface (M, mM, mol cm^{-3})
C_1 Constant
C_∞ Concentration of soluble ferrous ions in the bulk solution (M, mM, mol cm^{-3})
d Particle diameter (m)
da/dn Crack growth rate
D Diffusion coefficient (cm^2 s^{-1}) or average diameter of wear particle (m)
E Pure erosion (mg s^{-1})
$E_{applied}$ Applied potential (V)
E_k Kinetic energy of impacting particle (J)
E_r Erosion rate (mg kg^{-1})
f Contact frequency between interacting surfaces
f_a Stripping coefficient
F Faraday's constant (C or A s mol^{-1})
F_f Friction force (N)
$g_{cfs} \equiv (da/dn)_{cfs}$ Maximum attainable crack growth rate under given corrosion fatigue conditions
$g_r \equiv (da/dn)_r$ Average crack growth rate during the generation of a wear debris particle in the reference environment
$g_s \equiv (da/dn)_s$ Crack growth rate in the corrosive environment under sustained stresses
G Constant
h Wear depth (m)
H Hardness (Vickers number)
H^* Hardness with anodic dissolution (Vickers number)
i Current density (A m^{-2})
i_a Activation-controlled current density or anodic current density passing through apparent area (A m^{-2})
i_{act} Activated current density (A m^{-2})
i_{corr} Corrosion current density (A m^{-2})
i_d Diffusion-controlled current density (A m^{-2})
i_r Average anodic current density passing through the real area of contact (A m^{-2})
i_p Passive current density (A m^{-2})
i_{peak} Peak current density (A m^{-2})
i_s Anodic current density for steady state corrosion on the passivated metal surface (A m^{-2})
i_{th} Threshold current density to induce synergy (A m^{-2})
i_u Unaffected current density (A m^{-2})
I Anodic or cathodic current (A)
I_a Anodic current (A)
I_{ac} Activated current (A)
$I_{a,w}$ Anodic current under sliding wear (A)
I_p Passive current (A)
I_{peak} Peak current about I_p (A)
I_{lim} Limiting current (A)
i_0 Maximum current density on fresh metal surface at the start of repassivation (A m^{-2})
J Flux of species (mol s^{-1} cm^{-2})
k Dimensionless wear coefficient
k_a Anodic rate constant
k_B Rate constant
k_d Generation rate of wear particles
k_1 Forward reaction rate
k_{-1} Backward reaction rate
k_m Mass-transport coefficient (cm s^{-1})
K Constant
K_a Proportionality constant
K_w Proportionality constant
K_1 Constant
l Length of wear track (m)
L Characteristic length or sliding distance (m)
L_{ox} Oxide layer thickness (m)
m Mass of oxide removal (kg) or energy exponent
M_{ox} Mass of oxide (kg)
M_p Mass of sand impacting the surface per unit time (kg s^{-1})
M_R Metal dissolution rate (mol cm^{-2} s^{-1})
M_w Oxide molar mass (g mol^{-1})
n Velocity ratio exponent or bulk concentration exponent or number of contact cycles at a given crack site
n_f Scratch overpotential (V)

N_c Number of critical microcracks on a unit area of the wear surface that are ready to propagate
N_r Number of critical microcracks that are ready to propagate to generate wear debris in a reference environment
ΔN_c Increment of the number of critical microcracks due to the presence of corrosion
N Number of erodent impacts
$N_{o,u}$ Total number of surface sites available
N_u Number of moles of oxide formed (mol)
p Pressure (Pa)
q Charge passed $= \int I\,dt$(C)
Q Wear volume loss (m^3)
r_p Particle radius (m)
R Universal molar gas constant (= 8.31451 $J\,K^{-1}\,mol^{-1}$)
Re Reynolds number
R_f Roundness factor for particle (value 0–1)
s Total sliding distance (m)
S Synergy (mg or $mg\,s^{-1}$)
Sc Schmidt number
Sh Sherwood number
t Time (s)
t_c Critical time for crack to propagate a length D (s)
t_0 Time to produce a scratch (s)
T total wear or erosion loss ($mg\,s^{-1}$)
T_s Temperature of the surface (°C)
U Free stream velocity or relative sliding velocity ($m\,s^{-1}$)
U_{cr} Critical velocity for plastic deformation ($m\,s^{-1}$)
U_f Velocity component normal to surface ($m\,s^{-1}$)
U_p Particle velocity ($m\,s^{-1}$)
ν^* Friction velocity ($=\sqrt{(\tau_w/\rho)}$) ($m\,s^{-1}$)
ν_e Sliding velocity ($m\,s^{-1}$)
V Volume loss (m^3)
V_p Particle volume (m^3)
V_u Erosion rate (μm^3 per impact)
V_c Material loss due to 'pure' corrosion without the effect of mechanical wear (m^3)
V_{cw} Total material loss during corrosion wear due to both corrosion and wear (m^3)
V_w Pure wear loss in otherwise the same environment as in the corrosion wear process but without corrosion (m^3)
V_{wc} Material loss due to wear during corrosion wear (m^3)
W Normal load (N)
y Vertical height (m)
y^+ Dimensionless scale
z Number of electrons, average valance, charge on migrating species

α Angle of impact (°)
α_1 Symmetry constant
δ Boundary layer thickness or repassivated oxide thickness (m)
ΔC Change in concentration of species ($=(C_b - C_0)$) (M, mM, $mol\,cm^{-3}$)
ΔC Enhancement of C due to the presence of erosion (mg or $mg\,s^{-1}$)
ΔE Synergistic effect or enhanced erosion rate due to corrosion (mg or $mg\,s^{-1}$)
ΔH Change in hardness (Hv or Pa)
ΔV_c Wear-induced corrosion (m^3)
ΔV_s Synergistic effect (m^3)
ΔV_w Corrosion-induced wear (m^3)
$\Delta\varphi$ Potential difference across oxide (V)
ε_c Critical plastic strain ($\mu m\,m^{-1}$)
η Turbulent eddy viscosity (Pa s)
θ Surface coverage fraction
κ Specific wear rate ($mm^3\,N^{-1}\,m^{-1}$)
ρ_f Density of fluid ($kg\,m^{-3}$)
ρ_{ox} Density of oxide ($kg\,m^{-3}$)
ρ_p Density of particle ($kg\,m^{-3}$)
ρ_T Density of target material ($kg\,m^{-3}$)
σ Dynamic plastic flow stress for target ($N\,m^{-2}$)
τ Passive film recovery time (s)
τ_w Wall shear stress ($N\,m^{-2}$)
τ_0 Characteristic passivation time of a metal during repassivation from a fresh surface (s)
τ_i Characteristic reaction time for the reaction controlling a corrosion fatigue process (s)
ν Kinematic viscosity ($m^2\,s^{-1}$)
μ Dynamic viscosity (Pa s) or coefficient of friction

2.15.1 Introduction and Definitions

The term 'tribo' has origins in the Greek word 'tribos' (rubbing), which led to the sciences of lubrication, friction, and wear being called 'Tribology,' while the term corrosion has its origins in the medieval Latin word 'corrodere,' which means to gnaw through with 'cor' meaning intensive force and 'rodere' to gnaw.[1] For this chapter, corrosion will be used to relate to surface degradation involving electrochemical and chemical processes. Therefore, the term tribocorrosion refers to the surface degradation mechanisms when mechanical wear and chemical/electrochemical processes interact with each other. Many of the mechano–electrochemical interactions involved in tribocorrosion are time dependent and nonlinear. Recent activity in tribocorrosion research aims at addressing

the need to select or design new surfaces for future equipment as well as minimize the operating costs and extend the life of existing machinery and medical devices. The subject therefore includes the interaction of corrosion and erosion (solids, liquid flow, and droplet impingements or cavitation bubbles), abrasion, adhesion, fretting, and fatigue wear processes. Tribocorrosion is often linked to the synergy or antagonism resulting from the coupling of mechanical and environmental effects. For example, during friction the adhesive dissipation of energy is often influenced by chemical effects.[2] Similarly, stainless steels rely on a 1–10 nm thick passive film for their protection from aggressive and corrosive environments. This film forms essentially instantaneously by reaction with the environment but abrasion can lead to a local rupture or complete removal of the film and permanent deformation of the substrate. This can lead to areas of the substrate being exposed to the aggressive environment and unless repassivation (repair or self-healing) mechanisms reform the passive film to inhibit corrosion processes, accelerated anodic dissolution will occur within these sites.

The subject area does, however, suffer from rather loosely defined terms that sometimes relate to the same processes. Thus, to help readers, the terms used in this chapter and their definitions are given in the following paragraphs, but some overlap between these terms is inevitable.

The term flow-corrosion (sometimes referred to as flow-enhanced corrosion, flow-induced corrosion, or flow-accelerated corrosion (FAC) in some texts) will be used to describe corrosion that occurs on surfaces exposed to laminar or turbulent flow but is unable to disrupt passive surface films or corrosion layers.

Erosion-corrosion will be used where mechanical processes as well as flow effects are acting on the surface. This covers solid particle, liquid droplet, cavitation impingement, and some highly turbulent flows where energetic jetting or impingement flows onto the surface occur and can disrupt surface films or layers. Surface film disruption is by mechanical processes associated with contact stresses, stress wave reinforcement, surface shear stresses or pressure fluctuations generated on the surface through water hammer, dynamic fluid pressures, transverse momentum transfer, etc. These mechanical processes can lead to increased corrosion through electrochemical and chemical processes. Reference is normally made to the critical flow velocity above which erosion-corrosion occurs for specific flow geometries and materials.

Tribocorrosion will be used to cover the much broader area or all wear and corrosion interactions. This will cover, for example, abrasion-corrosion and sliding wear-corrosion. Wear is the term used to cover abrasive, adhesive, oxidative, fretting, fatigue-based processes that are part of mechanically induced surface degradation and material loss mechanisms.

Understanding tribocorrosion processes is important if costs associated with material losses are to be minimized. For example, modern fluid handling and propulsion systems face increasing demands for higher flow rates with the inherent risk of flow-corrosion, and if solids are entrained and impinge on the surface or cavitation occurs near the surface, erosion-corrosion can result and accelerate material loss significantly. This is especially true for industries that transport slurries and other particle-laden fluids in pipes or seawater propulsion systems such as offshore and marine technologies. These industries expend the equivalent of millions of pounds every year to repair erosion-corrosion damage caused by solid particle impingement and cavitation. Typical examples of this kind of material destruction are erosion-corrosion damage to pump impellers (see **Figure 1**), propellers, valves, heat exchanger tubes, pneumatic and hydrotransport systems, fluidized bed combustors, and many other types of fluid-handling equipment. The erosion-corrosion data and models published in the open literature aspire to allow informed surface selection for combined erosion and corrosion resistance. They are rarely based on substantive physical understanding and some approaches are simply designed to make numerical modeling easier. The erosion-corrosion data are very patchy and incomplete, and therefore hamper any improvement in the performance of machines in aggressive environments.

Figure 1 Cavitation-corrosion of a 150 mm diameter bronze centrifugal pump impeller from a swimming pool pump.

Tribocorrosion is also critical to long-term human health and patient mobility. The correct understanding of biotribocorrosion and joint design can offer huge savings for health services and impact on the quality of life. For example, current predictions suggest 268 000 total knee arthroplasty revisions will be required per annum in the United States by 2030 as against 38 000 now.[3] Estimating a cost of $30k per revision, this implies an additional annual cost of $6.9 billion at today's costs. Extending joint life through improved tribological design would therefore allow enormous savings in surgery costs besides allowing patients to remain active and in employment. Hence, there is growing need to understand the human cellular and biological reactions to metal ion release generated by tribocorrosion processes as well as surface biotribology performance of these joint implants.

On the positive side, tribocorrosion phenomena can be used as a manufacturing process such as in the chemical–mechanical polishing of silicon wafers. The coupling of mechanical and environmental effects can also create surfaces of specific reaction layers on materials that could inhibit corrosion and/or wear. Examples of this are self-lubricating and/or self-healing surface layers.[4]

The presence of well-adhered oxide surface films or loosely adhered layers of corrosion products and the ability to reform these films and layers are also critical in many engineering applications where control of friction and wear are important in sliding contacts. These contacts involve contact between the high spots on the interacting surfaces due to their roughness after machining or finishing. Therefore, the load on a contact is rarely distributed evenly over the apparent contact area but is rather distributed among high-spot (asperity) contacts between surfaces, often referred to as junctions. The actual contact area between surfaces is therefore the sum of the individual areas of junctions and will likely be proportional to the normal load. Friction and wear generally are developed by interactions between these asperities, which can be subjected to severe contact stresses that often result in temporary flattening (elastically) or, if yield stresses are exceeded, permanent plastic flow and deformation of the surface. The friction forces are typically lowered when oxide layers are present on asperities compared with asperity contacts between bare metal. For some systems, the friction forces can be modeled by summing the forces associated with adhesive and deformation processes relating to these asperities. Adhesive forces are often related to the shear strength of junctions made between contacting asperities of surfaces in relative motion and relate to the frictional work. The presence of oxide films or corrosion products on these asperities will alter the level of frictional work and the resulting surface temperatures. **Figure 2** illustrates how oxide-filmed asperities may interact during sliding, assuming that the oxides deform with the asperity. The asperities on the two interacting surfaces are represented as idealized cylindrical protrusions (i.e., flattened protrusions), thereby allowing a flat circular contact of radius 'a_p' in plan view, to visualize the actual contact area evolution as asperities slide over each other.

Unpredictable interactions can result between wear and corrosion when the surfaces in contact have complex, multiple-phase microstructures that can lead to microgalvanic activity and selective phase corrosion (localized attack) as well as wear modes involving wear debris or corrosion products along with the interacting surfaces that can roll or slide within the contact. Examples of such surfaces include composites or surfaces that undergo compositional changes induced by tribological interactions. For instance, the presence of carbides in a metallic surface, typically formed for improved wear resistance, establishes a microcorrosion cell as the carbide is likely to be cathodic with respect to the surrounding metallic matrix. This can result in the preferential anodic dissolution of the metallic matrix close to or at the matrix–carbide interface and thereby accelerate carbide removal from surfaces and reduce the antiwear properties of the surface.

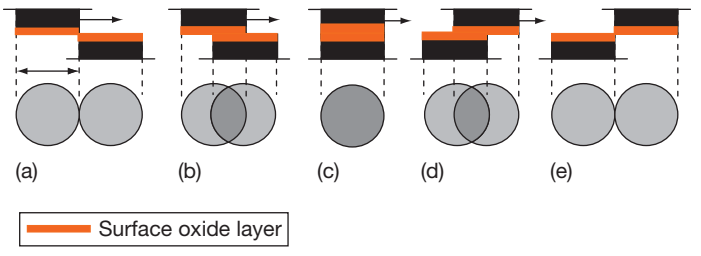


Figure 2 Schematic of asperity–asperity interactions of a passivated surface.

2.15.2 Tribocorrosion Fundamentals

2.15.2.1 Factors Affecting Tribocorrosion

The corrosion dependence on material properties is complex within tribocontacts. Reference to the galvanic series (see **Chapter 2.07, Galvanic Corrosion**) is useful to understand the potential differences possible when a depassivated metal is electrically coupled to a passive metal. This potential difference will drive charge transfer within the contact. This charge transfer is likely to be a function of the magnitude of surface degradation by mechanical processes and how this is affected by the material chemistry, the environmental chemistry, the microstructure, mass transport, and the rate and extent of depassivation, should the surface be passivated. The influence of material parameters on wear is discussed in the following section.

2.15.2.1.1 Material parameters and composition

A review of available wear models such as those used in solid particle erosion shows the diversity of material properties that have been included in published predictive wear equations. Meng and Ludema,[5] for example, have reviewed 28 erosion models and found that on average only five parameters are used per equation but in total 33 different parameters, not all material properties, are quoted. These are tabulated in Table 1.

Simpler approaches based on a few parameters are still in use and are shown to have wide applicability and therefore are useful. For example, the main theory for sliding wear of metals is based on the assumption that contact between two surfaces occurs where asperities contact and the local deformation is plastic (see Figure 2). The Figure shows the interaction between sliding asperties as one slides over the other; see Figures 2(a)–2(e). The true contact area is therefore given by the summation of individual asperity contact areas and is proportional to the normal load.

The overall wear volume loss, Q, per unit sliding distance is associated with the detachment of material from the asperity contacts and the volume of each particle of wear debris will relate to the size of the individual asperity contact. Assuming that the debris volume, δV, is proportional to a_r^3 gives the well-known Archard wear law[6]:

$$Q = \frac{kW}{H} \qquad [1]$$

The Archard wear law suggests, for a system with constant k, that the wear rate is directly proportional to load on the contact but inversely proportional to the surface hardness, H, of the wearing material. It can be written:

$$\kappa = \frac{V}{WL} = \frac{k}{H} \qquad [2]$$

where κ is the dimensional specific wear rate (SWR) (typically quoted as $mm^3\,Nm^{-1}$ or $m^2\,N^{-1}$), V the volume loss, W the applied load, and L the sliding distance.

A similar wear equation to the adhesive Archard wear law (eqn [2]) can be derived for abrasion.[7] It is

Table 1 Parameters selected in wear models[5]

Erodent/abradent	*Target*	*Fluid flow*
Density	Density	Impact angle
Hardness	Hardness	Impact angle max. wear
Moment of inertia	Flow stress	Kinetic energy (KE) transfer between particle and target
Roundness	Young's modulus	Temperature
Single mass	Fracture toughness	Corrosivity
Size	Critical plastic strain	
Velocity	Depth of deformation	
Rebound velocity	Incremental strain per impact	
KE of particle	Thermal conductivity	
	Melting temperature	
	Enthalpy of melting	
	Cutting energy	
	Deformation energy	
	Heat capacity	
	Weibull flaw parameter	
	Lamé constant	
	Grain diameter	

also interesting to note that models developed for predicting current from depassivated surfaces show a $(W/H)^{1/2}$ dependence, as discussed later in Section 2.15.3.1.3.

Turning now to solid particle erosion, it is tempting to assume that solid particle erosion rates would also be inversely dependent on the surface hardness, H, as predicted by simple plastic deformation erosion models, eqn [3], where α is the erodent impact angle. A more complex model, also given by Hutchings,[8] includes the critical strain concept, eqn [4] and this predicts the erosion should be dependent on $H^{3/2}$. However, these relationships rarely hold in practice for engineering materials.

Erosion models

$$E_r = \frac{K\rho_T U_P^n}{H} f(\alpha) \qquad [3]$$

$$E_r = \frac{K_1 \rho_T \rho_P^{1/2} U_P^3}{\varepsilon_c^2 H^{3/2}} \qquad [4]$$

where: K is a constant, ρ_T is the density of the target, ρ_P is the density of the particle, U_P is the particle velocity, and ε_c is the critical plastic strain.

Corrosion resistant alloys (CRAs) are widely used to resist flow corrosion. For example, chromium is added to improve the resistance to flow corrosion of carbon steel (CS), see Poulson.[9] Figure 3 shows the trend for such a passive system (where passive means that the surface has a propensity to form a protective surface film or layer with high impedance and thus suppresses charge transfer at the metal–liquid interface) as a function of flow velocity and its vulnerability to mechanical processes, which accelerate corrosion loss rates. In general, corrosion-resistant alloys do not resist erosion well and the interactions that exist between corrosion and erosion are not well understood. These interactions are sometimes referred to as synergistic or antagonistic. These terms refer to the possibility that simple summing of the mechanically induced material loss without a corrosive contribution and the electrochemical (corrosion) material loss without mechanical stimulus does not give an accurate material loss for tribocorrosion conditions. Some surface–environment conditions suffer accelerated material loss where the actual loss rates are higher than the simple summation of the pure mechanical and electrochemical levels. This case is termed 'synergistic,' while combinations of surface/environment that result in reduced material loss and are below the levels of the simple summation are termed 'antagonistic' (negative synergy). This presents a dilemma for those seeking erosion-corrosion-resistant surface selections. Also, cost reduction considerations favor replacing expensive solid alloy components with coatings on CS substrates, and hence, considerable effort is now being made to research coating performance under erosion-corrosion.

2.15.2.1.2 Environment factors

Corrosion is a reaction process on metallic surfaces that occurs as a result of interaction with the environment. The rates and modes of corrosion attack are

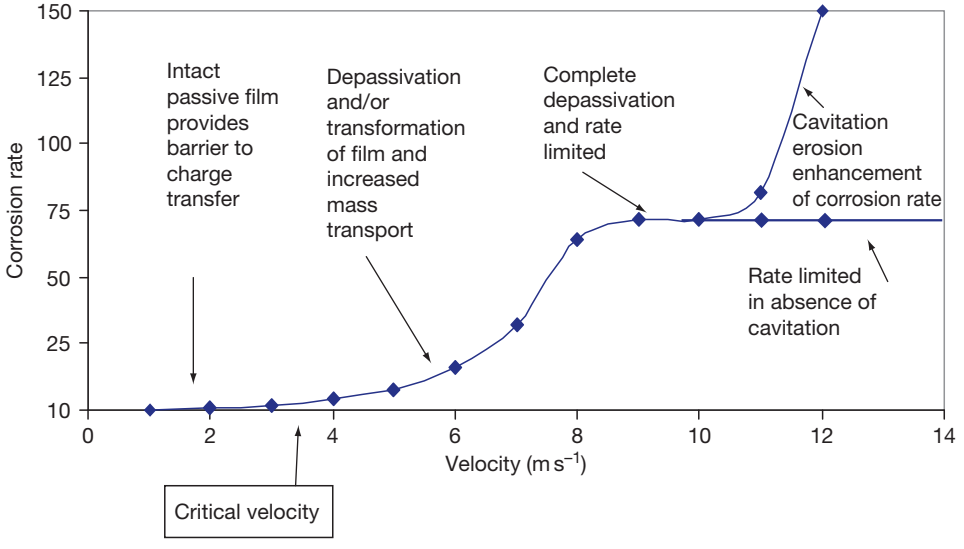


Figure 3 Arbitrary flow corrosion rates versus flow velocity.

related to the environmental conditions such as pH, dissolved oxygen, salinity, temperature, flow structure, and whether biofouling occurs. The latter point is often ignored but can have dramatic effects.

The influence of environmental factors on erosion-corrosion is likely to be very system dependent. However, the sensitivity of wall wastage rates to such environmental parameters can be found in the literature. Kim *et al.*[10] report how these parameters effect the galvanic corrosion of CS in alkaline-chloride solutions. They studied the aqueous corrosion characteristics of a CS coupled to type-304 (UNS S30400) stainless steel in deaerated solutions as a function of velocity, pH, and temperature. They found that the galvanic corrosion between CS and type-304 stainless steel was effected by the presence of an oxide layer formed on the CS. The galvanic current density increased with increasing flow velocity (between 0 and 0.6 m s^{-1}) and temperature (between 25 and 75 °C), but decreased with increasing pH from 8 to 10. Flow velocity effects were only significant at 50 and 75 °C. This effect can be explained by the increased solubility of magnetite (corrosion oxidation product) at the higher temperatures. The importance of passive films and their stability will be addressed in more detail later in this chapter.

Proteinaceous material present in natural joint lubricants can play an important influence within biomedical implants, particularly on metallic joint surfaces. Research has shown that proteins have different affinities for different metals, probably as a result of various ligands in their molecular structure that have different bonding energies and therefore different affinities for specific cations.[11] The effect of albumin (a protein found in serum) on the corrosion rate of biometallic materials is varied with Al and Ti being unaffected, Cr and Ni showing a slight increase, while Co and Cu dissolve at a much greater rate.[12] Williams *et al.*[13] found that the corrosion rate of stainless steel and commercially pure titanium increased when exposed to proteins in a static environment, while under fretting contact conditions, proteins were found to decrease the corrosion rate of the stainless steel but had little appreciable effect on titanium alloys. In a recent study on the effect of synovial fluid (SF) on corrosion properties of CoCrMo alloys,[14] proteins were found to be absorbed on the surface. It was hypothesized that protein (normally negatively charged) adsorbs to generate a film that reduces the corrosion rate by acting as a negatively charged barrier preventing corrosive anions (e.g., Cl^-) approaching the CoCrMo interface. Contu *et al.*,[15] on the other hand, attributed the reduction in corrosion rate of CoCrMo in bovine serum (BS) to the anodic dissolution process being under diffusion control resulting in positive shifts in surface potentials compared with inorganic solutions at the same pH. Goldberg *et al.*[16] carried out a series of scratch tests on CoCrMo (ASTM F75) samples and found a decrease in peak currents (decrease in corrosion rate). The authors attributed this to the barrier effects of the adsorbed protein preventing water from reaching the sample surface, or lubrication resulting in less material removed from the surface during scratching.

A guide to system performance for solutions with different pH can be found by reference to the relevant potential–pH (Pourbaix) diagram. This approach has been illustrated by Stack[17] in a regime mapping approach to erosion-corrosion and by Celis[18] for sliding wear-corrosion contacts.

The influence of the bulk oxygen concentration, $C_{b,}O_2$, in aqueous environments on the corrosion rate depends on the controlling corrosion mechanisms, which will be discussed in greater detail later in the chapter. The dramatic influence of oxygen levels on fretting wear-corrosion rates of a type-304L stainless steel sleeve fretting against a 304L tube in simulated nuclear plant conditions using boric acid with Lithine (an inhibitor) at a pH of 8.7 is illustrated in **Figure 4** from Deforge *et al.*[19] It is also interesting to see that most tribocorrosion occurs on the sleeve for systems

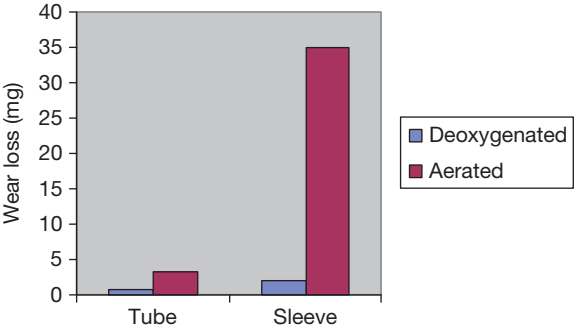


Figure 4 Fretting wear loss of stainless steel type-304L tubes in contact with a stainless steel type-304L sleeve after 1 week in boric acid + lithine, pH = 8.7 at room temperature. The OCPs for the deaerated (overpressure of hydrogen) and aerated conditions were −0.630 and −0.030 V versus SHE, respectively. Reproduced from Deforge, D.; Wenger, F.; Ponthiaux, P.; Lina, A; Ambard, A. Tribocorrosion of AISI 304L stainless steel sliding against AISI 304L stainless steel in nuclear environment, 207th Electrochemical Society Meeting, Quebec City, Canada May 2005, session E1 on Chemical, Electrochemical, and Mechanical Effects on CMP, Tribocorrosion, and Biotribocorrosion, abstract #313.

exposed to flow, and under mass transport control, the corrosion rate can be found to be directly proportional to $C_{b,}O_2$, or $C_{b,}O_2{}''$, where n is flow and system dependent.

Temperature of the surfaces is an important variable which can influence wear-corrosion rates and hence the temperature of the environment, be it air or lubricant, must be taken into account. In addition, many contacts between surfaces in relative motion result in an increased heating of the contact zones due to friction. In aqueous solutions, temperature not only influences the pH, solubility of gases (i.e., oxygen and carbon dioxide), and ionic species as well as diffusion rates of reactant species but can also affect the kinetics of oxidation reaction at the oxide–metal interface and the dissolution of oxides at the oxide–water interface. Added complications can occur when elevated temperatures increase hydrogen diffusion rates, and if this effect is coupled with a surface susceptibility to hydrogen embrittlement, then increased wall loss can result.

In nuclear power plant, the effect of water chemistry has been shown to strongly influence the wear rates of 304L stainless steel in pressurized high temperature water. The high temperature of the water promotes thicker oxide layers that are harder due to the formation of micron-sized magnetite crystallites. If these layers are detached and stay within the contact, they can create very abrasive (rolling abrasion where entrained oxide particles roll and indent the contact surfaces) wear conditions. For low-pH waters between 5 and 7, Kaczorowski[20] reports dramatic increases in material loss by the promotion of such three-body mechanisms and by decreasing the electrochemical potential in the electronegative direction.

Another major environmental factor with significant influence on erosion-corrosion rates is that of flow (see **Chapter 2.13, Flow Assisted Corrosion**), but care needs to be taken to use appropriate flow characterization parameters. For example, a single value of flow velocity referred to as the critical velocity is often quoted to represent a transition from flow corrosion to enhanced mechanical-corrosion interactive erosion-corrosion processes. It is also used to indicate the resistance of the passive and protective films to mechanical breakdown.[21] However, it should be remembered that a single value of free stream velocity can result in widely different flow field structures that are dependent on the Reynolds number of the flow geometry and type and thickness of boundary layer induced at the liquid–solid interface. These factors will directly affect the velocity profiles and transverse momentum transfer close to the solid–liquid interface, which will be shown later to influence the thickness of the boundary and subviscous layers and thereby dictate wall shear stress levels and mass transport efficiencies. Therefore, critical velocity values are very geometry or test rig specific and can not be readily applied to predict component service life in generic flow systems.

Poulson[9] has investigated the effects of flow on corrosion of CS and the effects of depassivation of the surface (see **Figure 5**). The corrosion of CS in neutral solutions is governed by the reduction of dissolved oxygen as shown in eqn [5].

$$O_2 + H_2O + 4e^- \rightarrow 4OH^- \qquad [5]$$

This cathodic electrochemical reaction drives the anodic dissolution of iron, in eqn [6].

$$Fe \rightarrow Fe^{2+} + 2e^- \qquad [6]$$

Poulson reported that the flow corrosion rate was closely related to the mass-transport coefficient. However, under detached/separated flow conditions (such as flows through sudden expansion pipe work), the rate can be independent of flow.

Al-Hosani *et al.*[22] have studied the galvanic corrosion induced by the coupling of copper-based alloys to Mo-stainless steels in static Arabian Gulf seawater under a variety of conditions including temperatures, cathode–anode area ratios and stirring. For static conditions, the corrosion rates were found to be proportional to the cathode–anode area ratio. This can be attributed to the fact that the cathodic reaction (the reduction of oxygen) controls the overall corrosion process, and this is diffusion controlled. However, the galvanic corrosion rate of the couples increased nonlinearly with a rise in temperature to reach a maximum at 50 °C, decreasing sharply at higher temperatures due to the decreased solubility of O_2 in solution at this temperature. The effect of stirring the electrolyte shifted the open circuit

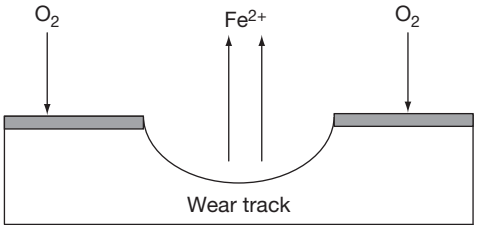


Figure 5 Corrosion of carbon steel (CS) activated by wear that depassivates surface in the wear track.

potential (OCP) of the stainless steel to a more electronegative value. Thus, the corrosion rates in the stirred solutions were found to be 1.33–3.0 times above those measured in stagnant solutions.

Erosion-corrosion conditions can themselves develop local microgalvanic cells between eroded and uneroded surface areas. Such microgalvanic cells have been reported between erosion-affected regions (depassivated) of a surface and noneroded regions (passive). For example, enhanced material loss rates due to such galvanic cells are reported by Hodgkiess.[23]

The environment can change the wear mechanism and thereby the wear rate and life of surfaces. This is illustrated by the tribological properties of plasma-sprayed Cr_2O_3 coatings, studied by Wei,[24] at room temperature, using distilled water, water–ethanol, *n*-monohydric alcohols, and polyhydric alcohols as lubricants. Testing indicated that both friction and wear were much higher (approximately twice) in water than in the *n*-monohydric alcohols. Analysis of the morphology and composition of the worn surfaces indicated that the wear of Cr_2O_3 coating was controlled by the combination of microfracture and tribochemical wear in water and was only controlled by microfracture in ethanol.

2.15.2.1.3 Friction

Friction is the resistance force encountered when surfaces in contact move relative to each other. A popular way to express this force in terms of a loaded sliding contact is by using the dynamic coefficient of friction, μ, defined as the frictional force, F_f, divided by the normal load, W, on the contact when the contact is sliding.

$$\mu = \frac{F_f}{W} \qquad [7]$$

For lubricated contacts in relative motion separated by a film of lubricant, the dynamic coefficient of friction, μ, is typically very low (<0.1) and can be as low as 0.1, even when the lubricant film is only a few nanometers thick. These values compare with values of $\sim$1.0 for unlubricated dry sliding. Friction between contacting surfaces in relative motion results primarily from forces acting perpendicular to the applied load. As the fluid film thins, the interaction of surface asperities can cause increases in friction. Water, however, is regarded as a poor lubricant, as it does not form a fluid film readily or support loads due to a minimal pressure–viscosity interaction. However, water-lubricated contacts reduce the contact temperature compared with dry contacts and aqueous lubricants are being considered as an environmentally acceptable alternative to hydrocarbon-based oils. It should be noted that the friction force required to be overcome to induce relative motion in a contact can be used in eqn [7] and μ is then termed the static coefficient of friction and is typically higher than the dynamic values.

Bowden and Tabor[25] proposed a model, which in its simplest form assumes that frictional forces arise from adhesive forces and deformation forces induced by the abrasive ploughing nature of the harder asperities interacting with the softer ones of the counter surface. The adhesive force is linked to the asperity–asperity contact, which makes up the real area of contact between surfaces and the junctions between them. The shear strength of these junctions influences the level of friction. Adhesive effects are thought to be due to the summation of interfacial intermolecular interactions that operate at the asperity contacts. There are also likely to be some chemical processes activated within the contact, such as oxidation, which can form coherent films and their composition can influence friction levels. The presence of oxide layers and adsorbed films on metal surfaces generally weakens the shear strength of the asperity junctions and thus lower adhesive forces resist motion resulting in lower friction. For example, a multilayered film can be formed on ferrous surfaces under dry sliding, typically with FeO close to the surface, which is covered by a layer of Fe_3O_4, which in turn is covered by Fe_2O_3. The Fe_2O_3 layer is associated with high friction ($\mu = 1.1$), but if this layer is penetrated exposing the sublayers, then lower friction ($\mu = 0.5$) can result.[26] However, for soft ductile metals, and where the oxide layers or adsorbed films are partially removed, appreciable adhesion between nascent (freshly formed or exposed material) surfaces can result. Adhesion can also be enhanced if similar surfaces are in contact, that is, stainless steel pairs.

2.15.2.2 Corrosion and Associated Electrochemical Reactions

This section reviews the relevant corrosion reactions related to tribocorrosion and the important role of the surface films formed under tribocorrosion conditions called tribofilms.

2.15.2.2.1 Tribofilms

Published research has tried to deal with the synergistic and antagonistic effects between erosion and corrosion processes, which result in accelerated material loss (synergism) or in some cases actually decelerate material loss (antagonism, sometimes called negative synergism).[27–31] A wide range of corrosion-resistant materials relies on a relatively thin surface film to provide a barrier (of high impedance) to charge transfer (and thus corrosion) between the relatively active bulk material and the corrosive environment. The thickness of such films is dependent on time, temperature, interfacial potential, pH, and the environment (i.e., species concentration and type). The films range from being tens of Angstroms to hundreds of Angstroms thick and are formed in relatively short times, typically <1 s. These films render the surface passive, but for fluid machinery handling flows where solid particles have been entrained or cavitation is induced, the passive film can be removed by mechanical wear or bubble collapse/microjet/shockwave impingement processes. Where the film is mechanically removed, charge transfer can occur at the interface without retardation from the barrier film.[32] The stability of these surface films will depend on their adherence to the parent bulk material and their cohesive strength when subjected to contact stresses. The solubility of the surface films within the changing environment of tribocontacts will also influence the stability and could instigate compositional changes in the surface film. Clearly, density and hardness of the surface films will influence the microcontact mechanics of tribological processes of wear and friction.

Similar mechano–electrochemical processes occur in sliding contacts where tribochemical wear occurs by detrimental chemical reactions within the contact, induced by the local environment, and in combination with the mechanical contact mechanisms and induced stress fields. Contact between the sliding surfaces can accelerate chemical and electrochemical reactions and material removal. Sliding contact results in increased surface temperatures and can induce surface cracks and the generation of nascent surfaces, which are highly reactive with their environment. However, the chemical formation of surface films can be advantageous, as they can have low friction properties and also promote material transfer to change the contact from base–base material to chemical film–chemical film and thus change the contact conditions.

The interaction between the tribofilms formed on interacting surfaces often results in nanometer-thick mechanically mixed layers that are formed in the shear zone between the contacting surfaces. These layers are often a mixture of components from both surfaces generated by material transfer but also contain corrosion products, wear debris, and remnants of any passive film, if the surfaces are passive metallic systems. This layer is often referred to as the tribofilm and its composition and thickness will control the level of tribocorrosion at the contact when sliding wear occurs in a corrosive environment. The environment clearly has an influence on the composition of the film. For example, the environment will determine whether oxides or sulfides can form or whether calcium carbonate (chalk)-based films can form.

Mischler *et al.*[33] show the importance of the passive film in controlling wear rates and friction when testing CS in borate solution (pH 8.4). They reported two concurrent mechanisms: (1) oxidative wear involving the detachment of passive film followed by repassivation; and (2) surface disruption by abrasion, subsurface deformation, and cracking. Interestingly, when held at cathodic potentials, there was no film present and asperity interaction resulted in debris being smeared over the wear track, subsequently decreasing the wear rate.

A mechanistic view given by Staehle[34] looks at the initial interaction of dislocations in the metal substrate with surface films in terms of single and cross-slip and conditions for perforating these films. This approach suggests that the significant properties of initial films in terms of interaction with dislocations are plasticity, strength, stacking fault energy, and epitaxy. Other factors to consider are the duration in which the surface is totally or partially film-free, the kinetics of film-free and repassivation regions, and the dependence of these film-free and repassivation regions on pH, potential, environment, alloy composition, temperature, mass-transport, and contact-induced stress. Staehle shows that film-free areas only represent 10% of the surface area affected, as repassivation occurs very quickly under high-speed/strain elongation tests. It was also found that the reaction rates on the mechanically affected surface areas were one to two orders of magnitude higher than nonaffected areas. This emphasizes the importance of film integrity in minimizing corrosion rates. Staehle concludes that increasing the chromium content or the presence of a Molybdate inhibitor increased the rate of repassivation of steels. However, increasing the concentration of sulphates and the presence chlorides reduced the rate of repassivation.

2.15.2.2.2 Depassivation and repassivation kinetics

Depassivation

Depassivation in tribocontacts involves the rapid removal of the surface layers and sometimes the underlying bulk material by mechanical action such as surface abrasion, adhesion, plastic deformation, cutting, thermal cycling, fatigue and particle erosion, cavitation bubble, or liquid droplet impact mechanisms. For example, Figure 6 illustrates the level of damage to passive films as a function of solid particle impingement angle. Mechanical removal and/or rupture of the passive layer enables charge transfer to proceed at varying rates and parent metal dissolution is likely. 'Recovery' or repassivation aspects of the passive film are therefore important. Figure 7 shows the current response over time for individual solid particle impacts on the naturally passivating system of stainless steel in a NaCl solution. Some impacts will only result in partial passive layer removal or cracking, which will influence repassivation kinetics and possibly the composition of the regrown layer.

Mechanical deformation that disrupts and perforates passive films takes place by dislocation movement. Dislocation interaction with the passive film can lead to rupture and/or perforation that results in a transient exposure of highly reactive parent material. This process is sometimes referred to as 'slip dissolution.' The degree of disruption to the passive film depends on whether the slip is concentrated on a single plane or is diffuse on multiple planes as with cross-slip. An illustration of a passive film disruption by slip in a single plane is shown in Figure 8. As Figure 8 (b) shows, if the displacement is great enough, corrosion occurs under the passive film, which will further compromise the integrity of the film if subjected to further deformation in a tribocontact.

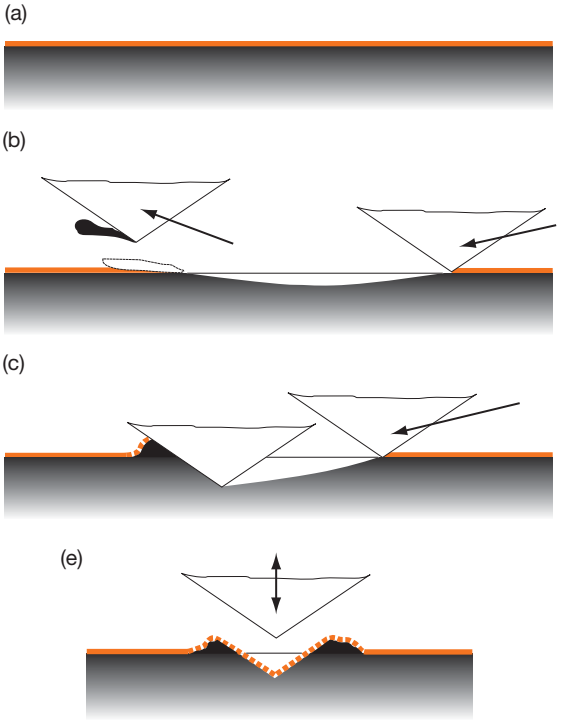


Figure 6 Schematic showing that the level of passive layer damage is dependent on particle impact angle.

Repassivation

Surface coverage models for passive film growth/recovery typically relate to steady state film formation and are not applicable to transient events of <1 s duration, which are often found in tribocorrosion. However, these models are useful and they consider the number of moles of oxide formed N_u as a ratio of the total number of surface sites available, $N_{o,u}$, in units of mol cm^{-2} (from Jemmely *et al.*[35]):

$$\theta = \frac{N_u}{N_{o,u}} \qquad [8]$$

The rate of lateral surface coverage during early stages of repassivation of nascent surfaces, assuming complete oxidation of the parent metal to oxide, can be given by

$$\frac{d\theta}{dt} = i\left(\frac{M_{ox}}{L_{ox}\rho_{ox}zF}\right) \qquad [9]$$

where i is the anodic current density, z is the charge number, L_{ox} is the oxide film thickness, ρ_{ox} is the oxide density, and M_{ox} is the molar mass of the oxide.

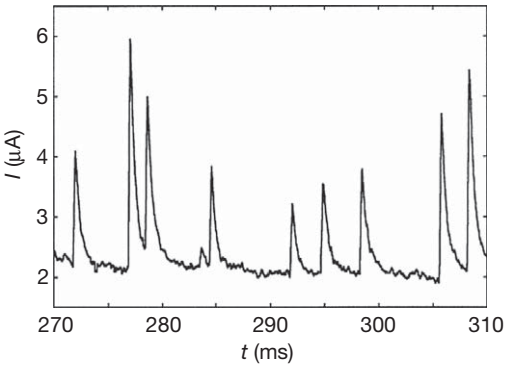


Figure 7 Current transients in erosion–corrosion due to depassivation/repassivation. Reproduced from Sasaki, K.; Burstein, G. T. *Philos. Mag. Lett.* **2000**, *80*(7), 489–493.

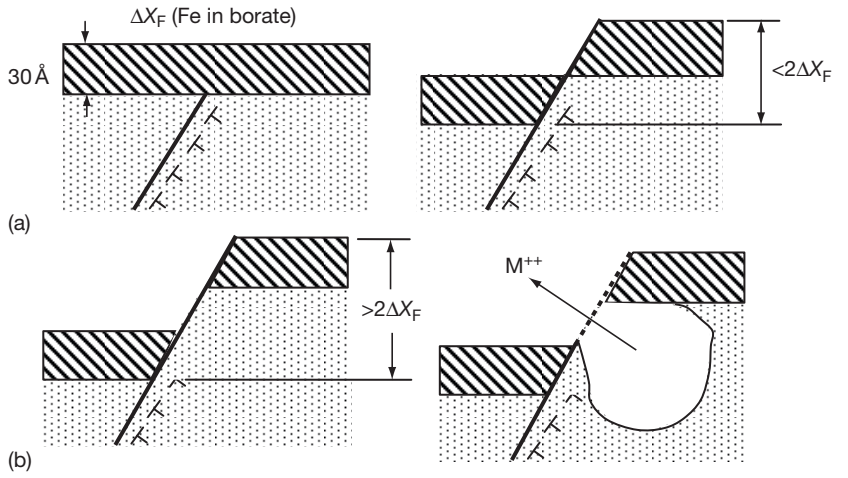


Figure 8 (a) Dislocation on a single slip plane with the surface covered by a thin passive film (b) dislocation is more extensive and film is perforated and corrosion occurs under passive film. Reproduced from Staehle, R. W. *Corros. Sci.* **2007**, *49*, 7–19.

If Tafel kinetics applies then the anodic current density i can be given by

$$i = k_a(1-\theta)\exp\left(\frac{E_{applied}}{b_a}\right) \quad [10]$$

where k_a is the rate constant and b_a is the Tafel slope ($=dE/d\ln\ i$). This assumes that oxidation occurs exclusively on bare parent metal.

Film growth models may then be used to explain the thickening of oxide layers. These models assume growth occurs by high field conduction which is thought to apply over the longer term and are therefore not applicable in the early stages of repassivation of nascent surface sites.

$$i = k_B\exp\left(\frac{B\Delta\varphi}{L_{ox}}\right) \quad [11]$$

where k_B is the rate constant and the migration kinetic constant for a single ionic species $B = \alpha_1 zFa_j/RT_s$. Here α_1 is a symmetry constant, z is the charge on the migrating ionic species, and a_j is the jump distance. The term $\Delta\varphi$ is the potential difference across the oxide layer and taken to be the difference between the electrode potential and the minimum potential for the initiation of film growth. The other term in the equation is the oxide layer thickness, L_{ox}.

This can be linked to the current–time transients by setting:

$$L_{ox} = \frac{M_{ox}q}{zF\rho_{ox}} \quad [12]$$

Where q is the charge passed or $\int i\,dt$.

The use of such models to predict wall loss rates under depassivation–repassivation conditions depends on the assumptions mentioned earlier and knowing the oxidation process and the associated charge number as well as the area of depassivation due to the wear processes. Some of the possible environmental factors that effect depassivation and repassivation are summarized for erosion-corrosion in **Table 2**.

Thus, if the contact conditions allow, the rupture of the passive film can be repaired as shown schematically in **Figures 9–11**. This is an ideal case where accelerated dissolution of the substrate is minimal. Obviously, more severe abrasion would result in accelerated corrosion by inhibiting effective repassivation.

2.15.2.2.3 Critical velocity concept

The velocity profile established between flowing electrolyte and a solid surface, see **Figure 12**, generates a maximum shear stress at the wall–liquid interface.

The wall shear stresses can be calculated from eqn [13].

$$\tau_w = (\mu+\eta)\frac{dU}{dy} \quad [13]$$

where μ is the dynamic viscosity of fluid, η is the turbulent eddy viscosity, and dU/dy is the velocity gradient at the wall. For laminar flows, the turbulent eddy viscosity can be set to zero.

These are stresses that could be applied to loosely adherent surface deposits/layers or passive films and change mass-transport conditions thereby influencing flow corrosion and erosion-corrosion. However, it is now thought that passive film removal and its associated critical velocity is related to mass-transport-induced

Table 2 Possible influences on the depassivation and repassivation kinetics, adapted from Wood[27]

Environmental Factor	*Depassivation processes that cause passive film removal or disruption*	*Repassivation processes that encourage regrowth (repair) of the passive film*
Solid particle impact	Plastic deformation or cutting	Enhanced repassivation on plastically extruded regions
Surface/liquid interface	High velocities/wall shear stresses	Film growth kinetics(monolayer coverage plus lateral film growth)
Occluded geometries/pits	Local environment changes such as low pH	Metastable pits
Contact stress/shear stress/thermal pulse by frictional heat	Destabilises and delaminates passive film	Thermal pulse enhances film regrowth
Temperature	Local temperature	Local temperature
Turbulence bulk flow structure	Turbulent flow features shear the film from surface	Rougher/larger surface area

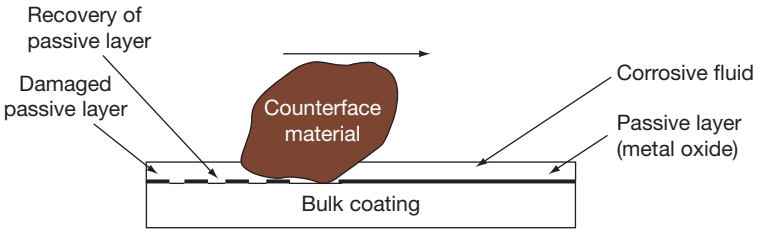


Figure 9 Wear-corrosion of hard coatings on metallic substrates where damage to the passive film results in repassivation and healing of the passive film on the coating. Reproduced from Wood, R. J. K. *J. Phys. D: Appl. Phys.* **2007**, *40*, 5502–5521. After Dearnley, P. A.; Aldrich-Smith, G. *Wear* **2004**, *256*, 491–499.

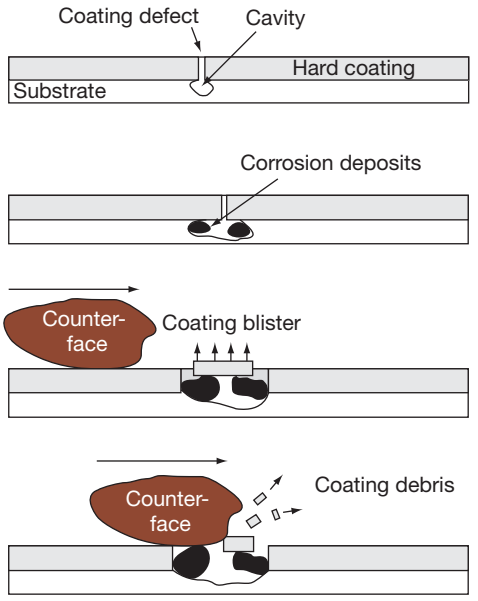


Figure 10 Corrosion-wear of hard coated metallic alloy. Pitting/blistering of the hard coating culminates in mechanical fragmentation and removal of the coating. Reproduced from Wood, R. J. K. *J. Phys. D: Appl. Phys.* **2007**, *40*, 5502–5521. After Dearnley, P. A.; Aldrich-Smith, G. *Wear* **2004**, *256*, 491–499.

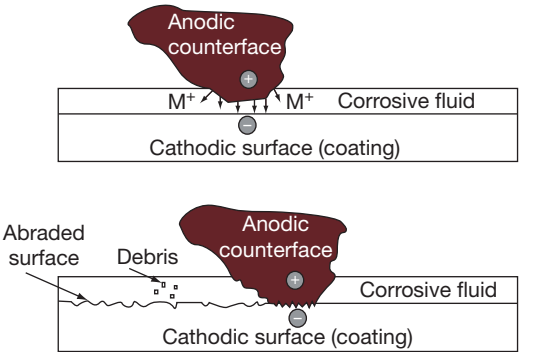


Figure 11 Wear-corrosion leading to increased coating material loss due to anodic roughening of the counterface. Reproduced from Wood, R. J. K. *J. Phys. D: Appl. Phys.* **2007**, *40*, 5502–5521. After Dearnley, P. A.; Aldrich-Smith, G. *Wear* **2004**, *256*, 491–499.

mechanisms (surface dissolution and polarization) and not wall shear stress, τ_w (see eqn [13]), as previously assumed.

For single-phase turbulent flows in pipes, a 'break-away' or 'critical velocity' above which damage occurs to pipe surfaces has been reported based on

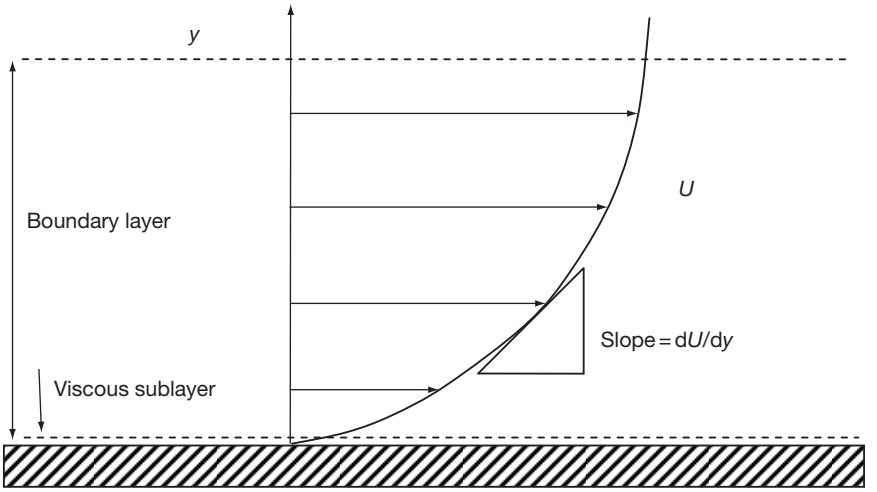


Figure 12 Near-wall velocity profile.

Table 3 Critical flow parameters for copper alloy tubing in seawater, after Roberge[36]

Alloy	*Critical shear stress ($N m^{-2}$)*	*Critical velocity in 25 mm diameter pipe ($m s^{-1}$)*	*Design velocity based on 50% critical shear stress ($m s^{-1}$)*	*Maximum tubular design velocity ($m s^{-1}$)*
Cupro-nickel with Cr	297	12.6	8.6	9.0
70–30 Cupro-nickel	48	4.6	3.1	4.5–4.6
90–10 Cupro-nickel	43	4.3	2.9	3.0–3.6
Aluminum bronzes			2.7	
Arsenical A1 brass	19	2.7	1.9	2.4
Inhib. admiralty				1.2–1.8
Low Si bronze				0.9
P-deoxidized copper	9.6	1.9	1.3	0.6–0.9

the principles of the critical wall shear stress required to strip the passive film off the pipe surface. Table 3 lists typical values for various copper alloys in pipe geometries handling seawater.

2.15.2.2.4 Flow corrosion

Flow corrosion is defined as the corrosion resulting from the effect of flow of a fluid that does not contain solid particles in sufficient concentration and/or size to impinge on the metal surface.[37] The reactants are transported to a metal surface by diffusion or convection, while the products must be removed by the same mass transport mechanisms or accumulate as scales.[38]

Corrosion models have been developed over the years that have attempted to establish the relationship between fluid dynamics, mass transport, and ultimately the chemical and electrochemical reactions that occur on the surface of a metal. These models have recently been discussed in a series of papers by Nesic *et al.*[39] The models describe the formation of protective scales and assess quantitatively the conditions (chemical and hydrodynamic) under which scales form, and only consider the dissolution of such scales. These models tend to neglect the lateral momentum transfer associated with turbulence, which are now considered important in understanding flow corrosion.

Mass transfer of reactants to and from the electrode surfaces plays an important role during flow corrosion where the dissolved oxygen or the dissolved metal ions have to diffuse from the solution to the metal or from the metal to the solution, respectively.[40] Mass-transport has been studied by several authors in different hydrodynamic systems that use convection to enhance the rate of mass transfer to the electrode. Such devices include free and submerged

jet impingement geometries, rotating disc electrodes (RDE), and rotating cylinder electrodes (RCE). However, recently, the microwall jet electrode (WJE) has been deployed to investigate the quantitative electrochemistry during process flows.

The standard model used to describe flow corrosion current densities for mixed controlled reactions (where activation and diffusion processes are present) is the modified Koutecky–Levich model[41] given in eqn [14], where k_m (the mass transport coefficient) can be derived from eqns [29]–[31], will need correcting. Equation [14] is the summation of activation current density, i_a, which is controlled by the charge transfer kinetics and the diffusion controlled current density, i_d, which is related to the mass transport of species to and from the reacting surface.

$$\frac{1}{i} = \frac{1}{i_a} + \frac{1}{i_d} = \frac{1}{zFk_1} + \frac{k_{-1}}{k_1 zFk_m} \qquad [14]$$

where i is the corrosion current density, n is the number of electrons, F is the Faraday's constant, k_1 is the forward reaction rate, and k_{-1} is the backward reaction rate.

The values of k_m are normally obtained from nondimensional correlations between the Sherwood number $Sh = (k_m L/D)$, the Schmidt number $Sc = (v/D)$, and the Reynolds number $Re = (UL/v)$.

$$Sh = aRe^x Sc^y \qquad [15]$$

Giving

$$k_m = aD^{1-y} L^{x-1} v^{y-x} U^x \qquad [16]$$

Taking typical values for the exponents as $x = 0.6$ and $y = 0.33$ for turbulent pipe flows where $2000 < Re < 5000$[9] gives

$$k_m = aU^{0.6}\left(\frac{D^{0.66}}{v^{0.27} L^{0.4}}\right) \qquad [17]$$

where D is the diffusion coefficient, U is the flow velocity, v is the kinematic viscosity, L is the characteristic length, and a is the scaling constant.

Flow corrosion rates induced by solids-free flows, as suggested by eqn [14], have been shown to depend on k_m^n where n varies depending on the surface–environment combinations and the controlling corrosion mechanism. For example, corrosion rates for copper alloys in seawater under partial activation or mixed control are shown to have $n < 1$. Steels in sulfuric acid and under simple mass transfer control have $n = 1$; while for steels in water where flow modifies the corrosion potential and also influences the solubility of the oxide, n lies between 1 and 3, see Poulson.[9] The rates of mass-transport processes also increase with increasing surface roughness. For pipes, different relationships apply depending on the surface finish. When surface roughness is developed on a pipe wall, the rate of mass transfer is proportional to the flow turbulence or the velocity distribution in the laminar sublayer.[36] For a fully eroded pipe with so-called scalloped surfaces where erosion scars resemble inverted scallop shells

$$Sh = 0.005 ReSc^{0.33} \qquad [18]$$

For a previously roughened pipe

$$Sh = 0.007 Re^{0.96} Sc^{0.33} \qquad [19]$$

The upper bound value is reached for extremely rough surfaces and given by

$$Sh = 0.01 ReSc^{0.33} \qquad [20]$$

However, a power-law relationship between erosion-corrosion rate and k_m is not always seen as in the case of steel where surface films are removed above a critical k_m resulting in a sudden increase in erosion-corrosion rate.

It has also been proposed that corrosion films/coatings can also be damaged and/or completely removed by mechanical forces resulting from the impact of either small fluid packages contained in eddies and/or bubbles created by highly turbulent flows.[42] Attempts to quantify the effective hydrodynamic forces in terms of wall shear stresses have been confronted with the problem that the measured wall shear stresses in flow corrosion susceptible systems are orders of magnitude too small,[43,44] even under highly disturbed flow conditions, to overcome the fracture mechanics of the surface scales. It has been calculated that the destruction of such layers would require shear forces of the order of 1 MPa, while measurable wall shear stresses, even in highly turbulent flow, are only of the order of 1 Pa,[45] and thus are too small to be the right parameter to quantify hydrodynamic forces responsible for the initiation of flow corrosion. Nevertheless, critical flow intensities are undoubtedly identified in terms of critical wall shear stresses above which corrosion products are hydrodynamically damaged. This leads to the significance of lateral momentum transfer by the near-wall turbulence structures in the hydrodynamic boundary layer that disturb scales and/or protective films.[42–45]

The near-wall region can be divided into three distinct regions: (1) the viscous sublayer, (2) the buffer layer, and (3) the fully turbulent region. Turbulent features are often reported with reference to the dimensionless scale $y^+ = v^* y/v$, where v^* is the

friction velocity ($=\tau_w/\rho$ or wall shear stress over fluid density), y is the vertical height above a horizontal surface, and v is the kinematic viscosity of the fluid. The viscous sublayer extends from the wall out to $\sim y^+ = 5$–8. In this region, the mean velocity profile is linear and it is assumed that this region is devoid of turbulence. However, flow visualization has revealed slow-moving oscillations known as near-wall or sublayer streaks.[46,47] These streaks are $\sim 100 y^+$ units across and $1000 y^+$ units long and are believed to occur as a result of elongated regions of alternating low- and high-speed fluid produced by streamwise vortices close to the wall. The near-wall streaks slowly lift up into the buffer region, where they undergo a distinct oscillation and finally break up violently $\sim y^+ = 40$ (see **Figure 13**).[48,49] The majority of the turbulence production occurs in the buffer region from $y^+ = 8$ to 30 to 50 during violent outward ejections of low-speed fluid and for continuity considerations, the ejections are followed by a sweeping inrush of high-speed fluid at a shallow angle toward the wall. An ejection transports low-momentum fluid from the wall region into the main flow. Conversely, sweep events transport high-momentum fluid from the main flow toward the wall (**Figure 13**) at velocities close to $0.9U$, where U is the free stream velocity. The result of these flows is fluctuating wall shear stresses, see eqn [13], and pressures which will influence erosion-corrosion rates and mechanisms locally.

2.15.2.3 Wear Fundamentals

2.15.2.3.1 Solid particle erosion

To understand erosion-corrosion, it is desirable to first understand the erosion mechanisms. Solid particle erosion in liquid systems is a process by which discrete small solid particles that are suspended in the flow, with inertia, strike the surface of a material causing damage or material loss to its surface. Erosion can be a problem for components such as gas and hydraulic turbine blades, propulsors, pipelines, valves and fluidized bed combustion systems. An example of the surface features erosion can develop in flow through a valve is shown in **Figure 14**. Erosion does have its beneficial applications, notably for cleaning and preparation of surfaces for subsequent coating/painting, descaling/dewaxing of production risers in the oil and gas industries, oil sand transportation, hydrotransport systems and quarrying rocks using abrasive cutting jets, or subsea steel oil structures using submerged abrasive water jets.

The flow conditions along with the properties of the target material or any film, if present, as well as the fluid and erodent properties will influence the contact mechanics of the erodent to component surface interactions. The main important target material properties are hardness, fracture toughness, coefficient of restitution elasticity, and the degree of surface roughness. The general expression for erosion rate, E, has been established empirically and can take the form

$$E = M_p K f(\alpha) U_p^n \qquad [21]$$

where M_p is the mass of particles impacting on the surface, U_p is the particle velocity on impact, α is the particle impact angle, K and n (typically between 2 and 3)[50] are constants assumed to be dependent on the characteristics of the erodent/target materials involved, and $f(\alpha)$ is a functional relationship for the dependence of the erosion rate on the impact angle.[51]

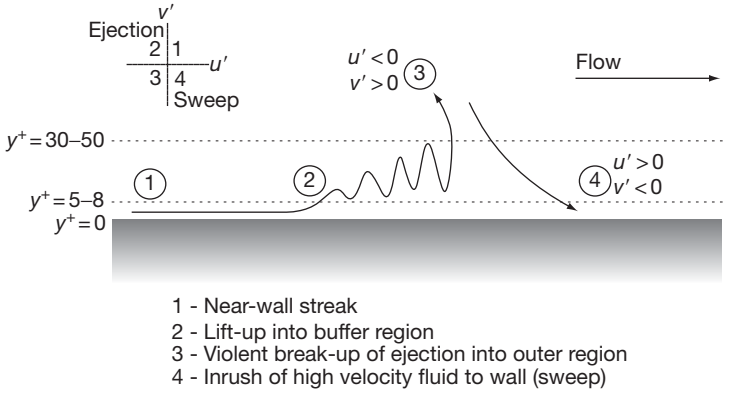


Figure 13 Schematic showing the near-wall turbulence structures within the viscous sublayer, buffer layer, and fully turbulent region.

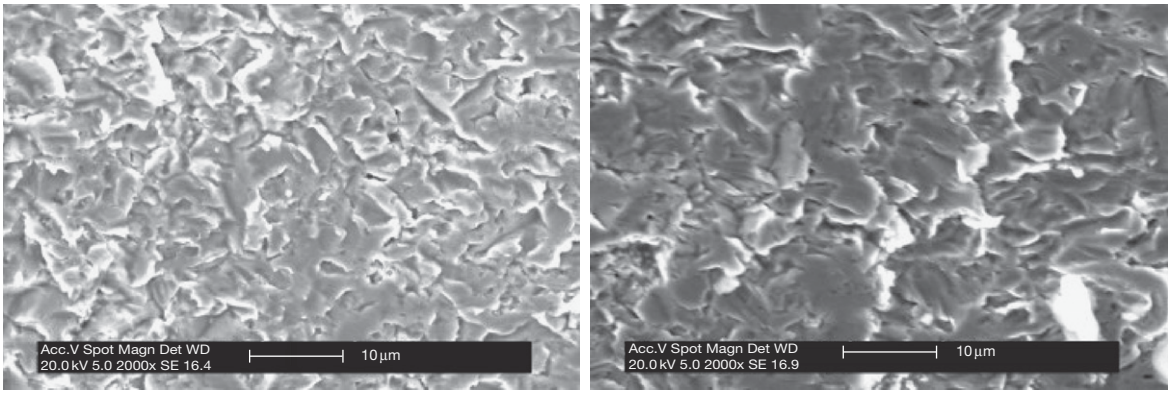
Figure 14 Sand erosion patterns of a Stellite tapered flow passage located downstream of a choke valve that induced swirl flow. Large bore is 75 mm in diameter.

This can be rewritten for the unit volume loss per impact, V_u, as

$$V_u = E_k^m f(\alpha) C_v G \qquad [22]$$

where V_u is the volume loss per impact in units of m^3 per impact, E_k is the particle kinetic energy (KE), m is the energy exponent with $m = 0.5n$ assuming $E \propto d^3$ where d is the particle diameter, C_v is the volume fraction, G is a constant, and $f(\alpha)$ is as before. From eqn [21] or [22], it can be seen that the erosion rate will be strongly dependent on the KE of the impacting particles, the number of impacting particles, and the impact angle. All three of these factors vary for most industrial components exposed to solid particle-laden flows, and therefore, to sustain long service life, the internal surfaces must perform over a wide range of solid impact conditions. The sand-to-wall impact conditions are likely to depend on the flow regime present, the orientation and/or geometry of the component, the ability of the flow to keep the particles in suspension[52] and the relative size of erodents with respect to boundary layer thickness (small particles can form protective clouds within subviscous layers). For suspended solids traveling in a horizontal liquid flow, both the solid particle impact velocities ($<3.0\ m\ s^{-1}$) and angles are relatively low for most practical purposes.[53] Typical impact erosion craters are shown in **Figure 15**.[54]

Plots of V_u against KE of impact, E_k can be used to map the erosion performance of surfaces. These maps enable comparisons to be made between materials



(a) (b)

Figure 15 SEM micrographs of 0.2% CS (a) $7\ m\ s^{-1}$ and (b) $12\ m\ s^{-1}$, 1 wt.% silica (296 µm) sand in 0.1 M NaOH at 40 °C, test time = 30 min. The eroded region observed was located approximately normal to the flow and in the center of the sample. Reproduced from Harvey, T. J.; Wharton, J. A.; Wood, R. J. K. *Tribology – Mater. Surf. Interfaces* **2007**, *1*(1), 33–47.

at a wide range of impact energies covering various particle sizes and velocities. However, it can be concluded that this plot format is not well suited to accommodating high levels of uniform corrosion in the presence of low energy and low concentrations of solid impacts. The approach assumes that all mass loss (mechanical and electrochemical) is generated by the impacting particle, and hence, all the erosion-corrosion (T) data lie well above the E power law fit (straight line on **Figure 16**). As erosion becomes more influential (i.e., higher KE or higher sand concentrations), the data begin to approach the E power law fit, that is, the data move from left to right for increasing KE and vertically down for increasing sand concentration. E_k is defined as

$$E_k = 2/3\pi r_p^3 \rho_p U_p^2 \qquad [23]$$

where r_p is the particle radius, ρ_p the erodent particle density, and U_p the impact velocity. Care should be taken as experimental evidence shows that erosion rates can be dependent on $U_p^n d^y$ where U_p is the particle velocity and d is the particle diameter, and n and y are far removed from 2 to 3 assumed in the simple energy approach seen earlier.[55]

The main parameters of concern for erosion relate to the solid particle–target interactions and hence the number of particles impinging, individual particle energies, particle impingement angles, particle-to-target hardness ratios, and the shape of the particles. Near wall particle–particle interactions can also severely influence erosion rates when the volume concentration of solid particles present is high.

Finnie[56] developed an erosion model based on cutting wear mechanisms of the form:

$$V_u = C_1 \frac{M_p U_p^2}{4\sigma} f(\alpha) \qquad [24]$$

where C_1 is an arbitrary constant denoting the proportion of particles that cut the surface. Gane and Murray[57] found a value of $C_1 = 0.5$ and gave reasonable predictions. Keating and Nesic,[58] in numerically predicting erosion-corrosion in bends and sudden expansions by two-phase flows (liquid–solid), used a modified Finnie approach based on earlier work by Bergevin.[59] This approach incorporated the concept of a critical velocity for plastic deformation, U_{cr}. They substituted $(U_p \sin\alpha - U_{cr})$ for the impact velocity in eqn [24] to give

For low angles ($\alpha \leq 18.5°$)

$$V_u = \frac{M_p(U_p \sin\alpha - U_{cr})}{2\sigma} \times \left[U_p \cos\alpha - \frac{3}{2}(U_p \sin\alpha - U_{cr})\right] \qquad [25]$$

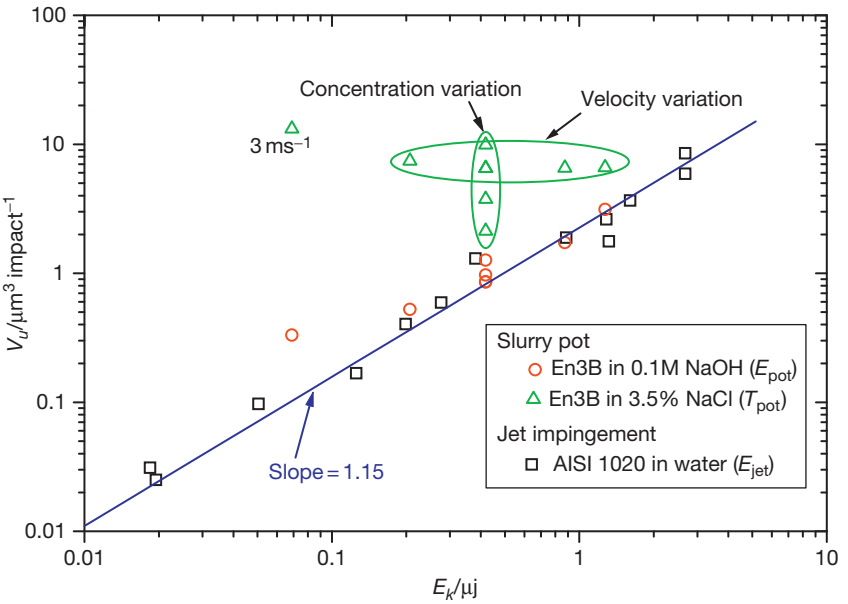


Figure 16 Comparison of volume loss per impact against particle impact KE (slurry pot)[54] and jet impingement studies.[31] The Δ indicates erosion–corrosion data, displaying the two sets of data (concentration and velocity) presented in previous plots. E_{pot} and E_{jet} are taken as mechanically nominated erosion from the slurry pot and jet impingement rigs respectively and T_{pot} is the erosion–corrosion data from the slurry pot.

For higher angles of impingement

$$V_u = \frac{M_p\left(U_p \sin\alpha - U_{cr}\right)^2}{12\sigma}\frac{\cos^2\alpha}{\sin^2\alpha} \qquad [26]$$

Bitter[60] quotes a value of $U_{cr} = 0.668\ \mathrm{m\,s^{-1}}$ for steel. Keating and Nesic used this value to successfully predict erosion rates in a sudden expansion and found the original Finnie model not to be so accurate. However, they modeled erosion-corrosion damage in a U-bend, using the original Finnie model as the modified version yielded no erosion due to the low particle velocities involved. Keating and Nesic concluded that their modeling needs more experimental validation before further refinements can be made.

Erosion models typically recognize that two erosion mechanisms act, cutting and deformation erosion, with discrete models representing each, such as the Neilson and Gilchrist erosion model used for rocket motor nozzles.[61] Such models have been successfully used by Forder[62] and Wood and Jones[63–67] to predict erosion of internal components within choke valves and slurry ducts. The cutting erosion model for low impact angles was first proposed by Finnie[56] and later modified by Hashish.[55] The deformation model was proposed by Bitter[60] and is thought to be applicable at higher impact angles (30°–90°). Particle shape and material properties for both particle and target have been included, which earlier simpler models have not considered. As the particle impingement angles are predicted to be below 10°, for critical components such as straights and bends, see Wood and Jones,[63–67] the contribution to the overall wear rate from deformation mechanisms can be ignored, and hence, the volumetric erosion per impact can be given by the modified Hashish model only

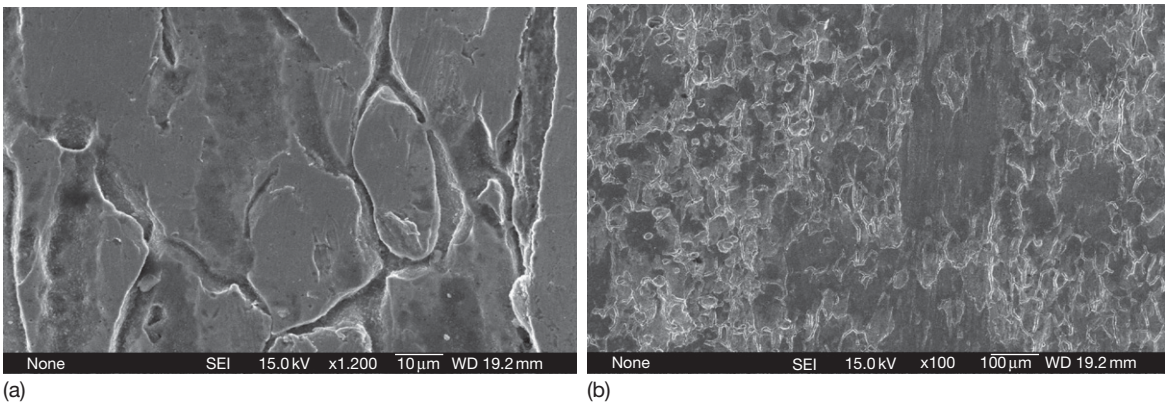
$$V_u = \left\{\frac{100}{2\sqrt{29}} r_p^3 \left(\frac{U_p}{C_k}\right)^n \sin 2\alpha\sqrt{\sin\alpha}\right\} \qquad [27]$$

where $n = 2.54$ and

$$C_k = \sqrt{\frac{3\sigma R_f^{0.6}}{\rho_p}} \qquad [28]$$

Experimental studies on pipe systems show variation of the erosion mechanism and location with flow regimes. Wood and Jones[63–67] show erosion damage of AISI 304L stainless steel pipe components with a nominal wall thickness of 5 mm from a horizontal pipe loop tests. The pipes were of 80 mm nominal bore, although the actual mean of a series of measurements was 77.8 mm. The erosion of the bend with the greatest curvature of the whole loop ($Rc/D = 1.2$) was analyzed along with its upstream straight.

Figure 17 shows the top internal surfaces (12 o'clock position) of the straight pipe after 210 h of slurry testing at $3\ \mathrm{m\,s^{-1}}$ with an asymmetric slurry flow of 10% (w/w) sand of 500–1400 µm in size in water. **Figure 18** shows the overall component details for the straight and bend investigated. The damage in **Figure 17** is consistent with that of the as-manufactured (scaled) surface finish. There appears to be no significant erosion impact damage, although a few individual impact craters are seen due to impacts from the relatively few sand particles that are entrained into the energetic upper fluid flow. The lack of damage seen in **Figure 17** is in contrast with the damage seen at the bottom of the straight pipe section as shown in **Figure 19**. This reflects



(a) (b)

Figure 17 SEM micrographs at (a) high and (b) low magnifications showing surface morphology of straight position (top A position) (after 210 h).

the asymmetry of sand burden in the flow with the majority of sand particles traveling in the bottom half of the pipe and only a few energetic particles saltating into the flow streams in the top half of the pipe.

In Figure 19, the impact features reveal that the erosion processes occurred by a simultaneous plastic deformation and microcutting of the stainless steel pipe surface, similar to type II cutting mentioned by Hutchings when the erodent rotates backwards on impact and efficiently machines the surface.[8]

Erosion of the bend outer wall at section two (shown in Figure 18) is even more intense as shown in Figure 20. The surface of the bend at position C (3 o'clock) has clearly been eroded with evidence of low-angle impingements with extensive plastic deformation and cutting (type II) resulting in the removal of the as-manufactured surface morphology. This level of damage is a result of impingement by the larger and energetic sand particles with the bend wall and reflects the repositioning of the particle burden due to the bend geometry and resulting change in flow direction. Erosion-corrosion models should concentrate on such impact scars and their repassivation.

2.15.2.3.2 Cavitation erosion

Cavitation bubble collapse close to a solid surface can damage the surface by two mechanisms: shock wave impingement and high-speed liquid microjet impingement. Shock waves are formed by the rapid collapse of the bubble, which generates very high internal temperatures and pressures that are relieved by shock wave propagation. The presence of the solid surface generates asymmetric bubble collapse and the formation of a microjet with a diameter of 10–50 μm that impinges onto the surface at normal incidence

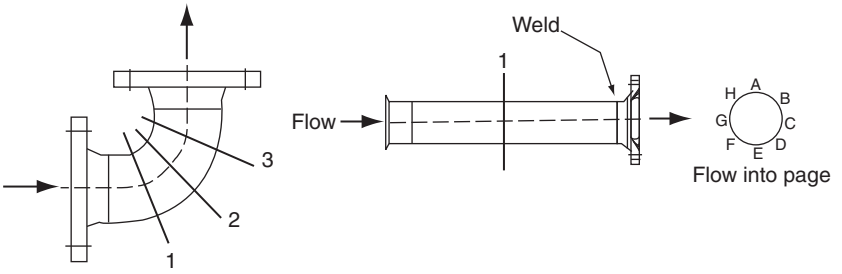


Figure 18 Drawings of bend and straight sections which were mounted horizontally and subjected to asymmetric slurry flow for 210 h.

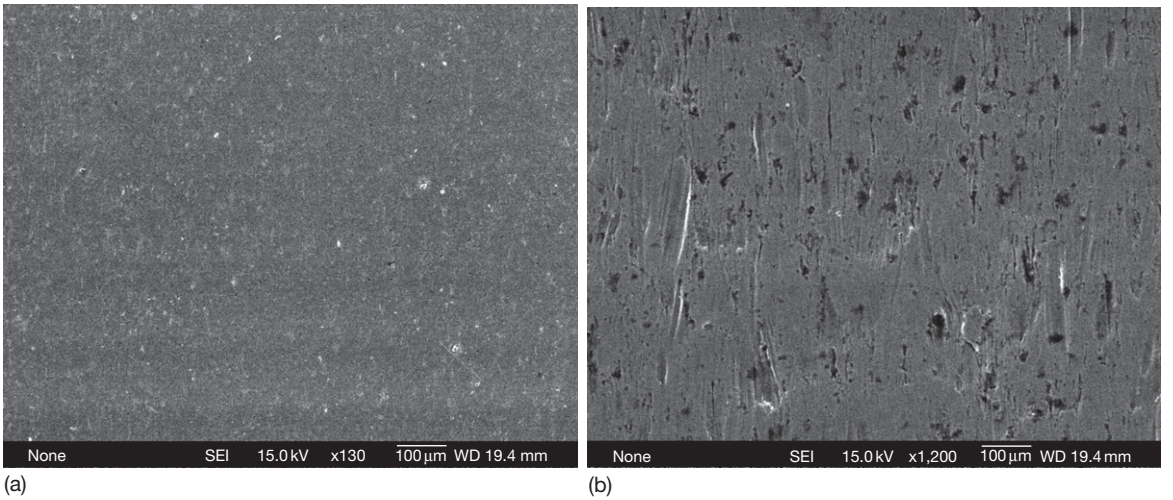
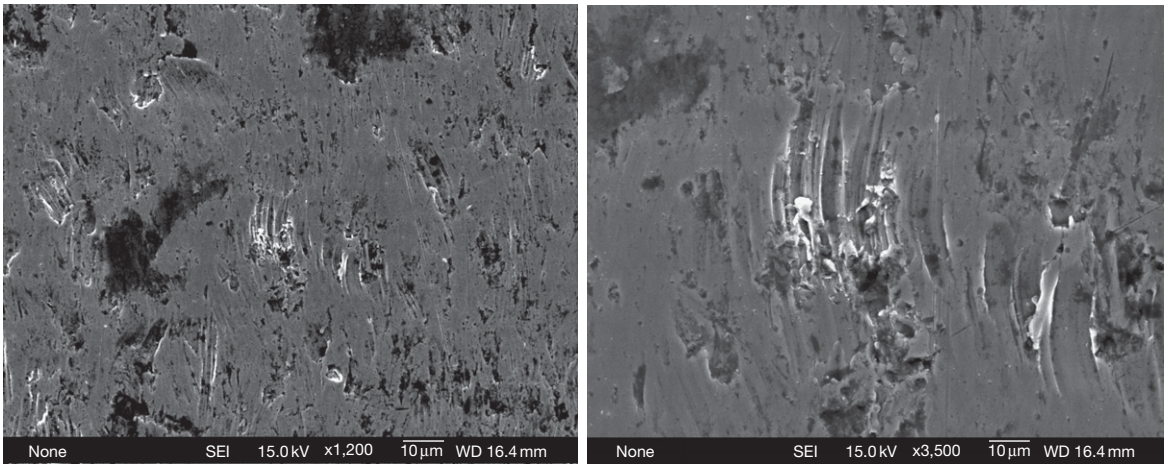


Figure 19 SEM micrographs showing the surface morphology of the straight pipe at the bottom (position E) after 210 h. (b) shows individual impact scars from sand microcutting at low angle impingements with the pipe surface.

at ~250–300 m s^{-1}. Both of these processes produce localized temperature rise, high intensity turbulence, and transient high pressures of the order of ρcV. These pressures are large enough to cause plastic deformation in most engineering surfaces. These pressures and temperatures enhance mass transport and electrochemical reactions and thereby also accelerate corrosion rates of metallic surfaces. The erosion produced by such near-wall bubble collapse is thought to be dominated by a fatigue process and thus cavitation erosion resistant surfaces tend to have high fatigue resistance.

The overall damage induced by cavitation and corrosion therefore involves the interaction of hydrodynamic, mechanical, metallurgical, and chemical factors. The presence of oxygen and aggressive ions in the electrolyte is thought to be important parameters in controlling cavitation-corrosion rates. Liu *et al.*[68] show that cavitation-corrosion of CS in NaCl solutions is mainly controlled by the diffusion of oxygen to the surface and is sensitive to temperature and NaCl concentrations, see **Figures 21** and **22**. The erosion process promotes anodic dissolution reactions, which are charge transfer (activation) controlled. The impingement of the microjets is also likely to contribute to surface activation by damaging the surface layer of CS corrosion products. They also



(a) (b)

Figure 20 Micrographs showing the surface morphology of bend erosion outer wall at section two and position C. Both images show impact craters with extensive plastic deformation and cutting damage.

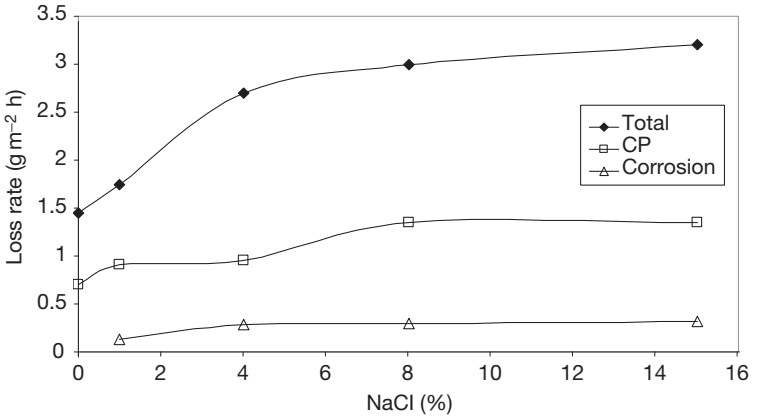


Figure 21 Effect of cathodic protection on cavitation-corrosion of CS. Total = cavatition-corrosion loss rate, CP = loss rate under cathodic protection (Cathodic overpotential of 450 mV), Corrosion = corrosion rate from *in situ* electrochemical measurements. Reworked data from Lui, J.; Lin, Y.; Yong, X.; Li, X. *Corrosion* **2005**, *61*(11), 1061–1069.

show that the application of cathodic protection dramatically reduced the surface wastage rates.

Little systematic research appears to have been conducted on cavitation-corrosion. However, surfaces that combine high fracture toughness and corrosion resistance have typically been studied. A cathodic arc plasma ion plated Ni–Al film on AISI 1045 CS has been reported by Chang *et al.*[69] to show good cavitation erosion-corrosion resistance. Three coating compositions were studied, Ni–48Al, Ni–40Al, and Ni–30Al, and all increased the cavitation erosion resistance of the substrate 10 times and the erosion-corrosion resistance to 3.5% NaCl solutions two times and to 3.5% HCl solutions three times. The potentiodynamic polarization data of the coated specimens in these two electrolytes indicated that the coated specimens exhibited a more electropositive corrosion potential and a lower corrosion current density than did the bare substrate material. However, cavitation erosion induced pitting corrosion of the substrate and this accelerated the mass loss rate of specimens (see Figure 23).

Chernega *et al.*[70] found that cyclic cavitation erosion-corrosion interactions increased the material loss of complex carbide coatings 1.5–2 times the continuous cavitation rates. They found that cyclic cavitation-corrosion wear processes on high velocity oxy-fuel (HVOF) carbide coatings based on Cr, Cr–Zr, Cr–Nb, Cr–Si, and Nb were enhanced by 10–20% in NaCl solutions compared with pure water. The cyclic cavitation-corrosion treatment enhanced the average corrosion processes over time by factors of between 2 and 4 compared with a continuous cavitation corrosion influence. This enhancement was related to the destruction of the protective surface oxide films, the removal of corrosion products, and the increase in microcracks in the diffusion coating as a result of microshocks produced by near-surface cavitation bubble collapse.

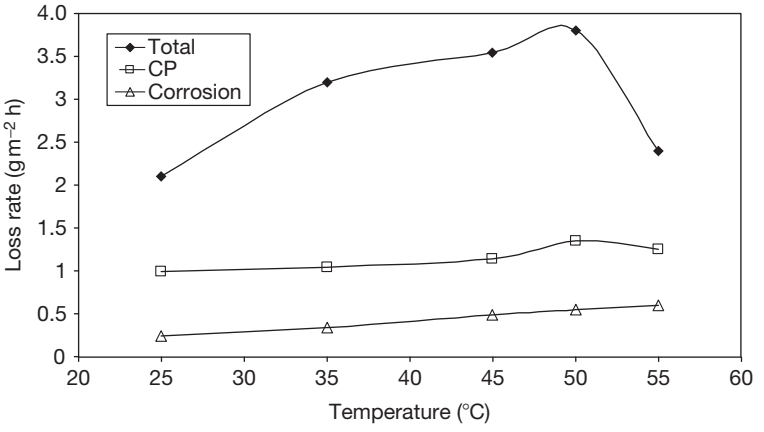


Figure 22 Effect of cathodic protection on cavitation-corrosion of CS. Total = cavatition-corrosion loss rate, CP = loss rate under cathodic protection (Cathodic overpotential of 450 mV), Corrosion = corrosion rate from *in situ* electrochemical measurements. Reworked data from Lui, J.; Lin, Y.; Yong, X.; Li, X. *Corrosion* **2005**, *61*(11), 1061–1069.

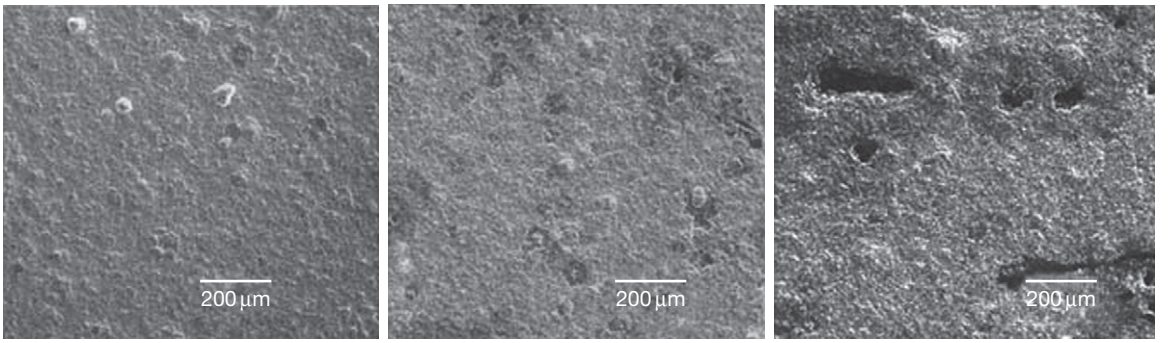


Figure 23 Surface morphology of a Ni–48Al intermetallic coating (a) untested, (b) tested for 30 min, and (c) for 60 min in 3.5% NaCl electrolyte at ASTM G32–98 ultrasonic cavitation test conditions. Reproduced from Chang, J. T.; Yeh, C. H.; He, J. L.; Chen, K. C. *Wear* **2003**, *255*, 162–169, Part 1 AUG-SEP 2003.

2.15.2.3.3 Abrasion

There are two types of abrasion: two-body or grooving abrasion as shown in Figure 24(a), while three-body or rolling abrasion is illustrated in Figure 24(b). There are various important parameters that control abrasion rates such as the hardness ratio between abradant and target, the load per particle, the role of the fluid film and if it supports the contact load, particle entrainment, size and shape, and friability. Wear rates of rolling abrasion are generally lower than those generated by grooving abrasion. Industrial surveys place abrasion as the most common wear mechanism with over 50% of wear problems being associated with both types of abrasion.[71] In abrasive wear, material is removed or displaced from a surface by hard particles, or sometimes by hard protuberances or asperities on a counterface or embedded hard particles within a surface, forced against and sliding along the surface. The sources of the hard particles, which can be entrained into the sliding contact, include contaminants from the outside environment, wear debris, and oxidation products formed with the tribocontact or other chemical processes.

Figure 24(c) is a schematic of three-body rolling abrasion of a passive surface showing a wear rate that strips the passive film completely.

2.15.2.4 Experimental Techniques

Table 4 lists the variety of experimental techniques used to study wear-corrosion interactions or aspects of these interactions. Most are either corrosion cells attached to wear testers or attempts to modify corrosion cells to induce some tribological contact (i.e., scratch test or adding slurry in a rotating disc/cylinder electrode). This is not ideal as experimental setups should reflect the best approach for wear and corrosions measurements and should be designed specifically and should not be just a modified existing tribotester.

2.15.3 Tribocorrosion Mechanisms

The interaction between tribological and electrochemical corrosive effects causes materials to corrode

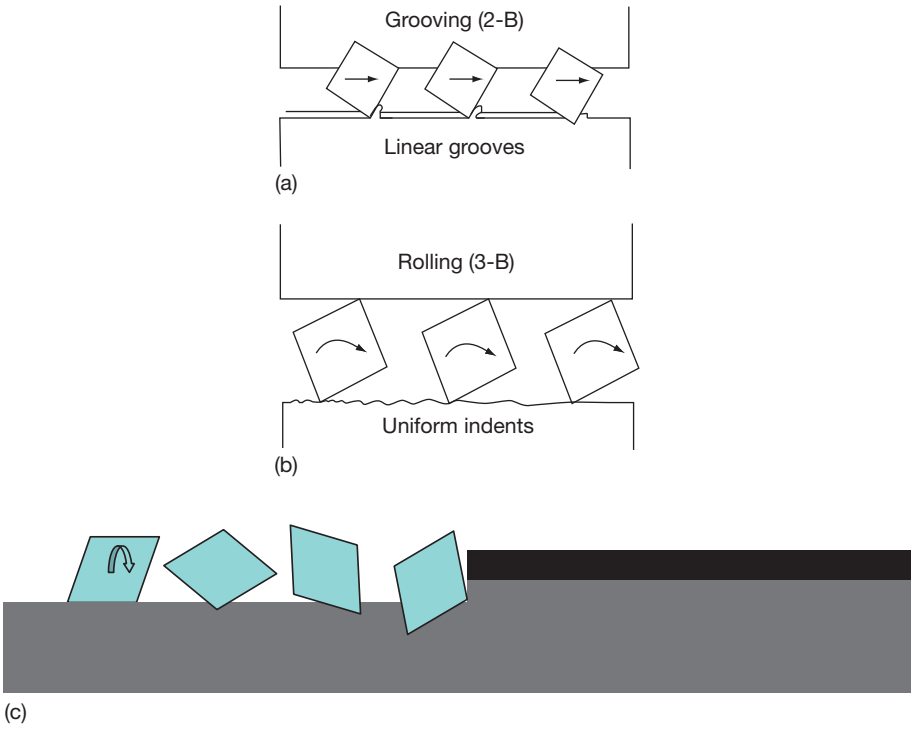


Figure 24 (a) and (b): Schematics of two-body grooving and three-body rolling abrasion, (c). Depassivation of a passive surface by three-body rolling abrasion removes the passive layer and some substrate material revealing bare metal surface.

Table 4 Experimental techniques used to study wear-corrosion interactions or aspects of these interactions

	Technique	*Parameters evaluated*	*Comment*
Engineering-based rig with added electrochemical cell	Pin-on-disc, block on cylinder	Friction, linear wear, temperature, wear scar, electrochemical parameters, current/potential versus time, polarization, *in situ* AC techniques (EIS)	Can generate complex electrochemical outputs and functionality of electrochemical cells – needs careful scrutiny
	Reciprocating pin on plate		
Electrochemical rig with mechanical elements added	Abrasion and microabrasion		
	Continuous grinding		
	Instantaneous fracture[72]		
	Scratch tester[73]		
	Rapid strain[74,75]		
	Guillotine cutting[76]		
	Triboellipsometry[77]		
	Cavitation hydrodynamic and vibrating horn)		Vibrating horn not representative of hydrodynamic cavitation erosion
	Impinging fluid jet/slurry jet	Peak current densities (passive and nascent), surface potential, repassivation rates, wear rate, volume loss per impact, friction, surface temperature	Uniform current distribution but shear rate varies with radius
	Pipe flow		Nonuniform flow/mass transfer distribution, rough surfaces, nonuniform current distribution
	Rotating disc (laminar)		Nonuniform current distribution, shear rate varies with radius of disc
	Rotating cylinder (turbulent)		Uniform current distribution but smooth electrode surface

at a substantially higher rate than those experienced under static or quiescent conditions.

2.15.3.1 Mechanical–Electrochemical Interactions

This section focuses on erosion-corrosion, abrasion-corrosion, and sliding wear-corrosion interactions as there is published work in these areas. The effects of the environment on the wear-induced fatigue life are not well researched and thus are not well understood and will not be reviewed further but the readers are referred to Sethuramiah[78] for a review of related work.

2.15.3.1.1 Erosion-corrosion

There are several possible regions of erosion-corrosion interaction between different flow regimes and corrosion. The most significant erosion-corrosion regions are where turbulent slurry flow and corrosion overlap and where cavitation interacts with corrosion, although it is possible that solids entrained into laminar flows could induce erosion-corrosion if their impact energy is sufficient to cause plastic deformation of the surface or puncture the tribofilm.

Figure 25 shows the mass loss rates for various sand concentrations for CS in 0.1 M NaOH solution for erosion rates and in 3.5% NaCl solution for erosion-corrosion rates.[54] These results were

obtained at 7 m s^{-1} fluid stream velocity and at 40 °C. The erosion rates E (0.1 M NaOH solution results) show a nonlinear increase with increasing velocity. The static corrosion rate C is shown on the far left and this plus the erosion rates at the various velocities should equate to the erosion-corrosion rates T at the various velocities if no synergistic effects are present. However, as seen, the synergy terms goes negative and the interactions between E and C are, therefore, antagonistic at 0.5 w/w% sand concentration ($T < (E+C)$), while a synergy is present for higher concentrations ($T > (E+C)$).

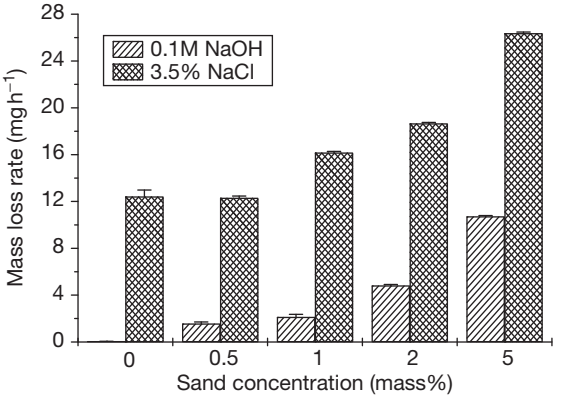


Figure 25 Experimental results for sand concentration variation at 7 m s^{-1} (294 μm silica sand at 40 °C). The exposed surface area of the sample to erosion was 3.4 cm^2. Reproduced from Harvey, T. J.; Wharton, J. A.; Wood, R. J. K. *Tribology – Mater. Surf. Interfaces* **2007**, *1*(1), 33–47.

Figure 26 shows results for the same material as **Figure 25** at 0% and 1% (w/w) sand concentrations for a variety of velocities at 40 °C.[54] This shows that erosion rate E is dependent on V^2 but that the erosion-corrosion rate T is not linearly dependent on either V or V^2. At 3 m s^{-1} there is an antagonism, while at higher velocities a synergy is seen due to the presence of surface residual stresses and scales (layers) generated by the drawing processes used to manufacture these cylindrical samples.

When corrosion rate is partially or wholly controlled by mass transfer of reactant to or product from the surface, then local conditions under erosion may well influence the mass transfer kinetics (measured by the mass transfer coefficient k_m). Under such conditions the corrosion will be controlled by the mass transfer (typically diffusion processes) and the driving concentration gradient (relative concentrations of active species near surface compared with free stream concentrations).[79] Both the mass transfer and concentration gradient will be affected by solid particle impingement that influences the local fluid flow field and increases surface roughness. Erosion will also increase the surface area wetted by the corrosive electrolyte and could establish microgalvanic cells on the surface with damaged areas being anodic to the passive and (cathodic) unaffected areas.

The relationships between fluid mechanics and induced flow patterns for both open and closed fluid systems are covered elsewhere.[80] It is important to

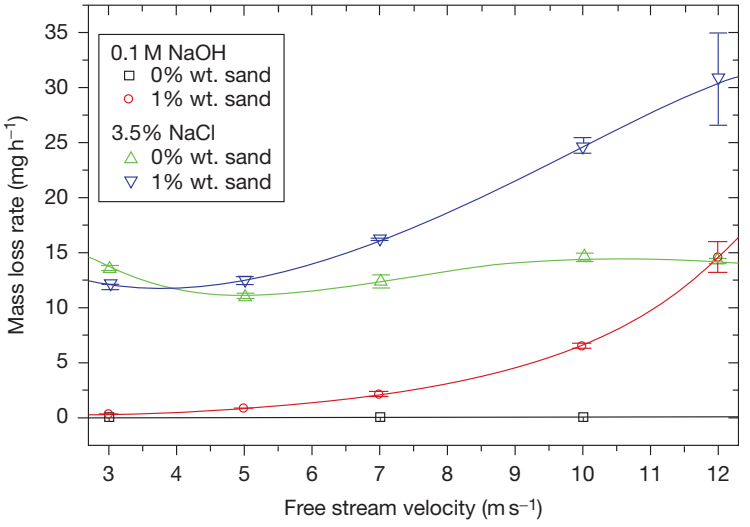


Figure 26 Experimental results for free stream velocity variation at 0% and 1% by weight sand concentration (294 μm silica sand at 40 °C). Reproduced from Harvey, T. J.; Wharton, J. A.; Wood, R. J. K. *Tribology – Mater. Surf. Interfaces* **2007**, *1*(1), 33–47.

understand the near-wall flow conditions if erosion-corrosion rates and locations are to be accurately quantified for engineering components. In particular, the development of eddy structures in turbulent flows through lateral momentum transfer between fluid elements, which in turn disrupt the turbulent boundary layers that are developed within near-wall flows,[81–83] requires particular attention as such eddies can develop localized damage patterns on solid surfaces that can lead to an accelerated loss of fluid containment and thus considerably decrease component life. An example is given in the following section, namely rippling in pipe flows

Rippling of pipe surfaces by erosion-corrosion induced by near-wall turbulence

For multiphase flows (solid–liquid), it is assumed that initially the turbulent eddy pattern in the flow is determined by the surface roughness generated by the finishing of the pipe (see **Figure 27(a)**). The ability of the sand particles to erode is largely determined by the bulk flow but will result in the modification of the surface roughness. The initial surface roughness is thus replaced by the roughness resulting from individual erosion events. In ductile materials, such as stainless steels, this roughness will be in proportion to the impact scar dimensions (**Figure 27(b)**). These in turn give rise to a new eddy pattern in the near-wall flow. During the erosion process, the sand particles have to pass through these eddies and are thereby deflected. The eddy pattern determines the angle of impact and areas of increased sand particle impact and also the corrosion rates over the surface. The concentration of damage at specific points results in the establishment of a ripple pattern, reflecting the turbulence pattern. The surface waviness increases until a steady state is reached, at which point the surface continues to wear but the wavelength and shape remain constant (**Figure 27(c)**). The final waviness is thus a reflection of the conditions to which the material is exposed. The more severe the condition, the greater the turbulence in the boundary layer and hence the larger the wavelengths and amplitudes of surface waviness.

2.15.3.1.2 Abrasion-corrosion

The interactions between two sliding metallic surfaces that entrain abrasives between them in a corrosive environment are often complex. **Figure 28** is an attempt to schematically represent some of the possible processes involved when abrasive (body 3) enters a sliding contact between bodies 1 and 2. The abrasive can depassivate the surface by penetrating or removing the oxide layer thereby enhancing corrosion at these sites. Abrasive wear particles can be formed from either or both surfaces and can be transferred and deposited on the opposite surface (4 and 5) resulting in galvanic coupling and accelerated corrosion. The ploughing and indentation action of the abrasives on the surfaces will also induce residual stresses (tensile and compressive) and induce strains that could result in microgalvanic cells being established between affected and unaffected surface areas. Abrasion of the original oxide layer could also result in better corrosion-resistant films by producing new surfaces of different composition. For example, the good abrasive corrosion behavior of plasma-nitrided 4140 low-alloy steel reported by Lee *et al.*[84] was linked to the presence of the nitrided layer with a homogeneous structure of ε and γ′ phases enriched with Cr and N atoms. Under abrasion-corrosion, this layer produced a thick and dense protective layer that

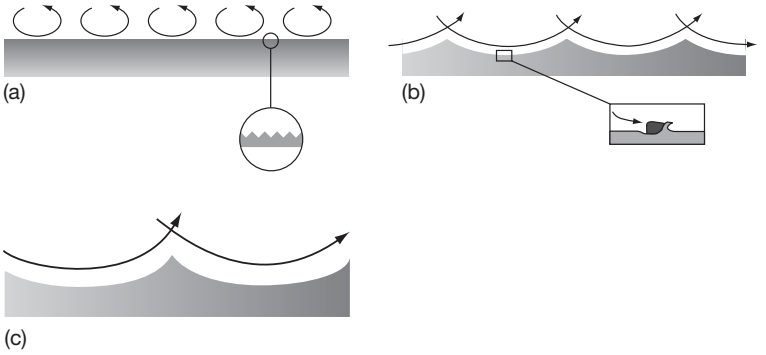


Figure 27 Ripple formation due to turbulence eddies and the influence between microroughness and eddy size: (a) initial stage with turbulence boundary layer, microroughness determined by prior machining operations; (b) intermediate stage with larger eddies due to changing surface roughness; microroughness determined by the individual erosion event; (c) final stage with steady ripple pattern, after. Reproduced from Karimi, A.; Schmid, R. K. *Wear* **1992**, *156*, 33–47.

resulted in a significant decrease in corrosion current in chloride solutions.

Figure 29 shows results from microabrasion and abrasion-corrosion tests showing a synergy for all but very low concentrations of abrasive. The lowest concentration shows a slight antagonism and under these conditions, two-body grooving abrasion dominates while at concentrations above 0.05 volume fraction the mechanism is three-body rolling abrasion. This shows that the type of abrasion can lead to different wear-corrosion interactions.

More complex and sometimes negative (antagonistic) interactions between abrasion and corrosion occur on composite surfaces.[85] Here, preferential corrosion at the carbide–binder interface resulted in a 20% reduction in abrasion-corrosion wear rate compared with purely abrasive conditions. The physical mechanism of this antagonistim is thought to relate to the preferential corrosion of metallic tungsten at the carbide–binder interface resulting in loose and flexible carbides and a lowering contact stresses under abrasion conditions that in turn reduces material loss, see **Figures 30** and **31**.

In situ electrochemical measurements made during abrasion-corrosion tests are useful as they record the kinetics of the reactions occurring and as well inform on the degree of depassivation and repassivation. **Figure 32** is a typical example of such measurements. It shows the OCP and current recorded during separate microabrasion-corrosion tests at pH 11 on a thermally

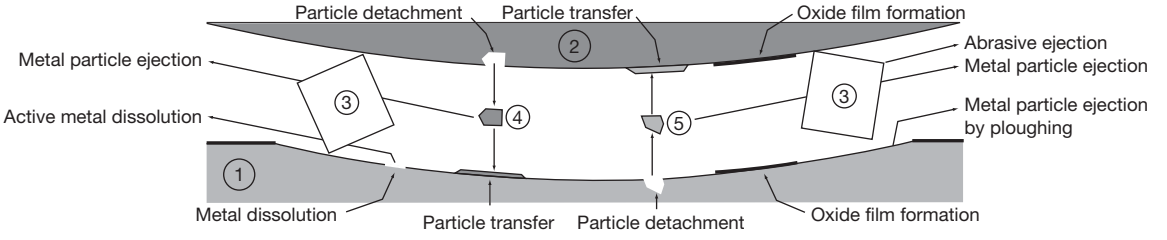


Figure 28 Schematic illustration of the tribocorrosion contact involving two metallic samples (first body and second body) and abrasive particles (third body). Wear particles originating from the first and second bodies (4 and 5) form additional contact components. The corresponding material fluxes are also indicated.

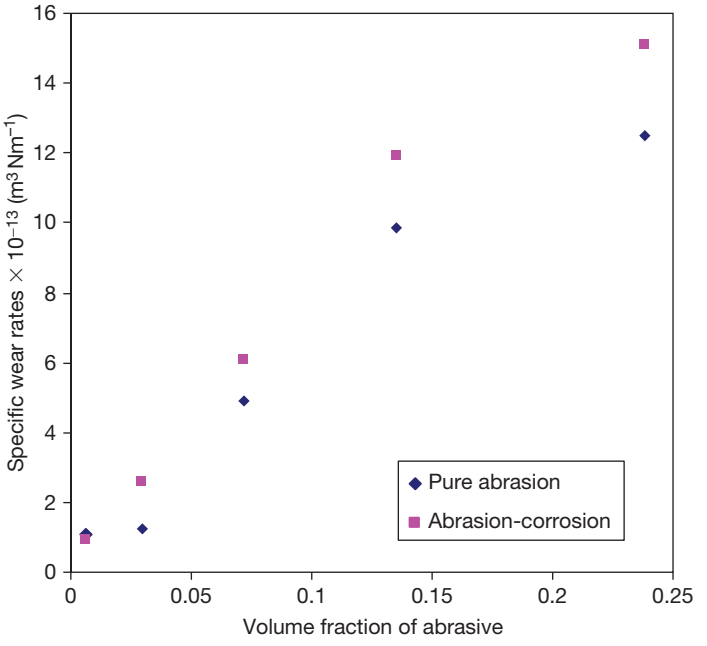


Figure 29 Microabrasion and abrasion-corrosion expressed as specific wear rates (SWR) of type-316 stainless steel as a function of abrasive concentration. The load was 5 N and the electrolyte was distilled water for the abrasion tests and 3.5% NaCl solution for abrasion-corrosion. The abrasive used was 4 μm SiC.

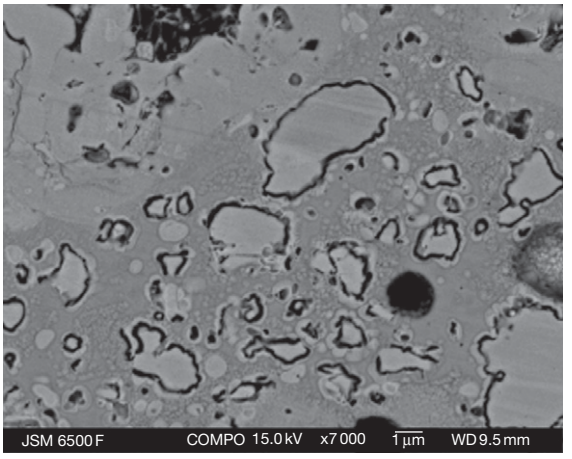


Figure 30 Selective corrosion of W and W_2C phases (which are formed due to decarburization of WC) around the periphery of the carbide grains in a HVOF sprayed WC–10Co–4Cr coating when exposed to an alkaline environment of pH 11 for 168 h.

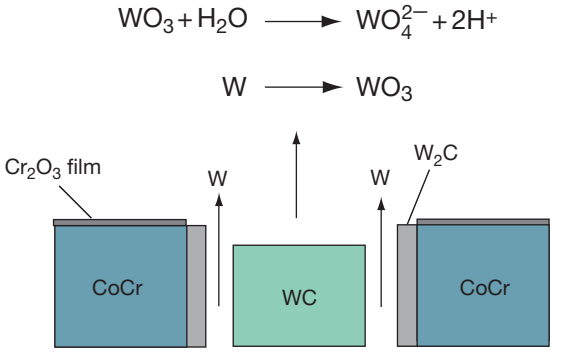


Figure 31 Schematic showing the selective corrosion of W and W_2C phases in a WC–10Co–4Cr coating.

sprayed tungsten carbide coating. The potential drops by ~60 mV over a 250 s period once ball rotation and abrasion commences, while the current instantly increases from low freely corroding levels to 2 μA. Both these trends show that depassivation is occurring in the wear scar area. On stopping the rotation of the ball, and thus abrasion, at 2200 s, the potential recovers in the electropositive direction and the current decreases toward the initial freely corroding levels, indicating repassivation of the wear scar.

2.15.3.1.3 Sliding wear-corrosion

As discussed earlier, typically passive oxide films are responsible for the corrosion resistance of metallic surfaces. However, these oxides are vulnerable when subjected to tribological contact conditions where mechanical effects can damage or completely remove these films. This typically leads to accelerated corrosion in regions of damaged or removed oxide films.[86]

Ponthiaux[87] suggests that it is important to understand the galvanic activity resulting from worn areas surrounded by unworn areas in order to predict the performance of tribocorrosion contacts. Garcia[88] developed a model for the current I that flows through an electrode of area A at a given anodic potential when subjected to a reciprocating wearing contact with a frequency of f as

$$I = fA_a \int_0^{1/f} i_a(t)\ \mathrm{d}t + f(A_s - A_a) \int_0^{1/f} i_p(t)\ \mathrm{d}t \cdots \quad [29]$$

where A_a is the worn area, $A_s - A_a$ is the passive area, $i_a(t)$ is the repassivation current density, and $i_p(t)$ is the passive current density at a given anodic potential.

A model that incorporates a time constant for repassivation has been developed by Goldberg[89] for currents generated by scratching passive surfaces. It has the form

$$I(t) = I_{peak} \exp\left(\frac{-(t - t_0)}{\tau}\right) + I_p \cdots \quad [30]$$

If Tafel behavior is applicable, then

$$I_{peak} = i_p A_r\ e^{[n_f/b_a]} + \frac{\delta \rho_{ox} z F A_r}{M_w \tau} \cdots \quad [31]$$

where I_{peak} is the peak current ($I_{peak}=I_{max}-I_p$), I_{max} is the maximum recorded current, τ is the time constant for repassivation, t_0 is the time to produce the scratch, I_p is the passive current at $t=\infty$, i_p is the current density for passivation, A_r is the scratch area, n_f is the scratch overpotential, b_a is the Tafel slope, δ is the repassivated oxide thickness, ρ_{ox} is the oxide density, z is the charge on the cation, F is Faraday's constant, and M_w is the oxide molar mass.

Mischler[90] developed a model for reciprocating contacts that includes the effects of load and material hardness. He modeled the anodic current $I_{a,w}$ generated by wear induced by a hard insulating counter body.

$$I_{a,w} = K_w l f \left(\frac{W}{H}\right)^{\frac{1}{2}} \int_0^{1/f} i\ \mathrm{d}t \ldots \quad [32]$$

where K_w is the proportionality factor, l is the length of wear track, W is the normal load, H is the hardness of metal, t is the instant time, f is the frequency of reciprocation of contact or rotation rate if for pin-on-disc, and i is the instantaneous current density.

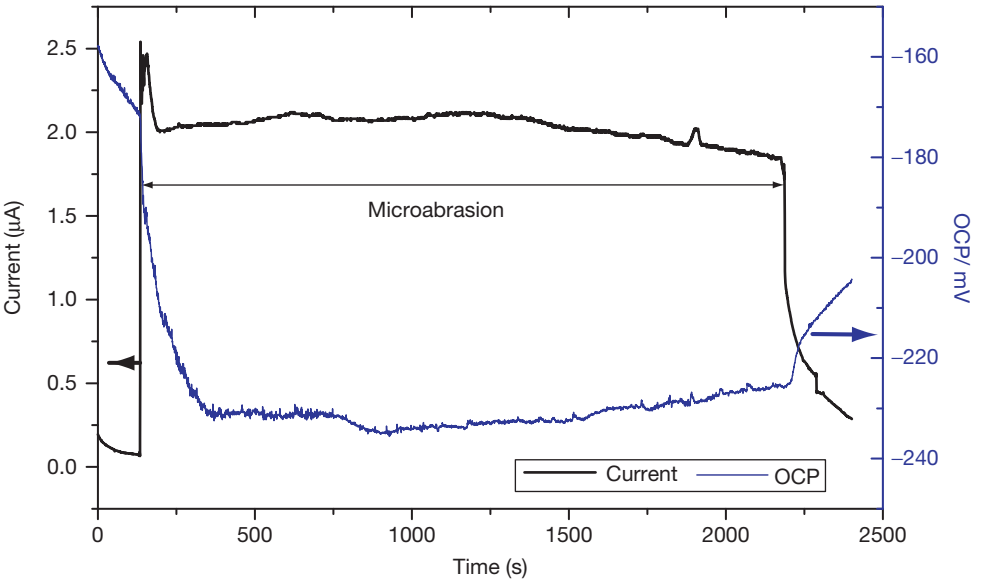


Figure 32 OCP and current as a function of time during microabrasion of a 250 μm thick D-Gun WC–10Co–4Cr coating on stainless steel 316 for a sliding distance of 90 m using 0.16 volume fraction slurry of 4.5 μm SiC abrasives suspended in NaOH solution of pH 11. The applied load was 0.2 N and a Zirconia ball of 25.5 mm dia was rotated to give a sliding velocity of 0.05 m s^{-1}. Reproduced from Wood, R. J. K. *J. Phys. D: Appl. Phys.* **2007**, *40*, 5502–5521.

These three models are essentially based on the same principle but involve different contact parameters. Their validity may be limited to the specific test conditions/rig geometry used in their development. More complex models are being developed for tribocorrosion interactions in aqueous sliding contacts,[91] and these will be explored in the following section. However, it is extremely important to recognize the roles of reactive species in the tribological processes. Jiang and Stack[92] discuss the mechanisms of wear debris generation and the roles of reactive species in the generation of wear debris during sliding wear in gaseous or aqueous environments. Low cycle fatigue was taken as an important mechanism for the generation of wear debris particles during sliding wear of metals. The generation of wear debris particles was described as a process of microcrack initiation and propagation. The presence of reactive species in the environment was taken to chemically activate reaction(s) at the crack tip that weaken(s) the bonding of atoms at the tip, increasing the rate of generation of wear debris particles. The ratio between the contact frequency and the corrosion passivation rate was found to be a very important factor in determining the dominant material loss regime for a given tribosystem. At low values, wear-induced corrosion tends to be more significant in causing the overall material loss. However, corrosion-induced wear tends to be the dominant regime for material loss at high values.

2.15.3.2 Wear-Corrosion Interactions

2.15.3.2.1 Mechanisms of enhanced surface degradation

These interactions can be defined as follows. The total damage under wear-corrosion, T, can be represented as

$$T = E + C + S \qquad [33]$$

where E is the pure wear material loss and C is the pure corrosion rate. The synergistic effect (interactive term), S, is referred to as ΔE or $(\Delta C + \Delta E)$ depending on the literature source, where ΔE is the enhanced wear loss due to corrosion and ΔC is the enhanced corrosion due to wear.

Synergy is defined as 'the difference between wear-corrosion and the summation of its two parts' and can be expressed by eqns [34] and [35].

$$\Delta E = T - (E + C + \Delta C) \qquad [34]$$

or

$$S = \Delta E + \Delta C = T - (E + C) \qquad [35]$$

where T, C, and E are typically gravimetric terms relating to wear-corrosion, electrochemical corrosion (*in situ*), and mechanical wear mechanisms respectively. C can be taken as the corrosion current under static conditions and ΔC measured *in situ* when wear processes are occurring. Under erosion, more confusion exists in the literature as static C or flow only corrosion rates C' are used leading to very different values of S as shown in eqns [36]–[38].

$$T = E + C + \Delta E + \Delta C \quad [36]$$

$$T = E + C' + \Delta E + \Delta C' \quad [37]$$

$$T = E + C'' + \Delta E \quad [38]$$

where C' is the solids free flow corrosion rate and C'' is the corrosion rate under wear-corrosion conditions. The ASTM G 119-93 standard is a useful guide for researchers wishing to evaluate synergy.[93] It should be noted that researchers working in the field of tribocorrosion are now moving away from evaluation of and using the terms synergy and antagonism in favor of quantifying the rates of electrochemical–mechanical interactions and identifying their mechanisms and physical meaning.

Wear can mechanically strip the protective corrosion film creating fresh reactive corrosion sites and produce ΔC,[94] depending on the rate of repassivation and the integrity of the film formed. Other possible wear-enhanced corrosion mechanisms include (1) local acidification at wear sites, accelerating corrosion rates, and prohibiting film formation; (2) increased mass transport by high turbulence levels; (3) lowering the fatigue strength of a metal by corrosion; (4) anodic wear scars can cathodically polarize the surrounding unworn surfaces and destabilize passive films in these regions enhancing corrosion[95]; and (5) surface roughening of the specimen during wear-enhanced mass transfer effects increasing the corrosion rate.[95] Corrosion-enhanced wear mechanisms are also possible (ΔE). The ΔE wear rate could be due to (6) the removal of work-hardened surfaces by corrosion processes, which expose the underlying base metal to erosion mechanisms[96]; (7) preferential corrosive attack at grain boundaries resulting in grain loosening and eventual removal[97]; (8) the increase in the number of stress concentration defects resulting from micropitting; and (9) detachment of plastically deformed flakes on the metal surface due to stress corrosion cracking. Corrosion can increase surface roughness, which in turn can increase wear processes. This is not a synergistic interaction.

Most of these mechanisms, if dominant, would be expected to lead to positive synergy (i.e., detrimental and an acceleration of wear-corrosion). However, in some instances, an antagonism can occur. Possible mechanisms that reduce wear rates ($-\Delta E$) are (10) microgalvanic (work-function difference) effect-induced inhomogeneous shot-peening[98] by high-velocity sand particle impacts or; (11) the presence of a soft or loosely adherent corrosion film or; (12) the blunting of the crack tips by lateral dissolution retarding the speed of crack propagation. The reduction in corrosion rates ($-\Delta C$) could result from (13) a rapid corrosion film growth, scaling, or the formation of a passive film reducing corrosion rates dramatically or; (14) the generation of thermally affected film.

Malka *et al.*[99] give an example of where positive synergies are dominant and have looked at whether erosion enhances corrosion and/or corrosion enhances erosion in pipe loop experiments on uncoated AISI 1018 CS. The tests solutions contained 1 wt% NaCl solutions purged with CO_2 (partial pressure of 1.2 bar) and 2 wt% silica sand (275 μm in size). They found that erosion enhances corrosion and corrosion enhances erosion but the dominant synergistic effect was that of corrosion on erosion, that is, ΔE. For roughly equal corrosion (C') and erosion rates (E), the total mass loss (T) was equal to two to three times the simple summation of E and C', ($E+C'$), showing significant synergy under 2 m s^{-1} flow velocities.

It has also been shown that for a variety of engineering materials the synergistic effects, which result in the damage due to separate corrosion and erosion processes, are normally greater than the sum of the individual damage processes and can accelerate material removal significantly 50 times in some cases (i.e., $T = 50 \times (C' + E)$), see Wood and Hutton.[31] Thus, the levels of S can be as high as 50 times the expected mass loss rate and would comprise most safety margins used in the design of fluid machinery and pipe systems.

An example of antagonism was seen in erosion-corrosion testing of naval grade cast nickel aluminum bronze (NAB) under jet impingement conditions with a 2.5% (w/w) sand/3.5% NaCl solution slurry. **Figure 33** shows the rig used and **Figure 34** shows the results from combined erosion-corrosion (T), erosion (E), and flow corrosion (C') tests for different sand particle kinetic energies generated under the conditions detailed in **Table 5**. As can be seen, the T values are less than $E+C'$ values.

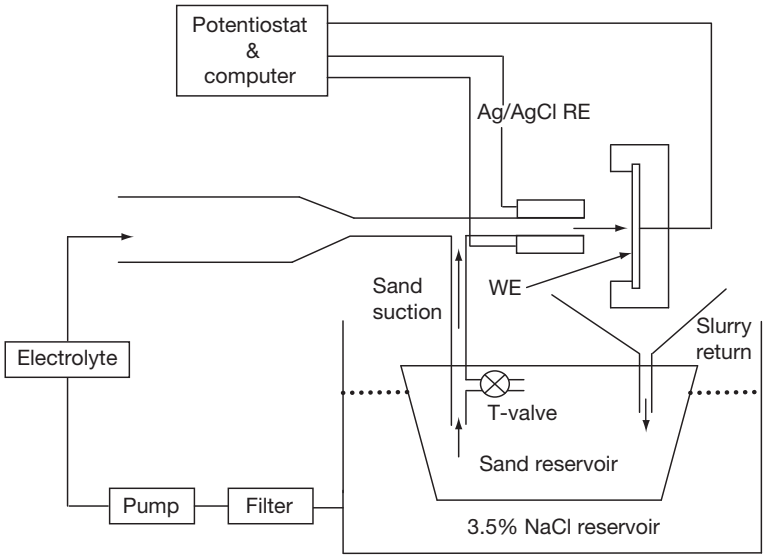


Figure 33 Modified jet impingement slurry rig used for electrochemical measurement under flow corrosion and erosion-corrosion.

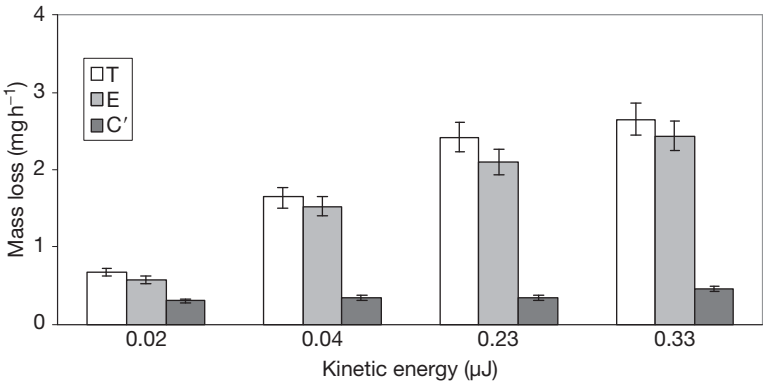


Figure 34 Mass loss rates for cast nickel aluminum bronze (NAB) at different kinetic energies. Conditions: 2.5–3.0 w/w% silica sand concentration, 3.5% NaCl solution.

The reason for this antagonistic effect was determined by profiling the microhardness of uneroded and eroded areas to reveal an increase in surface hardness in eroded areas as shown in Figure 35. Within the erosion scar, the hardness was doubled by hardening processes by the 10^5–10^6 strain rates induced by the sand impacts. Transverse sections revealed a work-hardened layer some 125 nm thick, while for erosion-corrosion the layer was 250 nm thick and the erosion mechanisms were subtly different. How corrosion influences the depth of work-hardening is not clear but it is likely that work-hardening will change the surface potential and thereby will decrease the corrosion reaction rate of the eroded surface as well as reduce erosion rates.

It is difficult to represent the interactive terms between wear and corrosion graphically to allow instant insight for designers and equipment operators. But some attempts at mapping are useful, particularly where quantitative terms are used for material loss rates. Such maps can also indicate the tribocorrosion regimes or dominant loss mechanism. This can be used to select more resistant surfaces or compositions. An example of such a plot for erosion-corrosion is given in Figure 36. Here the corrosion loss rates (C') have been plotted against the mechanical loss rates given by ($T-C'$). This assumes that $\Delta C'$ is negligible

Table 5 Conditions used for the erosion and erosion–corrosion experiments (CP = cathodic protection)

Experiments	*Jet velocity ($m\,s^{-1}$)*	*Sand diameter (μm)*	*Test solution*	*Re jet*
Corrosion (C')	3.1–6	–	3.5% NaCl	
Erosion (E)	3.1–6	135 and 235 (mean diameter)	3.5% NaCl with CP	18415–29702
Erosion–corrosion (T)	3.1–6	135 and 235 (mean diameter)	3.5% NaCl	

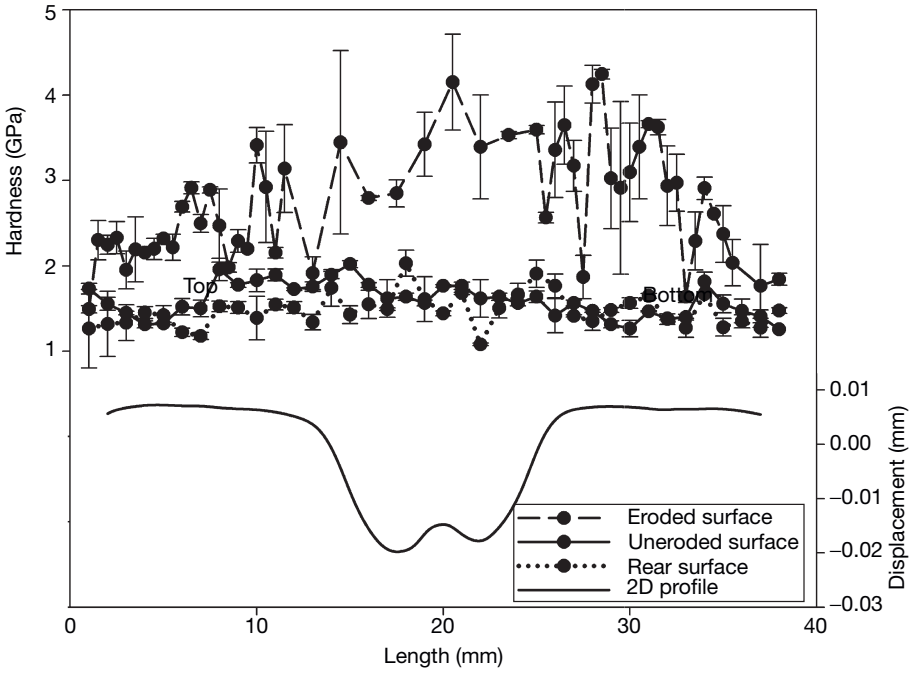


Figure 35 Hardness measurement for NAB on eroded and uneroded and back surface tested at 6 m s^{-1} with its 2D-profile. Positions (1) and (3) represent hardness measurements through the center of the erosion scar and far away from the scar respectively. Reproduced from Barik, R. C. Environmental factors affecting the marine corrosion performance of nickel aluminium bronze, PhD thesis, School of Engineering Sciences, University of Southampton, 2006.

and can be ignored. Results for CS, cast nickel aluminum bronze, and a sprayed nickel aluminum bronze coating are shown. The regimes are also shown defined by the ratio $(T-C')/C'$, where low values show electrochemical processes to be dominant, values between 0.1 and 1 are in the corrosion-erosion regime, while values between 1 and 10 are in the erosion-corrosion regime and values over 10 are in the mechanically dominated regime.

For abrasion-corrosion, a variation in synergy levels has been linked to the abrasion mechanism present in the contact. **Figure 37** shows results from microabrasion-corrosion tests on types 304 and 316 stainless steel (UNS S30403 and S31603) and super duplex stainless steel (UNS 32760).[100] The levels of synergy are represented as a percentage calculated by eqn [39] where C (passive corrosion rates) are assumed negligible.

$$\frac{S}{T} \times 100 = \frac{T-E}{T} \times 100 \qquad [39]$$

The synergy levels $(S/T \times 100)$ are plotted against volume fraction of abrasive. The abrasive used was 4.5 μm SiC particles. The abrasive mode is dependent on the abrasive concentration. For low concentrations of abrasive, a two-body grooving abrasion process dominates, while for higher concentrations three-body rolling abrasion processes dominate. It can be seen from **Figure 37** that the synergy resulting from a two-body wear mode was highly variable as evidenced by the large degree of scatter. Both S31603 and S32760 stainless steels produced negative values

of S under this wear condition, while S was positive for S30403. The large antagonisms produced by the two-body wear mechanism may be a reflection of the differences in repassivation kinetics and/or composition of the passive films reducing the overall level of two-body abrasion.[100] However, the mixed-mode region shows a significant reduction in scatter for each stainless steel. Likewise all the three stainless steel types within the three-body region showed an improved reproducibility with increasing volume fraction of abrasive. Rabinowicz *et al.*[101] in their study of abrasive wear under a three-body regime reported that three-body abrasion leads to more reproducible test results compared with two-body. Batchelor and Stachowiak[102] have also suggested that three-body abrasion is more effective in removing the corrosion products, and by implication any passive film present, during abrasion-corrosion. **Figure 37** also shows that under three-body wear, S generally converges with increasing slurry concentration to a uniform positive synergy of 18% for the three stainless steels.

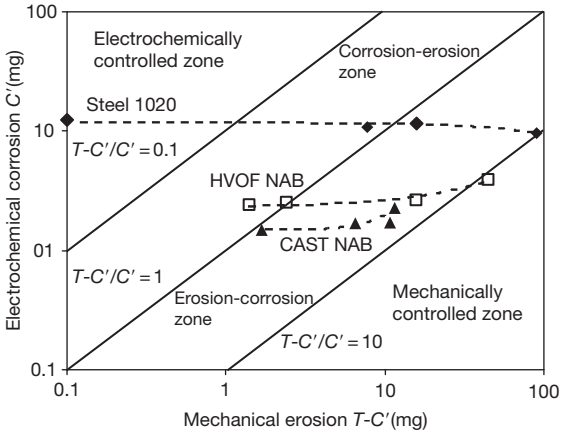


Figure 36 Erosion-corrosion regime map with electrochemical material loss versus mechanical material loss for erosion-corrosion of CS AISI 1020, HVOF Nickel Aluminum Bronze coating and cast Nickel Aluminum Bronze under jet impingement conditions with 2% (w/w) sand in seawater. Reproduced from Wood, R. J. K. *Erosion-Corrosion, Comprehensive Structural Integrity*; Elsevier: Oxford, UK, 2007.

2.15.3.2.2 Erosion-corrosion models

Ferng *et al.* [103,104] modeled erosion-corrosion in pipe flow by adding the erosion and corrosion contributions together. The erosion model was based on droplet impingement but used an oxide removal based on eqn [40] for predicting the mass of oxide removed, m,

$$m = C_s N f(\alpha) \frac{\rho_f U_f^2}{H} \qquad [40]$$

where C_s is a system constant, N is the frequency of impingement, f is a characteristic function, α is the impact angle, ρ_f the fluid density, U_f is the normal velocity, and H is the pipe wall hardness.

The metal loss rate, M_R, in mol cm^{-2}s^{-1} model is based on a mixed control corrosion process of dissolution of magnetite and the mass transfer of ferrous ions Fe^{2+} giving

$$M_R = \frac{C_{eq} - C_\infty}{(1/2k_1) + (1/k_m)} \qquad [41]$$

where k_1 is the reaction rate constant, C_{eq} is the

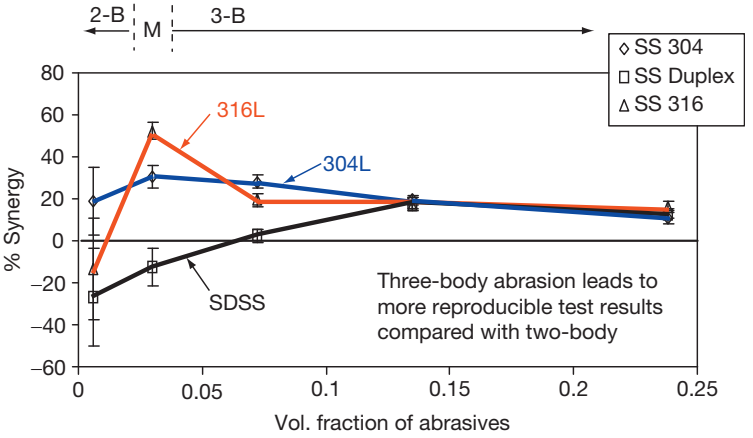


Figure 37 Synergy % ([S/T]×100) versus volume fraction of the abrasives for UNS S30403, S31603 and S32760. Low volume fraction of abrasive gives two-body grooving abrasion while higher concentration generate three-body rolling abrasion. Intermediate concentrations caused a mixed two- and three-body mode (M on figure).

soluble ferrous ion concentration in equilibrium with magnetite, C_∞ is the soluble ferrous ion concentration in the bulk solution, and k_m is the mass transfer coefficient. The combination of eqns [40] and [41] with 3-D fluid modeling has been successfully used to predict erosion-corrosion rates of pipes and fittings and the location of erosion-corrosion attack.

Erosion damage by solid particle impact or cavitation bubble collapse to an oxide or passive film will reveal the underlying nascent surface inducing a higher activity (higher corrosion current), for a limited duration, than for the intact oxide surface. Bozzini *et al.*[105] employed a simple approximate model using a 'recovering target' concept similar to those used in **Section 2.15.3.1.3** for sliding wear-corrosion. This erosion-corrosion model has the advantage that it can be applied to both passivating and actively corroding conditions. The impacting particles are modeled with rigid spheres of uniform radius r_p. Particle impact is assumed to be a Poisson process with parameter λ (impacts $m^{-2}\,s^{-1}$). It is assumed that each impact gives rise to an alteration of the corrosion rate through a localized change in corrosion current density for a period of time, relating to a recovery to its initial state. The effective corrosion current density, i_{corr} ($nA\,cm^{-2}$), at a given electrode potential (typically the corrosion potential) can be related to the mechanically influenced corrosion component of the synergistic damage through a coefficient f_a, (such that $0 \le f_a \le 1$) expressing the fraction of the corroding surface, which is affected by the erosive action of impinging particles, by eqn [42]

$$i_{corr} = f_a i_{act} + (1 - f_a) i_u \qquad [42]$$

where the subscripts 'act' and 'u' relate to 'activated' and 'unaffected' respectively. The current densities i_{act} and i_u are characteristic of the corroding material in the absence and in the presence of the erosive action and can be measured separately by means of suitable experiments. In general, the coefficient f_a can be defined in eqn [43]:

$$f_a = \left(\frac{\text{No. of impacts}}{\text{Control area}}\right) \times \left(\frac{\text{Damaged area}}{\text{Impact}}\right) \times \text{Recovery time} = \lambda A_a \tau \qquad [43]$$

where A_a is the affected surface area and τ is the passive recovery time.

Equation [42] indicates an accelerating element to corrosion caused by film damage or removal and is one element of the complex interactions between erosion and corrosion that needs to be understood. A similar approach that includes particle impingement angle and assumes an exponential repassivation reaction has been developed by Lu.[106]

$$i_{corr} = i_s + i_{peak} \frac{\tau \dot{A}_e}{1 + \tau \dot{A}_e} \qquad [44]$$

where $\dot{A}_e$ is the normalized rate of active surface area generated

$$\dot{A}_e = \left(\frac{C_p U \sin\theta}{\rho_p V_p}\right) A_{crater} \qquad [45]$$

where i_s is the stable passive current density, i_{peak} is the peak response in local current density due to impingement, τ is the exponential decay constant, C_p is the mass of solids per unit volume, U is the impact velocity, θ is the impact angle, ρ_p is the particle density, V_p is the particle volume, and A_{crater} is the crater area.

More recent erosion-corrosion models incorporate dynamic Hertzian impact area models and squeeze film effects in an attempt to evaluate more accurately actual areas of depassivation and actual particle impact velocities, Harvey.[54]

As mentioned earlier, the synergistic term S can be positive or negative (antagonistic). Bozzini *et al.*[105] showed that annealed CS had a more active corrosion potential before it was work hardened, resulting in $i_u > i_a$ and by using eqn [42] showed a reduction (i.e., an antagonism) in overall corrosion rate with erosion present.

The Bozzoni model or ones of that type go some way to explain experimental trends. For example, Neville *et al.*[107] have shown that the corrosion current density increases with increasing solid loading (200, 400 and 600 $mg\,l^{-1}$) for UNS S31603 under an impinging jet of 3.5% NaCl at 17 $m\,s^{-1}$. Synergistic effects have also been shown to be a function of impinging solid particle energy for various stainless steels (UNS S31603, S32100, S32250, and S32760).[108] Such results indicate that synergy is more pronounced at lower particle energies. For example, under an erosion energy of 0.02 μJ per impact, the volume loss under erosion-corrosion for UNS S32250 is ten times that for erosion alone. However, at higher energies, ~7.5 μJ, the erosion-corrosion rate is only 1.4 times that for erosion alone. Under very low energy impacts, little differences between erosion and erosion-corrosion rates are recorded for passive steels suggesting that impact energies above a critical value maybe required to initiate depassivation.

There are data reported in the literature that certain erosion-corrosion test conditions producing an antagonism, while others produce a positive synergy. The author in Wood[28] reports antagonisms for a superduplex steel UNS S32750 in 3.5% NaCl solutions at $E_k = 0.05\,\mu J$ and 90° impingement angle. However, Neville and Hu observed a positive synergy on two superaustentic stainless steels (UNS S31254 and S32654) and a superduplex steel UNS S32750, but these tests were conducted at a higher E_k of 4 μJ.[29] This suggests a complex relationship with synergy as a function of erosion conditions (energy) and repassivation tendency as this effects the crater size, crater shape, levels of strain rate, and passive film or oxide layer recovery time. **Table 6** shows some typical values for E_k that may be generated in pipes, valves, pumps, and propulsors. It assumes average sand particles of 100 μm in diameter and that these have velocities comparable to the average fluid velocities of the systems. Therefore, erosion-corrosion data are required at these energies if correct material selection is to be made.

Erosion-corrosion conditions can themselves develop local microgalvanic cells between eroded and uneroded surface areas. Such microgalvanic cells have been reported between erosion affected regions (depassivated) of a surface and noneroded regions (passive). For example, enhanced material loss rates due to such cells are reported by Hodgkiess.[23]

To summarize, **Table 7** is an attempt to list the mechanisms that may influence the sign of the interactive terms namely the ΔC and ΔE components.

Another possible synergy mechanism for ΔE has been suggested by Guo *et al.*, [110] who studied near-surface hardness changes induced by erosion strain hardening and corrosion softening on AISI 1045 CS and observed that the relative hardness degradation, $\Delta H_v/H_v$, could be correlated approximately with the anodic current density, i_a, as expressed in eqn [46]:

$$\frac{\Delta H_v}{H_v} = -B \log\left[\frac{i_a}{i_{th}}\right] \qquad [46]$$

Table 6 Potential particle kinetic energies and flow velocities for marine components

	Marine component			
	Pipe	*Valve*	*Pump*	*Propulsor*
Velocity ($m\,s^{-1}$)	0–3	5	10	25
Particle diameter (μm)	100	100	100	100
Particle KE (μJ)	0–0.007	0.02	0.07	0.44

where $\Delta H_v = H_v^* - H_v$ and is the change of hardness due to anodic dissolution at the surface, H_v and H_v^* the hardness without and with anodic current present at the electrode surface respectively, $B\,(>0)$ is an experimental constant, and i_{th} is the threshold anodic current density to induce the synergistic electrochemical–mechanical effect, that is, when $i_a \leq i_{th}$, $\Delta H_v = 0$. However, relatively large anodic currents were employed to achieve relatively minor changes in hardness (typically 0.1%). This could be an artifact of the experimental equipment employed as a microhardness tested was performed. Thus, surface layer properties were affected by the substrate as more than 10% of the coating thickness was penetrated during measurement.

Empirically based models of synergy

Wood and Hutton,[31] summarizing experimental data published on erosion wear over a wide range of solid particle and cavitation erosion experiments, found that the data fell within two groups: the medium- and high-synergy systems. The erosion-corrosion enhancement, S, was found to be closely linked to the pure erosion component E. Two expressions were derived by plotting S/C' against E/C' ratios to obtain the following expressions: For the medium-synergy group

$$1\,\frac{S}{C'} = \exp\left[1.277\,\ln\left(\frac{E}{C'}\right) - 1.9125\right] = 0.1477\left(\frac{E}{C'}\right)^{1.277} \qquad [47]$$

For the high-synergy group

$$2\,\frac{S}{C'} = \exp\left[0.755\ln\left(\frac{E}{C'}\right) + 1.222\right] = 3.3940\left(\frac{E}{C'}\right)^{0.755} \qquad [48]$$

Sliding wear models

A sliding wear model proposed by Jiang *et al.*[91] defined the total wear-corrosion loss V_{cw} as

$$V_{cw} = (V_w + \Delta V_w) + (V_c + \Delta V_c) \qquad [49]$$

where V_c is the pure corrosion loss, V_w is the material loss due to wear, ΔV_c is the wear-induced corrosion, and ΔV_w is the corrosion-induced wear.

They looked at material loss due to corrosion, V_c, during sliding in a corrosive environment and related this to the anodic corrosion current under the influence of wear by Faraday's law. For metals that can

Table 7 'Synergy' overview of processes that could lead to positive and negative interactive effects between mechanical and electrochemical processes present under erosion-corrosion conditions

Factor	*Positive interaction*	*Negative interaction*
Critical impact energy	Above critical impact energy to penetrate/damage passive/product film/coverage. Results in increased charge transfer at the liquid/metal interface	Below critical impact energy. Results in reduced charge transfer at the liquid/metal interface as film composition could be influenced by impact
Surface roughness, R_a	Roughening effects on mass transfer coefficients unknown but k_m is likely to increase with increasing R_a. The Silverman review[109] suggests that the mass-transfer dependence on Reynolds number is affected by surface roughness as $k_m \propto Re^{0.9}$ with solids eroding the surface compared to $k_m \propto Re^{0.65}$ for solids-free flows for rotating cylinder work due to roughening effects on the surface. Roughness may promote microturbulence, affecting the local double layer. Roughness could also promote local microelectrode behavior (potential field distortion) at the tips of impact craters	Roughness influences the contact mechanics of angular solid particle impingement. Increased roughness could reduce contact stresses and thereby the near and far field stress distributions
Plastic deformation/strain	Plastically deformed and stressed surfaces enhance corrosion processes. Corrosion causes premature detachment of plastically deformed or strain hardened impact carter lips	
Increased or unsteady hydrodynamics or turbulence	Unsteady hydrodynamics leads to instablility of the double layers (nonsteady state) and unsteady driving concentration gradients of active species	High concentrations of solid particles could block the surface from incoming particles. Particle–particle interaction alters impact conditions
Contact temperature	Local surface (flash) temperatures could be significantly higher around and within impact craters which could accelerate corrosion rates	Local surface (flash) temperatures could influence oxide film composition and microstructure and thus could produce more erosion resistant films
Localized corrosion	Pitting and microgalvanic corrosion cells due to localized defects in the passive layers induced by erosion or exposure of inclusions or voids. Crack systems would also be vulnerable to crevice corrosion attack and accelerated crack propagation. Microgalvanic corrosion cells could be formed between erosion carters and surrounding unaffected areas	Strain-induced shift in surface potential in the electropositive direction
Multiphase surfaces	Corrosion of interface – potential to lose bond integrity between hard and binder phases, binder recession due to corrosion leaves hard phase loose to be removed by mechanical action. These interactions could also apply to splat boundaries which exist after spraying	Substrate corrosion products can eject via interconnecting pores onto the coating surface, reducing erosion rate
Passive film state	Depassivation: removal of air-borne oxides or oxide layers produced by passivation. Oxide layer could increase friction between impacting solid particle and the bulk substrate material	The overall recovery times for repassivation and adherent oxide layers formed by passivation are important relative to depassivation events. Oxide layers could also decrease the friction between impacting solid particle and the bulk substrate material. Alternatively denser oxide could be formed that reduces corrosion or erosion or both

Source: Partly taken from reviews by Wood [31] and Wang.[30]

passivate in the corrosion environment, the anodic corrosion current is composed of two components: (1) the repassivation anodic current density of i_a passing through the area, A_r, where the passive film has been removed/damaged by the mechanical wear action; and (2) the passive current density of i_p passing through the rest of the apparent area of contact, $A - A_r$, where the passive film has been fully

developed. A is equal to the apparent area of contact, assuming that the area surrounding the wear track is insulated from electrochemical reactions. For nonpassive metals, the current density can be taken as i_p. It was assumed that a depassivation–repassivation event occurs in between two successive contact events during corrosion-wear sliding tests and that the removal of the passive films during sliding corrosion-wear occurs on the real area of contact, A_r, which depends on the mechanical properties of the material and the contact surface topographies. For sliding between two ductile rough bodies or a hard rough body sliding on a smooth ductile, the real contact area A_r is proportional to $l(W/H)^{1/2}$, where l is the width of the wear track.

Thus,

$$V_c + \Delta V_c = I_{ac} t A_{atom}/(zF\rho_t) \qquad [50]$$

where

$$I_{ac} = A_r i_r + (A - A_r) i_a \qquad [51]$$

Using eqn [32] due to Mischler

$$A_r i_r = I_{a,w} = K_w l f \left(\frac{W}{H}\right)^{\frac{1}{2}} \int_0^{1/f} i \, dt \qquad [52]$$

where

$$i = (i_0 - i_a) \exp(-t/\tau_0) + i_a \qquad [53]$$

The characteristic passivation time, τ_0, reflects how easily the metal surface passivates. The lower the value of τ_0, the more easily the metal tends to passivate. For nonpassive materials, $i_0 = i_a$, and the anodic dissolution current, $i = i_a$, is constant. This gives

$$V_c + \Delta V_c = [A_{atom} s/(zF\rho_t v_e)] \Big\{ A_a i_a + K_a l (W/H)^{1/2} (i_0 - i_a) f \tau_0 [1 - \exp(-1/f\tau_0)] \Big\} \qquad [54]$$

where s is the total sliding distance, v_e is the sliding velocity, K_a is a proportionality constant, and f is the contact frequency between interacting surfaces.

This approach is coupled to the concept of microfatigue crack propagation where a critical time, t_c, is required for an existing microcrack nucleus to propagate a length, D_a, taken as the diameter of a wear debris particle and is then related to the microcrack growth rate, da/dn. There are various scenarios of how corrosion affects crack propagation. First, corrosion can increase the number of potential microcrack initiation sites, N_c. Second, reactive or corrosive species at the crack tip can promote the propagation of the crack, increasing the generation rate of wear debris particles by reducing t_c. Third, the corrosive environment may also affect the average wear debris particle diameter, D_a, although it is difficult to predict how this factor will vary with the corrosion conditions. Finally, it is possible under certain conditions that wear is reduced as a result of crack-tip blunting or the formation of some wear-protective load-bearing areas from the corrosion products. However, Jiang *et al.*[91] considered the corrosion-enhanced microcrack propagation only.

The wear volume of a specimen is equal to the product of the average volume of a single particle and the total number of wear debris particles generated within a sliding time, t, which can be expressed by

$$V = \frac{\pi D_a^3}{6} k_d A_a t = \frac{\pi D_a^3 s}{6 v_e} k_d A_a \qquad [55]$$

where k_d is the particle generation rate.

They assumed that during sliding wear, a microcrack is initiated on the wear surface and will propagate for a certain distance after each contact with an asperity from the counter surface. A wear debris particle is generated after a certain number of cycles of contact at the same spot with the encountering surface.

The critical time, t_c, for generating a wear debris particle with an average size, D_a, from an existing microcrack nucleus is related to the microcrack growth rate, da/dn, by

$$D_a = \int_0^{t_c} \frac{da}{d\tau} d\tau = \int_0^{t_c} \frac{da}{dn} f \, d\tau \qquad [56]$$

If N_c active microcrack sites are present on a unit area of the wear surface that are ready to propagate, then the generation rate, k_d, of wear debris particles can be calculated by

$$k_d = N_c/t_c \qquad [57]$$

Several factors are considered: the number of cracks that are above a critical size and thus propagate; the influence of stresses induced during the wear process; as well as corrosion fatigue processes. This process considers mechanical aspects in a reference environment, interactions of cyclic mechanical loading and the attack by the environment and sustained load crack growth; the following terms were developed:

$$k_d = (N_r + \Delta N_c)\{g_r + g_{cfs}[1 - \exp(-1/f\tau_i)] + g_s\} f D_a^{-1} \qquad [58]$$

where N_r is the number of critical microcracks that are ready to propagate to generate wear debris

in a reference environment, ΔN_c is increment of the number of critical microcracks due to the presence of corrosion, and τ_i is the characteristic reaction time for the reaction controlling a corrosion fatigue process.

Substituting this expression into eqn [57] gives the wear contribution $V_{wc} = V_w + \Delta V_w$:

$$V_{wc} = \frac{\pi D_a^2 A_a s}{6l}[N_r g_r + \Delta N_c g_r + (N_r + \Delta N_c) \times \{g_{cfs}[1 - \exp(-1/f\tau_i)] + g_s\}] \quad [59]$$

Thereby the total wear-corrosion loss, given by eqn [49], can be predicted by summing eqns [54] and [59].

2.15.4 Present and Future Applications of Tribocorrosion

Previous sections have focused on 'heavy' engineering applications and solution and models appropriate to these conditions. This knowledge and understanding have matured over the past 20 years or so. However, it may be the more subtle tribocorrosion problems that will drive further development and understanding. For example, corrosion–wear interactions can be used to control material loss for manufacturing of silicon wafers. This process is called chemical–mechanical planarization (CMP) and is an important process for building multilevel interconnections for electronic devices. CMP techniques have been developed for machining tungsten, aluminum, and copper. The fundamentals of these processes are not well understood and are an active area of research at present. The supporting electrolyte used in manufacturing is acidic and oxidizing leading to material removal by a cyclic process of passivation-film abrasion followed by repassivation and then passivation-film abrasion and so on. These are similar to wear-corrosion processed discussed earlier in the chapter.

2.15.4.1 Chemical Mechanical Polishing (CMP)

Figure 38 illustrates the general process with the silicon wafer typically being loaded horizontally and pressed onto a rotating conditioning pad with the contact between the wafer and pad being lubricated by a slurry. The wafer holder is itself rotated as well.

The roles of mechanical abrasion and electrochemical dissolution during CMP of aluminum have been studied by Tsai and Huang.[111] They found, as expected, that the total metal removal rate and the electrochemical dissolution rate of aluminum in 5 vol%. phosphoric acid + 0.5 M citric acid solution (pH 4) increased with increasing contact pressure and specimen rotation speed. However, the contribution of electrochemical dissolution to the total removal rate of aluminum was less than that produced by mechanical abrasion. They also detected metallic aluminum in the solution, indicating that the passive film was not the sole material removed under the chemical–mechanical abrasion condition but some wear of the parent material had occurred.

The influence of the type of conditioning pad used has been studied by Tan *et al.*[112] They studied three

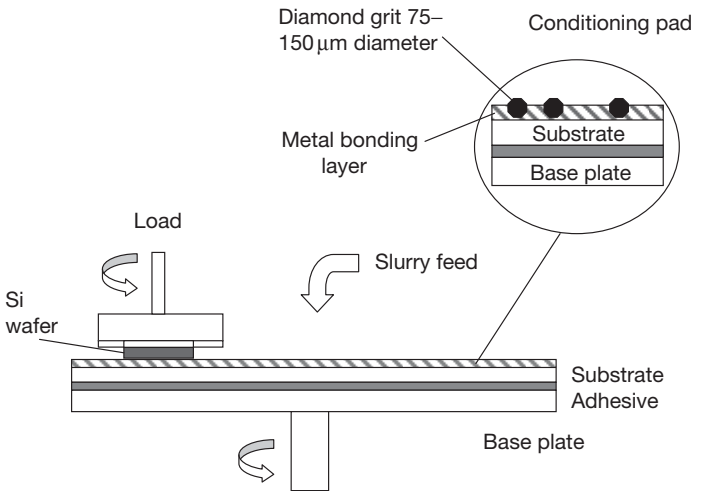


Figure 38 Schematic of a CMP set-up with a rotating wafer loaded against a rotating conditioner pad supported by a base plate.

types of conditioners: (1) bonded by electroplating nickel; (2) brazing Ni–Cr alloy; and (3) diamond-like carbon (DLC) film deposition on a brazed conditioner. They found that after wear-corrosion testing, the DLC film deposited on the conditioner surface of brazed Ni–Cr alloy exhibited better overall performance in terms of a higher removal rate. The pad suffered a lower wear loss and a better corrosion resistance compared with the other two conditioners.

Ziomek-Moroz *et al.*[113] researched into CMP of different metallic thin films, such as tungsten and copper in aqueous media. They emphasize the importance of constructing wear-corrosion maps for these complicated tribocorrosion-metallic thin film systems. The wear component followed the Preston wear equation,[114] namely

$$h = \int_{t_1}^{t_2} kpU\,\mathrm{d}t \qquad [60]$$

where h is the wear depth developed between time t_1 and t_2, k is a wear factor, p is pressure, and U is the relative velocity of the workpiece contact point with respect to the tool contact point. Equation [60] can be rearranged in terms of volume loss (V), normal applied load (W), and sliding distance (L), as shown in eqn [61].

$$k = \frac{h}{\int pU\ \mathrm{d}t} = \frac{h}{pL} = \frac{hA}{WL} = \frac{V}{WL} \qquad [61]$$

Ziomek-Moroz *et al.* found that the synergism between corrosion and erosion during CMP of thin metallic films is a key factor in controlling the polish rate and makes this process very complex. For the CMP of tungsten and copper, the presence of a passive protective film on the metal surface plays a critical role in successful planarization. Therefore, understanding the corrosion mechanisms that control both the polish rate and the passivation behavior of the thin films are critical in controlling CMP.

The direct planarizing of tungsten by wear processes is difficult because of its high hardness. Therefore, an effective approach has been developed to facilitate planarizing of a tungsten surface by removing a continuously growing passive film on tungsten when exposed to a low-pH potassium ferricyanide slurry. Since the passive film is softer than tungsten, this chemical mechanical planarization process is effective. Akonko *et al.*[115] have investigated the corrosion, wear, and corrosive wear behavior of tungsten in $K_3[Fe(CN)_6]$ slurries. The tungsten removal rate increased with increasing slurry pH and surface potential.

2.15.4.2 Erosion-Corrosion

In order to gain further understanding of the effects of erosion-enhanced corrosion, ΔE, electrochemical noise techniques are being used.[116] These measurements allow current and/or potential perturbations about the mean corrosion current to be analyzed without grossly polarizing the surfaces. These perturbations are thought to provide valuable information on the extent of passive film rupture and repassivation during flow-enhanced corrosion and erosion-corrosion. Although individual impacts or pitting events are difficult to resolve, the result of multiple impacts or pits as over time is useful as a diagnostic tool to detect possible surface degradation mechanisms.

Figure 39 shows a corrosion flow cell designed for electrochemical noise analysis for pipe flow with no solids present in the fluid. Typically, two 'nominally identical' working electrodes are used and their potential noise recorded relative to a reference electrode, together with the current noise between the two working electrodes. **Figure 40** shows typical potential and current responses at a pipe Reynolds number of 2000 (just around the transition from laminar to turbulent flow) on type-304 stainless steel in 3.5% NaCl solution.

Sasaki and Burstein observed current transients for single-particle impacts[117] on passive stainless steels, illustrating that monitoring such transients can provide direct information on the erosion-corrosion processes. A simple way to analyze such data is to take the standard deviation of the current traces under flow corrosion, σ_c, and under erosion-corrosion,

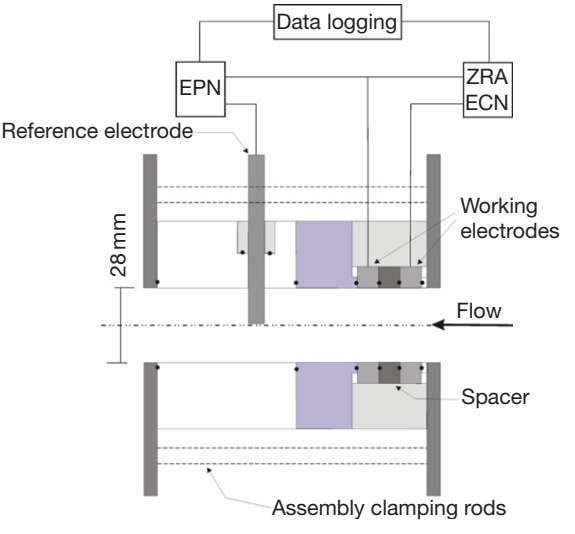


Figure 39 Schematic of electrochemical flow cell and measurement set-up.

σ_T. Attempts are currently being made to link the standard deviation ratio σ_c/σ_T to synergy, see Figure 41.

Caution should be applied to any electrochemical measurement under erosion-corrosion conditions. Other issues make electrochemical analysis and its comparison to synergy difficult. These include the possibility of local film currents between anodes/cathodes,[118] which will not be seen by ECN measurements or by potentiostats and the effects of charging/recharging double-layer currents due to fluctuating local events.

2.15.4.3 Biotribocorrosion

2.15.4.3.1 Dentistry

The use of artificial materials to replace natural dental tissue is common in dentistry. Due to health aspects, there is concern about the risks that the corrosion and wear of these materials may induce. Linear wear and the release of wear particles are especially of interest. The insight into wear-corrosion processes in dentistry is, however, still in its infancy. Lambrechts *et al.*[119] give an overview of the degradation processes that occur on artificial materials used in dentistry, especially in restorative dentistry. Light-cured composite resins of inorganic silica fillers in organic (polymer) resins are currently seen as viable alternatives to amalgam. Wear of dental composite resins involves adhesion, abrasion, attrition, chemical degradation, and fatigue, see Table 8. These mechanisms often act simultaneously and/or sequentially and interact with each other in a complex form, making the prediction of loss rates difficult.

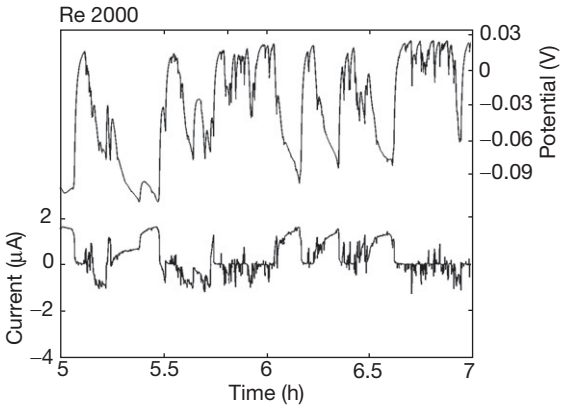


Figure 40 Electrochemical current and potential noise for AISI 304L in 3.5 wt.% NaCl solution.

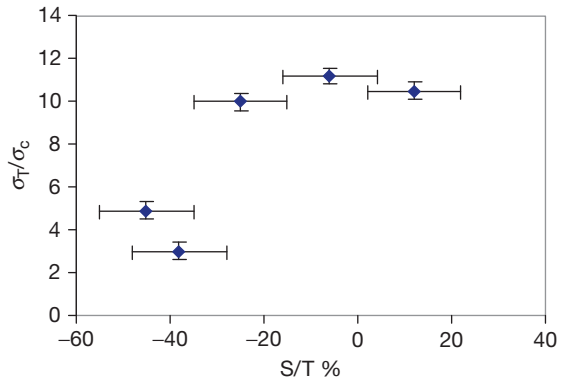


Figure 41 Standard deviation of erosion-corrosion current noise divided by the standard deviation of corrosion current noise as a function of the synergy as presented by gravimetric S/T%. Data from erosion–corrosion jet impingement tests on HVOF nickel aluminum bronze and aluminum alloy coatings in sand/3.5%NaCl slurries.

Wear in the filled composites occurs by a complex set of processes involving tribochemical reactions between the filler particles and water, formation of surface films containing a mixture of filler fragments and reaction products, and film delamination, as well as dissolution of the reaction products.[120]

The wear behavior of a dental ceramic composite containing 92 wt% silica glass and alumina filler particles in a polymeric resin matrix has been examined by Nagarajan *et al.*[121] Wear tests confirmed the presence of Al and other elemental constituents of the filler particles. It is proposed that three simultaneous processes occur at the sliding contact: tribochemical reactions and film formation, dissolution of the reacted products, and mechanical removal of the film by microfracture. At low loads, wear occurs primarily by a tribochemical mechanism, that is, formation and dissolution of the reaction products.

Lewis and Dwyer-Joyce[122] have reviewed the current understanding of wear of enamel and dentine in teeth and highlight the major influencing factors. They looked at various tribological conditions to which teeth are exposed. These include noncontact between opposing teeth, contact but no sliding, and contact with sliding between opposing teeth. They emphasize that the largest single influencing factor is that of the chemical environment and this interacts with a synergistic effect with all wear mechanisms to enhance material loss rates. Research has shown that increasing acidity reduces both hardness and elastic modulus of enamel in an appropriate linear relationship without load or sliding between teeth. Material loss for both enamel and dentine increases with decreasing pH generated by acids found in soft drinks. Citric acid adjusted by NaOH (to the required pH range between 2 and 6.5) and citric acid with titratable

Table 8 Dental wear and wear-corrosion processes

Wear mechanism	*Effect*	*Driving mechanism*	*Cause*
Abrasion	Occlusal material wear	Frictional surface interactions three-body abrasion	Toothbrush/paste food bolus
Attrition	Changes surface texture	Sharp asperities two-body abrasion	Antagonistic surface 50% harder than the wearing surface
Chemical wear	Degrades silane couplings in composites	Wetting	Food or saliva
Surface fatigue	Surface layer delamination and severe roughening	Rolling contacts and cyclic loads	Subsurface stress concentrations
Adhesive	Platelet-shaped wear transferred to antagonistic surface which can either reduce or enhance wear rates	Asperity–asperity interaction	Shear stresses

acidity give the highest loss rates. In the contact without sliding between teeth, the enamel loss increases with increasing load or cyclic load. For contacts with sliding, a nonlinear relationship is seen, with wear rates decreasing as the pH drops from 7 to 3 but increasing when the pH is further decreased to 1. Until now, these synergistic and antagonistic effects between the chemical and mechanical tooth wear have not been properly investigated.

Dental implants, on the other hand, are used to replace teeth lost due to decay, trauma, or periodontal diseases. These implants are subject to micromovements at the implant–bone interface or implant–porcelain interface due to the transmitted mastication loads and also subject to the aggressive oral environment. Such implants are thereby subject to tribocorrosion that can lead to implant failure. Vieira *et al.*[123] studied titanium grade 2 in contact with low-pH artificial saliva under fretting conditions and with different inhibitors. They found that the addition of citric acid or anodic inhibitor to artificial saliva resulted in a slight improvement in the tribocorrosion behavior of Ti. No significant differences were observed in the wear rate per dissipated energy, but a lower wear volume loss was obtained that can be attributed to the slightly lower corrosion rate observed in these solutions during the fretting tests. This increased protection was attributed to the oxidation and reduction reactions occurring in the contact area during fretting. They noted that the tribolayers formed in the contact became more stable after ~7000 cycles in solutions containing citric acid or anodic inhibitor, and resulted in lower coefficient of friction and a lower corrosion current. The addition of a cathodic or an organic inhibitor to the artificial saliva resulted in a deteriorating performance of the titanium. Both an increase in the wear volume loss per unit-dissipated energy and a significantly higher corrosion rate during fretting tests were observed in these solutions.

Figure 42 Total hip replacement joint in CoCrMo.

The tribocorrosion of orthodontic archwires is also a growing research area and various stainless steel archwires have been investigated by Berradja.[124] They found the fluctuations in potential and current during fretting-corrosion are more pronounced at increased sliding frequency than at increased applied normal loads.

2.15.4.3.2 Replacement joints

Researchers are becoming increasingly interested in metal-on-metal (MoM) joint implants, see **Figure 42**, with studies being focused on their wear and corrosion (ion release) behavior. The corrosion, wear, and wear-corrosion behavior for three materials (high-carbon CoCrMo, low-carbon CoCrMo, and UNS S31603) have been discussed by Yan *et al.*[125] In the steady state regime, 20–30% of the material

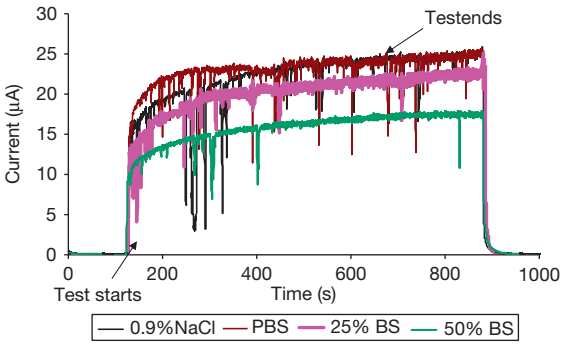


Figure 43 Electrochemical current noise obtained in microabrasion-corrosion test on as-cast CoCrCo (1.0 g cm^{-3} SiC) in 0.9% NaCl, phosphate buffered saline (PBS) solutions, 25% and 50% bovine serum (BS) at 37 °C. Reproduced from Sun, D. Wharton, J. A.; Wood, R. J. K.; Ma, L.; Raithforth, W. M. Microabrasion-corrosion of cast CoCrMo alloy in simulated body fluids, presented Leeds-Lyon 2007, Tribol. Int. submitted for publication.

degradation was attributed to corrosion-related damage. High-carbon CoCrMo showed excellent corrosion, wear, and corrosion-wear resistance and therefore delivered the best overall performance in terms of a lower wear rate, a lower friction coefficient, and a higher resistance to corrosion. Proteins in the joint fluids are thought to play a key role but this mechanism is not well understood.

Sinnett-Jones *et al.*[126] investigated the synergistic effects of corrosion and wear of surgical grade cast F-75 cobalt–chromium–molybdenum (CoCrMo) alloy. Microabrasion-corrosion identified depassivation and repassivation processes and showed strong synergistic effects ranging from negative to positive. The synergistic levels appear to be dependent on the integrity of the passive films and the repassivation kinetics.

The role of protein in the tribocorrosion of metal implants is not well understood but initial studies

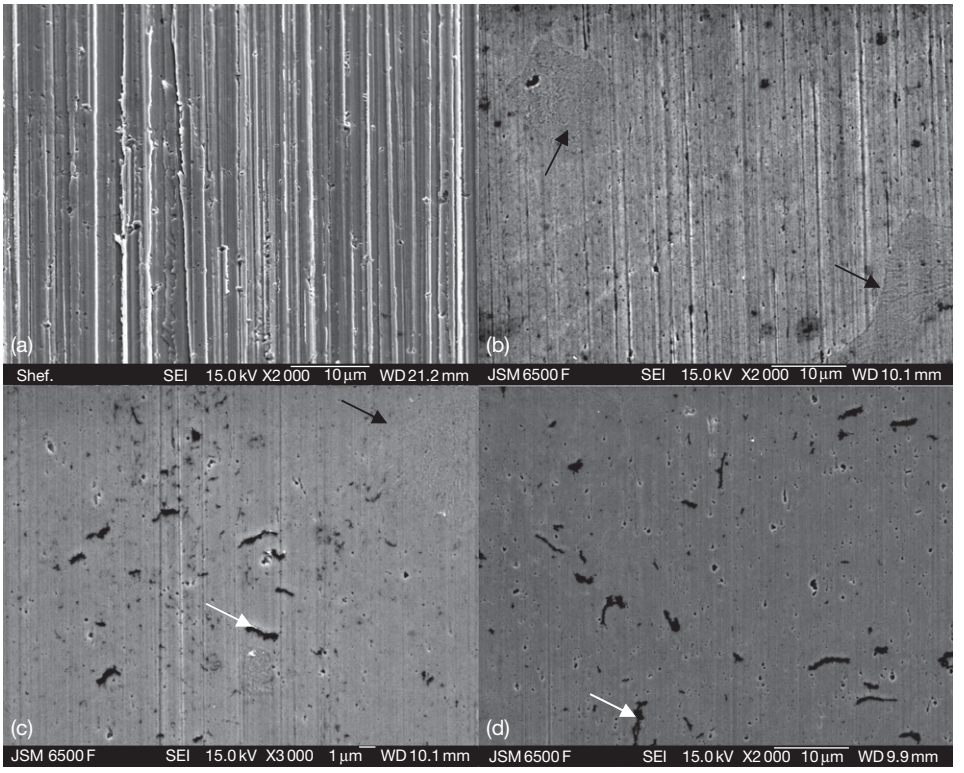


Figure 44 SEM micrographs of wear scars on as-cast CoCrMo under pure sliding conditions in (a) 0.9% NaCl, (b) PBS, (c) 25% BS, (d) 50% BS at 37°C. Black arrows show a carbide standing proud of the surface. The white arrows show possible protein and wear debris on the worn surfaces. Polished wear scars, as opposed to two-body abrasion seen in (a) and (b), were produced in proteinaceous solutions indicating proteins form a lubrication film which protect the surface. Reproduced from Sun, D. Wharton, J. A.; Wood, R. J. K.; Ma, L.; Raithforth, W. M. Microabrasion-corrosion of cast CoCrMo alloy in simulated body fluids, presented Leeds-Lyon 2007, Tribol. Int. submitted for publication.

show conflicting trends with corrosion being either enhanced or reduced when protein is present and adsorbed onto the articulating surfaces. Sun[127] has studied abrasion-corrosion of cast CoCrMo in saline (0.9% NaCl and phosphate buffer solution (PBS)) and protein containing solutions of BS. *In situ* current noise results, see **Figure 43**, show reduced current levels when proteins are present.

Using the same test geometry, but removing the abrasives, so that the alumina ball abrades the CoCrMo surface under sliding wear mode, again shows that the presence of protein alters and reduces the wear mechanism and rate as seen in **Figure 44**.

2.15.5 Conclusions

Tribocorrosion has only recently been recognized as an important underpinning area of corrosion and therefore is in its infancy. This chapter has defined tribocorrosion and presented the fundamental principles behind wear, flow corrosion, erosion-corrosion, cavitation erosion-corrosion, sliding wear-corrosion, and abrasion–corrosion interactions. It has shown the key material and environmental parameters that influence tribocorrosion and used examples to illustrate the performance of materials under a variety of tribocorrosion conditions. It has outlined the conditions under which tribocorrosion occurs in aqueous solutions and shown that the depassivation and repassivation kinetics within tribocorrosion conditions need to be understood or controlled if accelerated damage is to be avoided. It has discussed the motivation behind current research and presented some examples of where tribocorrosion research will play a major role in the future, such as in biotribocorrosion, manufacturing, and erosion-corrosion detection to aid condition monitoring of fluid machinery and chemical processes.

The influence of the corrosive environment is shown to be important in controlling wear and friction and the evolution of near-surface triboinduced layers, which are critical in both erosion-corrosion and wear-corrosion situations. The chapter has identified the key interactions between wear and corrosion mechanisms that occur along with some models and mapping techniques that aim at informing material selection and predict performance. However, these need major development before robust predictions can be made as our understanding of the mechano–electrochemical interactions in tribocorrosion are incomplete.

Acknowledgments

I thank all those in my research group who helped with this chapter and particularly Dr Julian Wharton for various figures and material and his constant encouragement. I also thank Drs Terry Harvey and Mandar Thakare as well as Dan Sun (PhD student) for their help with figures.

References

1. *Oxford Dictionary of English*, 2nd ed.; Oxford University Press, 2005.
2. Fischer, A.; Mischler, S. *J. Phys. D; Appl. Phys.*; **2006**, *39*(15); doi: 10.1088/0022-3727/39/15/E01.
3. Kurtz, S.; Ong, K.; Lau, E.; Mowat, F.; Halpern, M. *J Bone Joint Surg. Am.* **2007**, *89*, 780–785.
4. Celis, J.-P.; Ponthiaux, P. *WEAR* **2006**, *261*(9), 937–938.
5. Meng, H. C.; Ludema, K. C. *Wear* **1995**, *181–183*, 443–457.
6. Archard, J. F. *J. Appl. Phys.* **1953**, *24*, 981–988.
7. Archard, J. F. Wear Theory and Mechanisms. In *Wear Control Handbook*; Peterson, M. B., Winer, W. O., Eds.; ASME: New York, 1980; pp 35–80.
8. Hutchings, I. M. *Tribology – Friction and Wear of Engineering Materials*; Arnold: London, 1992.
9. Poulson, B. *Wear* **1999**, *233–235*, 497–504.
10. Kim, J. G.; Choi, Y. S.; Lee, H. D.; Chung, W. S. *Corrosion* **2003**, *59*(2), 121–129.
11. Jacobs, J. J.; Hallab, N. J.; Skipor, A. K.; Urban, R. M. *Clin. Orthop.* **2003**, *417*, 139–147.
12. Williams, D. F.; Clark, G. C. F. *J. Mater. Sci.* **1982**, *17*(6), 1675–1682.
13. Williams, R. L.; Brown, S. A.; Merritt, K. *Biomaterials* **1988**, *9*, 181–186.
14. Lewis, A. C.; Kilburn, M. R.; Papageorgiou, I.; Allen, G. C.; Case, C. P. *J. Biomed. Mater. Res. A* **2005**, *73*(4), 456–467.
15. Contu, F.; Elsener, B.; Bohni, H. *Corros. Sci.* **2005**, *47*, 1863–1875.
16. Goldberg, J. R.; Gilbert, J. L. J. *Biomed. Mater. Res.* **1997**, *37*(3), 421–431.
17. Stack, M. M.; Jana, B. D. *Tribol int* **2005**, *38*(11–12), 995–1006.
18. Celis, J.-P.; Ponthiaux, P.; Wenger, F. *Wear* **2006**, *261*(9), 939–946.
19. Deforge, D.; Wenger, F.; Ponthiaux, P.; Lina, A; Ambard, A. Tribocorrosion of AISI 304L stainless steel sliding against AISI 304L stainless steel in nuclear environment, 207th Electrochemical Society Meeting, Quebec City, Canada May 2005, session E1 on Chemical, Electrochemical, and Mechanical Effects on CMP, Tribocorrosion, and Biotribocorrosion, abstract #313.
20. Kaczorowski, D.; Combrade, P.; Vernot, J. Ph.; Beaudouin, A.; Crenn, C. *Tribol Int.* **2006**, *39*, 1503–1508.
21. Hu, X.; Neville, A. *Wear* **2005**, *258*(1–4), 641–648.
22. Al-Hosani, H. I.; Saber, T. M. H.; Mohammed, R. A.; Shams El Din, A. M. *Desalination* **1997**, *109*, 25–37.
23. Hodgkiess, T.; Neville, A.; Shrestha, S. *Wear* **1999**, *235*, 623–634.
24. Wei, J. J.; Xue, Q. J.; Wang, H. Q. *Lubr. Eng.* **1993**, *49*(12), 948–953.

25. Bowden, F. P.; Tabor, D. *The Friction and Lubrication of Solids*; Clarendon Press: Oxford, Part I, 1950; Part II, 1964.
26. Morris, S.; Wood, R. J. K.; Harvey, T. J.; Powrie, H. E. G. *Wear* **2003**, *255*, 430–443.
27. Wood, R. J. K. Challenges of living with erosion-corrosion, Advanced materials for fluid machinery, IMechE Conference Transactions; PEP London: UK **2004**; pp 113–132.
28. Wood, R. J. K. *La Houille Blanche* **1992**, 7–8, 605–610.
29. Neville, A.; Hu, X. *Wear* **2001**, *251*, 1284–1294.
30. Wang, H. W.; Stack, M. M. *J. Mater. Sci.* **2000**, *35*, 5263–5273.
31. Wood, R. J. K.; Hutton, S. P. *Wear* **1990**, *140*, 387–394.
32. Stemp, M.; Mischler, S.; Landolt, D. *Wear* **2003**, *255*(1–6), 466–475.
33. Mischler, S.; Spiegel, A.; Landolt, D. *Wear* **1999**, *225–229*, 1078–1087.
34. Staehle, R. W. *Corros. Sci.* **2007**, *49*, 7–19.
35. Jemmely, P.; Mischler, S.; Landolt, D. *Wear* **2000**, *237*, 63–76.
36. Roberge, P. *Erosion-Corrosion, Corrosion Testing Made Easy Series*; NACE International: Houston, USA, 2004.
37. Efird, D.; Wright, E. J.; Boros, J. A.; Hailey, T. G. *Corrosion* **1993**, *49*, 992–1003.
38. Kennelly, K. J.; Hausler, R. H.; Silverman, D. C. NACE Corrosion/1990; 1990.
39. Nesic, S.; Nyborg, R.; Stangeland, A. *Corrosion* **2003**, *59*, 443–456.
40. Sedamed, G. H.; Abdo, M. S. E.; Amder, M.; El-Latif, G. A. *Int. Comm. Heat mass Transfer* **1998**, *25*(3), 443–451.
41. Wood, R. J. K.; Hutton, S. P.; Schiffrin, D. J. *Corros. Sci. J.* **1990**, *30*(12), 1177–1201.
42. Schmitt, G.; Bosch, C.; Plagemann, P.; Moeller, K. Local wall shear stress gradients in the slug flow regime – Effects of hydrocarbon and corrosion inhibitor, NACE CORROSION/2002; 2002, Paper No. 02244.
43. Schmitt, G.; Gudde, T. Local mass transport coefficients and local wall shear stresses at flow disturbances, NACE CORROSION/1995, 1995, Paper No. 102.
44. Schmitt, G.; Gudde, T.; Strobel-Effertz, E. Fracture mechanical properties of CO_2 corrosion product scales and their relation to localised corrosion, NACE CORROSION/1996, 1996, Paper No. 9.
45. Wharton, J. A.; Wood, R. J. K. *Wear* **2004**, *256*, 525–536.
46. Donohue, G. L.; Tiedermann, W. G.; Reischman, M. M. *J. Fluid Mech.* **1972**, *56*, 559–575.
47. Kline, S. J.; Reynolds, W. C.; Schraub, F. A.; Rundstrandler, P. *J. Fluid Mech.* **1967**, *30*, 741–773.
48. Kim, H. T.; Kline, S. J.; Reynolds, W. C. *J. Fluid Mech.* **1971**, *50*, 133–160.
49. Blackwelder, R. F. Some ideas on the control of near-wall eddies, AIAA; 1989, Paper 89–1009.
50. Faddick, R. R. Wear in pipes, Short course on slurry pipelining technology, Camborne School of Mines, 1982, pp 1–15.
51. Haugen, K.; Kvernvold, O.; Ronald, A.; Sandberg, R. *Wear* **1995**, *186–187*, 179–188.
52. Turchaninov, S. P. *The Life of Hydrotransport Pipelines*; Nedra Press: Moscow, Russia, 1973.
53. Shimoda, K.; Yukawa, T. Erosion of pipe bend in pneumatic conveyer, Proceedings of the 6th International Conference On Erosion by Liquid and Solid Impact, University of Cambridge, 1983, Paper 59.
54. Harvey, T. J.; Wharton, J. A.; Wood, R. J. K. *Tribology – Mater. Surf. Interf.* **2007**, *1*(1), 33–47.
55. Hashish, M. An improved model of erosion by solid particles, Proceedings of the 7th International Conference on erosion by liquid and solid impact, 1998, Paper 66, published by Cavendish Laboratory.
56. Finnie, I. *Wear* **1972**, *19*, 81–90.
57. Gane, N.; Murray, M. S. The transition from ploughing to cutting in erosive wear, Proceedings of the 5th International Conference on Erosion by Solid and Liquid Impact (Cambridge, UK, 1979), paper 40, pp 1–8.
58. Keating, A.; Nesic, S. *Corrosion* **2001**, *57*(7), 621–633.
59. Bergevin, K. Effect of slurry velocity on the mechanical and electrochemical components of erosion-corrosion in vertical pipes, Master's thesis, University of Saskatchewan, 1984.
60. Bitter, J. G. A. *Wear* **1963**, *6*, 5–21.
61. Neilson, J. H.; Gilchrist, A. *Wear* **1968**, *11*, 123–143.
62. Forder, A.; Thew, M. T.; Harrison, D. *Wear* **1998**, *216*, 184–193.
63. Wood, R. J. K.; Jones, T. F.; Miles, N. J.; Ganeshalingam, J. *Wear* **2001**, *250*(1–12), 771–779.
64. Wood, R. J. K.; Jones, T. F.; Ganeshalingam, J.; Wang, M. Erosion modelling of swirling and non-swirling slurries in pipes, Hydrotransport 15, Banff, Canada, BHR Group, 3–5 June 2002, 497.
65. Wood, R. J. K.; Jones, T. F.; Ganeshalingam, J. Erosion in swirl inducing pipes, ASME Fluids Engineering Division Summer Meeting, Montreal, Canada, July 2002, paper FEDSM2002–31287, ASME International.
66. Wood, R. J. K.; Jones, T. F.; Ganeshalingam, J.; Miles, N. J. *Wear* **2004**, *256*(9–10), 937–947.
67. Wood, R. J. K.; Jones, T. F. *Wear* **2003**, *255*, 206–218.
68. Lui, J.; Lin, Y.; Yong, X.; Li, X. *Corrosion* **2005**, *61*(11), 1061–1069.
69. Chang, J. T.; Yeh, C. H.; He, J. L.; Chen, K. C. *Wear* **2003**, *255*, 162–169, Part 1 AUG-SEP 2003.
70. Chernega, S. M.; Loskutova, T. V.; Yantsevych, K. V. *Metallofiz. Noveish. Tekhnol* **2003**, *25*, 519–532.
71. Neale, M.; Gee, M. G. *Guide to Wear Problems and Testing for Industry, Tribology in Practice Series*; PEP Ltd.: London, 2000.
72. Beck, T. R. Electrochemical Models for SCC of Titanium. In *Theory of Stress Corrosion Cracking in Alloys*; Scully, J. C., Ed.; NATO: Brussels, 1971; pp 64–85.
73. Cho, E.; Kwon, H. *Corros. Sci. Technol.* **2002**, *31*, 448–453.
74. Shibata, T.; Staehle, R. W. Application oh high speed elongation technique to stress corrosion cracking of Fe-Cr-Ni Alloys Proceedings of the Fifth International Congress on Metallic Corrosion 1975; pp 487–492.
75. Hoar, T. P.; Scully, J. C. *J. Electrochem. Soc.* **1964**, 348–352.
76. Burstein, G. T.; Sasaki, K. *J. Electrochem. Soc.* **2001**, B282–B287.
77. Ambrose, J. R.; Kruger, J. *Corrosion* **1972**, *28*, 30–35.
78. Sethuramiah, A. *Lubricated Wear: Science and Technology, Tribology Series 42*; Elsevier: Amsterdam, 2003.
79. Silverman, D. C. *Corrosion* **2004**, *60*(11), 1003–1023.
80. Douglas, J. F.; Gasiorek, J. M.; Swaffield, J. A.; Jack, L. B. *Fluid Mechanics*, 5th ed.; Pearson: Harlow, UK, 2005, ISBN 0–13–129293–5.
81. Cantwell, B. J. *Annu. Rev. Fluid Mech.* **1981**, *13*, 457–515.
82. Robinson, S. K. *Annu. Rev. Fluid Motion* **1991**, *23*, 601–639.
83. Panton, R. L. *Incompressible Flows*; Wiley/Interscience: New York, 1984, pp 717.

84. Lee, C. K.; Shih, H. C. *J. Mater. Sci.* **2000**, *35*(9), 2361–2369.
85. Thakare, M. R.; Wharton, J. A.; Wood, R. J. K.; Menger, C. *Tribol. Int.* **2008**, *41*(7), 629–639.
86. Mischler, S.; Spiegel, A.; Landolt, D. *Wear* **1999**, *229*, 1078–1087, Part 2.
87. Ponthiaux, P.; Wenger, F.; Drees, D.; Celis, J. P. *Wear* **2004**, *256*, 459–468.
88. Garcia, I.; Drees, D.; Celis, J. P. *Wear* **2001**, *249*, 452–460.
89. Goldberg, J. R.; Gilbert, J. L. *J. Biomed. Mater. Res.* **1997**, *37*(3), 421–431.
90. Mischler, S.; Debaud, S.; Landolt, D. *J. Electrochem. Soc.* **1998**, *145*(3), 750–758.
91. Jiang, J.; Stack, M. M.; Neville, A. *Tribol. Int.* **2002**, *35*(10), 669–679.
92. Jiaren Jiang; Stack, M. M. *Wear* **2006**, *261*(9), 954–965.
93. ASTM Standard G 119–93, Standard Guide for Determining Synergism between Wear and Corrosion; 1993.
94. Zeisel, H.; Durst, F. Computations of erosion-corrosion processes in separated two-phase flows, *NACE Corrosion,* NACE, 1990, Paper No. 29.
95. Vignal, V.; Mary, N.; Ponthiaux, P.; Wenger, F. *Wear* **2006**, *261*, 947–953.
96. Zhou, S.; Stack, M. M.; Newman, R. C. *Corrosion* **1996**, *52*, 934–946.
97. Matsumura, M. *Corros. Rev.* **1994**, *12*, 321–340.
98. Neville, A.; Hodgkiess, T.; Xu, H. *Wear* **1999**, *235*, 523–534.
99. Malka, R.; Nešić, S.; Gulino, D. A. *Wear* **2007**, *262*(7–8), 791–799.
100. Bello, J. O.; Wood, R. J. K.; Wharton, J. A. *Wear* **2007**, *263*, 149–159.
101. Rabinowicz, E.; Dunn, L. A.; Russell, P. G. *Wear* **1961**, *4*, 345–355.
102. Batchelor, A. W.; Stachowiak, G. W. *Wear* **1998**, *123*, 281–291.
103. Ferng, Y. M.; Ma, Y. P.; Ma, K. T.; Chung, N. M. *Corrosion* **1999**, *55*(4), 332–342.
104. Ferng, Y. M.; Ma, Y. P.; Chung, N. M. *Corrosion* **2000**, *56*(2), 116–126.
105. Bozzini, B.; Ricotti, M. E.; Boniardi, M.; Mele, C. *Wear* **2003**, *255*, 237–245.
106. Lu, B. T.; Luo, J. L.; Ma, H. Y. *J. Electrochem. Soc.* **2007**, *154*(3), C159–C168.
107. Neville, A.; Reyes, M.; Xu, H. *Tribol. Int.* **2002**, *35*(10), 643–650.
108. Wood, R. J. K. *Wear* **2006**, *261*(9), 1012–1023.
109. Nesic, S.; Bienkowski, J.; Bremhorst, K.; Yang, K. S. *Corrosion* **2000**, *56*(10), 1005–1014.
110. Guo, H. X.; Lu, B. T.; Luo, J. L. *Electrochim. ACTA* **2005**, *51*(2), 315–323.
111. Tsai, W. T.; Huang, T. M. *Thin Solid Films* **2000**, *379*(1–2), 107–113.
112. Tan, A. H.; Cheng, Y. C. *Wear* **2007**, *262*(5–6), 693–698.
113. Ziomek-Moroz, M.; Miller, A.; Hawk, J.; Cadien, K.; Li, D. Y. *Wear* **2003**, *255*, 869–874, Part 2.
114. Preston, F. *J. Soc. Glass Technol.* **1927**, *9*, 214–256.
115. Akonko, S. B.; Li, D. Y.; Ziomek-Moroz, M.; Hawk, J.; Miller, A.; Cadien, K. *Wear* **2005**, *259*, 1299–1307, Part 2.
116. Wood, R. J. K.; Wharton, J. A.; Speyer, A. J.; Tan, K. S. *Tribol. Int.* **2002**, *35*(10), 631–641.
117. Sasaki, K.; Burstein, G. T. *Phil. Mag. Lett.* **2000**, *80*(7), 489–493.
118. Oltra, R.; Chapey, B.; Renuad, L. *Wear* **1995**, *186–187*, 533–541.
119. Lambrechts, P.; Goovaerts, K.; Bharadwaj, D.; De Munck, J.; Bergmans, L.; Peumans, M.; Van Meerbeek, B. *Wear* **2006**, *261*(9), 980–986.
120. Nagarajan, V. S.; Jahanmir, S.; Thompson, V. P. *Dent. Mater.* **2004**, *20*(1), 63–71.
121. Nagarajan, V. S.; Hockey, B. J.; Jahanmir, S.; Thompson, V. P. *J. Mater. Sci.* **2000**, *35*(2), 487–496.
122. Lewis, R.; Dwyer-Joyce, R. S. *P I MECH. ENG. J.-J. Eng* **2005**, *219*(J1), 1–18.
123. Vieira, A. C.; Ribeiro, A. R.; Rocha, L. A.; Celis, J. P. *Wear* **2006**, *261*(9), 994–1001.
124. Berradja, A.; Bratu, F.; Benea, L.; Willems, G.; Celis, J.-P. *Wear* **2006**, *261*(9), 987–993.
125. Yan, Y.; Neville, A.; Dowson, D.; *J. Phys. D: Appl. Phys.* **2006**, *39*(15), 3200–3205.
126. Sinnett-Jones, P. E.; Wharton, J. A.; Wood, R. J. K. *Wear* **2005**, *259*(2), 898–909.
127. Sun, D. Wharton, J. A.; Wood, R. J. K.; Ma, L.; Raithforth, W. M. *Tribol. Int.* **2009**, *42*(1), 99–110.

Further Reading

Flow-induced corrosion: Fundamental studies and industrial experience, NACE, symposium proceedings, Corrosion 1990, Houston, USA.
Engineering Tribology, John Williams; Cambridge University Press: Cambridge UK.
Ludema, K. C. *Friction, Wear, Lubrication: a Textbook in Tribology*; CRC Press: Boca Raton, FL, 1996; p 45.
Hutchings, I. M. Tribology: Friction and Wear of Engineering Materials; Edward Arnold: UK, 1992; co-published by CRC Press, Boca Raton, USA.
Engineering Tribology, Stachowiak, G.; Batchelor, A. W.; Elsevier, 2005.

2.17 Corrosion in Natural Waters

S. B. Lyon
Corrosion and Protection Centre, School of Materials, University of Manchester, Oxford Road, Manchester M13 9PL, UK

This article is a revision of the Third Edition article 2.3 by C.W. Drane and R.J. Oliphant, volume 1, pp 2:43–2:59,

Glossary

Calcite Calcium carbonate.

Gypsum Calcium sulfate.

Permanent hardness Water hardness that cannot be removed by boiling; for example, the precipitation of calcium sulfate.

ppm Strictly parts per million by mass; equivalent to 1 mg of solute per kg of solvent. In dilute solutions and for practical usage, it is equivalent to 1 mg of solute per litre (dm^3) of solvent.

Temporary hardness Water hardness that can be removed by boiling causing the loss of dissolved CO_2; for example, the precipitation of calcium carbonate.

Water hardness A measure of the tendency of water to precipitate a solid scale; generally related to the concentration of dissolved cations (in particular calcium and magnesium) with respect to dissolved anions (in particular bicarbonate and sulfate).

Abbreviations

HVAC Heating, ventilation and air conditioning
LSI Langelier saturation index
PSI Puckorius scaling index
RSI Ryznar stability index
TDS Total dissolved solids

Symbols

K_{sp} Solubility product
ΔG Free energy change
[*x*] The concentration of species '*x*' is conventionally denoted by placing it in square brackets

2.17.1 Introduction

Metals immersed or partly immersed in water tend to corrode because of their thermodynamic instability. Natural waters may contain some or all impurities such as dissolved solids and gases, colloidal or suspended matter, and biological materials (microbes, algae). All these may affect the corrosive properties of the water in relation to the metals with which it is in contact. The effect may be either of stimulation or of suppression of either the cathodic or the anodic reaction. Thus, some metals may form a natural protective film in a specific water composition and the corrosiveness of the water to these metals depends on whether or not the dissolved species that it contains assist in the maintenance of a self-healing film (i.e., whether they are inhibiting or not).

The metals most commonly used for water systems are iron and steel. These metals often have some sort of applied protective coating; galvanized steel, for example, relies on a thin layer of zinc, which is anodic to the steel except at high temperatures. Many systems, however, contain a wide variety of other materials and the effect of various water constituents on these must be considered. Although there is increasing use of polymeric materials in water systems, metallic components are still very common and include copper, brasses, bronzes, lead, aluminum, stainless steel, and solder.

The passage of natural water through a pipe may modify the composition of the water and hence its corrosive properties. For example, dissolved oxygen or carbon dioxide may be removed partially or wholly by reaction with metals or organisms. Dissolution of a metal into water may, on the other hand, make it more corrosive. An example of this is the attack of some waters on copper and the subsequent increased pitting corrosion of less noble metals such as iron, galvanized steel, and aluminum. This enhanced pitting is caused by the deposition of minute quantities of copper on the less noble metal, thereby setting up numerous local bimetallic corrosion cells.

Failure of the metal can be the most important effect of corrosive water, but other effects may arise from low concentrations of metallic ions produced by corrosion. Water passed through a lead pipe may contain a toxic concentration of that metal; with copper there is a greater tolerance from the toxicity point of view but staining of fabrics and sanitary fittings may be objectionable. With iron, similarly, discoloration of the water may be unpleasant and may cause damage to materials being processed. The effects of dissolved zinc in the environment have also recently become a concern.

The concentrations of various substances in drinking water or in industrial process waters in the dissolved, colloidal, or suspended form are relatively low but vary considerably. For example, a hardness of 300–400 ppm (as $CaCO_3$) is sometimes tolerated in potable (drinking) public water supplies, whereas dissolved iron to the extent of 1 mg dm^{-3} would be unacceptable. On the other hand, saline waters, including seawater, estuarine water, and brines from boreholes can contain very large quantities of dissolved species; however, these last mentioned types of water are not considered in this section.

2.17.2 Constituents or Impurities in Water

The important constituents in natural water can be classified as follows:

1. dissolved gases (e.g., oxygen, nitrogen, carbon dioxide, ammonia, sulfurous gases);
2. mineral constituents, including hardness salts (e.g., calcium, magnesium, carbonate/bicarbonate, etc), sodium salts (e.g., chloride, sulfate, nitrate, etc.), other salts, and dissolved silica;
3. organic matter, including that of both animal and vegetable origin, oil, trade waste (including agricultural) constituents and synthetic detergents;
4. microbiological forms, including various types of algae and slime-forming bacteria.

In water analysis, determinations (except occasionally for dissolved gases) are made on the basis of mass

Table 1 Typical water analyses (values in mg dm^{-3})

	Very soft water	*Moderately soft water*	*Slightly hard water*	*Moderately hard water*	*Hard borehole water*	*Very hard borehole water*
pH	6.3	6.8	7.4	7.5	7.1	7.1
Alkalinity due to	2	38	90	180	250	470
Calcium	10	53	120	230	340	559
Total hardness	5	36	85	210	298	451
Calcium hardness	6	20	39	50	17	463
Sulfate	5	11	24	21	4	149
Chloride	Trace	0.3	3	4	7	6
Silica	33	88	185	332	400	1670
Total dissolved solids						

Source: Drane, C. W.; Oliphant, R. J. *Natural Waters*. In Corrosion, 3rd ed.; Shreir, L. L., Jarman, R. A., Burstein, G. T., Eds., 1994; chapter 2.3.

of solute to volume of solvent (mg dm^{-3}). More strictly, parts per million (ppm) refers to mass of solute to mass of solvent but in dilute solutions, and to all practical purposes, the two units are interchangeable. In treated water for high-pressure boilers, impurities are in much lower concentrations and are measured in parts per billion (ppb, equivalent to 1 μg dm^{-3}).

Water analysis for drinking-water supplies is concerned mainly with pollution and bacteriological tests. For industrial supplies, a mineral analysis is of more interest. Table 1 includes a typical selection and gives some indication of the wide range of impurities that can be found.

2.17.2.1 Dissolved Gases

Of the dissolved gases occurring in water, oxygen occupies a special position as it is the main cathodic reactant and can control the corrosion reaction rate if mass transport limits the arrival of oxygen to the cathode. Carbon dioxide is scarcely less important as it dissolves in water to giving rise to equilibria with carbonic acid, bicarbonate and carbonate in solution, which are important in the control of solution pH and the solubility of calcium carbonate and, hence, the calcium hardness of the water. Therefore, the quantity of dissolved carbon dioxide in solution controls the solubility of the carbonate rocks that frequently comprise the geology close to boreholes.

Nitrogen is also generally present with oxygen although, as it is significantly less soluble, the mass ratio of dissolved nitrogen to oxygen in water is not the same as in air (it is about 4:3). It has little importance in connection with corrosion, but can be a nuisance if changes in physical conditions (e.g., reduction in local pressure) bring about its release from solution.

Other gases which are occasionally present in natural water usually arise from pollution. Ammonia, which in various forms may be present in waste waters, attacks copper and copper alloys; its presence in estuarine waters is one of the causes of corrosion of copper-based steam condenser tubes. Hydrogen sulfide may also be the result of pollution but often bacterial action (e.g., by sulfate reducing bacteria) is the main contributor. Both gases initiate and/or accelerate the corrosion of most metals.

2.17.2.2 Oxygen

Although other oxidants can occur in natural waters (in particular sulfur species) dissolved oxygen is probably the most significant constituent affecting corrosion; its importance lies in the fact that it is the most important cathodic reactant in neutral solutions. In surface waters, the oxygen concentration is in equilibrium with the atmosphere and, hence, approximates to saturation. Underground waters are more variable in oxygen content and some waters may be effectively oxygen-free. Contact with air, however, usually gives rise to an oxygen concentration similar to the figures in Table 2. The solubility is slightly less in the presence of solutes, but this effect is not very significant in natural waters containing less than 1000 ppm dissolved solids.

For many applications, notably feed-water for boilers, removal of oxygen is essential and is accomplished by physical or chemical de-aeration. For many other purposes (e.g., once-through systems, heating, cooling and ventilation, i.e., HVAC systems), however, full de-aeration is not applicable, as the water used is in continuous contact with air, from which it would rapidly take up more oxygen. In these circumstances, chemical treatment, with corrosion inhibitors for example, is necessary to restrict corrosion.

A major difficulty in chemical water treatment is that the oxygen content of the water may not be the same at all points. For example, in a thin layer of water,

between a deposit, crevice or other surface occlusion and the metal on which it is present, the oxygen can be depleted; the difference in oxygen content between the body of water and the stagnant water will then set up a corrosion current which is difficult to suppress.[1,2] Rather, similar conditions occur at the water line of a vessel containing water with air above it. Even if the water is conditioned to inhibit corrosion under submerged conditions, it is difficult to extend the protection to above the water line. Fluctuation in water level extends the area of localized attack.

2.17.2.3 Carbon Dioxide

The effect of dissolved carbon dioxide on corrosion is closely linked with the bicarbonate content of the water. In many waters divalent cations (such as calcium and magnesium) are present at significant concentration and, depending on the water chemistry and temperature may precipitate a scale onto surfaces (i.e., hard waters). The stability, or otherwise, of calcium and magnesium in solution depends on a number of factors, the most important of which are pH, dissolved total carbonate concentration (and its speciation as dissolved CO_2, carbonic acid and bicarbonate) and the concentration of Ca^{2+} and Mg^{2+}. Any external phenomenon that reduces the solubility of the divalent cations will result in their precipitation, typically as a scale containing either calcium carbonate or magnesium hydroxide (or both).

Calcium bicarbonate requires excess carbon dioxide in solution to stabilize it; the necessary concentration depends on the other constituents of the water and the temperature. Carbon dioxide dissolves in water to give carbonic acid in the following equilibria, with the equilibrium constants (K_x) indicated[3]:

$$H_2CO_3 \leftrightarrows H^+ + HCO_3^- \quad K_1 = 2.5 \times 10^{-4} \quad [1]$$
$$HCO_3^- \leftrightarrows H^+ + CO_3^{2-} \quad K_2 = 5.6 \times 10^{-11} \quad [2]$$

However, although carbonic acid is a stronger acid than acetic acid, its concentration in solution in equilibrium with atmospheric carbon dioxide (≈385 ppm) is relatively low and, therefore, its effect is considerably less. In the absence of dissolved carbonate species from other sources, the solution pH is therefore determined by the partial pressure of CO_2:

$$CO_2 + H_2O \leftrightarrows H^+ + HCO_3^- \quad K_a = 4.30 \times 10^{-7} \quad [3]$$

The speciation among carbonic acid, bicarbonate ion and carbonate ion depends on the above equilibria and, hence, the pH. **Figure 1** shows calculated mole fractions of carbonic acid, bicarbonate ion, and carbonate ion as a function of pH for a fixed total carbonate concentration.

2.17.2.4 Water Hardness (Carbonate Species)

The usual classification of water by hardness (**Table 3**) is as follows.

The hardness figures for natural water are very varied but most natural drinking water supplies fall into two well-defined groups:

1. Upland waters of low hardness commonly retained in storage reservoirs. In the United Kingdom, these are supplied to most localities in Scotland, Wales, and the North of England.

Table 2 Solubility of oxygen in distilled water

Temperature (°C)	*Oxygen content of saturated, distilled water*	
	mg dm^{-3}	*ml dm^{-3}*
0	14.16	9.90
5	12.37	8.65
10	10.92	7.64
15	9.76	6.83
20	8.84	6.18
25	8.11	5.67

Source: Sherwood, J. E.; Stagnitti, F.; Kokkinn, M. J.; Williams, W. D. *Int. J. Salt Lake Res.* **1992**, *1*, 1–6. Groisman, A. Sh.; Khomutov, N. E. *Russ. Chem. Rev.* **1990**, *59*, 1217–1250.

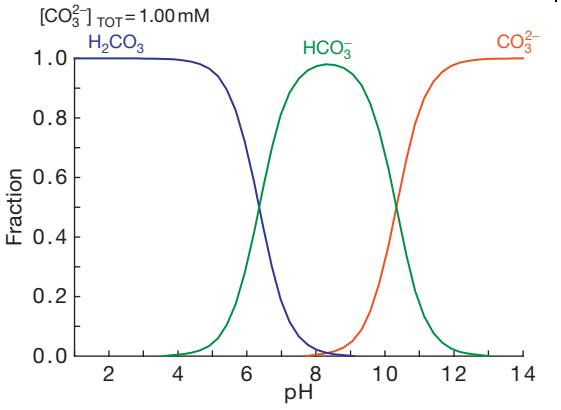


Figure 1 Calculated mole fractions of carbonic acid, bicarbonate ion, and carbonate ion as a function of pH at a total carbonate species concentration of 1 mM. Calculated with Puigdomenach, I.; Hydrochemical Equilibrium Constant Database (HYDRA/MEDUSA); Royal Institute of Technology: Stockholm. Available at www.kemi.kth.se/medusa.

2. Underground waters of high hardness, usually derived from aquifers in chalk, sandstone or limestone strata. In the United Kingdom, these are found mainly in the East and South of England.

A few sources, most of which are derived from rivers and vary according to season, are intermediate in composition between these.

The corrosive properties of natural waters are governed by many factors and cannot be related to hardness alone, but the following trends are apparent:

1. Soft upland waters are commonly corrosive to most metals, their behavior depending to some extent on pH. They are inevitably unsaturated with respect to calcium carbonate and it is not usually practicable to modify the carbonate equilibrium to make them nonaggressive.
2. Very hard waters are usually not very aggressive provided that they are supersaturated in calcium carbonate. Underground waters with a low pH value caused by excess carbon dioxide content are, however, aggressive unless corrective treatment is applied.
3. Waters of intermediate hardness frequently contain significant amounts of other constituents and there is often a tendency for the scale formed to be loosely attached to the metal surface, permitting corrosion to occur irregularly underneath. In most waters the bicarbonate content is less than the hardness, but a few natural waters are known where the reverse is the case.

Table 3 Range of calcium hardness for various natural waters

Calcium concentration	*Hardness classification*
<50 ppm $CaCO_3$	Soft
50–100 ppm $CaCO_3$	Moderately soft
100–150 ppm $CaCO_3$	Slightly hard
150–250 ppm $CaCO_3$	Moderately hard
250–350 ppm $CaCO_3$	Hard
>350 ppm $CaCO_3$	Very hard

Thresh, J. C.; Beale, J. F.; Suckling, E. V. *The Examination of Waters and Water Supplies*, 7th ed.; Taylor, E. W.; Churchill Livingston: London, 1958.

2.17.2.5 Other Inorganic Species

The principal ions found in natural water supplies include calcium, magnesium, sodium, bicarbonate, sulfate, chloride, and nitrate; also a few parts per million of iron or manganese may sometimes be present. From the corrosion point of view the small quantities of other anions present, for example, nitrite, phosphate, iodide, bromide and fluoride, have little significance. Although larger concentrations of some of these ions, notably nitrite and phosphate, may act as corrosion inhibitors, the small quantities present in natural waters will have little effect. Some of the minor constituents have other beneficial or harmful effects, for example, there is an optimum concentration of fluoride for control of dental caries and very low iodide or high nitrate/nitrite concentrations are objectionable on medical grounds (**Table 4**).

Dissolved ionic species will increase the ionic conductivity of water and, hence, tend to increase metallic corrosion by reducing the ohmic voltage loss between local anodes and cathodes. At low

Table 4 Limits for inorganic species in UK drinking water

Measured parameter	*Maximum permitted value*	*Units of measurement*	*Point of monitoring*
Conductivity	2500	$\mu S\,cm^{-1}$ at 20 °C	Supply point
pH	6.5–10		Customers' taps
Chloride	250	mg (Cl) dm^{-3}	Supply point
Sulfate	200	mg (SO_4) dm^{-3}	Supply point
Sodium	200	mg (Na) dm^{-3}	Customers' taps
Fluoride	1.5	mg (F) dm^{-3}	Customers' taps
Ammonium	0.5	mg (NH_4) dm^{-3}	Customers' taps
Nitrate	50	mg (NO_3) dm^{-3}	Customers' taps
Nitrite	0.5	mg (NO_2) dm^{-3}	Customers' taps
Copper	2	mg (Cu) dm^{-3}	Customers' taps
Iron	200	μg (Fe) dm^{-3}	Customers' taps
Manganese	50	μg (Mn) dm^{-3}	Customers' taps
Aluminum	200	μg (Al) dm^{-3}	Customers' taps
Lead (now)	25	μg (Pb) dm^{-3}	Customers' taps
Lead (after 2013)	10	μg (Pb) dm^{-3}	Customers' taps

Source: The Water Supply (Water Quality) Regulations 2000 (and as subsequently amended), UK Government, HMSO.

concentration the effect on ionic conductivity dominates, although chloride will always tend to promote pitting corrosion on passive materials (e.g., stainless steel) that are protected by a thin protective scale. In the latter case, the ratio of chloride to hydroxide (or bicarbonate) can be important.[4] Sulfate in general appears to behave very similarly; Hatch and Rice have shown that small concentrations in distilled water increase corrosion more than similar concentrations of chloride.[5] In practice, high-sulfate waters may attack concrete and the performance of some inhibitors appears to be adversely affected by the presence of sulfate. Sulfates have also a special role in bacterial corrosion under anaerobic conditions.

Regarding iron- and manganese-bearing waters, the small amount of deposit formed is not likely to have much effect on corrosion. However, most iron-bearing waters contain substantial amounts of carbon dioxide which may be troublesome. Manganese-bearing waters may be of a similar type but they sometimes contain complex organic compounds of manganese for which special treatment may be needed. Substantial iron and manganese content in water will tend to promote pitting corrosion of materials such as aluminum, stainless steel and copper. Another mineral constituent of water is silica, present both as a colloidal suspension and as a dissolved species.

2.17.2.6 Organic Species

The types of organic matter in water supplies are very diverse; they may be present in suspension, or in colloidal or true solution and may include microorganisms. The materials are largely derived from decaying vegetable matter but there are many other possible sources including run-off from fields and domestic and industrial wastes. Some of the worst corrosive effects ascribed to natural soft waters are attributed to a rather wide group of organic acids abstracted from peat and mosses, sometimes called humic acids.[6] Such waters have low pH values and are often discolored. They affect ferrous metals appreciably and also attack lead and copper. Corrosion control, either of steel or of copper, is rarely achieved solely by pH correction of such waters.

2.17.3 Scaling in Natural Waters

2.17.3.1 Controlling Factors in Scale Formation

The deposition of a scale from solution involves the nucleation, crystallization and growth of a new solid phase. Nucleation of deposits may take place either in solution (homogeneous nucleation) or on a suitable substrate where the *local* (i.e., adjacent to the substrate) chemistry may be particularly conducive to scaling (heterogeneous nucleation). Nucleation is a thermodynamically driven process and the elementary theory predicts that nucleation requires a driving force to overcome the new phase boundary (i.e., surface) created. Thus, the formation of a new phase in a solution *reduces* the overall free energy of the system by precipitating out a supersaturated species; however, it *increases* the local free energy by formation of a new surface. Both of these processes can be modeled (for homogeneous nucleation in solution) as follows[7]:

$$\Delta G_{\text{tot}} = 4/3\pi r^3 \Delta G_{\text{volume}} + 4\pi r^2 \Delta G_{\text{surface}} \qquad [4]$$

where ΔG_{tot} is the overall Gibbs free energy change during nucleation; ΔG_{volume} is the free energy change per unit volume of new species formed (which will be negative and favorable); $\Delta G_{\text{surface}}$ is the free energy per unit surface area of the nucleus (which will be positive and unfavorable). The above equation is easily altered for heterogeneous nucleation (i.e., on a surface). Nucleation theory predicts that nuclei do not grow below a specific, critical dimension, which can be seen clearly in **Figure 2**.

The above thermodynamic treatment can say nothing about the kinetics of the nucleation process except that at higher supersaturation, the driving force for nucleation is greater. Once nuclei have formed, their subsequent growth is determined by a number of interacting processes, particularly the diffusion of species in solution.

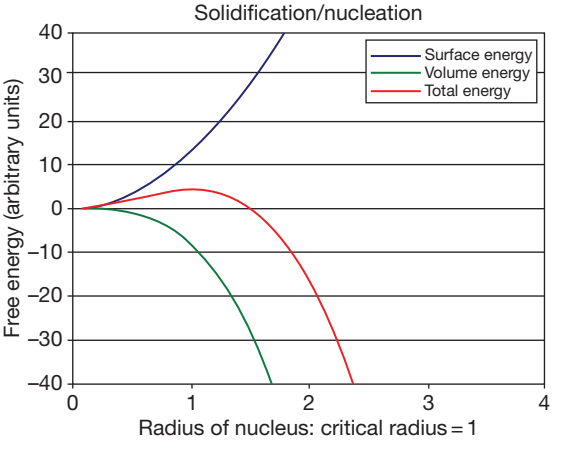


Figure 2 Variation of surface and volume free energy with nucleus size showing that a critical size is required before stable nucleation is possible.

The solubility of most scale-forming minerals decreases as the temperature increases. Therefore, such solutions will precipitate a scale directly on the heat-transfer surface because of their lower equilibrium solubility at these locations (carbonate scales that form from hard water in domestic kettles being the obvious example). Other soluble species may be scale-forming as a result of changes in pH, pressure, flow, etc. Generally, the crystallization of a scale-forming material from solution is influenced by several factors including the following[8]:

- Local and general supersaturation of the soluble species in solution.
- Solution temperature and particularly temperature differentials especially at surfaces where heat transfer is occurring.
- Solution pH and particularly local changes in pH at cathodic surfaces.
- Pressure, especially where the speciation is controlled by a soluble gas such as CO_2.
- Flow rate, as this influences local solution temperatures and concentrations, as well as transport of reacting species to the scaling surface.
- Scale nucleation and particularly scale growth are time-dependent processes.
- Surface condition: in particular particles of dirt and rough surfaces will encourage scale nucleation and adherence.

2.17.3.2 Calcium Carbonate Scales

The mathematical relationship between carbon dioxide, calcium bicarbonate and calcium carbonate was historically studied by several workers, including Langelier.[9,10] The simpler form of his original equation is (at constant temperature)

$$pH_s = pCa + pAlk + (pK_2 - pK_s) \qquad [5]$$

where pH_s is the pH at which calcium carbonate precipitates, pCa is the negative logarithm of the calcium concentration expressed as parts per million of $CaCO_3$, pAlk is the negative logarithm of the alkalinity to methyl orange expressed in parts per million of equivalent $CaCO_3$, pK_2 is the ionization constant of HCO_3^-, and pK_s is the solubility product of $CaCO_3$.

This formula does not apply to pH values over 9. Also, as the term $(pK_2 - pK_s)$ is a function of temperature and ionic strength (dissolved solids), high salinities will affect its accuracy. However, for the analysis of water at constant temperature much useful information can be obtained from the basic thermodynamic treatment of Langelier. In order to determine whether water is supersaturated with respect to calcium ions, the Langelier saturation index of water (LSI) is defined as

$$LSI = pH - pH_s \qquad [6]$$

where pH is the actual pH of the water and pH_s is the pH of saturation of calcium carbonate. If the LSI is positive the water will be supersaturated with calcium carbonate, which will, therefore, tend to precipitate as a scale. Conversely if the LSI is negative the water tends to dissolve calcium carbonate and is thus nonscaling. Nonscaling waters are expected to be corrosive while waters that scale are expected to be relatively less corrosive.

The LSI is not wholly satisfactory for practical prediction of the scaling and corrosion tendency. This is because it is possible for waters of varying hardness (and hence corrosivity) to have the same LSI value. The Ryznar stability index (RSI) is a modification of the LSI that overcomes this problem[11]:

$$RSI = 2pH_s - pH \qquad [7]$$

In this scheme, an index between 6 and 7 is effectively balanced with little tendency to scale. If $RSI < 6$, scaling is likely (and corrosion less likely) with the scaling tendency increasing as the index decreases. For $RSI > 7$ calcium carbonate formation probably does not lead to a protective corrosion inhibitor film and for $RSI > 8$ mild steel corrosion is likely to become an increasing problem.

Neither the LSI nor the RSI can account for the buffering capacity of the solution. Thus, water that is high in calcium ions, but low in alkalinity and buffering capacity can have a high saturation value for calcium carbonate. Typically, the water has a large thermodynamic driving force for scale precipitation, but does not have much scaling capacity because of the low quantity of carbonate ions present. To account for these effects an alternative index was devised[12] that is tied to the calculated equilibrium pH rather than the measured value. The PSI index is calculated in a manner similar to the Ryznar stability index. Thus:

$$PSI = 2pH_{eq} - pH_s \qquad [8]$$

where pH_s is as defined above; $pH_{eq} = 1.465 \times \log_{10}[\text{Alkalinity}] + 4.54$; and $[\text{Alkalinity}] = [HCO_3^-] + 2[CO_3^{2-}] + [OH^-]$. The index is interpreted in a similar manner to the RSI.

It is important to note that the above indices assume relatively dilute solutions and unit activity coefficients for dissolved species. None of the above indices can satisfactorily predict the scaling tendency of water that has significant quantities of other ions in solution, especially if they are present at high concentration. Other indices[13] have therefore been developed for such conditions encountered, for example, in the oil and gas industry.

There are many modern computational tools now available to determine solution speciation in a wide range of conditions. For example, **Figure 3** shows the effects of varying the calcium ion concentration, at constant total carbonate concentration, on the pH of precipitation of a calcium carbonate scale. Thus, as the water hardness (as calcium ion concentration) increases from 0.5 to 4 mM at a total carbonate of 1 mM, the pH for scale precipitation changes from about 8.5 to about 7.5. The domains of pH in which carbonic acid and bicarbonate ion dominate are also clearly evident.

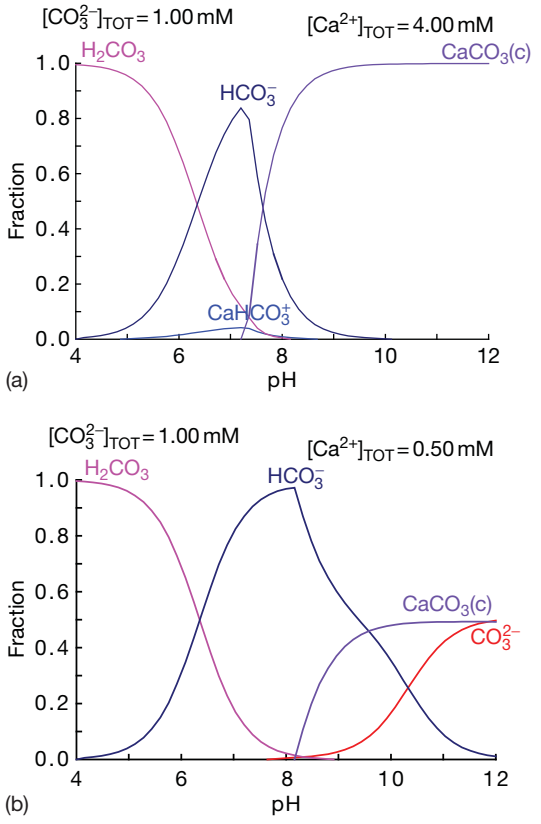


Figure 3 Mole fraction of stable species as a function of pH for (a) 4 mM and (b) 0.5 mM Ca^{2+} concentrations at a constant total carbonate concentration of 1 mM at 298 K. Calculated with Puigdomenach, I.; Hydrochemical Equilibrium Constant Database (HYDRA/MEDUSA); Royal Institute of Technology: Stockholm. Available at www.kemi.kth.se/medusa.

Depending on the circumstances, calcium carbonate scale formation may be detrimental or beneficial. The detrimental effects are summarized below:

- loss of heat transfer efficiency (calcium carbonate has a low thermal conductivity), including reduced life of heat transfer surfaces because of overheating (**Figure 4**);
- loss of pipe diameter leading to reduced flow capacity or an increased pressure requirement to achieve the same flow rate;
- blockage of valves, etc.;
- reduced efficiency of detergents and soaps for washing; and
- deposits left after evaporation of water, which spoils the appearance.

Conversely divalent metal cations can act beneficially, particularly as cathodic inhibitors, for corrosion. This is because of the alkalinity at the cathode caused by oxygen or water reduction:

$$O_2 + 2H_2O + 4e^- \leftrightarrows 4OH^- \qquad [9]$$

Thus, the local increase in hydroxide ion concentration will therefore tend to precipitate a scale, even in systems that are not intrinsically scale-forming; as is evidently shown in **Figure 3**. If the scale is sufficiently dense and nonporous, it will provide significant

Figure 4 Carbonate scale: dense generally protective scale on a steel water pipe with some corrosion evident as general thinning: Photo courtesy of the author.

corrosion protection and, hence, the calcium ions act effectively as inhibitors. The degree of protection afforded by calcium carbonate has been studied by McCauley.[14] The significance of carbon dioxide in corrosion is also discussed in some detail by Simmonds.[15]

2.17.3.3 Other Inorganic Scales

Calcium carbonate is by far the most common scaling substance in natural water systems but it is by no means the only one. Thus, all dissolved species will tend to form a scale when their solubility limit is exceeded. Also, although $CaCO_3$ is usually dominant in scales formed in potable water, waters used for other industrial purposes (e.g., heating and cooling) may form other deposits depending on the water chemistry. Indeed, many such scales comprise deposits of a range of compounds. Scales may also act to secure (adhere) other forms of fouling deposit (such as mud, silts, etc.) to a substrate.

2.17.3.3.1 Silica

Silica is present in many waters; in supplied waters it can be derived from clay and similar soil types. The solubility of silica is relatively high, at least 100 ppm below pH 7, increasing to above 300 ppm at pH 10; therefore, it rarely forms a scale on its own. However, the presence of dissolved silica can promote the precipitation of silicate species, particularly magnesium silicates.[16]

In some more extreme conditions, for example where the supply waters are relatively high in silica, the water requires treatment to prevent scale formation. For example, fouling of heat transfer surfaces and osmosis membranes in desalination plant is problematic[17] and geothermal water and power supply systems tend to readily form silica scales.[18]

2.17.3.3.2 Magnesium hydroxide and silicate

In the presence of significant dissolved magnesium ions and at high pH (typically greater than about 9–10) magnesium hydroxide may precipitate with, and participate in, overall carbonate scale formation[19]:

$$Mg^{2+} + 2OH^- \leftrightarrows Mg(OH)_2 \quad K = 7.0 \times 10^{-12} \qquad [10]$$

This is certainly an issue in the formation of scales from brines and marine sources, especially under cathodic protection (and is dealt with elsewhere). However, in the presence of significant dissolved silica, which may occur in re-circulating evaporative cooling systems for example, formation of magnesium silicate may occur either directly, or by transformation of prior magnesium hydroxide.[20] Deposition of magnesium silicate can generally be prevented by keeping the pH less than about 8 and, more generally, by ensuring that the dissolved silica levels do not exceed 200 ppm.

2.17.3.3.3 Calcium sulfate

Water containing relatively lower quantities of carbonate but a higher concentration of sulfate will tend to form a scale of calcium sulfate (gypsum: $CaSO_4{\cdot}2H_2O$; $K_{sp} = 1.95 \times 10^{-4}$).[21] Such a water composition may occur naturally but are more commonly found in cooling towers, where acidification with sulfuric acid (to prevent calcite deposition) is used as part of the water treatment system.

In order to avoid precipitation of gypsum on heat transfer surfaces in evaporative cooling systems when using acid treatment, alternative acids could be considered (although these are more corrosive and more expensive). Alternatively, proper blow-down procedures should be used, which involve removal of part of the water from the cooling circuit with replacement with fresh make-up water. Under the correct conditions, calcium sulfate scale formation should not be a problem.

2.17.3.3.4 Calcium phosphate

Phosphoric acid (or more commonly polyphosphate) is commonly added to heating and cooling waters to prevent calcium carbonate scale formation. Generally this functions not necessarily by prevention of precipitation, but by changing the nature of the precipitate such that it does not form as an adherent scale. However, scaling by calcium phosphate (which is less soluble that calcium carbonate) can occur if tricalcium phosphate ($Ca_3(PO_4)_2$) is formed as its solubility product is smaller than that of calcium carbonate.[22]

2.17.3.3.5 Iron and manganese oxides and hydroxides

Many scaling deposits may contain small quantities of dissolved iron or manganese. While the levels of these species are usually well below 200 ppb in most water supplies, occasionally higher levels may reach the customer due to corrosion and biological activity in the water supply system from iron and manganese oxidizing bacteria, giving rise to 'red' or 'black' water

quality problems.[23,24] However, in circulating heating and cooling systems, the iron plus manganese concentrations may reach several part per million; unless controlled, iron and manganese oxides are extremely likely to deposit within the system.

The other source of iron in heating and cooling systems is corrosion of the ferrous materials within the system. Indeed, a high level of dissolved iron is a good indicator of an ineffective inhibitor system. The most common scale in hot water systems is magnetite (Fe_3O_4) while haematite (Fe_2O_3) is more commonly found in cooling systems.

2.17.3.4 Fouling Deposits

The origins of fouling deposits are generally different in origin from scale deposits. The best way to differentiate them is that a scale deposit is a specific and important form of fouling deposit that can, to a large extent, be predicted from the water chemistry.[25] On the other hand, fouling is a general description of any kind of deposit that can form. It includes deposits that arise from inorganic or organic (biological) origins. Thus, in addition to scales, fouling deposits include the following:

- Sediments carried into the system with the make up water, such as silt.
- Products of reactions between the water and chemical additives.
- Insoluble corrosion products.
- Biofilms, including slimes, from bacterial, algal and other origins.

2.17.3.4.1 Corrosion

Corrosion of a metal, giving rise to an insoluble corrosion product on the metal is one of the most common forms of fouling deposit. One of the main purposes of water treatment is to limit the formation of such deposits by controlling or inhibiting the corrosion process. Corrosion (i.e., anodic dissolution) can occur by any of the normal processes and therefore requires a cathodic reactant (oxidizing agent). In open water systems the cathodic reactant is dissolved oxygen while in closed systems hydrogen generation from water reduction can occur. The amount of fouling due to corrosion product formation often increases on (hot) heat transfer surfaces in response to the general increase of corrosion rate with temperature.

2.17.3.4.2 Biofilms

Although natural waters are disinfected for potable (drinking) purposes, the amount of biocide (e.g., dissolved chlorine) present may decrease or be eliminated during service (e.g., by heating). Also, water exposed to open air will rapidly pick up all kinds of active biological material (e.g., bacteria, algae, etc.) that may have a direct or indirect effect on corrosion.[26]

The most common consequence of biological activity is the development of a biofilm on the metal substrate. Corrosion in the crevice formed effectively by the biofilm is possible generally by differential aeration or concentration cell mechanisms. Organisms that prefer anaerobic conditions (e.g., sulfate reducing bacteria) may preferentially colonize beneath bio-films, or other deposits, and generate local environments that are conducive to corrosion (e.g., containing sulfide species). Other bacterial organisms (acidophiles) may produce organic or inorganic (sulfuric) acids while yet other organism (e.g., algae) may respire consuming dissolved carbon dioxide and producing oxygen. If none of these issues is significant, the mere presence of an adherent biofilm will reduce heat transfer efficiency and, if it is removed by cleaning, may block valves, etc. Therefore, effective biocidal treatment is essential in both domestic and industrial settings.

2.17.3.5 Water Treatment

It is beyond the scope of this section to provide a detailed discussion of water treatment, however, for completeness brief comments are given below.[27]

2.17.3.5.1 Water softening and water hardening

For many applications, calcium carbonate scaling from hard waters may be effectively inhibited by partial exchange of calcium ions with sodium ions, which is the process of cation exchange (water softening). It is important to note that softening, by reducing the scaling tendency of water and by increasing its ionic conductivity, will generally increase the corrosivity of the water. Conversely, where supply waters are extremely soft, mild water hardening is applied by use of lime (calcium hydroxide) treatment. This is especially effective for reducing the general solubility of metals in soft water supplies for domestic purposes.

2.17.3.5.2 Treatment with acid

The traditional method used to control the formation of calcium carbonate scale, especially on heat transfer surfaces in evaporative cooling systems, is to add sufficient acid to create nonscaling conditions. The

carbonic acid–bicarbonate equilibrium is thereby moved in the direction of carbonic acid, which encourages loss of carbonate (as CO_2) from the water. The most commonly used addition is that of sulfuric acid and although calcium sulfate is more soluble than calcium carbonate, as noted above, scaling of gypsum can occur. In such circumstances other acids, such as phosphoric, may be used.

2.17.3.5.3 Scale inhibitors and dispersants

Scale inhibitors/dispersants function by modifying the kinetics of nucleation and growth of the scale, although the precise mechanisms of action are still a subject of research. The most cost-effective scale inhibitors are inorganic polyphosphates (e.g., sodium hexametaphosphate); however, at higher temperatures these are subject to reversion back to orthophosphate, which is ineffective.[28] Organic phosphonates are also effective and more stable at higher temperatures; however, they are more costly. Organic polymers are also often very effective at inhibition of scaling by limiting nucleation and growth of the precipitated material.

2.17.3.5.4 Surfactants

Conventional cationic or anionic detergents tend to be used in heating and cooling systems in order to limit the scope of contamination of heat transfer surfaces by oils or greases and function by emulsifying the hydrocarbon contaminants. In addition, nonionic surfactants are also useful for changing the wetting characteristics of surfaces. This can effectively reduce the tendency for fouling-type materials (silts, biomass) to deposit on surfaces.

2.17.3.5.5 Corrosion inhibition in potable water systems

Polyphosphates and silicates are suitable, at low concentrations, for use in drinking (potable) water systems for corrosion and/or scale control; they can also, in many cases limit deposition and staining from dissolved iron and manganese. Polyphosphates, used at levels from 2 to 10 ppm are particularly effective at ambient temperatures and at pH of less than 7.5 can effectively inhibit calcium scale growth. For effective corrosion control, some calcium ions in solution are required to assist in the deposition of a calcium phosphate film that limits corrosion of copper, iron and lead. Sodium silicate is less effective on its own but may be used in synergy with sodium hexametaphosphate at an optimum ratio of 4:3.[29]

2.17.4 Corrosivity of Natural Waters

The extent to which water is corrosive (its 'corrosivity') is a function of a considerable number of interacting variables (i.e., oxygen concentration, flow rate, quantity of dissolved species, scaling tendency, etc.).[30–32] In view of this, it is not possible to determine the corrosivity of water for an arbitrary water composition, although it remains a future goal.

2.17.4.1 pH

The pH of natural waters is rarely outside the fairly narrow range of 4.5–8.5. Within this range, steel corrodes at approximately constant rate. The form which the corrosion takes is, however, affected by the pH. At values between 7.5 and 9.0, there is a tendency for the corrosion products to adhere as a hard crusty deposit.[33,34] Sometimes there are separate 'tubercles,' but these are more usually joined up to form a more or less continuous layer. Attack under the deposit is, however, usually irregular. At lower pH values, adherent corrosion products are not so evident although a very hard form of deposit is sometimes seen in pipes which have been in service for some years. Loss of pressure (due to a reduction in pipe cross section because of the development of internal corrosion product and other scales) is more commonly found in the higher pH range. In the lower pH range corrosion products may stay in suspension giving rise to 'red water' complaints. Water treatment may reduce the amount of corrosion, but if inhibition is not complete or possible, the type of attack is unaltered. For this reason, it is difficult to prevent corrosion in the pH range (4.5–8.5), as even a low degree of attack produces an adherent corrosion product which puts a barrier between the inhibitor and the metal. Grey cast iron behaves in a manner similar to steel at alkaline pH values but at low pH values it is subject to graphitization, which is the preferential dissolution of the ferrous component (graphite being an effective cathode).

Copper is affected to a marked extent by pH. In acid waters, slight corrosion occurs and the small amount of copper in solution may cause green staining of fabrics and sanitary ware. In addition deposition of copper onto aluminum or galvanized surfaces results in pitting of these metals. In most waters the critical pH is about 7.0 but in soft water containing organic acids it may be higher.[35] Localized pitting corrosion of copper occurs as follows: (Type I) in cold water systems where a residual carbon film is present

from the metal drawing process; (Type II) in hot water systems generally where the bicarbonate to sulfate ratio is less than one; (Type III) in cold water systems but associated with an unfavorable bicarbonate to sulfate ratio as in Type II.[36]

Lead is affected by carbonate content, pH value and mineral constituents.[37] With soft waters the simplest method of corrosion control is usually to increase the pH value by adding lime to the water. Zinc coatings on steel (galvanized) are attacked in the same way as iron, but usually more slowly.[38] Very alkaline waters are usually aggressive to zinc and will often remove galvanized coatings; the corrosion products consist of basic zinc carbonate or other basic compounds and may take the form of a thick creamy deposit or hard abrasive particles.

2.17.4.2 Physical Processes

The effects of temperature and fluid flow are complex. The more general factors may, however, be summarized as follows[39,40] : (a) the velocity of corrosion reactions is greater at increased temperatures, (b) temperature changes may affect solubility of corrosion products or shift the position of such equilibria as that existing between calcium carbonate and carbon dioxide, (c) gases are less soluble at increased temperature, an effect which is, however, partly offset by greater diffusion rates and (d) modification of pH value. The overall effect is that corrosion is usually more rapid at higher temperatures, the corrosion products being often more objectionable in nature.

2.17.4.3 Water Chemistry

It is of considerable interest to the water supply industry to be able to predict with certainty the corrosion rate and condition of its assets from the corrosion tendency (corrosivity) of the water. Although the Langelier index is probably the most frequently quoted measure for the corrosivity of water, it is not a very reliable guide. All that the index can do (and all that its author claimed) is to provide an indication of the thermodynamic tendency to precipitate calcium carbonate. It cannot indicate if sufficient material will be deposited to completely cover all exposed metal surfaces; consequently a very soft water can have a strongly positive index but still be corrosive. Similarly the index does not account for the physical form of the precipitate: that is, a semiamorphous 'egg-shell' like deposit can spread uniformly over all the exposed surfaces rather than forming isolated crystals at a limited number of nucleation sites. Egg-shell type deposits have been shown to be associated with the presence of organic material which affects the growth mechanism of the calcium carbonate crystals.[41] Where a substantial and stable deposit is produced on a metal surface, this can be an effective anticorrosion barrier.

As well as the conventional chemical parameters generally useful in gauging the corrosivity of water, for example, pH, chloride, sulfate etc., various ratios of ions have been found to be significant for particular problems. Thus an increase in the corrosion rate of iron occurs when the chloride-to-carbonate ratio exceeds 3:1[42] and attack of the dezincification prone brasses arises when the chloride to carbonate hardness ratio exceeds 1:3.[43] More usually only very limited correlations can be made between water composition and corrosivity, and even where no multiple ion effects are involved, the response to a change in one parameter may be difficult to model mathematically, for example, the corrosion of iron which passes through a maximum between pH 7.5–8 in some natural waters.[44]

2.17.4.4 Bacteriological Effects

The traditional understanding of the effect of biological activity is outlined in the previous section and focuses on either (a) classical differential aeration or concentration cell effects caused by a biofilmed surface, or (b) biological activity that produces a species that influences adversely the corrosion process.[45] The classical example is thiobacillus ferroxidans, which (in the absence of oxygen) promotes solution phase oxidation of Fe^{2+} to Fe^{3+} and consequent deposition of iron (III) oxide throughout the water system.

Recently, however, a further adverse effect of such microbial activity has been increasingly recognized. Thus, the oxidation of dissolved reduced species of iron, manganese, copper, etc., to their higher oxidation states will tend to promote pitting corrosion, particularly on passive materials such as stainless steels.[46,47]

References

1. Evans, U. R. *The Corrosion of Metals*; Arnold: London, 1926; p 93.
2. Martins, J. I.; Nunes, M. C. *Electrochim. Acta* **2006**, *52*, 552–559.
3. Harned, H. S.; Davis, R. D. *J. Am. Chem. Soc.* **1943**, *65*, 2030–2037.

4. Turner, M. E. D. *Proc. Soc. Water Treat. Exam.* **1961**, *10*(2), 162; **1965**, *14*(2), 81.
5. Hatch, G. B.; Rice, O. *J. Am. Water Works Assoc.* **1959**, *51*, 719–727.
6. Eberle, S. H.; Feuerstein, W. *Naturwissenschaften* **1979**, *66*(11), 572–573.
7. Fletcher, N. H. *Eur. J. Phys.* **1993**, *14*, 29–35.
8. Gabrielli, C. C.; Maurin, G.; Poindessous, G.; Rosset, R. *J. Crystal Growth* **1999**, *200*(1–2), 236–250.
9. Langelier, W. F. *J. Am. Water Works Assoc.* **1936**, *28*, 1500–1521.
10. Langelier, W. F. *J. Am. Water Works Assoc.* **1946**, *38*, 169–178.
11. Ryznar, J. W. *J. Am. Water Works Assoc.* **1944**, *36*, 472–486.
12. Puckorius, P. *Power* **1983**, *127*(9), 79–81.
13. Stiff, H. A., Jr.; Davis, L. E. *Petroleum Trans. AIME* **1952**, *195*, 213.
14. McCauley, R. F.; Abdullah, M. O. *J. Am. Water Works Assoc.* **1958**, *50*, 1419.
15. Simmonds, M. A. *Proc. Soc. Water Treat. Exam* **1963**, *12*(4), 197; **1964**, *13*(1), 40.
16. Gill, J. S. *Colloids Surf. A* **1993**, *74*, 101–106.
17. Sheikholeslami, R.; Tan, S. *Desalination* **1999**, *126*, 267–280.
18. Gunnarsson, I.; Arnórsson, S. *Geothermics* **2005**, *34*, 320–329.
19. Dabir, B.; Peters, R. W.; Stevens, J. D. *Ind. Eng. Chem. Fund.* **1992**, *21*(3), 298–305.
20. Gunnlaugsson, E.; Einarsson, A. *Geothermics* **1989**, *18*(1–2), 113–120.
21. Hoang, T. A.; Ang, H. M.; Rohi, A. L. *Powder Technol.* **2007**, *179*, 31–37.
22. Gill, J. S. In *Calcium Phosphates in Biological and Industrial Systems*; Amjad, Z., Ed.; Kluwer Academic, 1998; Chapter 18.
23. Seth, A.; Bachmann, R. T.; Boxall, J. B.; Saul, A. J.; Edyvean, R. *Water Sci. Technol.* **2004**, *49*(2), 27–32.
24. Geldreich, E. G. *Microbial Quality of Water Supply in Distribution Systems*; CRC Press, 1996; pp 82–83.
25. Melo, L. F. Bott, T. R.; Bernardo, C. A. Eds. *NATO ASI Series 3: Applied Science* Springer, 1988; Vol. 145.
26. Videla, H. *Manual of Biocorrosion*; CRC Press, 1996.
27. Faust, S. O.; Alym, O. M. Eds. *Chemistry of Water Treatment*, 2nd ed.; CRC Press: Boca Raton, FL, 1998.
28. Hatch, G. B.; Rice, O. *Ind. Eng. Chem.* **1945**, *37*, 710–715.
29. Mehr, M. E.; Shahrabi, T.; Hosseini, M. G. *Anti-Corros. Meth. Mater.* **2004**, *51*, 399–405.
30. Geld, I.; McCaul, C. *J. Am. Water Works Assoc.* **1975**, *67* (10), 549–552.
31. *Internal Corrosion of Water Distribution Systems*, 2nd ed.; American Water Works Association, 1996.
32. Goddard, H. P. *Mater. Perform.* **1979**, *18*(5), 21–27.
33. Larson, T. E.; Skold, R. V. *J. Am. Water Works Assoc.* **1957**, *49*, 1294–1301.
34. Sarin, P.; Snoeyink, V. L.; Bebee, J.; Kriven, W. M.; Clement, J. A. *Water Res.* **2001**, *35*, 2961–2969.
35. Gilbert, P. T. *Proc. Soc. Water Treat. Exam.* **1966**, *15*, 165.
36. Mattson, E. *Werkst. Korros.* **1988**, *39*, 499–503.
37. Cassibba, R. O.; Fernandez, S. *J. Nucl. Mater.* **1989**, *161*, 93–101.
38. Pistofidis, N.; Vourlias, G.; Stergioudis, G. *J. Mater. Sci.* **2007**, *42*, 2878–2885.
39. Hackerman, N. *Ind. Eng. Chem.* **1952**, *44*, 1752–1755.
40. Mahato, B. K.; Voora, S. K.; Shemilt, L. W. *Corros. Sci.* **1968**, *8*, 173–180, 737–749.
41. Campbell, H. S.; Turner, M. E. D. *J. Inst. Water Engineers Scientists* **1983**, *37*(1), 55.
42. Lawson, T. E. Corrosion by domestic water, Illinois State Water Survey Bulletin #59; Urbana, IL, 1975.
43. Turner, M. E. D. *Proc. Soc. Water Treat. Exam.* **1961**, *10*(2), 162.
44. Larson, T. E.; Skold, R. V. *Corrosion* **1958**, *14*(6), 43.
45. Coetser, S. E.; Cloete, T. E. *Crit. Rev. Microbiol.* **2005**, *31*(4), 213–232.
46. Chamritski, I. G.; Burns, G. R.; Webster, B. J.; Laycock, N. J. *Corrosion* **2004**, *60*(7), 658–669.
47. Shi, X.; Acvi, R.; Lewandowski, Z. *Corrosion* **2002**, *58*(10), 728–738.

Further Reading

Thresh, J. C.; Beale, J. F.; Suckling, E. V. *The Examination of Water and Water Supplies,* 6th ed.; J. & A. Churchill: London, 1944.

Faust, S. O.; Alym, O. M. Eds. *Chemistry of Water Treatment*, 2nd ed.; CRC Press, 1998.

Stumm, W.; Morgan, J. J. *Aquatic Chemistry – Chemical Equilibria and Rates in Natural Waters*; Wiley, 1995.

Internal Corrosion of Water Distribution Systems, 2nd ed.; American Water Works Association, 1996.

PHREEQC: A Computer Program for Speciation, Batch-Reaction, One-Dimensional Transport, and Inverse Geochemical Calculations, US Geological Service. Available at water.usgs.gov/software/.

2.16 Atmospheric Corrosion

I. S. Cole
CSIRO Materials Science and Engineering, P.O. Box 56, 37 Graham Road, Highett, Victoria, 3190, Australia

Abbreviations

DRH Deliquescent relative humidity
ppt Part per trillion
ppm Part per million
RH Relative humidity
TCD Time of condensation
TCON Tendency to condensation

Symbols

C Concentration of aerosol particles in the atmosphere
C_C Cunningham correction factor
C_i Molar concentration of *i* (species)
C_s A shape factor of an object
D Deposition rate (per surface area) of aerosols
D_p Particle (or aerosol) diameter
H_a Henry's law constant for gaseous absorption
I Air turbulence
M Mass of droplet
M_i Molar weight of *i* (species)
p Air pressure
p_1 Pressure within a drop (on a surface) at the front extremity
p_2 Pressure within a drop (on a surface) at the back extremity
r Droplet radius (on a surface)
R The gas constant
R_a A parameter that characterizes the resistance to deposition across boundary layers
R_c Critical amount of rain required to guarantee run off
R_i The amount of rain in a given rain event
R_t Terrain roughness
r_1 Radius of curvature near the front of a drop
r_2 Radius of curvature near the back of a drop
R_s The reactivity of a surface
RH_c Relative humidity at which a surface wets
T Air temperature (°K)
T_s Temperature of the (metal) surface
S_i Initial pollutant load
S_f Final 'soluble' pollutant load
U_0 The approach velocity (of the air stream) at a distance x_0, from an object
u^* Average fiction velocity in the local boundary layer
V_a Vapor concentration of the air
V_d The deposition velocity of aerosols
V_s Saturation vapor concentration
β A constant used in calculating deposition efficiency
ξ Mixing ratio
η The capture efficiency (of aerosols) of an object
θ Plate or object angle (from the horizontal)
ρ_p The density of an aerosol (or particle) droplet
μ The viscosity of air
σ Surface tension (between metal surface and droplet)
ϕ The difference in contact angle between the front and back of a drop

2.16.1 Typical Atmospheric Conditions

The atmosphere is the layer of gases that surround our planet Earth and that are retained by gravity. The atmosphere is not a single entity and may be divided into several regions that are subject to varying physical and chemical conditions. The lowest layer, in which 'weather' occurs, is named the troposphere and is the subject of this chapter. Above this layer, there come, in sequence, the stratosphere, the mesosphere, the thermosphere, the ionosphere, and finally the exosphere. It is the events within the troposphere that are the subject of this chapter. In this section, the typical conditions that control the formation and chemistry of gases, particulate aerosols, or rain within the atmosphere (troposphere) are outlined.

2.16.1.1 Atmospheric Gases

The average composition of the atmosphere and the nature of atmospheric cycles for each gaseous constituent are given in Table 1 (from Seinfeld and Pandis[1]). The atmosphere is a dynamic system with its gaseous constituents being constantly exchanged by vegetation, the oceans, biological organisms, volcanic action, weathering, and human influences. For example, ozone levels are controlled by chemical processes in the upper and lower atmosphere. Methane, nitrous oxide, and carbon monoxide levels are also controlled by biological processes.

When in the atmosphere, a chemical substance may be altered either as a result of sunlight promoting the decomposition of molecules (photochemical reaction) or by chemical reactions with other substances, which can occur either homogeneously or heterogeneously. Such reactions may transform the state of a substance (gas to aerosol), allowing the substance to be removed from the atmosphere more effectively. As will be discussed later, heterogeneous reactions of ammonia and sulfur dioxide gases can lead to the formation of a range of ammonium sulfate particulates that deposit at a significantly faster rate than their gaseous precursors, dramatically affecting the course of atmospheric corrosion.

As evident in Table 1, the major constituents of the atmosphere are gases that are relatively inert such as N_2 and O_2; however, the atmosphere is an effective oxidizing medium. This is partially due to the presence of free radicals (e.g., the hydroxyl (OH) radical) and partially due to such oxidizing species as ozone (O_3). Thus, many species are converted into higher chemical oxidation states, which also have an increased polarity and solubility. For example, the partially soluble SO_2 may be converted into SO_3 and then to sulfuric acid, which acidifies rainwater, particulates, and moisture films, with significant consequences for atmospheric corrosion.

The average mixing ratio referred to in Table 1 is the average ratio of the amount (or mass) of a substance in a given volume to the total amount (or mass) of all constituents in that volume. Thus, for a particular species i, the volume mixing ratio is given by

$$\xi_i = C_i / C_{\text{total}} \qquad [1]$$

where C_i is the molar concentration of i and C_{total} is the total molar concentration of air.

The mixing ratio can be readily converted to mass concentration by the following equation:

$$\text{Concentration in } \mu\text{g m}^{-3} = \xi_i \times pM_i / RT \qquad [2]$$

where M_i is the molecular weight, p is air pressure, R is the gas constant ($8.314\,\text{J mol}^{-1}\,\text{K}^{-1}$), and T is temperature.

Table 1 Atmospheric gases

Gas	*Molecular weight (g)*	*Average mixing ratio (ppm)*	*Cycle*
Ar	39.948	9340	No cycle
Ne	20.179	18	No cycle
Kr	83.80	1.1	No cycle
Xe	131.30	0.09	No cycle
N_2	28.013	780 840	Biological, microbiological
O_2	32	209 460	Biological, microbiological
CH_4	16.043	1.72	Biogenic, chemical
CO_2	44.010	355	Anthropogenic, biogenic
CO	28.010	0.12 (NH); 0.06 (SH)	Anthropogenic, chemical
H_2	2.016	0.58	Biogenic, chemical
N_2O	44.012	0.311	Biogenic, chemical
SO_2	64.06	10^{-5}–10^{-4}	Anthropogenic, biogenic, chemical
NH_3	17	10^{-4}–10^{-3}	Biogenic, chemical
NO	30.007	10^{-6}–10^{-2}	Anthropogenic, biogenic, chemical
NO_2	46.007		
O_3	48	10^{-2}–10^{-1}	Chemical
H_2O	18.015	Variable	Physicochemical
He	4.003	5.2	Physicochemical

Source: Reproduced from Seinfeld, J.; Pandis, S. Atmospheric Chemistry and Physics: From Air Pollution to Climate Change; Wiley Interscience: New York, 1997.

2.16.1.2 Specific Atmospheric Gases

Our prime concern is with atmospheric gases that determine the interaction of moisture layers with the underlying metals through the following steps:

- Promoting the formation of moisture layers through the formation of hygroscopic particulates.
- Controlling the pH of moisture layers.
- Controlling the dissolution of oxide layers and promoting ligand exchange.

For these reasons, the key compounds are the following:

- Sulfur-containing compounds.
- Nitrogen-containing compounds.
- Strong oxidants (ozone and hydrogen peroxide).
- Organic acids.

2.16.1.2.1 Sulfur-containing compounds

The principal sulfur-containing compounds in the atmosphere are H_2S, dimethyl sulfide (DMS or CH_3SCH_3), CS_2, carbonyl sulfide (OCS), and SO_2,[2] and their mixing ratios (in parts per trillion) are given in **Table 2** (modified from Berresheim *et al.*[2]).

The primary sources of hydrogen sulfide are wetlands, plants and soils, and volcanoes. DMS is produced in oceanic waters by benthic and planktonic organisms, probably from the decomposition of dimethyl-sulfoniopropionate.[3] The oxidation of DMS (by OH and nitrate radicals) is a source of methane sulfonic acid and SO_2 in marine atmospheres. DMS is also the prime source of sulfate cloud condensation nuclei (CCN) in marine atmospheres[4] and is, therefore, a direct contributor to acidified rain. CS_2 is produced in oceans, wetlands, and plants and soils.[2]

OCS has a low reactivity and a long residence time of ~7 years, and 7–8% of it is transported through the troposphere and into the stratosphere.[5] This is important, as the photolysis of OCS or reaction with O atoms leads to the widespread production of SO_2 in the stratosphere that may reenter the troposphere across broad geographical zones. OCS has a number of sources including CS_2 oxidation, ocean emission, and biomass burning. While the transformation of OCS to SO_2 may give rise to a widespread but low level of tropospheric SO_2, the prime source of SO_2 is due to industrial activity.[2]

The oxidation state of OCS, DMS, CS_2, and H_2S is 2, that of SO_4 (and sulfate species) is 6, and that of SO_3 (and sulfate species) is 6. DMS, CS_2, and H_2S are rapidly oxidized (first to sulfur dioxide and then to sulfates) and, thus, have lifetimes of a few days or less, while oxidized compounds are stable and their residence times in the atmosphere depend on the deposition rates of the liquid entities (aerosols and raindrops) in which they reside. Thus, while in terms of atmospheric corrosion the critical compounds are sulfur dioxide and sulfates, the transport and reaction of reduced sulfur compounds need to be considered. For example, DMS and OCS may be transported significant distances through the troposphere where they may be oxidized to sulfur dioxide (and other compounds) and, thus, can give rise to low but significant SO_2 levels, even in isolated nonindustrial areas. Noller *et al.*[6] for example, found that rainwater in the Northern Territory of Australia (an isolated, nonindustrial zone) was acidic at least in part because of the oxidation of DMS.

Table 2 Average mixing ratio of a number of gases

Compound	*Location*	*Average mixing ratio (ppt)*
H_2S	Marine surface layer	3.6–7.5
	Coastal regions	65
	Forests	35–60
	Wetlands	450–840
	Urban areas	365
CH_3SCH_3	Marine surface layer	80–110
	Continental surface layer	8–60
CS_2	Marine surface layer	2–18
	Continental surface layer	35–120
OCS	Marine surface layer	500
	Continental surface layer	500
SO_2	Marine surface layer	20
	North American – clean continental	160
	Coastal Europe	260
	Polluted continental air	1500
NO_x	Urban–suburban	10–1
	Rural	0.2–10
	Remote tropical forest	0.02–0.08
	Remote marine	0.02–0.04
O_3	Urban–suburban	100–400
	Rural	50–120
	Remote tropical forest	20–40
	Remote marine	20–40

Source: Reproduced from Berresheim, H.; Wine, P. H.; Davis, D. D. Sulfur in the Atmosphere. In *Composition, Chemistry and Climate of the Atmosphere*; Singh, H. B., Ed.; Van Nostrand Reinhold: New York, 1995; pp 251–307.

2.16.1.2.2 Nitrogen-containing compounds

Key nitrogen-containing species in the atmosphere are nitrous oxide (N_2O), nitric oxide (NO), nitrogen dioxide (NO_2), nitric acid (HNO_3), and ammonia

(NH_3).[7] N_2O is emitted from biological sources in soils and waters[8] and has a residence time of more than 50 years. It may be transformed in the stratosphere to NO, which may then transform to nitrogen dioxide, which in turn may transform to nitric acid and reenter the troposphere. Thus, both nitrogen dioxide and nitric acid can be found at significant distances from nitrogen sources. The major sources of nitrogen oxides NO_x ($NO + NO_2$) are fossil fuel combustion, soil release, and biomass burning.[8] Table 2 indicates that NO_x concentrations fall away dramatically from urban sources.

NH_3 is the primary basic gas in the atmosphere, with its major sources being animal waste, NH_3-based fertilizers, and emissions from soils.[9] NH_3 is readily absorbed by surfaces (including water droplets) and therefore, its residence time is short. The concentration is highly variable (from 0.1 to 10 ppb), depending on the distance from the source. Further, ammonia may react directly with gaseous acids, as in the following equations:

$$NH_3(g) + HCl(g) \leftrightarrow NH_4Cl(s) \quad [3]$$

or

$$NH_3(g) + HNO_3(g) \leftrightarrow NH_4NO_3(s) \quad [4]$$

where, the equilibrium constant is 1×10^{16} atm^2 and 3.35×10^{19} atm^2 for the formation of ammonium chloride and ammonium nitrate, respectively. Thus, as NH_3 is rapidly absorbed in wetted aerosols, and its product NH_4^+ readily reacts to form solids, ammonia salts are a major aerosol component.

2.16.1.2.3 Strong oxidants, ozone, and hydrogen peroxide

Ozone plays a major role in catalyzing the oxidation of compounds, both in the atmosphere and on surfaces. Most of the Earth's ozone is in the stratosphere; however, it may be transported into the troposphere by so-called tropopause folding events.[10] Ozone may also be produced by industrial activity and from the burning of fuels. As it is highly reactive, its mixing ratio can vary greatly geographically and because of diurnal cycles. Indeed, ozone may be produced by the photolysis of nitrogen dioxide during daylight hours via the following steps:

$$NO_2 \xrightarrow{\text{Visible and ultraviolet light}} NO + O \quad [5]$$

$$O + O_2 \rightarrow O_3 \quad [6]$$

Ozone will also react with NO at night, so there is commonly a very strong diurnal cycle in ozone concentration levels. Base levels of ozone are ~20–25 ppb, but they may rise as high as 200–400 ppb close to NO_x sources.[1]

Hydrogen peroxide concentration is closely related to photochemical activity and has mixing ratios that range from 0.2 to 37 ppb but are centered ~2–4 ppb.[11,12] Hydrogen peroxide commonly forms from radical–radical reactions such as[13]

$$HO_2 + HO_2. \rightarrow H_2O_2 + O_2 \quad [7]$$

These radicals commonly form by oxidation of hydrocarbons; however, the radicals may also react with NO_x compounds so that hydrogen peroxide levels decrease in high NO_x environments.[12] In the gas phase, hydrogen peroxide may be destroyed by its reaction with the OH. radical or by its photolysis, while in the aqueous phase it reacts with SO_2 (eqn [20]). Nevertheless, hydrogen peroxide is the main oxidant found in clouds, fogs, or rain.[1] Other oxidants that are of importance are OH^- and HO_2. which may be formed by oxidation of hydrocarbons.

2.16.1.2.4 Organic acids

The main organic acids of concern are formic acid (HCOOH) and acetic acid (CH_3COOH). Such acids come from a variety of sources, including industrial activity, biogenic activity in soils, and volcanoes, but the most common source appears to be incomplete combustion. Organic acid concentrations are commonly in the 1–10 ppb range.[14]

2.16.1.3 Aqueous Phase Reactions

Prior to discussing particulate deposition and wet deposition, an understanding of aqueous phase reactions is required, as these control the chemistry and, to some extent, the formation of cloud and rain droplets, and aerosols, as well as the absorption of gases onto wet metal surfaces. The reaction of a gas with the liquid phase is controlled both by absorption and the immediate reaction of the absorbed phase, and subsequent reactions that may change the concentration of reaction products.

2.16.1.3.1 Absorption and immediate reaction

It is well established that the concentration of an absorbed gas may be related to the partial pressure of the gas according to the following relationship:

$$A(g) \leftrightarrow A(aq) \quad [8]$$

$$[A(aq)] = H_A\, p_A \quad [9]$$

where $[A(aq)]$ is the aqueous phase concentration (in

mol l^{-1}), p_A is the partial pressure in the gas phase (in atm.), and H_A is the Henry's law coefficient for species A, which is a function of temperature, as defined by Pandis and Seinfeld.[15] According to Seinfeld and Pandis,[1] species with $H_A < 1000$ will exist mainly in the gas phase, species with coefficients between 1000 and 10 000 are considered moderately soluble in water, and those above 10 000 are considered very soluble. Henry's law coefficients for common gases are set out in **Table 3** for a temperature of 298 K. While it is evident from **Table 3** that many common gases would be regarded as relatively insoluble (e.g., SO_2, CO_2), the absorption of these gases may be significantly increased by secondary reactions.

2.16.1.3.2 Secondary reactions

Dissociation and other reactions (e.g., oxidation) can decrease the concentration of dissolved species allowing further absorption. For example, absorbed SO_2 may dissociate as set out below:

$$SO_2(g) + H_2O \leftrightarrow SO_2 \cdot H_2O \quad [10]$$
$$SO_2 \cdot H_2O \leftrightarrow H^+ + HSO_3^- \quad [11]$$

while:

$$HSO_3 \leftrightarrow H^+ + SO_3^{2-} \quad [12]$$

The balance of HSO_3^- versus SO_3^{2-} depends on pH, with the former dominating at pH values of 2–7 and the latter above this. These dissociation reactions may be taken into account by defining an effective Henry's law coefficient that represents the equilibrium dissolved species when subsequent aqueous reactions are considered. The effective Henry's law coefficient for SO_2 goes from 1.23 M atm^{-1} at a pH of 1 to in excess of 10^7 M atm^{-1} at a pH of 8. This has a dramatic effect on S(IV) concentrations in solution, which for a mixing ratio of 200 ppb of SO_2 at 298 K would increase from 3×10^{-7} to 3×10^{-3} M as pH is increased from 0 to 6.[1]

In the case of NH_3, the Henry's law coefficient is 62 M atm^{-1} and, therefore, the gaseous species is absorbed more readily than SO_2, which then dissociates according to the following reactions:

$$NH_3(g) + H_2O \leftrightarrow NH_3.H_2O \quad [13]$$
$$NH_3H_2O \leftrightarrow NH_4^+ + OH^- \quad [14]$$

The equilibrium constant (K_{eq}) value for the dissociation of $NH_3.H_2O$ is 1.7×10^{-5} and, thus, in solutions with pH < 8, the dissolved NH_3 is predominantly in the form of ammonium ion; for pH < 5 practically all the available NH_3 will be dissolved in the aqueous phase. This high solubility of ammonia makes it a dominant species in industrial aerosols.

The Henry's law coefficient for CO_2 absorption is 3.4×10^{-2} M atm^{-1} and, therefore, its absorption would be expected to be quite low; it is indeed so under acidic conditions. The reactions are

$$CO_2 + H_2O \leftrightarrow CO_2.H_2O \quad [15]$$
$$CO_2.H_2O \leftrightarrow H^+ + HCO_3^- \quad [16]$$
$$HCO_3^- \leftrightarrow H^+ + CO_3^{2-} \quad [17]$$

The K_{eq} value for eqns [16] and [17] are 4.3×10^{-7} and 4.7×10^{-11} respectively, and so in basic conditions CO_2H_2O readily dissociates to bicarbonate and hydrogen ions, allowing continued absorption of CO_2. For absorption of CO_2 (at a mixing ratio of 330 ppm), the effective Henry's law coefficient rises from a value of 3.4×10^{-2} M atm^{-1} at a pH of 4–1.5 M atm^{-1} at a pH of 8.

2.16.1.3.3 Absorption of other gases

Nitric acid is extremely water soluble and would be expected to be fully dissolved in solution at all pH > 1. Although formic acid has a relatively high Henry's law coefficient, the percent dissolved in the aqueous phase will depend heavily on pH, so that for pH < 4 only ~10% will be dissolved in solution, while for pH > 7 almost all formic acid will be dissolved. Ozone is practically insoluble, as indicated by its very low Henry's law coefficient.

2.16.1.3.4 Oxidation of reduced species

In order to determine the balance between the gas phase and the absorbed phase for a number of compounds, oxidation reactions also need to be included (as noted earlier, the atmosphere is in general strongly oxidizing). A prime and important example is the absorption of SO_2, where the oxidation of SO_3^{2-} or other forms of S(IV) to SO_4^{2-} or other forms of

Table 3 Henry's law coefficients for common gases

Species	*H (M atm^{-1}) at 298 K*
O_2	1.3×10^{-3}
O_3	1.1×10^{-2}
NO_2	1.0×10^{-2}
CO_2	3.4×10^{-2}
SO_2	1.23
NH_3	62
OH	25
HCl	727
HCOOH	3.6×10^3
CH_3COOH	8.8×10^3
NO_3	2.1×10^5
HNO_3	2.1×10^5

S(VI) is critical, as it allows the continued absorption of SO_2, according to eqns [11] and [12].

This oxidation may occur via a variety of mechanisms, including reactions with O_3, H_2O_2, and O_2 (catalyzed by Mn (II), Fe (III), and NO_2), for example:

$$S(IV) + O_3 \rightarrow S(VI) + O_2 \quad [18]$$

$$S(IV) + 0.5O_2 \xrightarrow{Mn^{2+},Fe^{3+}} S(VI) \quad [19]$$

or

$$S(IV) + H_2O_2 \rightarrow S(VI) + H_2O \quad [20]$$

where S(IV) may be HSO_3^- or SO_3^{2-}, and S(VI) is HSO_4^- or SO_4^{2-}.

According to Hoffman and Calvert,[16] the reaction rate for eqn [19] is given by

$$1.2 \times 10^6[Fe(III)][S(IV)] \quad [21]$$

The sulfate that forms from the oxidation of S(IV) may exist as SO_4^{2-} or HSO_4^- (H_2SO_4 dissociates to HSO_4^- under pH ranges likely for droplets in the atmosphere or on surfaces). The balance is governed by

$$HSO_4^- \leftrightarrow H^+ + SO_4^{2-} \quad [22]$$

The K_{eq} value for eqn [22] is 1×10^{-2} and thus the concentration of HSO_4^- will increase with acidity.

Hypochlorous acid (HOCl) and hypobromous acid (HOBr) may also lead to rapid sulfur oxidation:[17]

$$S(IV) + HOCl \rightarrow S(VI) + HCl \quad [23]$$

$$S(IV) + HOBr \rightarrow S(VI) + HBr \quad [24]$$

The observation by a number of atmospheric corrosion scientists[18–20] that there is a synergistic effect among SO_2, NO_2, and O_3 gases in determining the extent of corrosion, is likely to arise, at least in part, from the role of NO_2 and O_3 in promoting the oxidation to S(VI) and, therefore, the continued absorption of SO_2 gas into moisture layers (be they surface moisture, rain water, or wet aerosols), as well as the acidification of these moisture layers.

2.16.1.3.5 Ionic reactions in the liquid phase

The ionic species formed by gaseous absorption (and subsequent oxidation reactions) may combine to form ionic solids that can act as nuclei for cloud drops or aerosol particles. Further, ammonium chloride and ammonium nitrate may form as a direct consequence of gaseous reactions, as detailed in eqns [3] and [4]. The stability of the different solids will depend on the liquid concentration of the various species and pH. Typical reactions and their equilibrium constants are given below:

$$Na_2SO_4(s) \leftrightarrow 2Na^+ + SO_4^{2-} \quad K_{eq} = 0.48 \quad [25]$$

$$NaHSO_4(s) \leftrightarrow Na^+ + HSO_4^- \quad K_{eq} = 2.44 \times 10^4 \quad [26]$$

$$NaNO_3(s) \leftrightarrow Na^+ + NO_3^- \quad K_{eq} = 11.97 \quad [27]$$

$$(NH_4)_2SO_4(s) \leftrightarrow 2NH_4^+ + SO_4^{2-} \quad K_{eq} = 1.425 \quad [28]$$

2.16.1.4 Analysis of Typical Systems

While some of the basic reactions governing gas/aqueous phase interactions are set out in eqns [25]–[28], a rigorous analysis requires the evaluation of the chemical potentials of the gas, liquid, and solid phases as a function of concentration, temperature, and pressure, and the effect of droplet geometry on these factors. This level of detail is outside the scope of this chapter and therefore, the results from detailed analyses of three systems representing industrial and urban environments particularly relevant to atmospheric corrosion are presented, as follows:

- Industrial environment containing H_2SO_4, NH_3, and water vapor
- Urban environment containing HNO_3, NH_3, and water vapor
- A mixed industrial/urban environment containing H_2SO_4, NH_3, HNO_3, and water vapor.

2.16.1.4.1 Sulfuric acid–ammonia–water systems

In this section, we follow the approach of Seinfeld *et al.*[1,21] and consider a system containing gaseous H_2SO_4, NH_3, and water. Aerosols will nucleate homogeneously or heterogeneously and, because of their high solubility, both H_2SO_4 and NH_3 will be readily absorbed; they will react according to eqns [13], [14], and [28] to form a series of solid and dissolved species, including ($(NH_4)_3H(SO_4)_2$), $(NH_4)_2SO_4$, NH_4HSO_4, NH_4^+, SO_4^{2-}, HSO_4^-, and NH_3(aq). This system is of prime importance in determining the chemistry of aerosols in industrial locations. **Figure 1** (from Seinfeld and Pandis[21]) illustrates this behavior at low RH (30%) and at a H_2SO_4 concentration of $10\,\mu g\,m^{-3}$.

At low ammonia levels, the aerosols consist primarily of H_2SO_4, with some NH_4HSO_4(s) and significant water (H_2SO_4 attracts a significant amount of water even at low RH).[21] When the ammonia/sulfuric acid ratio increases above a molar ratio of

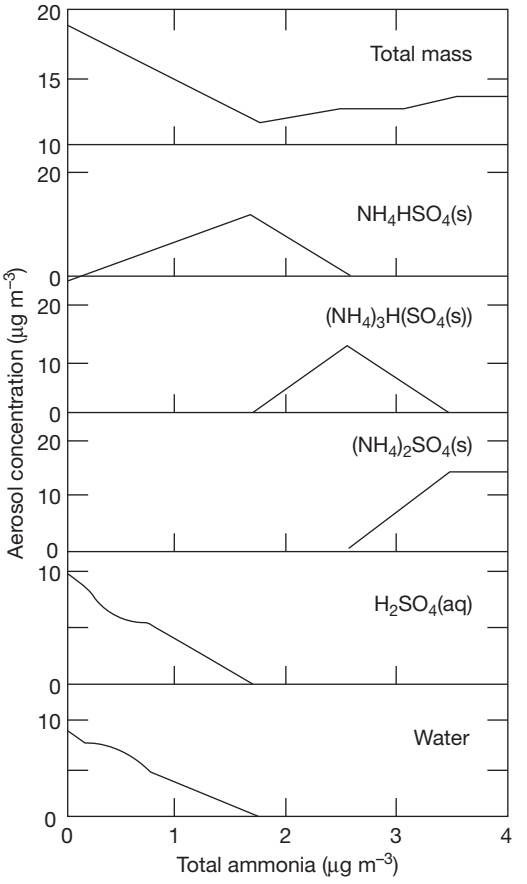


Figure 1 Aerosol composition for a system containing 10 μg m^{-3} H_2SO_4 at 30% RH and $T = 298$ K as a function of the total (gas plus aerosol) concentration. Reproduced from Seinfeld, J.; Pandis, S. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*; Wiley Interscience: New York, 1997.

0.5, NH_4HSO_4(s) becomes the dominant species, and H_2SO_4 and associated water levels fall to zero.[21] As the molar ratio approaches 1, the salt letovicite $(NH_4)_3H(SO_4)_2$(s) forms and gradually replaces NH_4HSO_4(s). At a molar ratio of 1.25–1.5, $(NH_4)_3H(SO_4)_2$(s) goes from being the dominant to the sole species and at molar ratios from 1.5 to 2 ammonium sulfate salts form. At molar ratios of 2 or greater, ammonium sulfate salt is the sole species.[15]

At 75% RH, the results are very different; the system is affected by the deliquescence of NH_4HSO_4 (which has a deliquescent relative humidity (DRH) of 40%) and $(NH_4)_3H(SO_4)_2$ (DRH = 69%), and so is in general a solution of NH_4^+ and a sulfate species (whose form depends on ammonia levels). Thus, the aerosol is a liquid solution of H_2SO_4 at low ammonia levels, HSO_4^- at moderate ammonia levels, and SO_4^{2-} at ammonia/sulfate ratios above 1.5, while the solid $(NH_4)_2SO_4$ forms at ratios above 2 when there is sufficient ammonia to completely neutralize the sulfate.[21]

It is evident from the above discussion that the composition and hydrogen ion concentration within a wet aerosol will depend critically on the relative balance of H_2SO_4 and NH_3 in the atmosphere. The hydrogen ion concentration will determine the extent of oxide dissolution promoted by the deposition of the aerosol and thus the extent of corrosion. This result can be generalized to environmental systems containing a range of S-containing gases and ammonia, and it emphasizes the need for an understanding of both acid and alkali precursors in understanding the corrosiveness of an environment.

2.16.1.4.2 Ammonia–nitric acid–water systems

According to Seinfeld and Pandis,[1] ammonia and nitric acid readily react in the atmosphere to form NH_4NO_3 (eqn [4]). This may either occur homogeneously or, if the ambient RH is lower than the DRH of NH_4NO_3 (62%), it may occur heterogeneously, with HNO_3 being readily absorbed in wet aerosols ($H_A = 2.1 \times 10^5$ M atm^{-1}). If the RH is above the DRH of NH_4NO_3, the compound may dissociate:

$$NH_3(g) + HNO_3(g) \leftrightarrow NH_4^+ + NO_3^- \qquad [29]$$

However, the equilibrium dissociation constant of NH_4NO_3 is highly dependant on temperature and so, for typical atmospheric levels of ammonia and nitrate, almost all of the species will be in the gas phase above 30 °C, while at temperatures approaching 0 °C most of the available species will be solid (below the DRH).[1] Therefore, in urban atmospheres that are low in sulfates but high in acid solutions, ammonium nitrate particulate will readily form.

2.16.1.4.3 Ammonia–nitric acid–sulfuric acid–water systems

This type of environment reflects a typical industrial/urban atmosphere that contains NH_3, HNO_3, H_2SO_4, and water vapor. According to Seinfeld and Pandis[1] a range of solid phases ($(NH_4)_2SO_4$, NH_4HSO_4, $(NH_4)_3H(SO_4)_2$, NH_4NO_3) and aqueous phases (NH_4^+, H^+, HSO_4^-, SO_4^{2-}, NO_3^-) can exist.

In situations where there is insufficient ammonia to neutralize the sulfate, the sulfate will tend to drive the nitrate to the gas phase and will tend to exist as bisulfate (or NH_4HSO_4 if ambient RH is below the DRH of this salt).[1] In ammonia-rich solutions there is competition between sulfate and nitrate for the

available ammonia, which results in a reduction in the formation of ammonia sulfate aerosol compared to systems free of nitric acid.[1]

The three cases discussed above highlight under what conditions different compounds in aerosols will form and what aqueous species will exist; they highlight, in particular, the roles of ammonia in neutralizing acid aerosols and nitrates in replacing sulfate aerosol formation. The same factors will also determine the chemistry of cloud nuclei and raindrops. It is apparent that to obtain an accurate understanding of aerosol (or raindrops) formation and chemistry, a full knowledge of gas species is required. From the 1980s onward, a significant body of experimental work[18–20] was carried out on defining the effects of various gas dosages on atmospheric corrosion. Most of these dose functions were on the basis of SO_2 levels, although some[19,20] incorporated the effects of oxidants (such as O_3). Unfortunately, no consensus was reached as to the appropriate dose function to use, possibly because the exposure regimes incorporated different levels of other gases (e.g., nitrates, ammonia). These earlier studies might have been of more value if the gas regimes at each exposure site had been studied in more depth and a refined characterization scheme developed.

2.16.1.5 Reactions in Salt Aerosols

Marine aerosols may react with a range of gaseous species to form products such as ammonium chloride, sodium nitrate, sodium sulfate, and sodium bisulfate; for example, $H_2SO_4(g)$ may be absorbed into an aerosol, and it may react with NaCl. This reaction may permit H_2SO_4 (or other strong acids such as HNO_3 that have lower vapor pressures relative to HCl) to accumulate in the aerosol, increasing acidity, and causing HCl to volatilize[22]:

$$H_2SO_4(g) + 2NaCl \leftrightarrow Na_2SO_4 + 2HCl(g) \quad [30]$$

or

$$H_2SO_4(g) + NaCl \leftrightarrow NaHSO_4 + HCl(g) \quad [31]$$

Alternatively, in urban environments $HNO_3(g)$ may react with marine aerosols:

$$NaCl(s) + HNO_3(g) \leftrightarrow NaNO_3(s) + HCl(g) \quad [32]$$

Thus, gaseous HCl may enter the atmosphere through the dechlorination of airborne sea salt particles; however, as HCl is relatively soluble ($H_A = 727$), it is readily reabsorbed into atmospheric water droplets. In turn, these HCl-rich droplets may then react with and promote the absorption of other strong acids:

$$HCl(aq) + HNO_3(g) \leftrightarrow HCl(g) + HNO_3(aq) \quad [33]$$

$$HCl(aq) + H_2SO_4(g) \leftrightarrow HCl(g) + H_2SO_4(aq) \quad [34]$$

Such cycling of HCl can thus lead to significant acidification of both aerosols and rain or cloud droplets in marine locations[23] with significant contamination by industrial or urban pollutants. The extent of chloride depletion is evident in the work of Keene *et al.*[24] who looked at the Na/Cl ratio of rainwater (the majority of Na and Cl in rain water will be derived from the scavenging of marine aerosol) as a function of distance from the nearest coast in Australia. They found that the ratio moved from close to 0.8 (the expected ratio of seawater) to up to 1.3, at 500 km from the coast, which implies that chloride loss in marine aerosols may be up to 40%.

2.16.1.6 Aerosols

Aerosols or particles in the atmosphere may either be emitted from a source directly as particles (primary aerosol) or be formed in the atmosphere by gas-to-particle conversion processes.[1] Primary aerosols may arise from natural sources such as soil dust, sea salt, biological debris, and volcanic dust, or from anthropogenic sources such as industrial dust and soot from fuels or biomass combustion.[8] Of these, soil dust and sea salt are the major sources. Aerosols range in size from tens of angstroms to several hundred micrometers and, as shown in **Figure 2**,[25] can be divided into two classes: fine particles ($<2.5\,\mu m$) and coarse particles ($>2.5\,\mu m$). In general, coarse particles other than water droplets correspond to the primary aerosols and fine particles correspond to the secondary aerosols. Coarse and fine particles are transformed and removed from the atmosphere separately and so they can be treated independently.

As evident in **Figure 2**, fine particles are divided into two modes – nuclei mode and accumulation mode – the nuclei mode of particles, being from the condensation of gases, grow by coagulation, while the accumulation range of particles largely form from coagulation of the nuclei mode and by condensation of vapors onto existing particles. Typically, the accumulation particles outnumber the coarse particles by three orders of magnitude or more and have up to ten times the surface area.[26] However, the deposition rate of the accumulation particles is two orders of

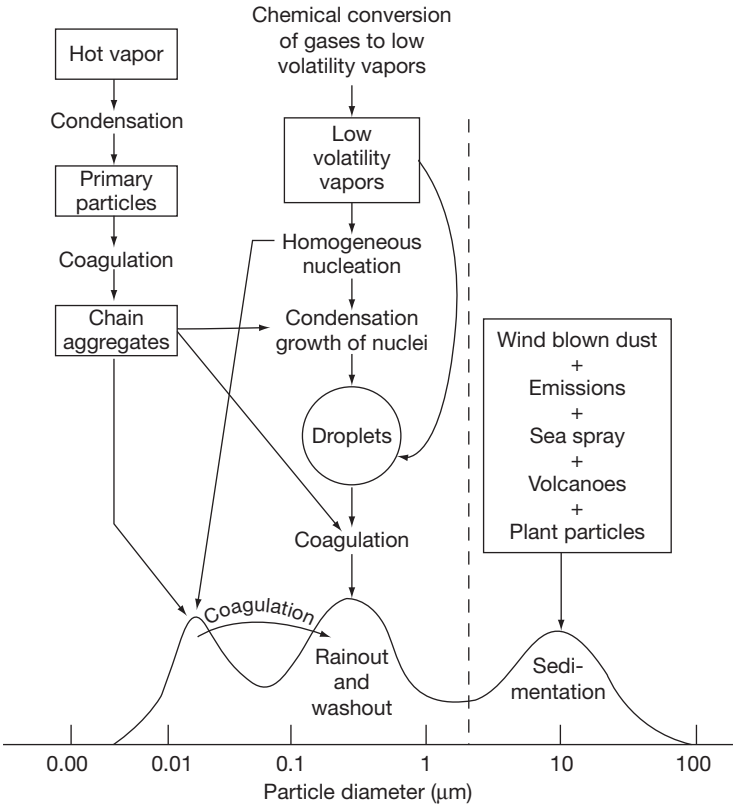


Figure 2 Idealized schematic of the distribution of aerosol size ranges and sources. Reproduced from Whitby, K. T.; Cantrell, B. Fine Particles. In Proceedings of the International Conference on Environmental Sensing and Assessment, Las Vegas, NV, 1976; Institute of Electrical and Electronic Engineers.

magnitude less than that of coarse particles; therefore, they would be expected to deposit on surfaces at similar rates (see **Section 2.16.3.1.2**).

The compositions of fine and coarse aerosols are reasonably distinct, reflecting their diverse origins. Fine aerosols commonly contain sulfate, ammonia, and organic and elemental carbon, while coarse aerosols contain sea salt, dust (silicon, calcium, magnesium, aluminum, and iron), and biogenic organic particles (pollen spores and plant fragments).[27] Nitric acid reacts with both aerosol fractions (according to eqns [4] and [32]) and thus nitrates are found in both size fractions.[27] Thus, the chemistry of the accumulation mode is influenced by the sulfuric acid and ammonia systems described above, while the coarse mode is influenced by the absorption of sulfuric and nitric acids into the marine aerosol.

Both coarse and fine aerosols may be hygroscopic and can absorb water from the atmosphere. The DRH values for common aerosol constituents are given in **Table 4**.[28,29] The DRH is the RH at which a solid aerosol will absorb enough water to become a saturated aqueous solution. Further increasing the RH will lead to additional condensation of water into the aerosol. DRH is dependent on temperature and for aerosols with more than one component, it is in general significantly less than that of the more wettable salt. Considering some marine and industrial aerosols, the DRH of a mixture of $NaNO_3$ and NaCl (which may occur if HNO_3 is absorbed into a marine aerosol) is 71.3%, while that of a mixture of NH_4NO_3 and $(NH_4)_2SO_4$ is 52.3%.[30] When wetted salts are dried, there is often a very strong hysteresis in recrystallization, so that recrystallization may occur at RH levels significantly below the DRH.

Wetted aerosols may also absorb organic molecules, with some studies[31] indicating that these may form an outer surface layer (with the polar part of the molecule in the aqueous solution and the nonpolar in air). The formation of such coatings may dramatically restrict evaporation from the aerosols.

Table 4 DRH of common aerosols (at 20 °C)

Salt	*DRH (%)*
Na_2SO_4	84.2
NH_4Cl	80.0
$(NH_4)_2SO_4$	79.9
NaCl	75.3
$NaNO_3$	74.3
$(NH_4)_3H(SO_4)_2$	69.0
NH_4NO_3	61.8
$NaHSO_4$	52.0
$(NH_4)HSO_4$	40.0
$MgCl_2$	35.0

Source: Tang, I. N. *Atmos. Environ*. **1980**, *14*, 819–828. Tang, I. N.; Munkelwitz, H. R. *Atmos. Environ*. **1993**, *27A*, 467–473.

2.16.1.6.1 Marine aerosol production

Salt aerosols may be generated either by the ocean or from surf, via bursting bubbles generated by ocean whitecaps (breaking waves that generate white foam),[32] particles torn from the crests of ocean whitecaps,[33] or breakers on the shore.[34] Surf waves tend to produce coarse aerosols, while ocean activity produces medium sized aerosols.[35] Thus, the magnitude of aerosol production from the ocean largely depends on the extent of whitecap coverage, which is related to latitude. Equatorial regions have low whitecap coverage and regions at high latitudes have high whitecap coverage.[33] Salt production from surf is dependant on local winds,[34] but may also depend on the seabed profile close to the shore (with steep slopes and mud flats reducing aerosol production).[36] An additional factor that may affect both ocean and surf wave heights (and thus aerosol production) is 'fetch.' To reach their maximum height, waves require a significant length of water where wind can blow across them. This length is defined as the fetch, and if fetch is below a critical value, waves will not reach their maximum height[37] and salt production will be reduced. Along a coast, islands, capes, and other features may reduce the fetch distance.

The implications of these aerosol production dynamics are significant. First, salt production and, thereby, salt deposition and marine aerosol-induced corrosion would be expected to be lower in tropical compared to temperate latitudes. Second, salt production from shallow or relatively narrow bays (<300 km) will be reduced and so will salt deposition and marine aerosol-induced corrosion. Experimental studies by Cole *et al.*[38,39] have demonstrated both the latitude and bay effects on salt production.

2.16.1.6.2 Types of aerosol

The nature of aerosols clearly depends on location. In marine and industrial locations, aerosols occur in both the coarse and the fine modes, with the fine mode being dominant adjacent to particular sources (e.g., highways) but falling rapidly with distance from the source.[1] The fine mode is dominated by sulfates, nitrates, and ammonia with the exact chemistry controlled by the processes outlined in **Sections 2.16.2.4.1–2.16.2.4.3**. The coarse mode consists of particles of soil dust, sea salt, fly ash, and particles generated by mechanical wear such as tire wear particles. Aerosol fractions are often derived from particles of less than 1 μm (PM_1) or 10 μm (PM_{10}) in diameter. Typical mass values[1] of PM_1 and PM_{10} are 30–150 and 100–300 $\mu g\,m^{-3}$ for urban locations.

The definition of the acidity of aerosol is quite complex. Three measures are often used, the pH of the aerosols (sometimes referred to as the *in situ* free acidity in aqueous aerosols), strong acidity and free acidity. Strong acidity is the total $[H^+]$ derived when the collected aerosol in dissolved in an excess of water (nmol H^+ per m^3 of air) while free acidity is the $[H^+]$ concentration of collected aerosols (without dissolution in water). The pH of aerosols depends heavily on the RH as it controls the aqueous content of the aerosol and, thereby, the concentration of ions. In sulfate containing aerosols, the strong acidity includes H^+ released by bisulfate dissociation which is not included in free acidity. The immediate reaction of an aerosol with a surface would be most affected by the pH of the aerosol; however, the overall impact of the environment on the surface as a whole would be affected by the strong acidity.

Pathak *et al.*[40] looked at the acidity of aerosols in the PM2.5 range collected in Hong Kong. These aerosols were rich in ammonia and sulfate with their composition controlled by the NH_4^+:SO_4^{2-} ratio which controls the neutralization of acid sulfate (as outlined in **Section 2.16.2.4.1**). A differentiation was made between samples collected when the wind blew off the ocean and that when it blew off the land. Land derived aerosols were ammonia-rich and marine derived aerosols were ammonia poor. Thus, the continental aerosol had a higher strong acidity than the marine derived aerosol (62 vs. 35 nmol H^+ per m^3 of air). However, the free acidity was higher for the marine aerosol (because of the higher water fraction in the marine aerosol). There was a very wide range in pH from −0.62 to 2.35 for all aerosols, with the pH

increasing as the RH and the ammonia: sulfate ratio increased.

Other studies have shown similar low pH values in urban locations. Yao *et al.*[41] found that in Hong Kong the fine aerosol primarily consisted of NH_4^+ and SO_4^{2-} in a ratio of 0.7 which is consistent with letovicite ($(NH_4)_3H(SO4)_2$) and had an estimated pH from −1 to 1.5. Takeuchi *et al.*[42] sampled aerosol from an urban site in Yokohama, Japan and found that the major constituents of the aerosol were NH_4^+, SO_4^{2-}, and NO_3^-, with some Na^+ and Cl^-, while the average aerosol pH was estimated to be between 2 and 2.2. Some chemical and pH differentiation occurred as a function of size, with the PM_7 aerosol having a slightly lower pH (1.9–2.1) and higher concentration of sulfate relative to the PM_2 fraction (pH 2–3). NH_4Cl and NH_4NO_3 are the major ammonia species reflecting processes that occur in a NO_2 rich environment (see **Section 2.16.2.4.3**).

In a marine location, the coarse fraction dominates (typical[1] PM_1 and PM_{10} ranges are 1–4 and 10 $\mu g\,m^{-3}$, respectively) and is associated with marine aerosols. Keene *et al.*[23] found that in the air over the ocean, in addition to the coarse fraction, there were also submicron sulfur-containing aerosols. Freshly ejected aerosol will have a pH similar to that of seawater (~8), which will rise as the aerosol evaporates and equilibrates with the environment, increasing the ionic concentration. In fact according to Sander and Crutzen[43] the concentration of ions in seawater and marine aerosol (at a RH of 76.2%) are respectively $[Cl^-]$, 0.55 and 5.4 M, $[HCO_3^-]$, 2.3×10^{-3} and 2.7×10^{-2} M, for seawater pH of 8.7 and an aerosol pH of 9.5. The aerosol pH was calculated for a RH of 76.2% assuming the solution was saturated with chloride and HCO_3 was in equilibrium with the gas phase. Subsequently, the pH may change as a result of aqueous phase reactions (particularly the absorption of H_2SO_4) and volatilization of the products of these reactions (e.g., HCl).[44] Models developed by Chameides and Stelson[45] indicate that the original alkaline pH may be reduced to below 5 within 15 min. The submicron S-containing aerosols will be primarily formed by the condensation of H_2SO_4 in clean environments, or by the absorption of H_2SO_4 by fine particulates in polluted environments. Subsequently, these wet aerosols may absorb other gases, including ammonia, which will lead to particle neutralization, so that submicron aerosols in marine locations will have a mixture of H_2SO_4, NH_4HSO_4, $(NH_4)_2SO_4$, and NH_4NO_3.[46]

Historically, there has been considerable variation in the estimated pH of marine aerosols. Recent work by Keene *et al.*[47] measured the pH of the larger sea-salt fractions (geometric mean diameter (GMD) $\geq$2.9 μm) and found that the median pH ranged from 3.1 to 3.4, while the median pH for submicron fractions was $\leq$1.6. This was supported by previous work[48] on marine aerosols off Bermuda, where pH values of aerosols were found to vary from 0 to 4–5 (1.3 μm aerosols had a pH of 2, 2.4 μm ~3, and 11 μm ~4.6). It is notable that the pH of these aerosols was strongly buffered by bisulfate. Similar results were found for aerosols measured off New England, although even the coarsest aerosols (GMD $\geq$ 10 μm) exhibited pH values less than 4. Acidified marine aerosols may be particularly corrosive as the acid component may promote oxide dissolution, while the chloride will attack a metal surface.

Ocean-produced aerosols will have residence times of hours to days and, therefore, could be rendered acidic by gaseous reactions. In contrast, most surf-produced aerosols (see **Section 2.16.2**) will be deposited within the first few hundred meters of a coast and within minutes of their production, and therefore, there may be insufficient time for enough gaseous absorption to render them acidic.

In rural locations, the coarse aerosol fraction is dominant (typically PM_1 varies from 2.5 to 8 and PM_{10} from 10 to 40) and aerosols are of natural sources but with a small anthropogenic source.[49] In remote continental[50] locations, the coarse fraction dominates (typically PM_1 varies from 0.5 to 2.5 and PM_{10} from 2 to 10) with source of particles being natural sources such as dust, pollen, plant waxes, or the products of atmospheric oxidation processes outlined in **Section 2.16.1.3**.

2.16.1.7 Cloud Nucleation and Rain Chemistry

Cloud nucleation occurs when aerosols undergo rapid growth in the presence of a supersaturation of water vapor. These aerosols are called CCN. The supersaturation of water arises because of the steady decrease in temperature with distance from the ground, so that when a hot mass of air rises, the RH of the air steadily increases (or alternatively the saturation vapor pressure decreases). CCN will tend to be the larger sized aerosols (from coarse aerosols into the accumulation mode). Once cloud droplets have formed, they may continue to grow by scavenging the remaining aerosols in the clouds or by coalescence. Once a cloud droplet reaches a critical size and falls as rain, it may continue to scavenge aerosols or absorb

gases (as detailed in previous sections) until it reaches the ground.

Because of the processes detailed previously, considerable concentrations of ionic species can build up in rainwater. Typically observed concentrations of the major species are given in **Figure 3**.[14] However, the formation and deposition of rain drops occur over significant time and spatial scales so that the composition of rainwater will reflect pollutant levels in a general area and not only at the point of deposition. Fog arises when CCN form and grow close to the ground. A number of studies have identified the problem of acid fog, both within cities and in rural and forested areas in their vicinity. North American studies of fog water have indicated pH levels from 2.9[51] to 3.6[52] with high nitrate, ammonia, and sulfate levels.

The variation of chloride concentration in rainwater, shown diametrically in **Figure 3** shows a strong dependence on the distance from the location where the rainwater was sampled. Keywood *et al.*[24] have shown that chloride concentration falls in an approximately exponential fashion, with peak values between 500 and 1000 $\mu eq\,l^{-1}$ at the coast, $\sim$10 $\mu eq\,l^{-1}$ at 500 km from the coast, and less than 1 $\mu eq\,l^{-1}$ in the center of the Australian continent. Given that the prime source of chloride in rainwater is the scavenging of marine aerosols, it is not surprising that they found that the factors that control aerosol concentration (sea state, rainfall) also influenced chloride concentration in rain water.

In **Table 5**, typical rainwater chemistry and pH values from locations around the globe are given. While there is a tendency for rainwater in or near industrial or urban areas to have low pH values, some industrial sites have relatively high pH values. In the mid-1980s[66] lower pH values were reported (pH of 4.2–4.7) close to the industrial belt in central and eastern Europe. The relatively moderate pH values reported in such industrial centers as Nanjiing (China), Dhanbad (India) or Mugla (Turkey) are

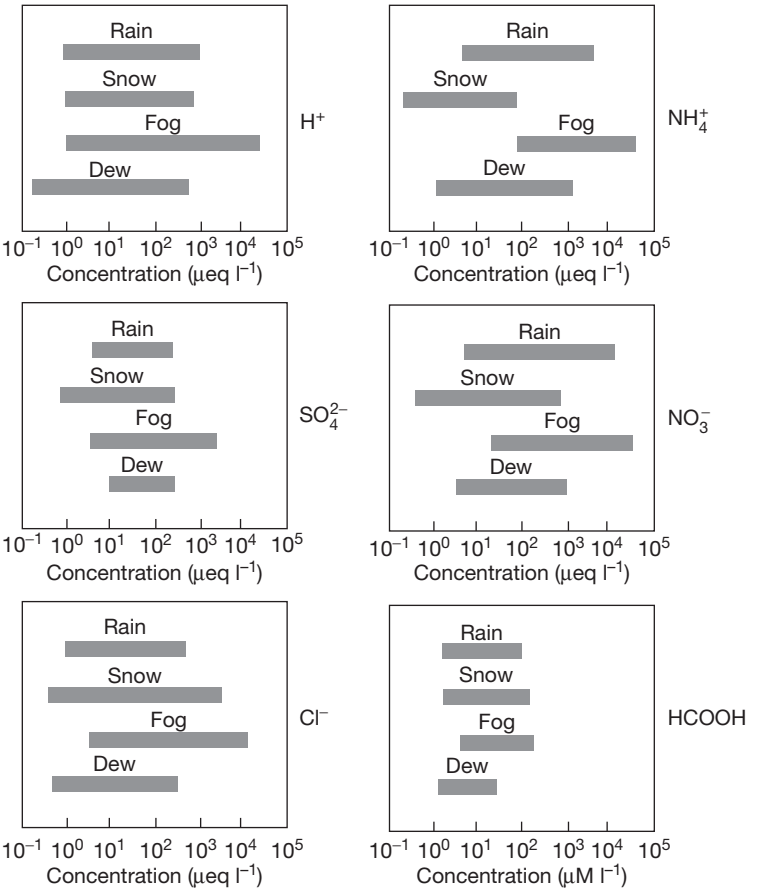


Figure 3 Concentration ranges of ionic and molecular constituents of different types of atmospheric water. Reproduced from Leygraf, C.; Graedel, T. *Atmospheric Corrosion*; Wiley and Sons: New York, 2000.

Table 5 Selected rain water compositions (concentration in $\mu eq\,l^{-1}$)

			pH mean	*pH range*	Cl^-	NO_3^-	SO_4^{2-}	*Na*	NH_4	*K*	*Mg*	*Ca*	*formate*	*Ac*	*Ref*
Mt.Rokko	Urban	Japan		4.5–5.1	25	13	12	21	10	0.8	1.8	1.6			53
Mt Awaga	Rural	Japan		4.5–5.3	25	9.5	12	28	11	3.7	3.1	2.6			53
Nanjiing	Urban/Industrial	China	5.15	4.9–5.4	142	39	241	23	193	12	32	295			54
Londrina	Urban/Industrial	Brazil	5.8	5.5–6.1	66	137	92	0	32	0.1	–	–			55
Mugla	Urban/Industrial near power station	Turkey	6.9	4.5–7.7	–	23	124	17	30	3.5	–	–			56
Mexico City	Urban/Industrial	Mexico	4.9	4.7–5.3	9	43	77	5	95	2	4	35			57
Amerfoort	Rural but nearby Urban/Industrial	S. Africa	4.4		10	25	59	9	22	5	7	19	8	6	58
Bangalore	Urban/Industrial	India	4.8		59	27	89	55	29	15	13	89			59
Dhanbad	Industrial-coal	India		4.4–6.9	27	10	63	18	35	13	16	70			60
Louis Trichardt	Rural	S. Africa	4.9		10	8	15	9	10	4	4	12	13	8	58
Jabiru	Rural	Australia		3.6–5.2	7.5	3.2	2.6	3.8	1.7	0.7	0.6	0.3	6.3		61
Katherine	Rural	Australia	4.8		11.8	4.3	3.2	7.0	2.0	0.9	1.0	1.3			61
Dorrigo	Rural	Australia	5.5		65	9.7	15	54	3.3	1.9	11	8.2	4.6	3.8	62
Barrington	Rural	Australia	5.8		24	7.3	8.8	19	7.1	2.2	6.1	11	6.8	5.7	62
Latrobe Valley	Industrial-coal power stations	Australia	5.8		137	11	35	116	–	1	34	26			63
Cape Grim	Isolated-Tasmania	Australia	6.0		1372	2.7	152	1167	1.6	16	247	78			64
Lithgow	Inland – coal power station	Australia	4.8		6.8	7.9	7.8	5.3	11.7	0.9	1.2	2.0	6.9	6.3	65

associated with the neutralizing role of ammonia and the alkali metals (particularly Ca^{+}), despite the high sulfate and nitrate contents of rain water. In urban areas, such as Mt Rokko in Japan, ammonia may play a significant neutralizing role, while nitrate levels tend to be equal in importance to sulfates.

Surprisingly, acid rainwater and cloud droplets have been found in areas remote from any industrial activity. Charlson *et al.*[67] proposed that the concentration of CCN may be controlled by DMS emissions from phytoplankton. Gillet and Ayers[61] found that DMS may indeed promote the formation of acid CCN (pH 4.8–5.8) in Tasmania, a largely rural island off the south-east coast of Australia. The same workers[64] also found acidic cloud droplets (average pH 3.8) and rainwater (average pH 4.9) in a number of inland sites in the remote Northern Territory of Australia. This acidity could be attributed to sulfates from DMS decomposition (the decomposition of DMS is complex, but may be triggered by OH radicals or NO) and to formate and acetate in cloud water and rainwater. Rainwater chemistry in rural and remote areas may also be influenced by the transport of particulate and gaseous matter. Work by Schwartz[68] indicates that industry-produced acids may be deposited up to 1000 km from their source. This is confirmed by Poste *et al.*,[62] who studied rainwater and fog acidity at two locations in northern New South Wales ~200 and 500 km from any industrial location, and ~40 and 60 km from the coast, respectively. Both sites exhibited high levels of salts of marine origin and significant sulfates from anthropogenic sources. Rainwater and fog water had similar compositions, but fog water was more concentrated (by a factor of 3).

2.16.1.8 Redox Potential and pH Diagrams

Graedel and Frankenthal,[69] on the basis of the evidence of Graedel[70] and Garrels,[71] have proposed distribution of potential and pH for dew rain water and fog. Garrels also proposed redox–pH ranges for 'environments in contact with the atmosphere.' Graedels and Frankenthals[69] and Garrels[71] distributions along with the potential–pH variations for the redox couples representing the stability of water, hydrogen peroxide, and the reduction of oxygen and ozone are given in **Figure 4**.

Line A and line B enclose the stability region of water; line A representing

$$2H^{+} + 2e^{-} \leftrightarrow H_2 \qquad [35]$$

and line B representing

$$2H_2O \leftrightarrow O_2 + 4H^{+} + 4e^{-} \qquad [36]$$

Line C and line D enclose the stability region of hydrogen peroxide; line C representing

$$H_2O_2 \leftrightarrow O_2 + 2H^{+} + 2e^{-} \qquad [37]$$

and line D representing

$$2H_2O \leftrightarrow H_2O_2 + 2H^{+} + 2e^{-} \qquad [38]$$

Line E is the variation in potential for the reduction of oxygen

$$O_2 + 2H_2O + 4e^{-} \leftrightarrow 4OH^{-} \qquad [39]$$

and line F is the variation in potential for the reduction of ozone

$$3H_2O \leftrightarrow O_3 + 6H^{+} + 6e^{-} \qquad [40]$$

The similarity of Garrels[71] and Graedel[70] analyses with the H_2O_2 stability boundary (eqn [37]) implies that natural waters behave as if they contain a small amount of dissolved oxygen. It is notable that 'waters' not in contact with oxygen such as ground water lose their oxidizing potential.[71] If the redox potential of dew/rain/aerosols is primarily controlled by the oxidizing species present, those with high ozone content would be expected to have higher potentials than predicted by Graedel and Frankenthal.[69]

2.16.2 Transport of Gases and Aerosols

2.16.2.1 General Scale of Movement of Gases and Particles

In the atmosphere, a series of processes occur in parallel – gaseous species, aerosols, and drops are constantly reacting, while the air is in constant motion on scales of a centimeter or less (tiny eddies) to continental dimensions. At the same time, species are entering the atmosphere from a variety of natural and man-made sources, or are leaving the atmosphere through scavenging by raindrops or by impact with the ground or ground-based objects. The period during which a species stays in the atmosphere is referred to as its residence time (which is closely related to the reactivity of the species) and residence time is closely related to the distance the species may travel from its source. In **Figure 5**, the temporal scales (residence times) and spatial scales are plotted for a range of chemical species.[1] As the OH radical (formed by photolysis of O_3 and H_2O_2) is highly

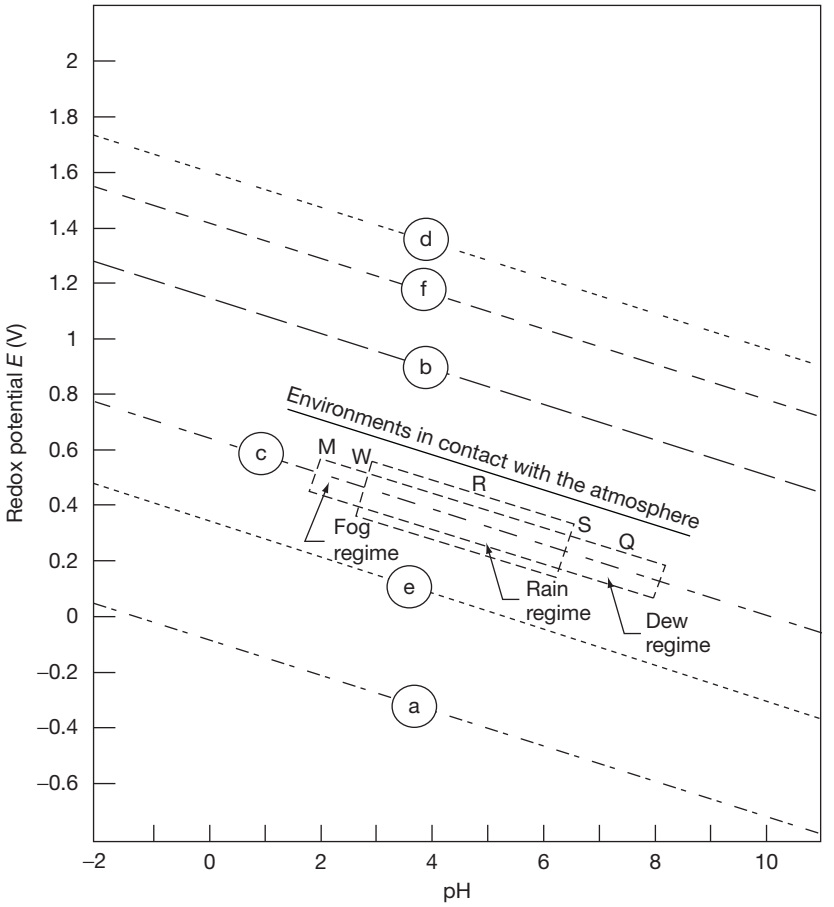


Figure 4 Redox diagram. Adopted from Pourbaix, M. *Atlas of Electrochemical Equilibria in Aqueous Solutions*, 2nd English Edition; National Association of Corrosion Engineers: 1440 South Creek Drive, Houston Texas 7704, 1974; Graedel, T. E.; Frankenthal, R. P. *J. Electrochem. Soc.* **1990**, *137*, 2385–2394.; Garrels, R. M. *Mineral Equilibria at Low Temperature and Pressure*; Harper: New York, 1960. MW represents mine waters, R represents rain waters, S represents streams, and Q represents normal ocean water.

reactive, it has a residence time of less than 0.01 s, and it will travel not more than centimeters from its source. In contrast, N_2O has lifetimes in tens of years, and it will be mixed throughout the global atmosphere (the major pathway for N_2O loss is photodissociation in the stratosphere).

The spatial impact of a species may be increased by its reaction with other species. For example, DMS has a residence time of a few hours and would not be expected to move more than several hundreds of meters from its source, yet as outlined earlier[64] the acidification of cloud droplets and raindrops at inland locations was attributed in part to DMS. As indicated earlier, the decomposition of DMS may produce both SO_2 and H_2SO_4. The residence time of SO_2 is significantly greater (days rather than hours) and subsequently its dispersal from source will be greater (kilometers rather than hundreds of meters). Finally, both gaseous SO_2 and H_2SO_4 are readily absorbed in aerosols that may have residence times of months and which can be dispersed distances ranging from tens to hundreds of kilometers from their source, depending on size. Such aerosols can then either form CCN or be scavenged by either these nuclei or by raindrops and thus acidification of cloud droplets or rain may occur at a significant distance from the original source of DMS.

Figure 5, therefore, has significant implications for atmospheric corrosion, as it highlights that highly reactive species such as hydroxide radicals will travel extremely short distances, oxidants such as NO_x and H_2O_2 limited distance (although tropospheric O_3 may travel further), gaseous SO_2 short distances, and aerosols will travel moderate distances from

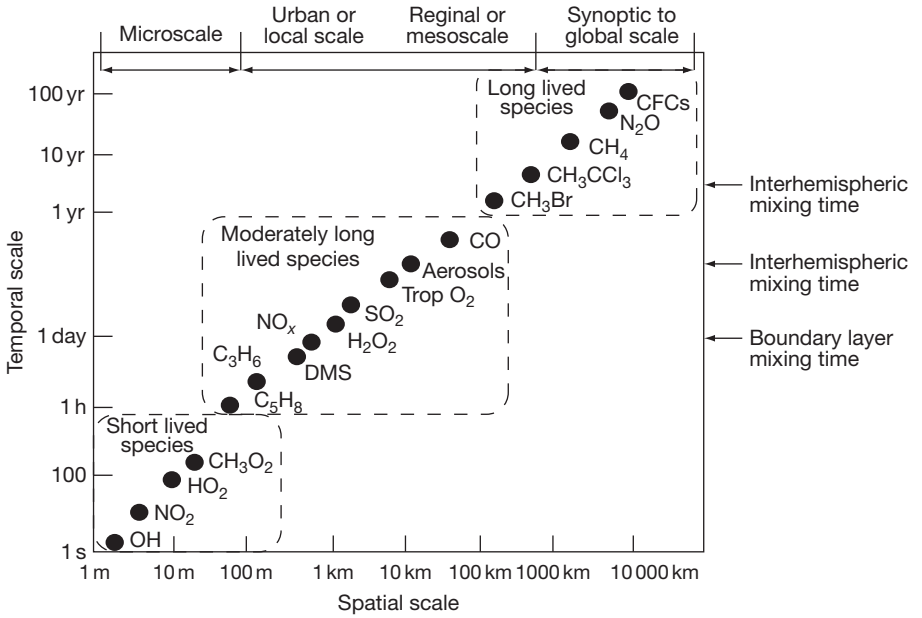


Figure 5 Spatial and temporal scales of variability for atmospheric constituents. Reproduced from Seinfeld, J.; Pandis, S. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*; Wiley Interscience: New York, 1997.

their source. Thus, direct impacts of gaseous species on metallic corrosion will be limited to tens of kilometers from source, while indirect effects via aerosols or rain deposition may be at appreciable distances from source.

2.16.2.2 Transportation Processes

More detailed studies of transport have been undertaken either by using an analytical approach (Eulerian or Langrangian) or by a numerical approach (using computational fluid dynamics (CFD)). In this section, the movement of gaseous species will be illustrated using the analytical approach and that of aerosols using the numerical approach.

2.16.2.2.1 Analytical approach to gas transport

The transport of gases in a turbulent fluid can be modeled[1] using a Gaussian distribution, with the plume equation having the form below when evaluated along the centerline ($y=0$) and at ground ($z=0$).

$$\langle c(x,0,0\rangle = q/(\pi u\sigma_y\sigma_z) \times \exp\left(-h^2/2\sigma_z^2\right) \qquad [41]$$

where q is the mean concentration from a point source, u is the wind speed, x is distance from the source, h is the effective source height, and σ_y^2,σ_z^2 are the Gaussian variances.

This equation has a maximum value for some intermediate value of x and of wind speed. The position of this maximum depends on the stability class of the plume, that is, A (extremely unstable), B (moderately unstable), C (slightly unstable), D (neutral), E (slightly stable), and F (moderately stable). Also, two case scenarios exist – that when the plume has not reached its maximum height and that when it has.[1]

The critical downward distance at which the surface concentration is at a maximum is given in **Figures 6** and **7** (from Seinfeld and Pandis[1]) as a function of source height and plume stability.

It is evident from **Figures 6** and **7** that for low-height sources (<10 m) such as short stacks, the maximum pollutant levels will occur at 100–500 m of the source, while for plumes emitted from sources at a height of 100 m the maxima may occur up to 10 km from the source.

Thus, gaseous pollutants may directly impact on metal structures for distances up to tens of kilometers from their source. However, they may also be absorbed in aerosols and transported significantly greater distances.

2.16.2.2.2 Aerosol transport

The corrosion science literature has concentrated on the transport of marine aerosols and within corrosion science there has been little work on the deposition of

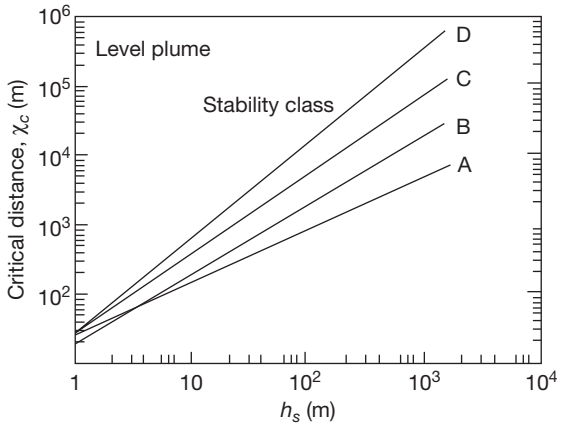


Figure 6 Critical downwind distance x_c as a function of source height h_s, and stability class for a plume that has reached its final height. Reproduced from Seinfeld, J.; Pandis, S. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*; Wiley Interscience: New York, 1997.

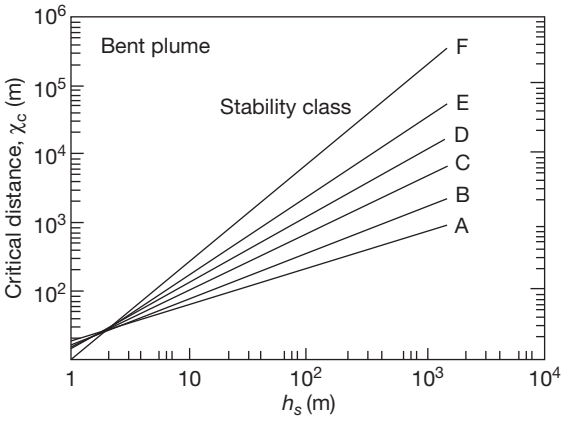


Figure 7 Critical downwind distance x_c as a function of source height h_s, and stability class for a plume that has not yet reached its final height. Seinfeld, J.; Pandis, S. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*; Wiley Interscience: New York, 1997.

industrial aerosols, which, as outlined previously, are significantly finer (generally <2.5 μm diameter). The factors influencing the transport of marine aerosols will also influence industrial aerosols, but industrial aerosols would be less affected by the various deposition processes and would, therefore, have significantly larger residence times, and would be transported longer distances.

Ohba *et al.*[72] proposed that aerosols should be divided into categories: salt produced on the shoreline, which is relatively coarse and whose concentration showed an exponential decrease with distance from the coast; and salt produced at sea, which is finer and whose concentration showed a slower decrease tending towards an asymptotic value that is relatively constant. Other works[73–79] have confirmed this double dependence of aerosol concentration on distance from the coast and have also shown that aerosol concentration depends on wind strength and direction and, in particular, on the strength and frequency of winds arising over oceans.

Cole *et al.*[80] developed a CFD model of marine aerosol transport. Aerosols are convected by wind, lifted by diffusion, and dragged down by gravity. Gravity depends on aerosol mass and, therefore, diameter, with the wet particle diameter being critically dependant on local RH for salt produced by the sea, but not for salt generated at the shoreline which has had insufficient evaporation time to reach equilibrium with the local RH. Salt is removed from the atmosphere by both wet (rainout and washout) and dry deposition. In rainout, salt aerosols become the nuclei of raindrops that agglomerate and fall to earth (thus, there is insignificant aerosol above the cloud layer). Washout refers to the scavenging of salt aerosols by raindrops and its effectiveness depends on the radius of the aerosol droplet. Dry deposition includes deposition on sloping as well as horizontal surfaces. Trees are particularly effective at scavenging salt from the atmosphere, as too are man-made structures. While medium to large aerosols will be readily trapped, small salt aerosols will follow the airstreams around tree leaves and man-made structures. Therefore, separate models are run for the coarse surf-produced aerosols and the relatively fine ocean-produced aerosols.

Figure 8 (from Cole *et al.*[80]) shows the effect of wind speed on the salt concentration (% volume × 10^{10}) with distance from the coast for surf-produced aerosols (ground roughness 0.5 m) at a height of 2.0 m above ground level. **Figure 9** (from Cole *et al.*[80]) shows the effect of wind speed on the transport of salt produced in the open ocean over a distance of 50 km from the coast at a height of 2.0 m above ground level (ground roughness 0.5 m, ground RH 70%, rainfall 1200 mm year^{-1}). **Figure 10** (from Cole *et al.*[80]) shows the effect of ground roughness on the salt concentration of surf-produced aerosols (ground roughness does not appear to have a significant effect on the transport of ocean-produced aerosols). **Table 6** gives some typical roughness values of various terrain types (from AS/NZ 1170).

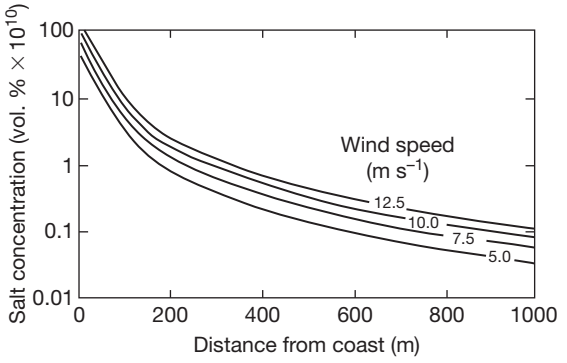


Figure 8 Effect of wind speed on salt concentration with distance from coast of surf-produced aerosols. Note that this excludes ocean-produced aerosols. Cole, I. S.; Chan, W. Y.; Trinidad, G. S.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2004**, *39*(1), 89–96.

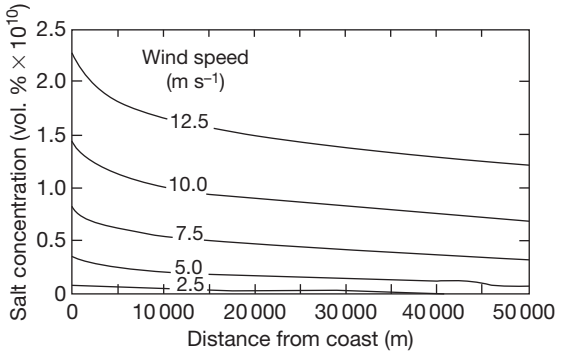


Figure 9 Effect of wind speed on salt concentration with distance from coast for ocean-produced salt. Reproduced from Cole, I. S.; Chan, W. Y.; Trinidad, G. S.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2004**, *39*(1), 89–96.

Rainfall also has a significant influence on the transport of ocean-produced aerosols. For example, given a wind speed of 7.5 m s^{-1}, a ground roughness of 0.5 m and 70% RH at a distance of 200 km from the coast, uniform rainfalls of 200 and 1600 mm $year^{-1}$ will reduce aerosol concentrations by 3% and 23% respectively, relative to the case of no rain. The effect of RH is even more marked. This can be illustrated from the ratios of the aerosol concentration when the surface RH is 20%, 70%, and 90% to that when it is 50%, which are 1.4, 0.5, and 0.2 respectively (under the same conditions as above).

The strong effect of humidity on salt concentration is related to two effects. The first arises as the model assumes (following Seinfeld and Pandis)[1] that there is a linear decrease in temperature with height and thus a corresponding increase in RH; therefore, the surface RH controls the cloud height. The cloud height is the upper limit for salt aerosol dispersion (aerosols act as scavengers or are scavenged by CCN); therefore, the higher the cloud level, the greater the vertical dispersion of salt and the greater the residence time and transport of the average aerosol. The second effect is on the particle size (or mass). The main constituents of marine aerosols are NaCl and $MgCl_2$, with each salt wetting at 75% and 35% RH (at 20 °C) respectively, so that above these humidities the aerosol increases its mass, leading to a higher terminal velocity and a greater sink effect of gravity. As the major mass change occurs after the wetting of the NaCl constituent, a significant decrease in transport is observed when the surface RH is increased from 50 to 70% (recall that the RH at a height will be greater than that at the surface). This model has been validated by experimental studies of aerosol deposition across Australia.[38] Experimental and numerical studies indicate that while coarse surf-produced aerosols are deposited close to the coast, those ocean-produced aerosols that have a diameter less than 2.5 μm can be transported significant distances (hundreds of kilometers) from the coast. As industrial aerosols tend to be finer than marine aerosols, it is probable that these too will be transported significant distances from their sources.

2.16.3 Deposition

2.16.3.1 Forms of Deposition

Traditionally, deposition onto surfaces has been classified as wet or dry deposition, with deposition by rain and snow being defined as wet deposition, and deposition by particulates and aerosols (even when these are wet) being defined as dry. Deposition modes can also be defined as follows:

- Deposition of snow, rain, or fog.
- Deposition of particulates or aerosols.
- Deposition of gases into wet surfaces.
- Depositions of gases onto dry surfaces.

2.16.3.1.1 Rain

Raindrops are generally assumed to have a 100% efficiency of deposition onto a surface. However, once a raindrop falls onto a surface, it may splash off the surface and break into smaller drops, which could reenter the airflow and not redeposit. Analyses by Cole and Paterson[81] indicated that although splash does occur, it does not lead to significant loss

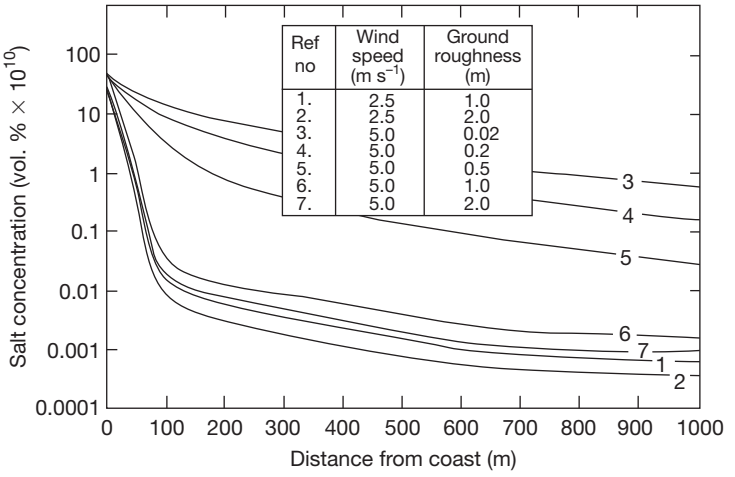


Figure 10 Effect of ground roughness on salt concentration for surf-produced aerosols. Cole, I. S.; Chan, W. Y.; Trinidad, G. S.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2004**, *39*(1), 89–96.

Table 6 Typical roughness values (from AS/NZ 1170)

Roughness (m)	*Terrain*	*Roughness (m)*	*Terrain*
2	City buildings (10–30 m)	0.2	Level wooded country; suburban buildings
1	Forests	0.06	Isolated trees, long grass
0.8	High-density metropolitan	0.02	Uncut grass, airfields
0.4	Center of small towns	0.008	Cut grass

of deposition efficiency for rain falling onto stationary objects. However, it can be quite significant for fast-moving objects (e.g., airplanes in flight).

2.16.3.1.2 Gas and particle deposition

Similar principles can be applied to both the dry deposition of gases and to the deposition of particles. A generic formulation, common in atmospheric physics, but pioneered in corrosion studies by Haynie,[82,83] and developed by Spence *et al.*,[84] is given by

$$D = CV_d \qquad [42]$$

where V_d is the deposition velocity, and C is the upstream concentration of depositing species. If C is measured in kg m^{-3}, D is measured in kg m^2 s^{-1}, and therefore, D is multiplied by the surface area to get the total deposition.

There are a number of approaches to defining the deposition velocity. One approach, developed for atmospheric corrosion by Spence *et al.*[84] is the resistance factor approach, where the deposition velocity can be expressed as follows:

$$V_d = 1/(R_a + R_s) \qquad [43]$$

where R_a characterizes the deposition across boundary layers to the surface, and R_s characterizes the reactivity of the surfaces. If R_s is low, the value of R_a determines V_d. V_d is determined from an analogy with momentum transport, which gives

$$V_d = u_*^2/U \qquad [44]$$

where U is the upstream velocity, and u is the average 'friction velocity' in the local boundary layer. u_*^2 is proportional to the local turbulent kinetic energy in the boundary layer and thus depends on the size of the obstacle. This approach recognizes the importance of air turbulence, but assumes that the turbulence is generated solely by flow over the obstacle, ignoring the effect of turbulence in the air upstream of the obstacle. Given that R_s is taken as zero, the work of Spence *et al.*[84] and Haynie[83] leads to a dependence of deposition velocity on wind speed, where

$$V_d \propto U^n \qquad [45]$$

Here the value of the exponent n varies from 0.6 to 0.7 for sheets and 0.5 for panels and wire.

A more recent approach is that of Klassen and Roberge[85] who considered the effect of inertial

impaction and turbulent diffusion and introduced the following equation for application to CLIMAT units and salt candles:

$$D = CU\eta \qquad [46]$$

where η is the capture efficiency of the target. This is multiplied by the projected target area to get the total deposition. This approach implicitly includes turbulence due to the target, but also assumes that there is no upstream turbulence.

In summary, all approaches presented above assume smooth flow upstream of the objects, which effectively mean that they are treating objects isolated from the surrounding environment. In the real environment, upstream turbulence due to ground roughness will vary significantly and may have a major impact on deposition onto salt candles, plates, or buildings.

A variation to the resistance approach is that of Cole *et al.*,[86] in which the overall deposition velocity v can be calculated from the deposition velocities for the individual deposition mechanisms. A mechanism may act in parallel or in series and when the mechanisms occur together or in parallel (at the same distance from the surface), the deposition velocities can simply be added together:

$$v = \sum_i v_i \qquad [47]$$

If the deposition mechanisms occur in different layers and the depositing aerosol must pass through all layers (act in series), deposition velocities must match between layers:

$$v_i = v_j \qquad [48]$$

The main deposition mechanisms for aerosol particles are the following:

- Gravitational settling.
- Turbulent diffusion.
- Laminar diffusion (also called Brownian deposition and diffusiophoresis).
- Thermophoresis (migration from high temperatures to low).
- Electrostatic attraction.
- Momentum-dominated impact.
- Vortex shedding (transport by transient laminar flows).
- Filtering (flow past or through raised fabrics).
- Photophoresis (motion generated by an intense beam of light).

However, while all these mechanisms can be important in interior spaces, in open spaces the prime mechanisms are gravity, momentum-dominated impact, and turbulent diffusion. Figure 11 compares the efficiency of deposition of particles of increasing sizes onto a cylinder (typical of a salt candle, as described in ISO 9205) via momentum-dominated impact and turbulent diffusion. Efficiency is defined as the percentage of aerosol flux that would pass through the air space in the absence of an object. For very small particles, smaller than 0.1 or 0.01 μm in diameter, electrostatic attraction and laminar diffusion play a significant role in deposition. For small particles with a diameter of ~1 μm, turbulent diffusion dominates, while for larger particles with a diameter of ~5 μm momentum-dominated impact dominates. For particles with a diameter above 20 μm gravitational settling becomes the prime deposition mechanism. As the particle diameter increases past 2.5 μm the deposition efficiency increases dramatically.

An approximate equation for the deposition velocity for momentum dominated impact is[86] as follows:

$$v_i = -x_0/\tau + \sqrt{(x_0/\tau)^2 + U_0^2} \qquad [49]$$

where U_0 is the approach velocity at distance x_0 from the object. If U_0 is the upstream velocity, x_0 is of similar

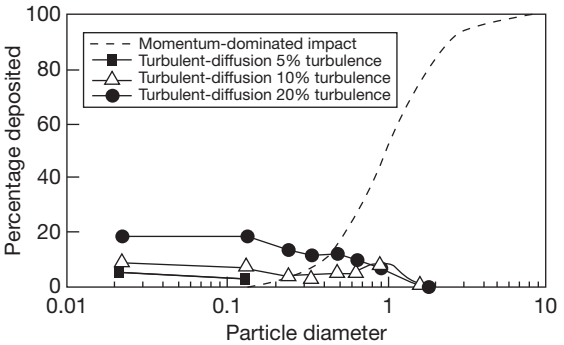


Figure 11 Influences of momentum-dominated impact and turbulent diffusion on the deposition of aerosols onto a cylinder of arbitrary diameter at arbitrary wind speed, as a percentage of the aerosol flux that would pass through in the absence of the cylinder. The percent turbulence is the turbulence intensity (rms velocity/mean velocity) upstream. The reference diameter is the particle diameter for which 50% of particles are deposited by momentum-dominated impact. For a salt candle the reference diameter is 40 microns. Reproduced from Cole, I. S.; Paterson, D. A.; Lau, D. Holistic Modeling of Gas and Aerosol Deposition and The Degradation of Cultural Objects. In *Physical Techniques in the Study of Art, Archaeology and Cultural Heritage*, to be published by Elsevier, 2007.

size to the object diameter. The relaxation time τ is given by

$$\tau = \frac{D_p^2 \rho_p C_c}{18\mu} \quad [50]$$

with particle diameter D_p, density ρ_p, air viscosity μ, and Cunningham slip correction factor C_c.

In fact, Cole *et al.*[87] found that if the deposition modes are combined, the deposition onto common objects (plates, salt candles) in the open environment can be estimated by a formula analogous to that of Klassen and Roberge[85]:

$$D = CVA\eta \quad [51]$$

where

$$\eta = 100 \times I^{\beta} \times C_s \quad [52]$$

and I is turbulence; β is a constant (determined numerically to be 0.38), and C_s is the shape factor.

The shape factor, as well as the deposition efficiency, is reasonably robust, being reasonably independent of wind speed; however, it does show a dependency on particle size. In Table 7 (from Cole *et al.*[87]), deposition efficiencies are given for particles in the size range 1–229 μm for salt candles and exposure plates (at 45° to the airflow). It is apparent that deposition onto salt candles is less efficient than onto exposure plates for small and medium size particles (<21 μm diameter). For large particles (>21 μm diameter), the efficiency of deposition onto salt candles is marginally higher than that onto exposure plates.

In practical terms, this implies that in severe marine locations where large aerosols will make up a large fraction of the aerosol distribution, it is reasonable to assume that the deposition onto an exposure plate will be approximately the same as that onto a salt candle. However, in milder locations the aerosol distribution will be dominated by aerosols of less than 15 μm and so the salt candle measurements of deposition will be ~71% of that on exposure plates.

The average shape factors for particles (assuming the particle distribution given by Fitzgerald)[35] of <20 μm for plates and salt candles are given in Table 8 (from Cole *et al.*[87]). These factors can be used in conjunction with eqn [48] to calculate deposition efficiencies and thus deposition rates. Deposition also depends on turbulence, which in turn depends on the upstream roughness of the terrain. However, the latter dependence is relatively weak

Table 7 Comparison of salt deposition efficiency on plate and salt candle

	Percentage deposition (%)		
Particle diameter (μm)	*Salt candle*	*Plate*	*Ratio*
1	9.2	12.2	1.33
6	7.6	12.2	1.61
11	7.2	12.7	1.76
15	9.8	14.4	1.47
21	18.6	16.7	0.90
29	31.0	19.0	0.61
39	50.4	33.2	0.66
72	77.0	66.8	0.87
134	91.8	88.8	0.97
249	96.4	94.4	0.98

Source: Reproduced from Cole, I. S.; Lau, D.; Chan, F.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2004**, *39*(4), 333–338.

Table 8 Summary of shape coefficients

Object	C_s
Salt candle	0.85
Exposure plate at 45° to airflow	1.2
Plate perpendicular to airflow	0.5
Plate parallel to flow	1.67

Source: Reproduced from Cole, I. S.; Lau, D.; Chan, F.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2004**, *39*(4), 333–338.

and can be approximated by

$$I = 0.069 + 0.5R_t \quad [53]$$

where R_t is terrain roughness.

Thus, as roughness increases from 0.06 to 0.4 m, the deposition efficiency onto a salt candle of aerosols of 20 μm diameter increases from 20% to 30%. Typical values of terrain roughness are 0.06 m for isolated trees and long grass, and 0.4 m for the center of small towns.

This model of deposition onto simple objects can be extended to more complex objects such as buildings or rows of buildings. The deposition onto an isolated building (10 m high and 20 × 20 m in plan) was computed using CFD and the result is given in Figure 12 (from Cole *et al.*[87]). The randomness in Figure 12 is caused by the stochastic nature of the simulation. The aerosol deposition rate is lowest near the ground, high near the upper edge of the front face and on the sides, and is highest on the upper corners and upper sides. The deposition rate is highly influenced by local wind turbulence, which is highest on corners and sides. Table 9 (from Cole *et al.*[87]) gives the average deposition rates for particles of varying diameter onto a building compared to deposition of

the same sized particles onto a salt candle. It is apparent that the average deposition rate onto a building face is less than half that on a salt candle, although deposition onto the edges of a building could be up to twice that on a salt candle.

In Figure 13 (from Cole *et al.*[87]), the calculated deposition onto a row of 2D buildings is shown (the buildings were 3.32 m high, 12.5 m wide and were spaced 30 m apart). The plot demonstrates the rapid fall-off in deposition with increasing number of houses. Such a fall-off was observed experimentally and is reported in Cole *et al.*[39] The data presented in Figure 13 has been transformed to derive a theoretical urban facade factor, so that the deposition onto the facade of a dwelling close to the coast in an urban (low-rise) setting would be only 23% of the deposition onto an equivalent facade of a dwelling where there were no dwellings between this dwelling and the coast.

The above examples have been calculated for marine aerosols; however, similar effects would also occur for industrial aerosols, although the exact magnitude of the effects would be dependent on particle size.

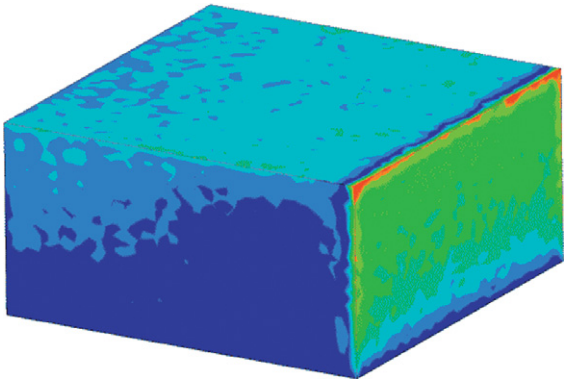


Figure 12 Aerosol deposition on a building 10 m high and 20 × 20 m in plan. The flow is from right to left. Blue is low concentration and red is high concentration. Reproduced from Cole, I. S.; Lau, D.; Chan, F.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2004**, *39*(4), 333–338.

Table 9 Deposition ratio on a building relative to an isolated salt candle

Face	*Deposition ratio (mg m^{-2})*
Front	0.42
Side	0.35
Top	0.49
Back	0.08

Source: Reproduced from Cole, I. S.; Lau, D.; Chan, F.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2004**, *39*(4), 333–338.

2.16.3.1.3 Deposition of gases onto wet and dry surfaces

In estimating both the wet and dry deposition of gases onto surfaces, it can be assumed that the gas concentration near the surface approximates the local composition and that any absorbed gas is readily replaced. The absorption of gases into moisture layers on metal surfaces is governed by the same factors that control absorption of gases into aerosols (outlined in Section 2.16.1.3) and will in part be controlled by the Henry's law coefficient. In contrast, the surface absorption of gaseous species onto dry surfaces is controlled by normalized reactivity. Table 10 (from Seinfeld and Pandis[1]) presents the effective H_a and the normalized reactivity values for various gas species.

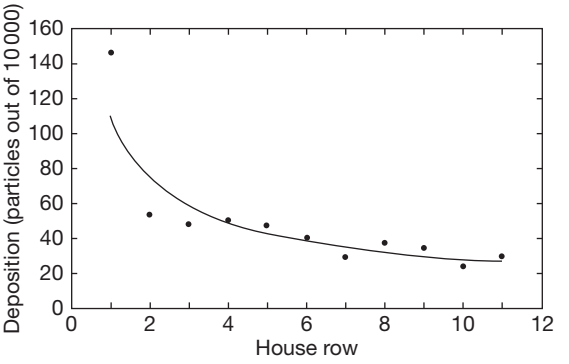


Figure 13 Results from CFD modeling of deposition onto a row of houses. Reproduced from Cole, I. S.; Lau, D.; Chan, F.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2004**, *39d*(4), 333–338.

Table 10 Relevant properties of gases for dry deposition calculations

Species	*H_a[a] ($M\,atm^{-1}$) at 298 K*	*Normalized reactivity*
Nitric oxide	2×10^{-3}	0
Ozone	1×10^{-2}	1
Nitrogen dioxide	1×10^{-2}	0.1
Hydrogen sulfide	0.12	–
Ammonia	2×10^{4}	0
Nitrous acid	1×10^{5}	0.1
Sulfur dioxide	1×10^{5}	0
Hydrogen peroxide	1×10^{5}	1
Formic acid	4×10^{6}	0
Acetic Acid	4×10^{6}	0
Hydrochloric acid	2.05×10^{6}	0
Nitric acid	1×10^{14}	0

[a]Effective H_a assuming a pH of 6.5.

Source: Reproduced from Seinfeld, J.; Pandis, S. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*; Wiley Interscience: New York, 1997.

The effective H_a in Table 10 takes into account reactions of the aqueous species. When H_a is high and the normalized reactivity is low (e.g., sulfur dioxide), deposition will be primarily through absorption into a moisture droplet (be it a wet aerosol or a surface moisture film). When H_a is low and the normalized reactivity is high (ozone), direct gaseous absorption onto dry surfaces will dominate. Table 10 indicates that those species of prime importance to atmospheric corrosion (ammonia, sulfur dioxide, formic acid, and acetic acid) will all primarily be absorbed through the moisture layer. Of the nitrogen species, nitric acid will be readily absorbed into the moisture layer, while nitrogen dioxide will undergo limited interactions with both a dry and a wet surface. While the oxidizing catalyst O_3 and hydrogen peroxide will be readily absorbed on a dry surface their residence times in the atmosphere are extremely short and, therefore, the source would have to be very close to the metal surface for the concentrations of these gases to be significant. Thus for practical purposes, gaseous deposition onto a dry surface is of marginal importance to atmospheric corrosion.

2.16.3.2 Comparison of Deposition Modes

The relative significance of particulate deposition and gaseous deposition in an industrial environment clearly depends on the particulate and gas concentrations (aerosol deposition will dominate in marine environments). However, a comparison of deposition modes in given typical compositions is useful. Consider the deposition onto a plate exposed at an angle of 45° in an exterior environment where the total concentration of particulates is 20 μg m^{-3} (dry weight), which may be ammonium sulfate, ammonium bisulfate, or hydrogen sulfate, while the wind speed is 3 m s^{-1} and gaseous concentrations are CO_2 at 400 ppm, SO_2 at 75 ppb, NH_3 at 20 ppb, O_3 at 200 ppb, and H_2O_2 at 10 ppb. Work by Sehmel[88] indicates that ammonium sulfate particles typically have a bimodal size distribution with peaks at ~0.7 and 3 μm, with the smaller particles having three times the concentration (by volume) of the coarse particles. For this exercise, let us simplify the distribution and assume that the mass of fine particles (0.7 μm diameter) is 15 μg m^{-3} and the mass of coarse particles (3 μm diameter) is 5 μg m^{-3}.

Deposition of particles traveling horizontally will be dominant and will be defined by eqn [46]. Sehmel[88] indicates that typical efficiencies for 0.7 and 3 μm particles would be ~0.26% and 18%, and thus the flux at a typical wind speed of 3 m s^{-1} will be 0.12 and 2.7 μg m^{-2} s^{-1} for the fine and coarse particles, respectively. This then gives a deposition rate per hour of 10 150 μg m^{-2} or 108 μg/dry weight per hour onto a plate of size 100 × 150 mm at an angle of 45° to the vertical.

If this amount of particulate were to deposit into a moisture film, it would then change the composition of that film to an extent that depends on the particulate composition and thickness of the moisture film. The change in $[H^+]$ and $[SO_4^{2-}]$ in M l^{-1} per hour is given in Table 11.

Modeling by Cole *et al.*[89] estimated that for the gaseous compositions given above, the aqueous phase concentrations will be 1.7×10^{-6} M of $[H^+]$ and 1.2×10^{-4} M of $[SO_4^{2-}]$, or 0.63×10^{-4} M of $[H^+]$ and 0.34×10^{-4} M of $[SO_4^{2-}]$ in the absence of NH_3. Thus, the deposition of particulate onto a moisture film will substantially increase the ionic content and acidity of the film. If the moisture film persists for a number of hours, particulate deposition is likely to dominate over gaseous absorption.

A second comparison can be made between the hydrogen ion flux to the surface of a similar metal sample from acidified rain and from particulates. In Figure 14, the 'free' hydrogen ion flux per day onto a metal surface is estimated from $(NH_4)_3H$

Table 11 Effect of particulates on moisture film composition

Particulate concentration (μg m^{-3} dry weight)	*Particulate type*	*Moisture film thickness (μm)*	*$[H^+]$ (M)*	*$[SO_4^{2+}]$ (M)*
20	$(NH_4)_2SO_4$	100	–	7.5×10^{-4}
20	$(NH_4)_3H(SO_4)_2$	100	4×10^{-4}	8×10^{-4}
20	NH_4HSO_4	100	8.5×10^{-4}	8.5×10^{-4}
20	H_2SO_4	100	20×10^{-4}	10×10^{-4}
20	$(NH_4)_2SO_4$	500	–	1.5×10^{-4}
20	$(NH_4)_3H(SO_4)_2$	500	0.8×10^{-4}	1.6×10^{-4}
20	NH_4HSO_4	500	1.7×10^{-4}	1.7×10^{-4}
20	H_2SO_4	500	4.0×10^{-4}	2.0×10^{-4}
5	H_2SO_4	500	1×10^{-4}	0.4×10^{-4}

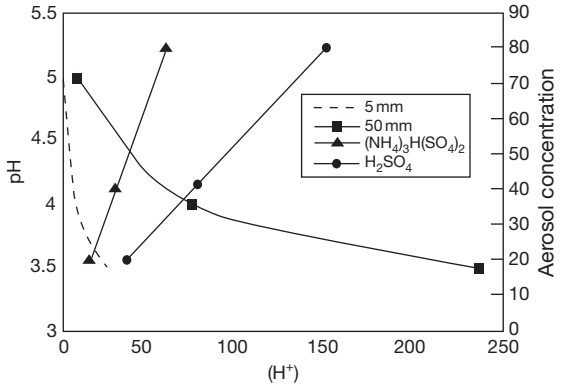


Figure 14 Hydrogen flux onto a metal surface by particle deposition and rain deposition.

$(SO_4)_2$ and H_2SO_4, assuming that the particulate dissociates on the surface in accordance with standard equilibrium constants. This is compared to the free hydrogen ion flux from rain water, assuming an unbuffered solution for a daily rainfall of 5 and 50 mm. It is apparent that fluxes are of similar orders of magnitude, but acidified rain is the major source of H^+ if the rain water pH is low and rainfall high, while particulate deposition is of prime importance if the rainfall is low and the concentration of acidified particulates is high.

The analysis indicates that while dry deposition of gases onto metal surfaces is likely to be of marginal importance, rain deposition, particulate deposition, and gaseous absorption into moisture films are all likely to be significant sources of both acidity and ion concentration in moisture on metal surfaces in industrial environments. The evaporation of moisture layers formed and enriched by the processes above will lead to the precipitation of a range of ionic compounds (many of which are listed in **Table 4**). Being hygroscopic, these compounds will promote further wetting if they remain on the surface. The next two sections discuss the cleaning, retention, and wetting of these salts on metal surfaces.

2.16.4 Role of Rain and Wind in Surface Cleaning

The products that accumulate on a surface control future hygroscopic wetting and the nature of the moisture films that develop. Surfaces may be cleaned by either rain or wind; however, the effectiveness of these processes is highly dependent on both the characteristics of the wind and rain, and the state of the surface, as discussed below.

2.16.4.1 Role of Rain in Surface Cleaning

In this section, a simple model for the movement of raindrops, and thus their ability to clean a 'fresh' surface, is developed. In this context, a fresh surface refers to a surface that remains relatively impervious to moisture (i.e., not a porous surface that absorbs moisture). According to Cole *et al.*,[81] during rain, drops will grow on the surface as they coalesce and absorb impacting droplets until they reach a critical size, above which they begin to move. At the critical size, the gravitational force directed along the plate ($gM \sin\theta$) is balanced by the shear force at the base of the drop (τA) and a force due to surface tension. The shear force tends to zero when the drop is not moving. For the force due to surface tension, let 2ϕ be the difference in interfacial contact angle between the front and the back of the drop. Then, for a drop of radius r, the radii of curvature near the front and back of the drop (r_1 and r_2 respectively) are roughly given by $r_1/r = r/r_2 = 1 - \tan\phi$. The difference in pressure within the drop between the front and back extremities of a drop is then

$$p_1 - p_2 = \sigma/r_1 - \sigma/r_2 \qquad [54]$$

If the drop is not moving and the gravitational force is balanced by the surface tension, this pressure difference is hydrostatic and is given as follows:

$$p_1 - p_2 = 2\rho g r \sin\theta \qquad [55]$$

By combining eqn [54] and [55], a relationship can be derived for the maximum radius of a stationary drop in terms of the plate angle and the difference in interfacial angle:

$$r^2 = \frac{\sigma}{2\rho g \sin\theta}\left[\frac{1}{1-\tan\phi} - (1-\tan\phi)\right] \qquad [56]$$

For air and water $g = 9.81\ \mathrm{m\,s^{-2}}$, $\rho = 1000\ \mathrm{kg\,m^{-3}}$, and $\sigma = 0.07275$ Pa m.

A diagram of the maximum drop radius as a function of plate angle and difference in interfacial contact angle is given in **Figure 15** (from Cole *et al.*[81]). It is apparent that both the plate angle and the difference in interfacial angle are of prime importance in determining the critical size. In particular, it is notable that the stable drop size increases dramatically as ϕ approaches 45°.

The concept of single-drop movement has been embedded in a more complex model of drop movement, created using a fluid dynamics passage in which 71 million droplets were deposited as simulated rain

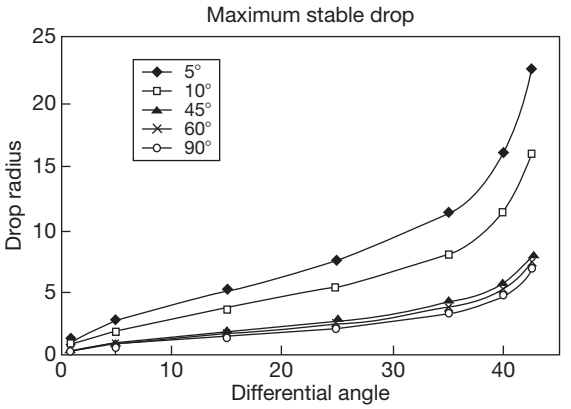


Figure 15 Variation of maximum stable drop size with ϕ, where 2ϕ is the difference in interfacial contact angle between the front and the back of the drop; at five plate orientations ($\theta = 5°$ to 90°). Reproduced from Cole, I. S.; Paterson, D. A. *Corros. Eng.. Sci. Technol.* **2007**, *42*(2), 91–104.

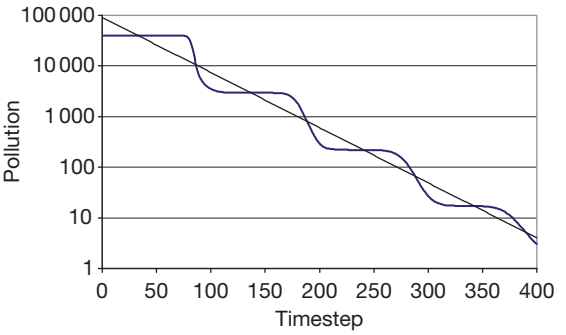


Figure 16 Pollution levels on the rain-washed part of the surface, together with a fitted exponential decay rate based on simulations of 71 million droplets depositing on a 2×2 m surface. Each timestep represents the deposition of 0.0157 mm of water. Reproduced from Cole, I. S.; Paterson, D. A. *Corros. Eng.. Sci. Technol.* **2007**, *42*(2), 91–104.

onto a 2×2 m surface and their subsequent movement down the surface was studied.[44]

During the studies,[81] drops on the surface grew by absorbing raindrops impacting on them and by merging with other drops. When drops rolled off the surface, the water they left behind was more than made up for by the collection of smaller drops on the way down. As each drop grew, it rolled faster and so it is reasonable to assume that once a drop reaches a critical size it will remove itself from a surface. Overall, it was observed that over time, the amount of water on the surface built up to a peak and then tended to release at once, giving periods of almost no runoff followed by periods of strong runoff. This is shown in the stepped shape of the pollution level curve in **Figure 16** (from Cole *et al.*[81]). In the experiments,[81] 0.0157 mm of water was deposited per timestep and therefore, the duration of a timestep depended on the simulated rain intensity. Loss of water through evaporation and substrate porosity was not included. The pollutant was assumed to have the solubility of sea salt.

The significance of **Figure 16** can be illustrated with regard to 'average' rainstorms in the Australian cities of Adelaide and Darwin, which are chosen as they span the rain intensity variation across mainland Australia, having 4.6 and 28.8 mm h^{-1} respectively as average rainfall intensities for a one-hour rain shower with a one-month recurrence interval. Therefore, the average shower timesteps in Adelaide and Darwin would be 13 and 2 s respectively. As shown in **Figure 16**, there is a factor of 10 decrease in soluble pollutant levels on the surface between timesteps 80 and 120. By timestep 200 (representing 43 min in Adelaide and 6.7 min in Darwin), the amount of pollution left was <1% of the initial value.[81] This only applies to soluble pollutants.

To the first order, the effect of rain on cleaning a fresh plate can be approximated by[81]

$$S_f = S_i \times e^{-\alpha R} \text{ if } R_i - R_c > 0 \qquad [57]$$

or

$$S_f = S_i \text{ if } R_i - R_c < 0 \qquad [58]$$

where S_f and S_i are the final and initial 'soluble' pollutant load, respectively, R_i is the amount of rain (in mm) in a particular rain shower or event, and R_c is the critical amount of rain required to guarantee runoff and thus cleaning. The value of R_c depends on drop size, surface parameters (e.g., surface energy), and plate slope, and can be estimated from an analysis such as that presented in **Figure 16**; for conditions relevant to the data set defined above, it can be approximated by $r/2$ (where r is the maximum radius of a stationary drop). Experimental studies[87] have verified this formulation and value of R_c.

2.16.4.2 Removal of Salts by Wind

While in theory the removal of salts by wind could be significant, in practice it is only significant in climates where rain is very limited or when the wind carries abrasive material. In wind tunnel tests, Muster and Cole[90] found that significant amounts of salt were removed only when wind speeds were ≥ 9 m s^{-1} and RH was $\leq 40\%$. These wind speeds represent <10% of the year in Australian cities. Work by Cole *et al.*[87] is in agreement with these results in the case where salts

were dry when deposited, but demonstrated that if salts were wet when deposited, wind speeds of up to $14\,m\,s^{-1}$ did not remove such salt particles. In the case where salt is deposited as a wet aerosol onto reactive surfaces (steel, zinc, aluminum–zinc), the aerosol reacts with the surface as it evaporates and forms a reaction product that binds the crystallized salt to the surface.

2.16.4.3 Transfer of Pollutants in Raindrops to a Surface

This section is concerned with defining the conditions where raindrops will evaporate on a metal surface, leaving precipitated pollutants. In tests where stainless steel, galvanized steel, and zinc plates were exposed to marine environments for periods of 2–24 h, Cole *et al.*[91] observed fine-scale crystals with compositions indicative of marine origin and with geometries consistent with deposition in raindrops. If raindrops are to evaporate, they must not be removed either by the motion of other drops (total rainfall must be less than R_c defined previously) or by their own motion. If the drop radius is less than that given by eqn [56], stationary drops will not begin to move; however, raindrops will maintain some of their preimpact velocity after contacting the metal surface. For such moving drops, the critical factor is whether they will evaporate before they roll off the metal surface. Work by Cole *et al.*[87] indicates that this is a function of the hydrophilic or hydrophobic nature of the surface and of course the distance they have to travel to roll off the plate. For strongly hydrophilic and strongly hydrophobic surfaces, the difference in interfacial angle (2ϕ) is small and so drops will tend to roll off in both cases. The presence of corrosion products on a surface can increase ϕ. For example, all but the finest drops (<10 µl) will run off an exposure plate (typical dimensions 150×150 mm at an angle of 45°), while quite large drops (up to 36 µl) may evaporate on an exposed roof (dimensions 3×3 m, 30° angle) on a summer day. This again indicates the difference between exposure plates and actual structures. The residue of raindrops observed by Cole *et al.*[91] might be associated with light rain and fine raindrops.

These studies, on the effectiveness of rain in cleaning surfaces and the nature of rain that can evaporate on a surface, highlight the fact that raindrop size and rainfall intensity are critical factors in determining the impact of rain on pollutant levels on metal surfaces. In general, locations closer to the equator will have more intense rain with larger raindrops, which will be more effective in cleaning surfaces, while locations at high or low latitudes will have less intense rain with smaller raindrops, which will be less effective in cleaning surfaces and may in fact evaporate on metal surfaces, thereby increasing the pollutant concentration on such surfaces.

2.16.5 Forms of Moisture on Surfaces

This section will discuss how moisture may form as a result of the atmospheric conditions previously discussed. Moisture may form on a surface because of the local RH, via the deposition of rain, fog, and wet aerosols, or because of wetting of hygroscopic salts on the surface.

2.16.5.1 Effect of RH on Clean Surfaces

A surface exposed to 'humid' air will adsorb moisture and build up monolayers on the surface. Immediately upon exposure, OH layers will develop on the surface.[14] With further exposure, additional water layers are adsorbed, with the number of layers being dependent on the hydrophobic or hydrophilic nature of the surface, the density of defects on the surface, and the RH of the air. Estimates of the number of monolayers vary from 1 to 2 layers at 40% RH, 2–5 at 60% RH, to 6–10 at 100% RH.[92] It is generally accepted that below 60% RH the monolayers are too thin to sustain the electrochemical corrosion process.

2.16.5.2 Effect of Aerosols/Raindrops

Aerosols show a significant variation in size range depending on the RH of the air and the history of the aerosol. Wet aerosols of marine origin[91] will have a size range of 5–120 µm and after impact on the surface they will reform into roughly spherical-cap shaped droplets (with the contact angle defined by the surface energy). In general, aerosol droplets falling onto a surface will only cover a fraction of the surface and will, therefore, remain isolated with significant areas of 'dry' metal separating the droplets. In contrast, when rain impacts a surface, a high fraction of the surface will become covered by drops, which will undergo coalescence until they reach the critical size for drop movement, or they may be scavenged by a moving drop.

2.16.5.3 Rewetting of Surfaces and Condensation

On a clean surface, condensation will occur when the air close to the surface is at 100% RH. The temperature of the surface may vary from that of the ambient air and, therefore, the RH close to the surface may be different from that of the ambient air.

The evaporation of aerosols or rain on surfaces may leave crystallized salts, many of which may be hygrosopic. These salts will wet when the surface RH is greater than the DRH of the contaminant salts. The DRH of a number of common pollutants is given in **Table 4**. A study by Cole *et al.*[93] compared the nature of surface contaminants with the wetting behavior of wetness sensors attached to metal plates and found that there was a high correlation between the DRH of the salts present and the wetting point of the sensor. Further, the nature of the salts present (and hence the wetting point of the sensor) was highly dependent on location, with severe marine sites contaminated with $MgCl_2$ and NaCl, marine sites contaminated with NaCl, severe industrial sites contaminated with $(NH_4)HSO_4$, and moderate industrial sites with a marine influence contaminated with $(NH_4)_3H(SO_4)_2$, $(NH_4)_2SO_4$, $NaHSO_4$, and NH_4NO_3.

The wetting of such crystalline salts will occur when

$$RH_s > \text{DRH of any contaminating salt} \quad [59]$$

where RH_s is the RH at the metal surface.

An alternative but equivalent expression can be expressed in terms of time of condensation (TCD), tendency to condensation (TCON), and vapor pressure, so that TCD is defined as

$$\text{TCD} = \text{time when TCON} > 0 \quad [60]$$

TCON is simply the difference between the vapor concentration of the air (V_a) and the saturation vapor concentration at the metal surface (V_s) required for condensation (vapor pressure is in Pascals):

$$\text{TCON} = V_a - V_s \quad [61]$$

and

$$V_s = (RH_c/100) \exp\left(22.565 - 2377.1\,T_s^{-1} - 33623\,T_s^{-1.5}\right) \quad [62]$$

where T_s is the surface temperature (K) and RH_c accounts for the hygroscopicity of contaminating salts, which can be approximated by the DRH of any contaminating salt:

$$V_a = V_{as} \times \text{RH}/100 \quad [63]$$

where V_{as} is the saturation vapor pressure of the adjacent air, which is given by the following equation:

$$V_{as} = \exp\left(22.565 - 2377.1\,T_a^{-1} - 33623\,T_a^{-1.5}\right) \quad [64]$$

where T_a is the air temperature.

2.16.5.4 Evaporation

Evaporation of moisture from a surface covered with hygroscopic salts can be understood using either the surface RH concept or partial vapor pressure concept. Thus, wetted salts will begin to dry when the surface RH is less than the DRH of the contaminating salt or when the TCON is negative. However, to determine either of these parameters it is necessary to know the temperature of the metal surface.

2.16.5.4.1 Surface temperature of plates

Metal surfaces exposed to the environment will undergo a diurnal cycle, heating up during the day (due to solar heating) and cooling down at night (because of the emission of radiation to the night sky), with both processes modulated by convective interaction with air. The extent of both heating and cooling will depend on a range of environmental factors, including cloud cover, wind speed, and the state of the surface.[94,95]

Condensation and evaporation will tend to reduce this diurnal cycle – condensation will release latent heat restricting further undercooling, while evaporation of moisture films after sunrise will cool surfaces, reducing solar heating. As a result, it may take significant time for metal surfaces to dry after sunrise. Predicted undercooling of a zinc plate openly exposed to the night sky is presented in **Figure 17** as a function of temperature. It is apparent that undercooling does not exceed 2.5 °C (still air, clear sky, 50% RH). The RH affects the extent of condensation and the latent heat released during condensation, so that on a clear night with 75% RH, a plate would undercool less than on a clear night with 50% RH. A breeze of 1 m s^{-1} significantly reduces undercooling, as does cloud cover.

The rate at which zinc plates heat up after sunrise is presented (for dry plates) in **Figure 18** as a function of incident solar radiation which in turn is a function of the time of day. Thermal equilibrium is assumed. **Figure 18** can be matched with the solar radiation level for a given location. For example, measured data for Melbourne (latitude 38° S), Australia, indicates that on a clear day close to the September solar equinox (spring), the solar radiation

levels will be 171, 376, and 696 W m^{-2} at 1, 2, and 4 h respectively after sunrise, while on a cloudy day they will be 28, 68, and 222 W m^{-2} at 1, 2, and 4 h respectively after sunrise. From Figure 18, it is clear that, as a function of solar radiation, the increase in temperature on a cloudy and sunny day is the same. Of course, much less radiation reaches the surfaces on a cloudy day, so the temperature rise on a plate is much lower. Wind dramatically decreases the temperature rise because of forced thermal convection, with a wind speed of 5 m s^{-1} more than halving the rise. Figure 19 (from Cole *et al.*[96]) compares the temperature rise of a dry plate with that of a wet plate. The cooling effect of evaporation dramatically reduces surface heating. This reduction is closely tied to the ambient conditions (temperature and RH) that control evaporation.

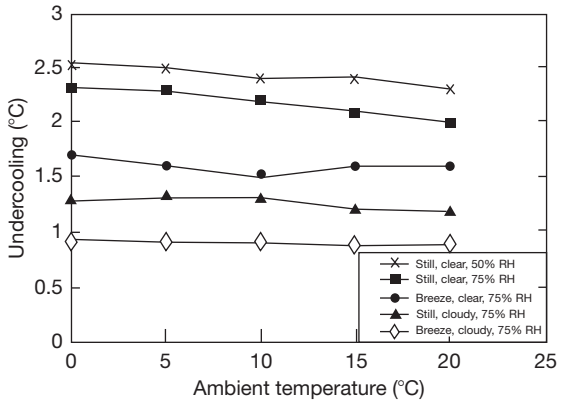


Figure 17 Predicted undercooling of a metal plate at different RHs. The breeze is 1 m s^{-1}. Reproduced from Cole, I. S.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2006**, *41*(1), 67–76.

2.16.5.4.2 Evaporation rate

Figure 20 (from Cole *et al.*[96]) shows the evaporation rates on a metal surface (assuming a continuous film) for a variety of environmental conditions as a function of time after sunrise. The rate of evaporation is clearly influenced by surface temperature, as discussed above. With clear skies, a modest wind (5 m s^{-1}) and not excessive humidity (75%), a moisture film of 200 μm will have evaporated an hour after sunrise. Evaporation actually starts before sunrise as the minimum temperatures (under these conditions) generally occur an hour before sunrise. In contrast, on a cloudy, still day of high humidity (95%), the same moisture film will still not have fully evaporated 4 h after sunrise.

A more precise understanding of the relevance of these effects to corrosion can be obtained if we calculate the time for evaporation of common aerosols. Using the data of Fitzgerald,[35] it can be estimated that a wetted aerosol on the metal surface will have a size range of 2–120 μm with a median of 12 μm and that a nonequilibrium aerosol (with a solids concentration equal to that of seawater) will have a size range of 5–230 μm with a median of 23 μm. The time for aerosols with median and extreme dimensions to evaporate is given in Table 12 (from Cole *et al.*[96]). It is apparent that, at least for some aerosols,

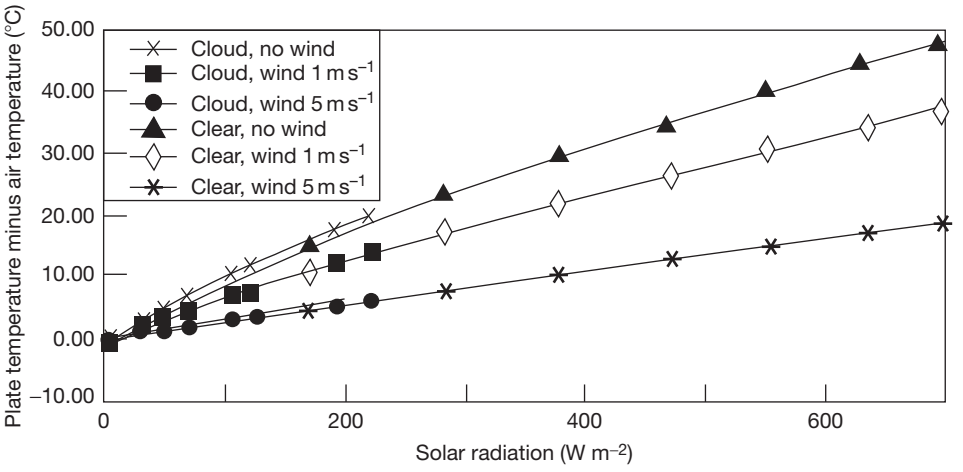


Figure 18 Predicted heating rate as a function of solar radiation for dry plates. Thermal equilibrium is assumed and the solar radiation is a function of time of day. Reproduced from Cole, I. S.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2006**, *41*(1), 67–76.

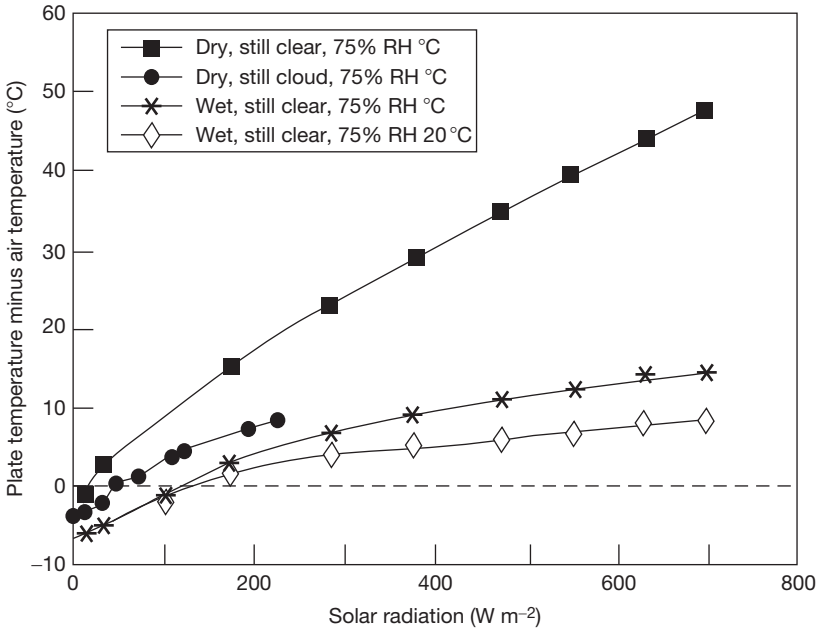


Figure 19 Comparison of temperature rise on a dry plate with that of a wet plate. Reproduced from Cole, I. S.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2006**, *41*(1), 67–76.

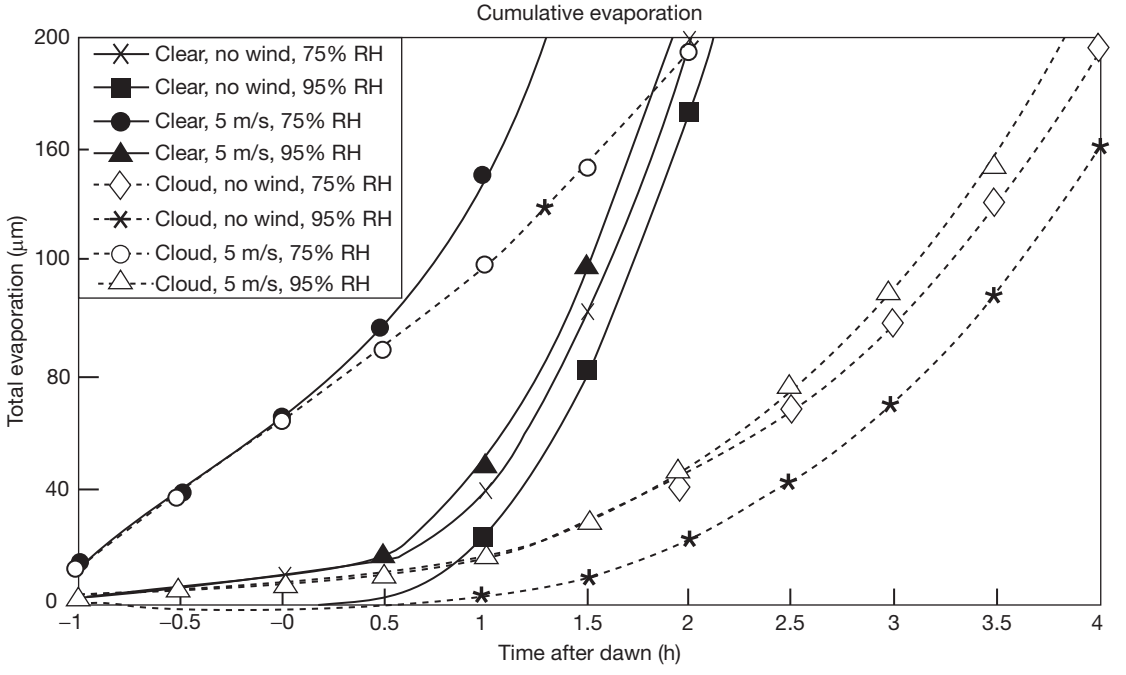


Figure 20 Evaporation rates on a metal surface assuming a continuous film of water. The assumed solar radiation 4 h after dawn is 696 W m^{-2} on a clear day and 222 W m^{-2} on a cloudy day. Reproduced from Cole, I. S.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2006**, *41*(1), 67–76.

Table 12 Time to dry (hours) for aerosols of given surface radii

Conditions	*Surface radius (μm)*			
	12	*23*	*120*	*230*
Cloud, no wind, 75% RH	Just after dawn	1	3	4
Clear, no wind, 75% RH	Just after dawn	0.6	1,5	2
Cloud, 5 m s^{-1}, 95% RH	Just after dawn	1.1	2.8	3.7
Cloud, 5 m s^{-1}, 75% RH	Prior to dawn	Prior to dawn	0.6	1.7

Source: Reproduced from Cole, I. S.; Paterson, D. A. *Corros. Eng. Sci. Technol.* 2006, *41*(1), 67–76.

evaporation will take several hours. Many workers suggest that a peak in corrosion rate occurs during the evaporation phase, as salt concentrates and oxygen diffusion paths shorten.

Cole *et al.*[65] have documented experimentally the duration of periods of wetness that occur on wetness sensors attached to zinc plates that are openly exposed to the atmosphere in a range of locations around Australia.[97] They found that periods of wetness could be divided in three modes which they labeled Mode A, Mode B and Mode C. Mode A are short daytime wetness events with an average duration of ~2 h (they result from either hygroscopic wetting of deposited salts or rainfall but are of short duration as daytime evaporation is rapid). Mode B are nighttime wetness events that occur when condensation occurs in the evening and evaporation the next morning. The duration of the wetness period varies depending on the severity of the location (and thus the nature of deposited salts) from an average of 10 h in benign (inland) sites to 20 h in severe marine locations. Mode C is wetness periods that last more than 1 day; they are generally associated with the deposition of salts with low DRH (e.g., MgCl) so that surface RH may not dry out the wetted salts on some days. Again, the duration depends on the severity of the site raging from an average of 72–120 h in marine and severe marine sites.

2.16.5.4.3 Evaporation and condensation on a porous surface

The above analysis refers to clean surfaces; however, on a corroded surface the situation may be more complex. First, many oxides are themselves hygroscopic, and second many oxides are porous and, therefore, when aerosols or raindrops are deposited onto such porous oxides, some of the solution will be absorbed into the pores and will subsequently evaporate, leaving contaminants within the pore structure. In these circumstances, capillary forces may alter the conditions for evaporation and condensation in the pores, making it possible for condensation to occur in an atmosphere of less than 100% RH. These effects are best explored using the differential vapor pressure approach, where the relative lowering of the saturated vapor pressure in pores is given by Thomson's equation.[98]

Table 13 Effect of pore size on tendency to condensate formation

Pore size (Å)	*RH for condensation (%)*	*%TCD (100)*	*%TCD (75)*
Flat surface	100	4.6	43.5
360	98	7.5	45.3
94	90	19.8	52.1

% TCD (100) is the % time the tendency to condensation (TCON) is positive when the surface wetting occurs at 100% RH while % TCD (75) is the % time the TCON is positive when the surface wetting occurs at 75% RH.

$$p = p_o \, e^{-2\sigma M/\rho RTr} \qquad [65]$$

where p and p_o are the saturated vapor pressures above a concave meniscus of radius r and a plane surface, respectively, σ is the surface tension of the liquid at absolute temperature T, ρ is its density and M is its molecular weight of the condensing gas, and R is the gas constant. It can be estimated that capillary radii of 360, 94, and 30 Å will reduce the RH for condensation to 98, 90, and 70% respectively. It has been argued[99] that this effect may explain the observation that steel still corrodes at ~70% RH, as ferric oxide gel has pore sizes ~30 Å,[100] and therefore, pore condensation may occur. Correspondingly, the period that moisture remains in pores may be considerably greater than that when it is free on the surface. The percent time that the tendency for condensation both on a clean surface and on a contaminated surface is greater than zero is given in **Table 13**, and compared to that for a clean pore and one contaminated with NaCl when both surfaces are exposed for a year in Canberra. It is apparent that even coarse pores significantly increase

the time when condensation is favored, both with and without chloride contamination, although the effect is more marked in the absence of chloride. Pore condensation and delayed evaporation may be relevant for metals with porous oxide, such as zinc, and could significantly change the wetness periods for these materials.

2.16.6 Summary of Environmental Conditions

The analysis in Section 2.16.3.1 indicated that pollutant deposition onto a surface may occur by gaseous absorption into a moisture film, or by deposition of aerosol or of rain. As outlined in Section 2.16.1.3, the composition of gases, aerosol, and rain drops will be affected as they are transported by atmospheric processes. However, the scale of these transport processes is likely to be quite different. Gases, particularly reactive gases, will not be transported far from their sources; aerosols may form relatively close to the pollutant source but may be transported significant distances, while the processes of rain drop formation are complex and will not bear a direct relation to pollutant sources. Thus, it is to be expected that the concentration of atmospheric gases will be highly influenced by local pollutant sources, and aerosol concentration will also be influenced by local sources but concentrations will spread significantly from these sources, while rain water chemistry will be affected more by the pollutants in a region rather than in one locality.

In Table 14 some key parameters defining gaseous concentrations, aerosols, and rain water chemistry are given for marine, industrial, urban, rural, and remote locations from data presented in this chapter. The table should only be taken as indicative and does not attempt to define the effect of particular local sources on pollutant levels.

2.16.7 Surface Reactions

2.16.7.1 Equilibrium Considerations

The conditions within the rain droplet or aerosol may determine the stability of both common metal oxides and the underlying metal. Consider zinc, which in dry environments is covered with a zincite (ZnO) film and which rapidly transforms to zinc hydroxide in the presence of moisture films. The potential–pH diagram of zinc–CO_2–H_2O system[101] (at 25 °C) with Graedels[101] regimes for dew, rain, and fog and the stability lines for H_2O, H_2O_2, and O_3 is presented in Figure 21 (modified with the original from Graedel[101]). At the typical rain regime or fog regime

Table 14 Summary of gaseous concentration, aerosol type and rain water composition ($\mu eq\, l^{-1}$) for marine, industrial, urban, rural, and remote locations

Location type	*Marine*	*Industrial*	*Urban*	*Rural*	*Remote*
Gaseous (mixing ratio)					
SO_2 (ppt)	260	1500	160–1500	160	20
H_2S (ppt)	65	365	365	35–60	3.6–7.5
NO_x (ppb)	0.2–1000	10–1000	10–1000	0.2–10	0.02–0.08
O_3 (ppb)	20–40	100–400	100–400	20–40	20–40
Aerosols					
pH	0–9.5	−1–2.4	1.9–3		
Major species	NaCl, MgCl	H_2SO4, $NH_4)_3H(SO_4)_2)$, $(NH4)_2SO_4$, NH_4HSO_4	NH_4NO_3, NH_4Cl $(NH4)_2SO_4$		Dust, pollen plant waxes
Secondary species	Na_2SO_4, H_2SO_4, NH_4HSO_4, $(NH_4)_2SO_4$ and $NH_4NO_3.NaHSO_4$		$NaNO_3$		
Rain water					
pH	4–5.6	4.2–7.3	4.4–6.1	3.6–5.8	5.6–6
Cl^-	100–1300	9–142	10–27	3–25	1–1300
NO_3^-	3–10	40–140	13–140	3–25	3–10
SO_4^{2-}	3–10	70–240	12–60	4–60	3–10
Na^+	100–1200	5–60	20–60	3–30	100–1200
NH_4^+	2–10	30–200	10–30	2–20	2–10
Ca^{2+}		20–300	2–35	2–20	80

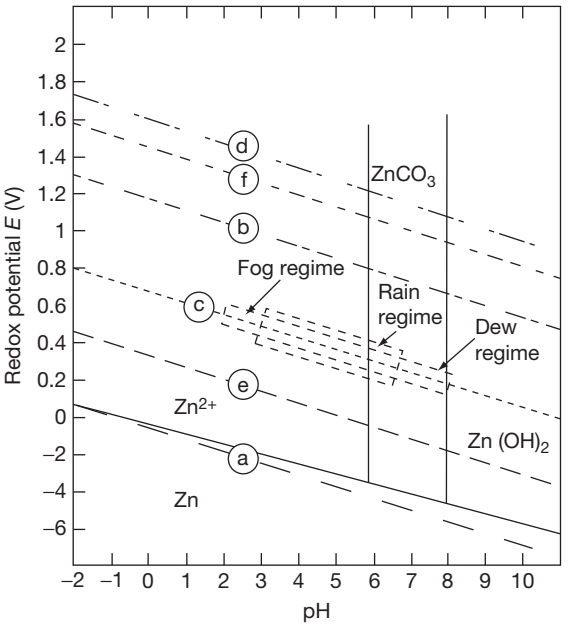


Figure 21 Potential–pH diagram for the system Zn–CO_2–H_2O at 25 °C for a concentration of zinc ions of 0.1 M and a concentration of H_2CO_3 of 1 *10^{-5} (this is the concentration in equilibrium with gaseous concentration of CO_2 of 340 ppmv). Modified from Graedel, T. E. *J. Electrochem. Soc.* **1989**, *136*, 19C–2003C.

the surface $Zn(OH)_2$ film and the underlying Zn metal are not the stable phases, and rather Zn^{2+} is the stable form. Therefore, dissolution of the zinc hydroxide is expected with the enrichment of the zinc ion concentration in solution. Stable oxide growth would not be expected. This enrichment of zinc ions in solution would then lead to precipitation of solid species.

A more complex situation occurs in the corrosion of iron. In **Figure 22**, a potential–pH diagram[69] is given for the Fe–H_2O system (at 25 °C) with the addition of the stability lines for H_2O, H_2O_2, and O_3. The diagram indicates that in many of the conditions commonly observed in dew rain and fog, the stable phase is either FeOOH or Fe^{2+} at low pH values. These low pH values could occur either in fog or during the evaporation phase of dew or rain. However although the initial corrosion product may be Fe^{2+}, the intermediate corrosion products are mixed Fe(II) and Fe(III) oxides and the final products contain only Fe(III).[102] The development of these oxides can in part be attributed to the reactive oxidants, hydrogen peroxide (H_2O_2), the OH, and the hydroperoxyl radical (HO_2). It will be recalled that the work of Graedel and Goldberg[103] demonstrated that small but significant concentrations of these oxidants could develop in rain droplets. The free radical and the peroxide reaction are likely to be[103]

$$Fe^{2+} + OH. \leftrightarrow Fe^{3+} + OH^- \quad [66]$$

$$Fe^{2+} + HO_2. + H_2O \leftrightarrow Fe^{3+} + H_2O_2 + OH^- \quad [67]$$

and

$$Fe^{2+} + H_2O_2 \leftrightarrow Fe^{3+} + OH. + OH^- \quad [68]$$

If solutions contain oxidants, the redox potential of the solutions will be influenced by the reduction/oxidation of these oxidants which may influence the stability of oxides and the passivation of metal surfaces. For example, according to Pourbaix[102] in a dilute solution of hydrogen peroxide the iron may be in the domain of the corrosion of iron (formation of Fe^{2+}) and the reduction of hydrogen peroxide (point A on **Figure 22** at $E = -0.200$ V and pH = 5.7), so that the following reactions could occur:

$$Fe \leftrightarrow Fe^{2+} + 2e^- \quad [69]$$

$$H_2O_2 + 2H^+ + 2e^- \leftrightarrow 2H_2O \quad [70]$$

The overall reaction will be

$$Fe + H_2O_2 + 2H^+ \leftrightarrow Fe^{2+} + 2H_2O \quad [71]$$

In contrast with a concentrated solution of hydrogen peroxide, iron would be in a passive domain with stable oxides (point B on **Figure 22**) and hydrogen peroxide is in its domain of double instability. In this region (below line c and above line d in **Figure 22**), hydrogen peroxide may decompose into both water and oxygen according to the following equations:

$$H_2O_2 + 2H^+ + 2e^- \leftrightarrow 2H_2O \quad [72]$$

$$H_2O_2 \leftrightarrow O_2 + H^+ + 2e^- \quad [73]$$

The overall reaction will be

$$2H_2O_2 \leftrightarrow 2H_2O + O_2 \quad [74]$$

Thus, the metal surface may facilitate the decomposition of hydrogen peroxide and the evolution of gaseous oxygen, but the iron will remain in the passive state.

However, hydrogen peroxide can act to oxidize sulfite to sulfates as outlined in eqn [19]. This will decrease the pH and the H_2O_2 concentration so that the critical parameter in determining if the peroxide concentration is such that Fe is in a passive or active region of the potential–pH space is the relative concentrations of sulfite and peroxide. Thus, if $[H_2O_2] > [HSO_3^-]$, H_2O_2 will not be completely

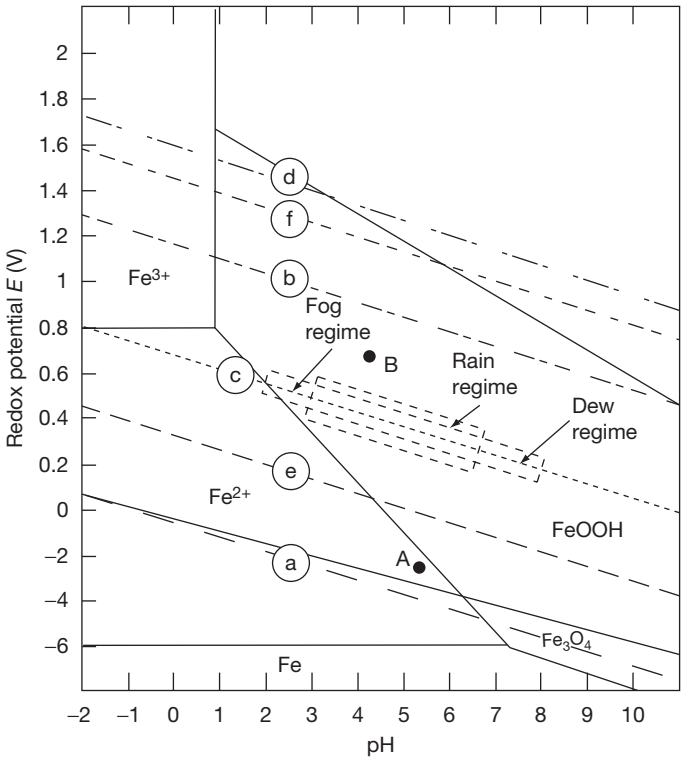


Figure 22 Potential–pH diagram for the system Fe–H_2O at 25 °C. Derived from Pourbaix, M. *Atlas of Electrochemical Equilibria in Aqueous Solutions*, 2nd English Edition; National Association of Corrosion Engineers: 1440 South Creek Drive, Houston Texas 7704, 1974; the approximate regimes for fog, rain, and dew are derived from Graedel, T. E.; Frankenthal, R. P. *J. Electrochem. Soc.* **1990**, *137*, 2385–2394.

depleted and hence may be available to passivate the metal surface.[69]

In contrast with the marked changes with the stability of the Fe phases because of the action of the oxidants, they have little effect on the equilibrium condition for the zinc system. On **Figure 21**, the potentials for reduction of ozone or hydrogen peroxide (at pH values likely to occur in rain water or aerosols) fall in the stability range for Zn^{2+}. Hence, the presence of these oxidants will further promote the corrosion of zinc.

Although the *E*–pH diagrams are useful, they only define the equilibrium conditions for relatively simple systems. In reality the kinetics of the corrosion processes are critical in determining the corrosion rate and the corrosion product while both rain drops and aerosols are multicomponent systems involving a wide range of aqueous species (chlorides, sulfates, carbonates, oxidants) and continuous gaseous absorption from the environment. Part of the dynamic nature of this system comes from the electrochemical processes themselves which can induce separation of anodic and cathodic areas, differential pH at these zones and ion migration driven by local potentials.

2.16.7.2 Influence of Electrochemical Processes on Moisture Film Stability and Chemistry

The establishment of separated anodic and cathodic areas may have significant impact on the distribution of species within a droplet. Two examples will illustrate this effect: the first, secondary spreading on zinc is a relatively long range effect, while the second, the development of anodic and cathodic zones associated with particular intermetallic zones containing aluminum 2000 and 7000 series is more local. Neufeld *et al.*[104] studied the effect of wetting fine NaCl crystals on zinc surfaces. It was found that in high humidities (>75% RH) deposited salt crystals rapidly wet forming roughly hemispherical moisture droplets. After a relatively short period, a fine moisture film spreads out from a central droplet. Scanning Kelvin

probe studies indicated an electrochemical cell developed, with the central drop becoming anodic and the secondary spread region cathodic. High pH developed in the secondary spread region and slightly acidic pH values in the central droplet. Neufeld *et al.*[104] demonstrated that significant ion migration occurred with cations (particularly Na^+) migrating into the secondary spread region and anions into the central anodic zone. Correspondingly, after the drop had dried out at the completion of the experiment, quite different corrosion products were found in the two zones with the main product in the secondary spread zone being zinc hydroxycarbonate and the main products in the inner anodic region being zinc hydroxychloride (simonkolleite) and sodium zinc hydroxysulfate chloride (gordiate).[91]

Secondary spreading is significant for a number of reasons. Primarily, the establishment of a very large cathodic zone removes any possibility of cathodic limitations for the corrosion process and thus may increase the initial corrosion processes. However, the high pH in the secondary spread area provides ideal conditions for the absorption of carbon dioxide (via eqns [15]–[17]). At pH 9 the carbonate ion concentration in solution is 10^6 times that at a pH of 6. High carbonate concentrations will lead to the precipitation of zinc hydroxy carbonate. Of course, the absorption of carbon dioxide will also tend to neutralize the alkaline pH. The development of significant coverage of the zinc surface by zinc-hydroxy carbonate will limit corrosion in subsequent wetting cycles. Muster *et al.*[105] have demonstrated that the carbonate blocked chloride penetration to anodic sites on the metal surface to a much greater extent than zinc hydroxycarbonate.

Such secondary spreading has also been observed in steel[106] and copper.[107] However, the phenomenon was not observed on aluminum when exposed to the identical condition.[108] The work on copper also shows interesting synergies between secondary spreading and the absorption of gaseous species in an analogous manner to that observed for zinc. Chen *et al.*[108] wetted a NaCl crystal on a copper plate in an atmosphere containing 150 ppb SO_2 and two different levels of CO_2 (<5 ppb and ambient 350 ppb). They observed that a narrow band of secondary spreading occurred with an alkaline pH occurring in the spread area, with sulfate and dithonate ($S_2O_6^{2-}$) species observed in the secondary spread zone, and the corrosion products nantokike (CuCl) and pantacamite ($Cu_2(OH)_3Cl$) observed in the central droplet region. It is known that Cu^{2+} may catalyze the oxidization of S(IV) to S(VI) in an analogous manner to that of the other transition metal ions Mn^{2+} or Fe^{2+} (refer to eqn [16]). Thus, the absorption of SO_2 and its dissociation via eqns [11] and [12] are promoted by the high pH in the secondary spread zone. Cu^{2+} catalyzed oxidation of the sulfite may occur by one of the following reactions[76]:

$$2Cu^{2+} + SO_3^{2-} + H_2O \leftrightarrow 2Cu^{+} + SO_4^{2-} + 2H^{+} \quad [75]$$

or

$$2Cu^{2+} + SO_3^{2-} \leftrightarrow 2Cu^{+} + S_2O_6^{2-} \quad [76]$$

The absorption of SO_2 and the formation of sulfate will lower the pH in the secondary spread area and secondary spreading will cease. The separation of the droplet into cathodic zones and anodic zones in the central area will lead to a higher level of $[H^+]$ in the anodic zone which will encourage copper dissolution and the precipitation of nanotokite and pantacamite (rather than NaCl).[108]

Thus, secondary spreading can lead to significant variation in pH and species concentration (both by variations in gaseous absorption and by migration), and so it can lead to the formation of different oxides with the consequence of further corrosion.

Dynamic interactions also occur involving solution chemistry, local pH, corrosion product formation, and electrochemistry at a finer scale. The trenching observed around cathodic intermetallics in aluminum AA2xxx and AA7xxx series alloys can illustrate these effects (this chapter will not attempt to cover the interesting and involved interactions of intermetallics and electrochemistry of these alloys will be discussed elsewhere[109]). To achieve the required strength, the formation of intermetallics is required in these structural aluminum grades. Some of these intermetallics – notably the Al–Cu–Fe–Mn group,[110,111] including ($Al_6(CuFeMn)$, Al_2Cu, Al_7Cu_2Fe, and $Al_{20}Cu_2(FeMn)_3$) – have a very strong ability to sustain a cathodic current. Birbilis and Bucheit[110] found that at the corrosion potential of AA7075-T651 in NaCl (i.e., $-965\,mV_{sce}$) Al_7Cu_2Fe and Al_2Cu could support cathodic currents in excess of 310 and 470 $\mu A\,cm^{-2}$, respectively while Schneider *et al.*[111] showed the cathodic current on $Al_{20}Cu_2(FeMn)_3$ to be ten times greater than that measured on AA2024-T3 alloy surfaces. The high cathodic current sustained by these intermetallics can lead to localized pitting corrosion and to trenching around Al–CuFe–Mn phases.[112–116] The exact mechanism of trenching is still subject to debate with some researchers[112] suggesting that it may be interpreted as galvanic corrosion between the particle and the matrix, and others[113,114] indicating that it is a form of

etching under the high pH conditions occurring at the cathodic intermetallics. In fact Park *et al.*[113] have found pH values ~9.5 at around these particles. Under these high pH conditions, it is suggested[114] that the aluminum matrix undergoes cathodic corrosion, together with dissolution of the oxide film producing aluminate ions and hydrogen gas according to the equation

$$Al + 3H_2O + OH^- \rightarrow Al(OH)_4^- + 3/2\,H_2 \quad [77]$$

The rate of eqn [77] is highly pH dependent. In fact, Park *et al.*[113] and Vukmirovic *et al.*[115] have found that increased buffering of the solution decreases the rate of pitting around the intermetallic particles which they attribute to the inability to sustain high localized pH values in a buffered solution. As highlighted previously in atmospheric conditions, highly buffered solutions can arise from various acid/salt couples such as the bisulfite/sulfite reactions (eqn [12]). The formation of colloidal alumina gels have been observed[113,116] between anodic and cathodic regions on the surface and these may help maintain both potential and pH differences between these regions. In fact Ilevbare *et al.*[116] have reported that as high pH develops at cathodic sites and low pH at anodic sites, increased galvanic potential differences are created.[116] Thus, under atmospheric conditions, an interplay between local electrochemical activity and local pH may arise, with corrosion product or gel formation controlled by the general chemistry of the droplet solution.

2.16.7.2.1 Oxide precipitation and stability

The fact that a surface is corroding implies that there is metal ion formation, as well as diffusion or migration through the oxide layer, and in many cases, these processes are highly influenced by oxide stability, dissolution, and growth. Oxide stability in its turn is affected by the local pH and ion concentration, which as discussed above, are influenced not only by the original pH of the deposited moisture and subsequent aqueous reaction but also by the establishment of local anodes and cathodes. These interactions will be discussed in relation to the aluminum and zinc systems.

The typical form of an oxide film on aluminum exposed to water or water vapor consists of a thin layer of γ-Al_2O_3 covered with a thin layer of boehmite (γ-AlOOH) which in turn is covered with bayerite ($Al(OH)_3$ which can be written as $Al_2O_3.3H_2O$.[117] The stability data of boehmite and bayerite, along with that of CuO and hydrated copper oxide ($Cu(OH)_2$) are presented in **Figure 22** from the data of Pourbaix.[102] It is evident that very low pH values are required for aluminum oxides to show significant solubility (less than 2.4 for boehmite). In fact only in acidic fogs and very acidic aerosols (for example, referring to **Figure 1**, the aerosols in the NH_3–SO_4 system, where the ammonia/sulfur dioxide ratio is less than 0.5 so that H_2SO_4 is the dominant and NH_4HSO_4 the secondary species) will boehmite dissolve directly. However, these low pH values may occur in the evaporation phase of rain drops or condensate (i.e., dew) or if localized pH changes occur in association with pitting corrosion or oxygen differentials. However, at a pH of 4 some dissolution of the oxide may occur, particularly in association with defects or cracks in the oxide film. Thus when acid fogs, acidified rain, or acid aerosols (either industrial or marine – but affected by absorption of acids) fall on an aluminum surface some dissolution of the oxide layer is expected, with the electrolyte solution and the underlying metal substrate being in contact through cracks and pores in the surface. Aluminum ions may diffuse through the pores and in sulfate rich environments may form amorphous aluminum sulfate hydrate – $Al_x(SO_4)_y(H_2O)_2$, which will gradually transform into aluminate, $Al_2(SO_4)\ (OH)_4\ 7(H_2O)_2$. The pitting may commence beneath defects in the oxide and the progression of the pitting (and further dissolution of the oxide) may be supported by the autocatalytic nature of the pitting process, when anodic regions occur at the bottom of the pit and because of the strong oxygen differential between the bottom and the top or exterior of the pit, cathodic zones also occur in these regions. The formation of gel structures by aluminum sulfate hydrate or other amorphous materials will tend to occur at the mouth of the pit (where the pH is higher) stabilizing the pH and oxygen differentials. In the chloride rich solutions generated by marine aerosols, chloride ions may complex with Al^{3+} ions at anodic sites and, thereby, prevent the formation of aluminum hydroxides. Further, aluminum hydroxide chlorides may form either by chlorination of the surface $Al(OH)_3$ or in solution. Further reactions with chloride ions can lead to the formation of Cadwaladerite $AlCl(OH)_2.4H_2O$ and then aluminum chloride $AlCl_3$.

$$Al(OH)_3 + Cl^- \leftrightarrow Al(OH)_2Cl + OH^- \quad [78]$$
$$Al(OH)_2Cl + Cl^- \leftrightarrow Al(OH)Cl_2 + OH^- \quad [79]$$
$$Al(OH)Cl_2 + Cl^- \leftrightarrow AlCl_3 + OH^- \quad [80]$$

Clearly, this reaction is favored in acidic and chloride

rich environments. The dissolution of the boehmite layer will then permit further access of the solution to the substrate and thus enhanced corrosion.

The situation above is very similar for structural alloys with a large range of intermetallics. First, the oxide layers above the intermetallics tend to be thinner with more defects, so that the electrolyte can more easily move through to the substrate. Second, as discussed in a previous section, certain intermetallics may establish large cathodic currents which can reinforce a strong pH difference between the cathodic intermetallic and the anodic matrix, increasing the corrosion rate and promoting more aggressive pitting in the matrix. These potential and pH differences can be supported by the formation of gels as highlighted above.

A second aspect is illustrated by the role of Cu in the corrosion of aluminum alloys. Copper is a more noble metal than aluminum and, therefore, during the corrosion processes will accumulate at the metal oxide boundary; it then migrates through the oxide to the surface where it may form clusters and act as cathodic sites. Copper that has entered into solution or any original copper in solution may be subsequently reduced back to copper metal on cathodic copper containing intermetallics. In the presence of Cl^- ions the formation of $CuCl^+$ will lower the energy barrier for the deposition of copper onto aluminum that leads to secondary pitting. Thus, copper plating and secondary pitting may be promoted in solutions with moderate pH and high chloride content such as acidified marine aerosols, **Figure 23**.

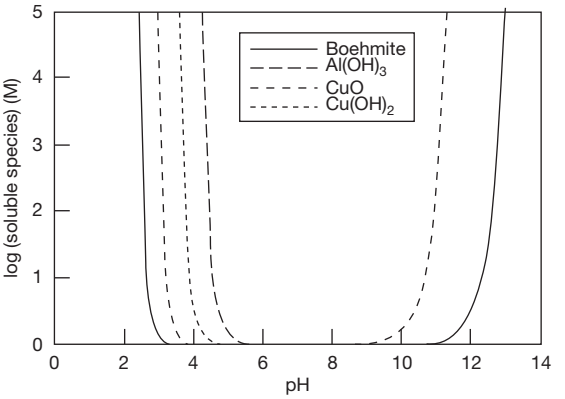


Figure 23 Stability of boehmite, bayerite and CuO and $Cu(OH)_2$. Reproduced from Pourbaix, M. *Atlas of Electrochemical Equilibria in Aqueous Solutions*, 2nd English Edition; National Association of Corrosion Engineers: 1440 South Creek Drive, Houston, Texas 7704, 1974.

In the zinc systems, it is meaningful to discuss corrosion that occurs as a result of secondary spreading and that occurs when conditions are not suitable for secondary spreading. If secondary spreading occurs, there is a wide separation of anodic and cathodic zones. In the cathodic regions, alkaline pH values develop and zinc hydroxycarbonate is observed to form both in the laboratory conditions and in the field. **Figure 24** from Graedel[118] gives the zinc carbonate stability diagram with the regime for marine aerosol added. It is evident that under typical conditions of rain, dew, or marine aerosol, zinc carbonate would not be expected to form. Therefore, the high pH and carbonate levels (up to 10^3–10^4 μM) in secondary spreading zones favor the formation of the carbonate whereas it would not form in a 'stable' drop.

In the central anodic zone of the droplet, slightly acidic pH values develop and common corrosion products are gordiate, zinc hydroxy chloride, and zinc hydroxy sulfate. Stability diagrams from Graedel indicated that both the hydroxy sulfate and hydroxy chloride have relatively limited stability regions that require sulfate and chloride concentrations in excess of 10^2 and 10^3 μeq l^{-1} respectively and zinc ion levels greater than 0.1 M; both these are centered around a pH of 6.2. However, as the anion concentration increases, the pH stability range also increases so that at concentrations of 10^5 μeq l^{-1} the hydroxy anions are stable from a pH of just below 6 to nearly 8. On metal surfaces exposed to a marine environment, the concentration of chloride and sulfate in deposited aerosols will range from 0.55 to 5.4 M, that is, from the concentration in seawater (in a freshly formed aerosol that has not yet equilibrated to the atmospheric conditions) to that at saturation point, when the RH is just above the deliquescent RH of the constituents of seawater. Thus, in the solutions formed by the deposition of wetted aerosols, although there will be sufficient chloride to promote the formation of zinc hydroxide chloride, there will not initially be sufficient zinc ions in solution. This may allow the formation of mixed cation OH anions and in fact Cole *et al.*[91] have observed the formation of gordiate after short periods of exposure. Thus, within the central zone of an aerosol droplet, subject to secondary spreading, gordiate and zinc hydroxychloride will form. As the droplet evaporates, the concentration of both anions and actions will increase allowing the formation of a wider range of salts such as zinc hydroxysulfate. In droplets that do not undergo secondary spreading similar processes will

occur in the central droplet; however, Cole *et al.*[91] found no evidence of the formation of zinc hydroxycarbonates. Traditional theories to explain how zinc hydroxycarbonate may form in marine chloride rich environments have suggested that zinc hydroxychloride may transform to the hydroxycarbonate with time; however, because of the high pH and carbonate concentration that occur in secondary spreading, the zinc hydroxycarbonate phase may form directly.

However, the zinc OH anions and particularly the hydroxychlorides are not stable under the common pH and chloride concentrations that occur in dew, rain, or fog (see **Figure 24**). Thus, it is quite possible that zinc hydroxychloride may dissolve under these types of moisture. In **Figure 25**, the concentrations of chloride and zinc ions that occur when zinc hydroxychloride powder (mass of 1.78 g) is placed in 1 l of a dilute solution of nitric acid with a pH of 4.5 and 6.5 are presented. It is evident that over a period of 500–1500 min (or ~8–24 h) significant chloride levels will build up in solution. Work by Cole *et al.*[119] indicated that moisture layers or droplets might exist on corroded surfaces for equivalent periods (dew may form early in the evening and last until after sunrise; likewise rain that falls in the late evening may not evaporate until the following sunrise) and, therefore, significant chloride concentration could build up from the dissolution of these oxide layers, enhancing corrosion.

Both these examples indicate that there is a very strong interaction between the atmospheric conditions, forms of deposition, and electrochemical and surface reactions controlling corrosion. This is of course controlled by the pH and chemical composition of the deposited moisture layers, but it is also influenced by drop size and residency time of the drop.

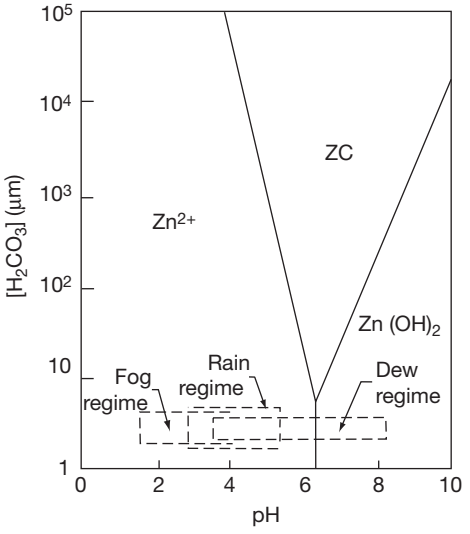


Figure 24 Stability diagram for zinc carbonate (ZC) in aerated aqueous solutions with varying H_2CO_3 content and pH at 25 °C and with zinc ion concentration of 0.1M. Reproduced from Graedel, T. E. *J. Electrochem. Soc*. **1989**, *136*, C193–C203.

2.16.7.2.2 Specific features of the atmospheric corrosion of steel

While the interactions between the environment and steel are similar to those discussed for aluminum and zinc, the atmospheric corrosion of steel has a number of specific features associated with the following:

- The additional reduction and oxidation pathways provided by the existence and ready transformation between Fe(II) and Fe (III) oxides.
- The high porous and often quite thick physical nature of 'rust' on ferrous surfaces.
- The interactions of Fe with sulfur dioxide and its oxidation products, where Fe^{2+} may catalyze the reduction of sulfite to sulfate, while the sulfate ion may characterize the dissolution of iron.

The standard reactions that occur during the atmospheric corrosion of steel are the anodic partial reaction[97] of iron (eqn [69]) and a number of possible cathodic reactions,[120] including oxygen reduction (eqn [39]) and hydrogen reduction (eqn [35]). Hydrogen reduction should dominate over oxygen reduction at pH values less than 4; however, because of the dissolution of iron oxides the electrolyte on top of iron-based metals is well buffered[121] so that H^+ reduction is negligible. However, a third cathodic reaction is the reduction of Fe^{3+} ions within the oxide scale, as first proposed by Evans[122]:

$$8FeOOH + Fe^{2+} + 2e^- \rightarrow 3Fe_3O_4 + 4H_2O \quad [81]$$

Stratmann[123] suggested that this reaction is critical in controlling and accelerating (relative to the corrosion rate without the redox reaction) the corrosion rate of ferrous alloys during wet and dry cycles. During the wetting of an existing oxide layer, Stratmann observed a corrosion rate much greater than that would be expected by oxygen reduction alone and therefore, proposed that the increased corrosion rate was due to the reduction of Fe^{3+} ions. He indeed demonstrated the phenomenon using *in situ* spectroscopy (showing that γ-FeOOH had in part transformed to

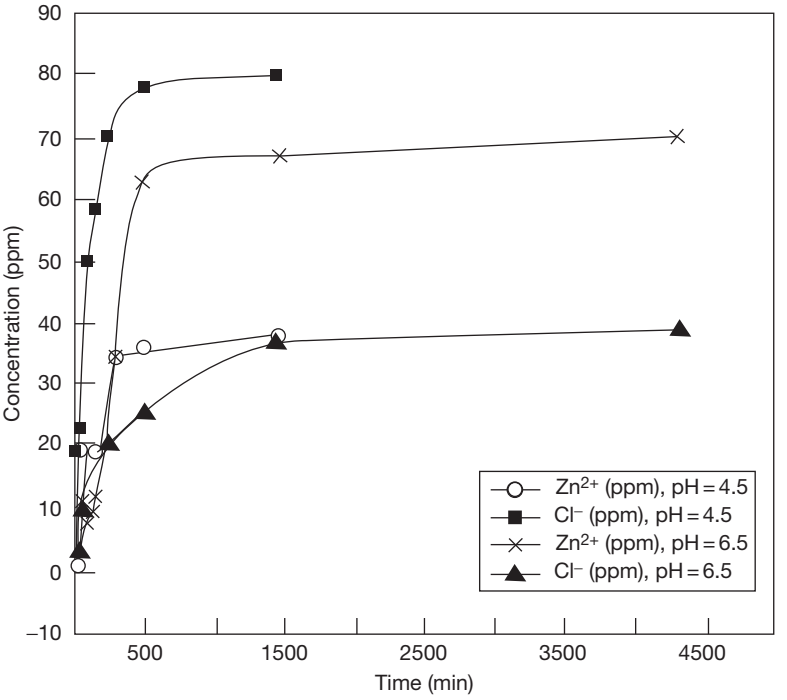


Figure 25 Concentration in solution of Zn^{2+} and Cl^{-} ions due to the dissolution of a zinc hydroxychloride table (1.78 g) in a aqueous solution (1 l) with nitric acid added to give a pH of 4.5 and 6.5.

an $Fe(OH)_2$ phase).[123] Further, the oxide scale has a very large area and the formation of Fe^{2+} states within the FeOOH lattice will control the potential and can dramatically increase the rate of electron transfer reactions, notably oxide reduction, both during the later stages of wetting and during the drying cycle (the reduction of Fe^{3+} to Fe^{2+} being constrained to the initial wetting period). During the drying cycle, the corrosion rate shows a marked peak while the corrosion potential shifts to more anodic values. The increased corrosion current is associated with the increased availability of oxygen as the pathway from the atmosphere to the oxide layer shortens dramatically as the moisture layer thins down, while Stratmann proposed that the anodic shift is due to the precipitation of corrosion products in the concentrated thin moisture films, which block anodic metal dissolution and eventually dramatically slow the corrosion rate. Further, the Fe^{2+} states within the oxide are oxidized during the drying cycle, which reduces the rate of oxygen reduction (reversing the previous effect), and it is only when the oxide is rewetted and new Fe^{2+} states are created that electron transfer is again enhanced.

Therefore, in atmospheric corrosion it would be expected that the nature of the climatic cycle and the nature of oxide layers would have a very pronounced effect on the corrosion rate. At a gross level, atmospheric sequences with a high number of wet/dry cycles are likely to promote a higher corrosion rate than more constant conditions. However, the duration of wet/dry cycles is also important, as Stratmann's work indicates that the extent of Fe^{3+} to Fe^{2+} reduction in the wetting stage will control the extent of corrosion during the drying stage. Precise details of the kinetics of this transformation are not available in the literature; however, it is possible that during short wet/dry cycles, such as those attributed to wetting or aerosol/rain deposition and subsequent drying during the day (referred to as Mode A wetting in **Section 2.16.5.4**), a lesser quantity of Fe^{3+} is reduced and thus the corrosion rate is decreased relative to longer wet/dry cycles (overnight and multiday wetting events, as detailed in **Section 2.16.5.4**).

The volume and porosity of the oxide layers will also be critical in controlling the corrosion rate. In this regard, there are two prime effects. First, as outlined in **Section 2.16.5.4.3**, fine pores will increase the duration of wetness by reducing the RH level at which condensation will occur and also lowering the RH level at which evaporation will occur. Further, increasing the porosity of the oxide will also increase

the cathodic surface area that develops during a wet/dry cycle, thus increasing the corrosion rate. Therefore, it is probable that processes that increase the porosity of the rust layer (such as localized dissolution under acid aerosols) will also increase the overall corrosion rate. However, again there are no systematic studies of this currently reported in the literature.

The last significantly different effect (relative to aluminum and zinc) is the impact of gaseous SO_2 and sulfate compounds on the corrosion of ferrous metals. As outlined in **Section 2.16.1.3.4**, Fe(III) states may catalyze the oxidation of S(IV) compounds (e.g., HSO_3^- or SO_3^{2-}) to SO_4^{2-} or other forms of S(VI). This in turn will promote greater absorption of SO_2 and further acidification of moisture layers. Within moisture layers containing sulfate compounds, sulfur is not consumed in the corrosion reactions, but may act as a catalyst according to the following reactions[124,125]:

$$Fe + SO_2 + O_2 \leftrightarrow FeSO_4 \quad [82]$$

$$4FeSO_4 + O_2 + 6H_2O \leftrightarrow FeOOH + 4H_2SO_4 \quad [83]$$

$$4H_2SO_4 + 4Fe + O_2 \leftrightarrow 4FeSO_4 + 4H_2O \quad [84]$$

Typically, one SO_4^{2-} ion can catalyze the dissolution of more than 100 atoms of iron before it is removed by leaching, spalling of rust, or the formation of basic sulfate.[125] The combined effect of the enhanced absorption of SO_2 and the catalytic role of SO_4^{2-} ion may be responsible for the very high dependence and rate of ferrous corrosion in industrial environments.[126,127]

There is extensive literature on the nature of ferrous rust products and it is beyond the scope of this review to detail this. A key point, however, is that as the redox reaction from Fe(III) to Fe(II) has a pronounced effect on cyclic corrosion, the development of Fe(III) compounds of greater stability should result in a decrease in the corrosion rate, while compounds of lower stability should result in higher corrosion rates. Partially for this reason, there has been great interest in the development of green rusts (GRs), which were initially identified by Stampfl.[128] A typical carbonate-based GR[129] is $Fe_4^{2+}Fe_2^{3+}(OH)_{12}CO_3{\cdot}2H_2O$. Reffass *et al.*[130] mapped the sequence of corrosion products that formed when carbon steel corroded in $NaHCO_3$/NaCl electrolytes and found that pitting corrosion induced by anodic polarization of the surface in $NaHCO_3$ led to the initial precipitation of $FeCO_3$, and that this $FeCO_3$ was then transformed to a species identified as a carbonate-based GR, which later transformed to α-FeOOH. Further, in a study of the oxidation of Fe(II) compounds formed by mixing $FeCl_2{\cdot}4H_2O$, NaOH and $NaHCO_3$ in solution, Reffass *et al.*[130] found that three redox couples developed: $FeCO_3$/GR(CO_3^{2-}), GR(CO_3^{2-}/α-FeOOH, and $FeCO_3$/α-FEOOH. If such reaction sequences are reproduced in atmospheric corrosion, then the above redox reactions may have a major role in driving the corrosion rate in an analogous way to that discussed for the Fe(II)/Fe(III) redox couples.

2.16.8 Conclusion

An understanding of atmospheric corrosion requires a wide understanding of the transfer of pollutants and moisture to metal surfaces and the form and residence time of the pollutants and moisture on the surface.

Despite the enormous range of the species in the atmosphere, surprisingly few species play an active role in atmospheric corrosion. For a chemical species to promote corrosion, it must be transported and deposited onto the metal surface and it must also be reactive to the surface. These conditions are somewhat contradictory as species that are reactive will readily react in the atmosphere and, therefore, will not be available for surface reactions. Gaseous deposition into surface moisture layers, aerosol deposition, and rain deposition are the prime processes for pollutant deposition onto surfaces. Deposition of reactive gases will only be important close to the gas source, while aerosols will be of higher concentration close to the source but may also be transported significant distances from the source. The formation of rain and absorption of gases or aerosols into rain drops require a complex process including the formation of an aerosol, transformation of the aerosol into CCN, growth of the CCN to form a cloud droplet, and the falling of the rain drop and scavenging of gaseous and aerosol species during the fall. Thus, rain drop chemistry will reflect the regional level of pollutants rather than just the local level. Species such as ammonia, sulfates, chlorides, and (to a lesser extent) nitrates, that are readily absorbed into forms of atmospheric moisture such as aerosols or rain drops will be readily transported and deposited on surfaces. Oxidants such as ozone or hydrogen peroxide are of critical importance as they will both facilitate the oxidation of reduced species (such as sulfites), and thus promote further absorption, as well as controlling the redox potential of the deposited solutions. The pH of aqueous phases in the atmosphere will depend critically on the

balance between acidic (primarily oxides of sulfur and nitrogen) and alkali precursors (predominately the alkali metals and ammonia) and, therefore, can be very variable. Although relatively acidic aerosols may develop, rainwater solutions tend to be highly buffered and, hence, extremes in pH are rare.

Once deposited, pollutants, aerosols, and rainwater may have a number of effects which are mitigated by local micro climatic effects. The microclimate of exposed metal components is in general subject to strong diurnal cycles which are amplified by radiative effects (heating during the day and cooling at night by the emission of radiation of the surface to the night sky) so that low surface RH may occur during daylight hours and high RH at night. Many of the salts deposited in aerosols or rainwater are hygroscopic and so deliquescence of such salts can lead to a cyclic pattern of localized wetting and drying. Although rain can promote the deposition of pollutants and transfer of acidity to the surface, it may also remove pollutants from the surface. The effectiveness of pollutant removal is controlled by the rain intensity and the nature of the surface. Electrochemical and surface reactions may influence the form and chemistry of droplets on a surface. Perhaps the most dramatic example of surface processes inducing droplet instability is secondary spreading. In secondary spreading, a fine moisture film spreads out from the centralized drop, with the central drop forming the anode and the secondary spread region the cathode of an extended electrochemical cell. Alkaline pH develops in the secondary spread region which may promote particular gaseous absorption reactions. For example, in secondary spreading on zinc, the absorption of carbon dioxide and the enrichment of the secondary spread region in carbonate occur and the subsequent precipitation of zinc hydroxycarbonate is promoted. However, pH variations and local electrochemical cells may occur on a much finer scale. In aluminum alloys, copper-containing intermetallics may be preferred sites for cathodic reactions, with high pH in the vicinity of the intermetallics promoting the dissolution of the aluminum matrix. However, the ability of the solutions to sustain such local variations in pH is heavily influenced by their buffering capacity (highly buffered solutions will not support these local variations) which in turn is determined by the history of aqueous and gaseous reactions prior to deposition.

The solution chemistry will be controlled by the precipitation of new oxides and the stability of existing oxides. There can be sharp 'cycles' in ion concentration and pH of deposited moisture. For example in the zinc system, the zinc hydroxy anions are stable at near neutral pH values and appreciable anion concentrations. These conditions can readily occur during the evaporation of aerosol or rain droplets and, therefore, precipitation of such zinc OH anions is readily observed under these conditions. However, when an aerosol or rain drop is first deposited, it will commonly be acidic with relatively low anion concentrations. Under these conditions, some of the existing surface precipitates are not stable and may dissolve.

Thus, atmospheric corrosion is influenced by a wide range of linked processes, both within the atmosphere and on the surface. However, as in general these processes relate to a relatively narrow band of chemical species, an adequate knowledge of the dynamics of atmospheric corrosion can be constructed.

References

1. Seinfeld, J.; Pandis, S. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*; Wiley Interscience: New York, 1997.
2. Berresheim, H.; Wine, P. H.; Davis, D. D. Sulfur in the Atmosphere. In *Composition, Chemistry and Climate of the Atmosphere*; Singh, H. B., Ed.; Van Nostrand Reinhold: New York, 1995; pp 251–307.
3. Dacey, J. W. H.; Wakeham, S. G. *Science* **1986**, *233*, 1314–1316.
4. Bates, T. S.; Charlson, R. J.; Gammon, R. H. *Nature* **1987**, *329*, 319–321.
5. Chin, M.; Davis, D. D. *J. Geophys. Res.* **1995**, *100*, 8993–9005.
6. Noller, B. N.; Currey, N. A.; Ayers, G. P.; Gillett, R. W. *Sci. Total Environ.* **1990**, *91*, 23–48.
7. National Research Council *Rethinking the Ozone Problem in Urban and Regional Air Pollution*; National Academy Press: Washington, DC, 1991.
8. Intergovernmental Panel on Climate Change (IPCC) *Climate Change 1994: Radiative Forcing of Climate Change and an Evaluation of the IPCC IS92 Emission Scenarios*; Cambridge University Press: Cambridge, UK, 1995.
9. Dentener, F. J.; Crutzen, P. J. *J. Atmos. Chem.* **1994**, *19*, 331–369.
10. Danielson, E. F. *J. Atmos. Sci.* **1968**, *25*, 502–518.
11. Van Valin, C. C.; Luria, M.; Ray, J. D.; Boatman, J. F. *J. Geophys. Res.* **1990**, *95*, 5689–5695.
12. Kleinman, L. *J. Geophys. Res.* **1991**, *96*, 20721–20733.
13. Stockwell, W. R. *J. Geophys. Res.* **1995**, *100*, 11695–11698.
14. Leygraf, C.; Graedel, T. *Atmospheric Corrosion*; Wiley and Sons: New York, 2000.
15. Pandis, S. N.; Seinfeld, J. H. *J. Geophys. Res.* **1989**, *94*, 1105–1126.
16. Hoffman, M. R.; Calvert, J. G. *Chemical Transformation Modules for Eulerian Acid Deposition Models: Vol. 2 – The Aqueous-Phase Chemistry*; EPA/600/3–85/017, US Environmental Protection Agency: Research Triangle Park, NC, 1985.
17. Troy, R. C.; Margerum, D. W. *Inorg. Chem.* **1991**, *30*(18), 3538–3543.

18. Chen, Z. Y.; Zakipour, S.; Persson, D.; Leygraf, C. *Corrosion* **2005**, *61*(11), 1022–1034.
19. Zakipour, S.; Tidbald, J.; Leygraf, J. *J. Electrochem. Soc.* **1995**, *142*(3), 757–764.
20. Svensson, J. E.; Johansson, L. G. *J. Electrochem. Soc.* **1993**, *140*(8), 2210–2218.
21. Pilinis, C.; Seinfeld, J. H. *Atmos. Environ* **1987**, *21*, 2453–2466.
22. Clegg, S. L.; Brimblecombe, P.; Wexler, A. S. *J. Phys. Chem. A* **1998**, *102*(12), 2137–2154.
23. Keene, W. C.; Sander, R.; Pszenny, A. P.; Vogt, R.; Crutzen, P. J.; Galloway, J. N. *J. Aerosol Sci.* **1998**, *29*(3), 339–356.
24. Keywood, M. D.; Chivas, A. R.; Fifield, L. K.; Creswell, R. G.; Ayers, G. P. *Aust. J. Soil Res* **1997**, *35*, 1177–1189.
25. Whitby, K. T.; Cantrell, B. Fine Particles. In Proceedings of the International Conference on Environmental Sensing and Assessment Institute of Electrical and Electronic Engineers: Las Vegas, NV, 1976.
26. Schutz, L.; Kramer, M. *J. Atmos. Chem.* **1987**, *5*, 173–184.
27. Wall, S. M.; John, W.; Ondo, J. L. *Atmos. Environ.* **1988**, *22*, 1649–1656.
28. Tang, I. N. *Atmos. Environ.* **1980**, *14*, 819–828.
29. Tang, I. N.; Munkelwitz, H. R. *Atmos. Environ.* **1993**, *27A*, 467–473.
30. Wexler, A. S.; Seinfeld, J. H. *Atmos. Environ.* **1991**, *25A*, 2731–2748.
31. Murphy, D. M.; Thomson, D. S.; Middlebrook, A. M.; Shein, M. E. *J. Geophys. Res. Atmos.* **1998**, *103*(D13), 16485–16491.
32. Spillane, M. C.; Monahan, E. C.; Bowyes, P. A.; Doyule, D. M.; Stabeno, P. J. In *Oceonic Whitecaps and Their Role in Air-Sea Exchange Processes*; Monahan, E. C., Macnioaill, G., Eds.; D. Reidel Publishing Company: Dordrecht, Holland, 1986; pp 209–218.
33. Monahan, E. C.; Davidson, K. L.; Spiel, D. E. *J. Geophys. Res.* **1982**, *87*, 8898–8904.
34. McKay, W. A.; Garland, J. A.; Livesley, D.; Halliwell, C. M.; Walker, M. I. *Atmos. Environ.* **1994**, *28*, 3299–3309.
35. Fitzgerald, J. W. *Atmos. Environ.* **1991**, *25A*, 533–545.
36. Cole, I. S.; Muster, T.; Paterson, D.; Furman, S.; Trinidad, G. S.; Wright, N. A holistic model of atmospheric corrosion: Extending a microclimate model into a true corrosion model. Paper 12-9-2004. Eurocorr 2004: Long Term Prediction and Modelling of Corrosion, 12–16 Nice: France, September 2004.
37. Bretschneider, C. L. Beach Erosion Board Technical Memorandum No. 118, US Army Corps of Engineers: Washington, DC, 1959.
38. Cole, I. S.; Ganther, W. D.; Paterson, D. A.; King, G. A.; Furman, S. A.; Lau, D. *Corros. Eng. Sci. Technol.* **2003**, *38*(4), 259–266.
39. Cole, I. S.; Paterson, D. A.; Ganther, W. D.; Neufeld, A.; Hinton, B.; McAdam, G.; McGeachie, M.; Jeffery, R.; Chotimongkol, L.; *et al. Corros. Eng. Sci. Technol.* **2003**, *38*(4), 267–274.
40. Pathak, R. K.; Louie, P. K.; Chan, C. K. *Atmos. Environ.* **2004**, *38*(19), 2965–2974.
41. Yao, X.; Ling, T. Y.; Fang, M.; Chang, C. K. *Atmos. Environ.* **2007**, *41*(2), 382–393.
42. Takeuchi, M.; Okochi, H.; Igawa, M. *Atmos. Environ.* **2004**, *38*(28), 4701–4708.
43. Sander, R.; Crutzen, P. J. *J. Geophy. Res.-Atmos.* **1996**, *101*(D4), 9121–9138.
44. Vogt, R.; Crutzen, P. J.; Sander, R. *Nature* **1996**, *383*, 327–330.
45. Chameides, W. L.; Stelson, A. W. *J. Geophys. Res.* **1992**, *97*(20), 565–580.
46. Yamamoto, N.; Fujii, M.; Endo, O.; Kumagai, K.; Yanagisawa, Y. *Aerosol Sci.* **2002**, *33*, 1667–1679.
47. Keene, W. C.; Stutz, J.; Pszenny, A. A. P.; Maben, J. R.; Fischer, E. V.; Smith, A. M.; von Glasow, R.; Pechtl, S.; Sive, B. C.; Varner, R. K. Inorganic chlorine and bromine in coastal New England air during summer *J. Geophys. Res.* **2007**, *112*, D10S12; doi.10.1029/2006JD007689.
48. Keene, W. C.; Pszenny, A. P.; Maben, J. R.; Sander, R. *Geophys. Res. Left* **2002**, *29*, 1101.
49. Hobbs, P. V.; Bowdle, D. A.; Radke, L. F. *J. Climate Appl. Meteorol.* **1985**, *24*, 1344–1356.
50. Deepak, A.; Gali, G. *The International Global Aerosol Program (IGAP)*; Deepak Publishing: Hampton, VA, 1991.
51. Jacob, D. J.; Waldman, J. M.; Munger, J. W.; Hoffman, M. R. *Environ. Sci. Technol.* **1985**, *19*, 730–736.
52. Isaac, G. A.; Daum, P. H. *Atmos. Environ.* **1987**, *21*, 1587–1600.
53. Aikawa, M.; Hiraki, T.; Tamaki, M. *Sci. Total Environ.* **2006**, *366*, 275–285.
54. Tu, J.; Wang, H. S.; Zhang, Z. F.; Jin, X.; Li, W. Q. *Atmos. Res.* **2005**, *73*(3–4), 283–298.
55. Pelicho, A. F.; Martins, L. D.; Nomi, S. N.; Solci, M. C. *Atmos. Environ.* **2006**, *40*(35), 6827–6835.
56. Demirak, A.; Balci, A.; Karaoglu, H.; Tosmur, B. *Environ. Monit. Assess.* **2006**, *123*(1–3), 271–283.
57. Baez, A. P.; Belmont, R. D.; Garcia, R. M.; Torres, M.; Padilla, H. *Atmos. Res.* **2006**, *80*(1), 67–85.
58. Mphepya, J. N.; Pienaar, J. J.; Galy-Lacaux, C.; Held, G.; Turner, C. R. *J. Atmos. Chem.* **2004**, *47*(1), 1–24.
59. Shivashankara, G. P.; Ranga, K.; Ramalingaiah, M. *Indian J. Environ. Health* **1999**, *41*, 229–238.
60. Singh, A. K.; Mondal, G. C.; Kumar, S.; Singh, K. K.; Kamal, K. P.; Sinha, A. *Environ. Monit. Assess.* **2007**, *125* (1–3), 99–110.
61. Gillet, R. W.; Ayers, G. P. *Clean Air* **1989**, *23*(3), 106–109.
62. Poste, D.; Bridgman, H. A. *J. Atmos. Chem.* **1991**, *13*, 83–95.
63. Fitzgerald, W. R. The measurement of acidity in the Latrobe Valley, Victoria In Proceedings of the 7th World Clean Air Congress; Hartmann, H., Ed.; Clean Air Society of Australia and New Zealand: Sydney, 1986; pp 236–244.
64. Ayers, G. P.; Gillett, R. W. *Clean Air* **1988**, *22*(2), 53–57.
65. Cole, I. S.; Ganther, W. D. *Corros. Eng. Sci. Technol.* **2006**, *41*(2), 161–167.
66. Schaug, J.; Hanssen, J. E.; Nodop, K.; Ottar, B.; Pacyna, J. M. Co-operative program for the Monitoring and Evaluation of the long range transmission of air pollutants in Europe (EMEP) EMEP-CCC-Report 3/87 Norwegian Institute for air Research Lillestrom: Norway, 1987.
67. Charlson, R. J.; Lovelock, J. E.; Andrease, M. O.; Warren, S. G. *Nature* **1987**, *326*, 665–661.
68. Schwartz, S. E. *Science* **1989**, *243*, 753–763.
69. Graedel, T. E.; Frankenthal, R. P. *J. Electrochem. Soc.* **1990**, *137*, 2385–2394.
70. Graedel, T. E. *J. Electrochem. Soc.* **1986**, *133*, 2476.
71. Garrels, R. M. *Mineral Equilibria at Low Temperature and Pressure*; Harper: New York, 1960.
72. Ohba, R.; Okabayashi, K.; Yamamoto, M.; Tsuru, M. *Atmos. Envir.* **1990**, *24A*(6), 925–935.
73. Gustafsson, M. E. R.; Franzen, L. G. *Atmos. Envir.* **1996**, *30*(6), 977–989.
74. Johnson, K. E.; Stanners, J. F. *The Characteristics of Corrosion Test Sites in the Community*; Report EUR 7433; Commission of The European Communities: Luxembourg, 1981.
75. Strekalov, P. V.; Panchenko, Y. *Prot. Met.* **1994**, *30*, 254.

76. Corvo, F.; Betancourt, N.; Mendoza, A. *Corros. Sci.* **1995**, *37*(12), 1889–1901.
77. Corvo, F.; Hayes, C.; Betancourt, N.; Maldonado, L.; Veleva, L.; Echeverria, M.; De Rincon, O.; Rincon, A. *Corros. Sci.* **1997**, *39*(5), 823–833.
78. Morcillo, M.; Chico, B.; Otero, E.; Mariaca, L. *Mater. Perf.* **1999**, *38*(4), 72–77.
79. Morcillo, M.; Chico, B.; Mariaca, L.; Otero, E. *Corros. Sci.* **2000**, *42*(1), 91–104.
80. Cole, I. S.; Chan, W. Y.; Trinidad, G. S.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2004**, *39*(1), 89–96.
81. Cole, I. S.; Paterson, D. A. *Corros. Eng.. Sci. Technol.* **2007**, *42*(2), 91–104.
82. Haynie, F. H. In *Atmospheric Corrosion of Metals*; Dean, S. W., Rhea, E. C., Eds.; ASTM STP 767; American Society for Testing and Materials: Philadelphia, PA, 1982; pp 286–308.
83. Haynie, F. H. In *The Degradation of Metals in the Atmosphere*; Dean, S., Lee, T. S., Eds.; ASTM STP 965; American Society for Testing and Materials: Philadelphia, PA, 1988; pp 282–289.
84. Spence, J. W.; Haynie, F. H.; Lipfert, F. W.; Cramer, S. D.; McDonald, L. G. *Corrosion* **1992**, *48*(12), 1009–1019.
85. Klassen, R. D.; Roberge, P. R. In Proceedings of NACE 2000 Conference Orlando, Florida, 2000; Paper No. 272.
86. Cole, I. S.; Paterson, D. A.; Lau, D. Holistic Modeling of Gas and Aerosol Deposition and the Degradation of Cultural Objects. In *Physical Techniques in the Study of Art, Archaeology and Cultural Heritage*; Creagh, D., Bradley, D., Eds.; Elsevier: Amsterdam, The Netherlands, 2007; Vol. 2; ISBN-13: 978-0-444-52856-8, ISSN 1871–1731.
87. Cole, I. S.; Lau, D.; Chan, F.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2004**, *39*(4), 333–338.
88. Sehmel, G. A. *Atmos. Environ.* **1980**, *14*, 983–1011.
89. Cole, I. S. Mechanism of Atmospheric Corrosion in Tropical Environments In *Marine Corrosion in Tropical Environments*; Dean, S. W., Delgadillo, G. H. – D., Bushman, J. B., Eds.; ASTM STP 1399; American Society for Testing and Materials: West Conshohocken, PA, 2000; pp 33–47.
90. Muster, T. H.; Cole, I. S. *J. Electrochem. Soc.* **2005**, *152* (3), B125–B131.
91. Cole, I. S.; Lau, D.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2004**, *39*(3), 209–218.
92. Nagano, H.; Doi, T.; Yamashita, M. *Mater. Sci. Forum* **1998**, *289*, 127–134.
93. Cole, I. S.; Ganther, W. D.; Sinclair, J. D.; Lau, D.; Paterson, D. A. *J. Electrochem. Soc.* **2004**, *151*(12), B627–B635.
94. Dean, S. W.; Reiser, D. B. Time of Wetness and Dew Formation: A Model of Atmospheric Heat Transfer. In *Atmospheric Corrosion*; Kirk, W. W., Lawson, H. H., Eds.; ASTM STP 1239; American Society for Testing and Materials: West Conshohocken, PA, 1995; pp 3–10.
95. Cole, I. S.; Ganther, W. D.; Paterson, D. A.; Bradbury, A. *Corros. Eng. Sci. Technol.* **2005**, *40*(4), 328–336.
96. Cole, I. S.; Paterson, D. A. *Corros. Eng. Sci. Technol.* **2006**, *41*(1), 67–76.
97. Dwyer, D. J.; Keleman, S. R.; Kaider, A. *J. Chem. Phys.* **1982**, *76*, 1832.
98. Thomson, W. Proceedings of the Royal Society Edinburgh 1928.
99. Skorchellanti, V. V.; Tukachinsky, S. E. *J. Appl. Chem. (USSR)* **1955**, *28*, 615.
100. Broad, D. W.; Foster, A. G. *J. Chem. Soc.*; 1946, 446.
101. Graedel, T. E. *J. Electrochem. Soc.* **1989**, *136*, 19C–2003C.
102. Pourbaix, M. *Atlas of Electrochemical Equilibria in Aqueous Solutions*; National Association of Corrosion Engineers: 1440 South Creek Drive, Houston Texas 7704, 1974.
103. Graedel, T. E.; Goldberg, K. I. *J. Geophys. Res.* **1983**, *88* (C15), 10865–10882.
104. Neufeld, A. K.; Cole, I. S.; Bond, A. M. *Corros. Sci.* **2002**, *44*(3), 555–572.
105. Muster, T. H.; Cole, I. S. *Corros. Sci.* **2004**, *46*(9), 2319–2335.
106. Tsuru, T.; Tamiya, K. I.; Nishikata, A. *Electrochimica Acta* **2004**, *49*(17–18), 2709–2715.
107. Chen, Z. Y.; Persson, D.; Samie, F.; Zakipour, S.; Leygraf, C. *J. Electrochem. Soc.* **2005**, *152*(12), B502–B511.
108. Nazarov, A.; Thierry, D. *Electrochimica Acta* **2004**, *49*(17–18), 2717–2724.
109. Muster, T.; Hughes, A. E.; Thompson, G. E. Copper in Al alloys. In *Corrosion Research Frontiers*; Nova Publishing: Hauppauge NY, USA, 2009; ISBN-978-1-60741-201-4.
110. Birbilis, N.; Buchheit, R. G. *J. Electrochem. Soc.* **2005**, *152*(4), B140–B151.
111. Schneider, O.; Ilevbare, G. O.; Scully, J. R.; Kelly, R. G. *J. Electrochem. Soc.* **2004**, *151*(8), B465–B472.
112. Buchheit, R. G.; Grant, R. P.; Hlava, P. F.; McKenzie, B.; Zender, G. L. *J. Electrochem. Soc.* **1997**, *144*(8), 2621–2628.
113. Park, J. O.; Apik, C. H.; Huang, Y. H.; Alkire, R. C. *J. Electrochem. Soc.* **1999**, *146*, 517–523.
114. Moon, S. M.; Pyun, S. I. *Electrochim. Acta* **1999**, *44*, 2445–2454.
115. Vukmirovic, M. B.; Dimitrov, N.; Sieradzki, K. J. *J. Electrochem. Soc.* **2002**, *149*(9), B428–B439.
116. Ilevbre, G. O.; Schneider, O.; Scully, J. R.; Kelly, R. G. *J. Electrochem. Soc.* **2004**, *151*(8), B453–B464.
117. Graedel, T. E. *J. Electrochem. Soc.* **1989**, *136*, C204–C212.
118. Graedel, T. E. *J. Electrochem. Soc.* **1989**, *136*, C193–C203.
119. Cole, I. S.; Ganther, W. D.; Lau, D. *Corros. Eng. Sci. Technol.* **2006**, *41*(4), 310–320.
120. Schikorr, G. *Werkst. Korros* **1963**, *14*, 69.
121. Cermakova-Knotkova, D.; Vlckova, J. *Werkst. Korros* **1970**, *21*, 16.
122. Evans, U. R. *Trans. Inst. Met. Finish.* **1960**, *37*, 1.
123. Stratmann, M. *Ber. Bunsenges. Phys. Chem.* **1990**, *94*, 626–639.
124. Schikorr, G. *Werk. Korr.* **1964**, *15*, 457.
125. Schikor, G. *Werk. Korr.* **1963**, *14*, 69.
126. Oesch, S.; Faller, M. *Corros. Sci.* **1997**, *39*(9), 1505–1530.
127. Butlin, R. N.; Coote, A. T.; Devenish, M. *Atmos. Environ.B – Urban Environ.* **1992**, *26*(2), 199–206.
128. Stampfl, P. P. *Corros. Sci.* **1969**, *9*, 185.
129. Hansen, H. C. B. *Caly Miner.* **1989**, *24*, 663.
130. Reffass, M.; Sabot, R.; Savall, C.; Jeanin, M.; Creus, J.; Refait, P. *Corros. Sci.* **2006**, *48*, 709–726.

2.18 Marine Corrosion

B. Phull
308 Humphrey Drive, Wilmington, NC 28411, USA

Abbreviations

ALWC Acclerated low water corrosion
BOD Biological oxygen demand
CCT Critical crevice temperature
CP Cathodic protection
CP Commercial purity (of titanium)
CPT Critical pitting temperature
DO Dissolved oxygen
H Heat treatment term in aluminum alloys (aged condition)
HE Hydrogen embrittlement
IGC Intergranular corrosion
L Low carbon
Max. Maximum
MCA Multiple crevice assembly
MSF Multistage flash (distillation)
O Heat treatment term in aluminum alloys (annealed condition)
P Perforation
PH Precipitation hardening
PREN Pitting resistance equivalent number
PTFE Poly tetra fluoro ethylene
RH Heat treatment term in PH stainless steels
SCC Stress corrosion cracking
SCE Saturated calomel electrode
SEM Scanning electron microscope
SRB Sulfate reducing bacteria
T Temper heat treatment term in aluminum alloys
T Temperature
TDS Total dissolved solids
YS Yield strength

Symbols

C_0 Bulk concentration of oxygen (moles cm^{-3})
d Diameter
D Diffusion constant
Pr Prandtl number
U_0 Bulk flow rate (cm s^{-1})
α Alpha phase
β Beta phase
ν Kinematic viscosity
v Velocity

2.18.1 Introduction

This chapter focuses on corrosion and behavior of important engineering materials in seawater. As entire textbooks have been devoted to seawater and marine corrosion, it is clearly not possible to cover this subject in great detail in this chapter.

It is customary to define marine environments in terms of 'zones' in which materials might be exposed, *viz.*, atmospheric, splash and spray, tidal, immersion or submerged, and bottom-mud or sediment. This chapter addresses corrosion issues primarily pertaining to the immersion, tidal, and splash zones. Corrosion in marine atmospheres and bottom-mud is covered elsewhere in this book.

Materials are subjected to seawater conditions in numerous applications, for example, ships, pleasure boats, submarines, offshore platforms, subsea pipelines and telecommunications cables, wharfs, seawater-cooled power and chemical plants, desalination plants, fishing gear, and so on. Seawater is a critical resource for food (primarily fish), table salt, and conversion to freshwater (desalination). It is also the primary source for extraction of magnesium and subsequent production of its alloys.

As in so many other environments there is a multitude of factors, which can affect corrosion of materials in seawater. Many of these factors are interrelated, resulting in complicated behavior of materials in specific applications. The important influencing factors are reviewed first, followed by an analysis of the corrosion mechanisms and the behavior of common engineering materials in seawater. Finally, test methods in seawater environments are discussed briefly.

2.18.2 Environment Characteristics

Seawater covers ~70% of the earth's surface. It is a highly complex medium that contains almost every element in the periodic table and a multifarious plethora of living organisms. Hence, it is sometimes described as a 'living' medium and considered to be the most corrosive of the natural environments.

2.18.2.1 Seawater Constituents

Conventional wisdom is that the corrosivity of seawater is because of its salt (primarily NaCl) content. However, as discussed later in this section, it is not the salt *per se*, but other constituents – especially the dissolved oxygen (DO) and pollutants such as sulfides that often play critical roles. The total salt content is typically expressed in terms of salinity and/or chlorinity. Salinity is the total solid matter (grams) in one kilogram of seawater (after all the organic matter has been oxidized, carbonate converted to oxide, and bromine and iodine replaced by chlorine). From an analytical viewpoint, it is easier to determine the chlorine content, which is expressed as chlorinity – the total amount of chlorine (grams), plus equivalent chlorine assumed to replace bromine and iodine, in 1 kg of seawater. Salinity is related to chlorinity by the following empirical expression:

$$\text{Salinity} = (1.80655 \times \text{Cholorinity}) + 0.03$$

Salinity and chlorinity are usually expressed as grams per kilogram or parts per thousand ($g\,kg^{-1}$, ppt or ‰). Multiplying by 1000 converts $g\,kg^{-1}$ or ppt to parts per million (ppm). Salinity can also be considered as total dissolved solids (TDS), stated as $mg\,l^{-1}$ (ppm). The properties of seawater can vary considerably by geographical location, seasons, and water depth. Table 1 shows a comparison of the relative approximate salinities of a number of seas and oceans in the world. River outflows, evaporation, rainfall, melting ice, wind, and wave action – all can exert a strong influence on salinity. Littoral waters are often rendered brackish (i.e., less salty) due to significant river outflow and may contain considerable sediments and pollutants.

The 'average' salinity of open ocean waters (i.e., well away from land) is in the range ~32–36 ppt, and the corresponding average chlorinity in the ~18–20 ppt range. The salinity of the landlocked Dead Sea is an order of magnitude higher. Table 2

shows the 'average' concentrations of ionic constituents in open seawater of 35‰ salinity. From a corrosion viewpoint, all the ocean waters can be considered as sufficiently corrosive toward carbon steel, typically the most commonly used constructional material, such that corrosion control in some form or another is essential for extended structure and component life and functionality in most cases.

In terms of salt content, the 'average' seawater composition corresponds to ~3.5% NaCl solution. Thus, many researchers who do not have easy access to natural seawater often use this composition as a synthetic substitute. Others use artificial seawater solutions[2] that attempt to more closely mimic real seawater, at least in terms of dissolved salts. The chloride (Cl^-) ion is considered to be the most aggressive specie in seawater, particularly for localized corrosion. Sulfate (SO_4^{2-}) ions play an important role in the activity of sulfate-reducing bacteria (SRB) which proliferate under anaerobic conditions, and can cause serious corrosion problems for certain materials and applications as discussed later in this section.

Calcium (Ca^{2+}) and bicarbonate (HCO_3^-) ions in seawater contribute toward calcium carbonate ($CaCO_3$) scale formation, which is especially enhanced on heat transfer surfaces because of the inverse solubility of $CaCO_3$ with temperature. Such scales can impede transport of DO to the metal surface, and hence, reduce general corrosion of active metals and the effects of seawater velocity so long as the scale is not dislodged. However, the scale reduces heat transfer, and can significantly increase the probability of localized (under-deposit) corrosion attack, particularly on materials with passive films that are susceptible to such conditions (e.g., 300-series stainless steels, aluminum alloys, nickel–copper alloys, etc.) in aerated seawater. Of course, Ca^{2+} and HCO_3^- are also critical species in the formation of undersea corals and protective shells by numerous marine organisms such as clams, oysters, barnacles, crabs, lobsters, star fish, urchins, and so on.

Organic matter necessary for biological processes contains carbon, oxygen, hydrogen, nitrogen, and phosphorus. Carbon, oxygen, and hydrogen (e.g., as polysaccharides and lipids) are available in abundance, but usable nitrogen and phosphorus compounds (primarily nitrates and phosphates) are generally limiting nutrients. The main sources of these latter nutrients include, river runoff into the sea, decaying organic matter, and upward water flow from deeper ocean areas. In littoral waters, these nutrients can reach very high concentrations where water runoff from land (e.g., in rivers) discharges large amounts of sewage and fertilizers from anthropogenic sources. Nutrient-rich waters encourage excessive plant growth, which depletes oxygen and kills animal life – a process known as eutrophication. An example of this is 'red tides,' which are associated with phytoplankton (algae) blooms.

Industrial waste from anthropogenic sources that is discharged into any body of water is technically a source of pollutants, which can consist of, for example, heavy metals and inorganic and organic compounds. The introduction of high levels of pollutants can alter the corrosivity of seawater toward certain materials, for example, copper-ion effects on aluminum alloys, sulfides and ammonia on copper alloys, sulfur compounds on steels, and so forth.

Table 1 Salinity of various ocean waters expressed as approximate total dissolved solids (mg l^{-1} or ppm)

Seawater	*Approximate TDS (ppm)*
Arabian Gulf	47 000
Red Sea	43 000
Mediterranean Sea	41 000
Atlantic Ocean	36 000
Indian Ocean	35 000
Pacific Ocean	34 000
North Sea	33 000
Bering Sea	32 000
Black Sea	22 000
Caspian Sea	13 000
Baltic Sea	8 000
Dead Sea	322 000

Table 2 Major ionic constituents in seawater of 35‰ salinity (35 000 ppm TDS)[1]

Ion	*Concentration (ppm)*
Chloride, Cl^-	19 353
Sodium, Na^+	10 760
Sulfate, SO_4^{2-}	2 712
Magnesium, Mg^{2+}	1 294
Calcium, Ca^{2+}	413
Potassium, K^+	387
Bicarbonate, HCO_3^-	142
Bromide, Br^-	67
Strontium, Sr^{2+}	8
Borate, $B_4O_7^{2-}$	4
Fluoride, F^-	1

2.18.2.2 Dissolved Gases

DO represents the most important species that controls the corrosion behavior of many materials in seawater, even more so than chloride. DO concentration is reported either as $ml\,l^{-1}$ or $mg\,l^{-1}$ ($ml\,l^{-1} \times 1.429 = mg\,l^{-1}$). The primary factors that control DO content of seawater include temperature, salinity, biological activity, and mixing (flow). The inverse solubility of dissolved oxygen versus temperature and dissolved oxygen versus salinity is depicted in **Figures 1** and **2**, respectively. The DO concentration of the Dead Sea, which has a salinity of ~322 ppt, is apparently ~0.1 $mg\,l^{-1}$.

Oxygen dissolves in seawater either from the atmosphere and/or from plant photosynthesis, which is most prevalent over a depth ~100 m. Thus, surface waters at normal atmospheric pressures tend to be saturated or supersaturated in oxygen, including the thin liquid environment in the splash zone. However, quiescent flow conditions that favor algal growth and proliferation of decaying matter can reduce DO content markedly because of biological oxygen demand (BOD).

The principle reaction in seawater that controls corrosion is oxygen reduction. For any given DO concentration, the corrosion process is often strongly influenced by seawater velocity. This is discussed under corrosion mechanisms later in this chapter. The role of carbon dioxide is discussed in the following section. Other gases such as hydrogen sulfide and ammonia can be generated by bacteria from decaying matter and can influence corrosion of certain materials as discussed later.

2.18.2.3 Seawater pH

Clean surface seawater is typically slightly alkaline. This is related to the carbon dioxide (CO_2) equilibrium with the atmosphere. Seawater contains bicarbonate

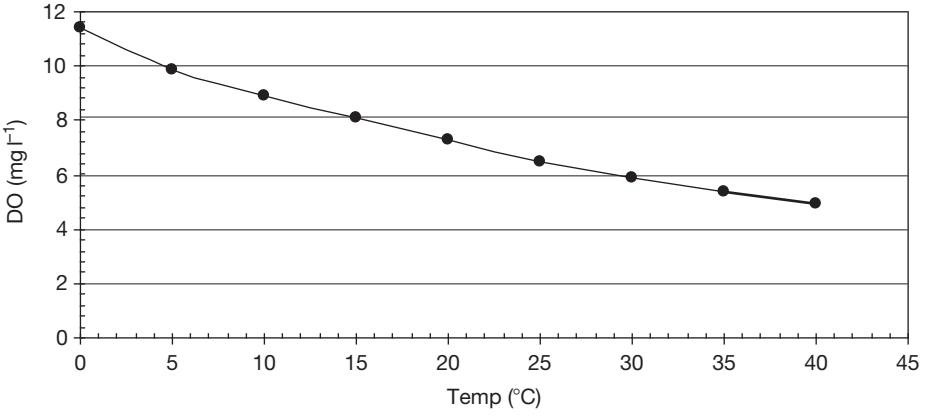


Figure 1 Dissolved oxygen (DO) versus temperature in natural seawater of 35‰ salinity at 1 atm pressure.

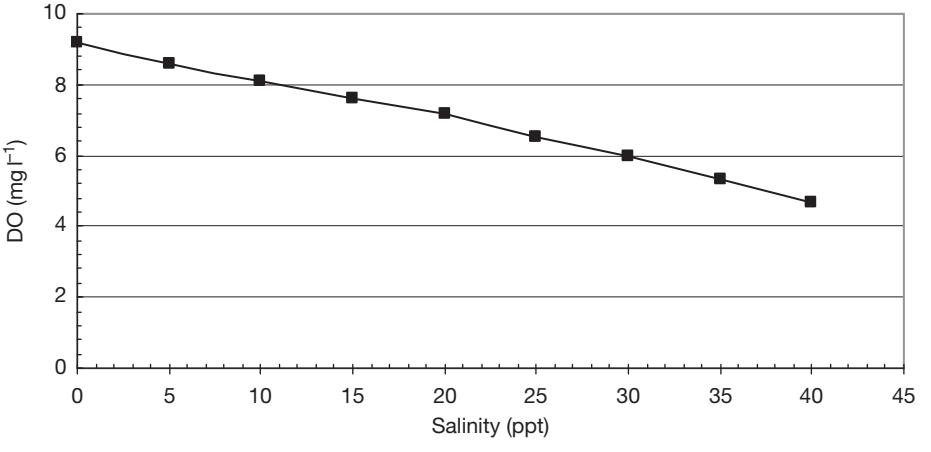


Figure 2 DO versus salinity in synthetic seawater at 20 °C and 1 atm pressure.

(HCO_3^-) ions, carbonate (CO_3^{2-}) ions, undissociated carbonic acid (H_2CO_3), and dissolved CO_2. The pH is influenced by temperature, CO_2 partial pressure (which affects carbonate in solution), photosynthesis, and to a lesser extent by boric acid. These factors interact to maintain the pH in a surprisingly alkaline range (8.1–8.3), and hence, seawater is said to be highly buffered. Hydrogen sulfide, for example, from sulfate-reducing bacterial (SRB) activity, can lower the pH considerably and affect the performance of certain materials.

2.18.2.4 Calcareous Deposits

Calcium, bicarbonate, and magnesium in seawater are important constituents that can result in the formation of calcareous deposits on metal surfaces. Calcareous deposits form on cathodic surfaces and can slow down the corrosion process by inhibiting diffusion of DO and, consequently, its reduction at the metal surface. Calcareous deposit formation is particularly important when large, bare structures (e.g., steel offshore platforms) are cathodically protected in seawater. Without calacareous deposits, current demand for cathodic protection would remain at a high level, consequently, sacrificial anode life would be considerably reduced and the structure may not polarize to adequate protection levels. Calcareous deposit morphology and thickness are affected by current density at the cathode, seawater composition, temperature, and velocity.

2.18.2.5 Temperature

Seawater temperature is also influenced by many geographical factors. Temperature generally decreases with depth. It also varies seasonally at the surface, ranging from ~35 °C at the equator to ~−2 °C at the poles. Increasing temperature lowers solubility of dissolved gases and scale-forming calcium compounds such as calcium carbonate, but generally increases biofouling. All these factors can have a complex effect on corrosion behavior as discussed later in this chapter.

2.18.2.6 Effect of Depth

Figures 3 and **4** illustrate how various parameters can vary with depth[2] and geographic location.[3] Generally, DO, temperature, and pH decrease steeply with depth over the first several hundred meters while salinity increases. Hydrostatic pressure increases with depth while light penetration decreases. It has been postulated that the increase in DO at several 1000 m in the Atlantic Ocean is due to the flow of cold seawater through the 'funnel' in the north; but in the Pacific Ocean, this effect is prevented because of a minor flow through the Bering Strait.[4] It is important to recognize that the data shown in **Figures 3** and **4**, for example, are very site specific. Thus, quite wide variations in parameters can occur from one location to another, which can only be determined from actual measurements. Consequently, these variations can influence corrosion behavior of certain materials.

2.18.2.7 Fouling, Macro- and Microorganisms

Fouling constitutes the attachment of macro- and microorganisms, most prevalent on stationary surfaces. Biofilms apparently start forming on all metallic and nonmetallic surfaces within hours of exposure to seawater.[5] Biofilms which are associated with organic matter and molecules (exopolymers) in seawater are colonized by various bacterial microorganisms. As the films become established, heavier slime formation is followed by attachment of macro species,[6] for example, barnacles, oysters, clams, sea squirts, and vegetable species such as seaweed, bryozoa, hydroids, and codium. The macrofouling species predominate in the tidal zone and can slow down diffusion of dissolved gases, and hence overall corrosion of certain materials. However, shielding of the surface can lead to local anaerobic conditions, anaerobic bacterial activity, and subsequently, localized attack. Some materials (e.g., copper-rich alloys) are more resistant but not entirely immune to macro fouling compared with other materials that foul copiously (e.g., carbon steel, stainless steel, titanium, and even nonmetallics).

2.18.2.8 Effect of Velocity

Velocity can influence corrosion behavior in seawater in multiple ways. If the velocity is too low (quiescent conditions), solids may settle out, subsequently leading to under-deposit attack (e.g., on susceptible grades of stainless steel such as Type 410, 430, 304, 316, etc.). At intermediate velocities, such deposits may be suspended in the flow (once deposits have formed, much higher flow rates may be necessary for their removal than preventing settlement in the first place). As velocity continues to increase, biofilm and calcareous deposit formation diminishes. However,

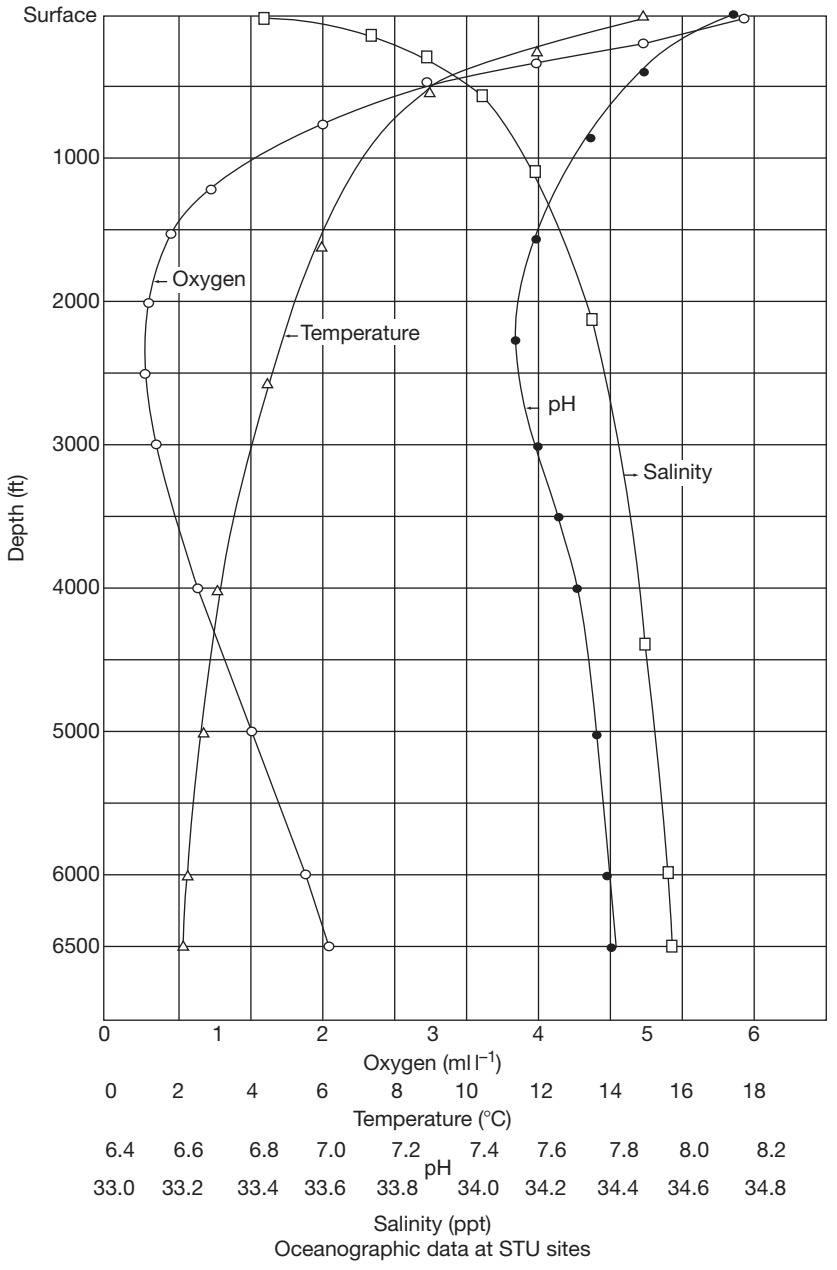


Figure 3 Variations in DO, temperature, pH, and salinity at a specific Pacific Ocean corrosion test site. Reproduced from Dexter, S.C.; Culberson, C. *Meter. Perform.* **1980**, 19(9), 23.

increased wall shear stress, turbulence, and impingement create conditions conducive for the removal of protective films on solid surfaces – thus, increasing the risk of erosion–corrosion of susceptible materials (e.g., carbon steel, copper-base alloys, aluminum-base alloys). At extreme velocities, even otherwise highly corrosion-resistant materials become increasingly susceptible to cavitation damage. Generally, for active materials, increasing the velocity increases corrosion rate as more DO is transported to the metal surface. For passive materials, the enhanced diffusion of DO due to higher velocities is usually beneficial as long as the attack (e.g., crevice corrosion) has not been already initiated. Macroorganisms, calcareous deposits, and coatings generally reduce velocity effects, partly by slowing diffusion of oxygen to the metal surface.

However, at higher velocities these protective mechanisms may be negated as discussed earlier.

2.18.2.9 Seawater Resistivity

Seawater typically has low electrolytic resistivity (resistivity is reciprocal of conductivity) due to its significant dissolved salt content. Low resistivity generally corresponds to greater local cell action, larger anode–cathode interaction distances, and consequently, higher corrosion rates (particularly for active materials such as carbon steel), higher galvanic corrosion rates, and higher current output from cathodic protection anodes. Table 3 shows that resistivity decreases with increasing salinity and temperature; the resistivity values shown in Table 3 were converted from the conductivity data.[7] Thus, higher resistivities can be expected for seawater affected by fresh water outflows, for example, in coastal areas and for colder waters.

2.18.3 Corrosion and Protection Mechanisms

From a corrosion standpoint, chloride ions are considered to be the most aggressive constituent in seawater. The precise role of chloride in the corrosion process is still not fully understood. For instance, chloride contributes to increased electrolytic conductivity by ion transport. Higher conductivity means that current between anodic and cathodic areas can flow over larger distances, for example, in galvanic couples; the magnitude of local cell currents can also be higher. Thus, the overall effect is typically higher general and/or localized corrosion rates. Corrosion attack at anodic areas is supported by reduction reaction(s) at cathodic sites – typically oxygen reduction in aerated seawater.

Corrosion products can provide some degree of protection if they impede oxygen diffusion to cathodic sites. Reduction in general corrosion rate of steel in seawater due to formation of corrosion products and macrofouling is illustrated in Figure 5.

Oxygen reduction is obviously a critical reaction for metallic corrosion in seawater. Deoxygenation of seawater is a recognized method of corrosion control for carbon steel, for example, in multistage flash distillation plants and water injection in offshore oil/gas production during secondary recovery. Thus, chloride ions are rendered relatively harmless in the complete absence of oxygen. Typical polarization curves for steel in aerated and deoxygenated seawater are shown in Figure 6. It is apparent that although the corrosion potential is nearly 400 mV more 'active' in the absence of oxygen, the corrosion rate is significantly lower when compared with aerated seawater.

Differences in oxygen levels between local areas produce potential differences that give rise to concentration cells – commonly referred to as differential oxygenation or differential aeration cells. The low oxygen or anoxic (oxygen-free) area becomes anodic because of its more active (electronegative) potential; conversely, the area of higher oxygen becomes

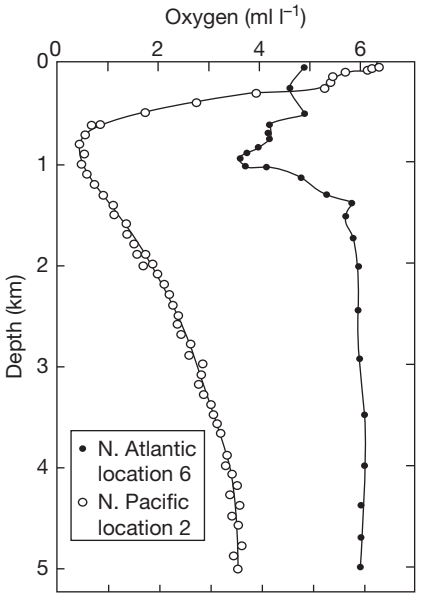


Figure 4 DO-depth profiles at two specific open ocean test sites. Reproduced from Compton, K.G., *Corrosion* **1970**, *26*, 448.

Table 3 Resistivity (Ω cm) as function of salinity and temperature

Salinity (‰)	*0 °C*	*5 °C*	*15 °C*	*20 °C*	*25 °C*	*30 °C*
10	107.1	92.5	80.9	71.6	64	57.7
20	57.3	49.6	43.5	38.5	34.5	31.1
30	39.6	34.4	30.2	26.8	24	21.6
31	38.5	33.4	29.3	26	23.3	21
32	37.4	32.4	28.5	25.3	22.6	20.4
33	36.3	31.5	27.7	24.6	22	19.9
34	35.3	30.7	27	23.9	21.4	19.4
35	34.4	29.9	26.3	23.3	20.9	18.9

Source: Whitfield, M., Jagner, D., Eds. *Marine Electrochemistry, A Practical Introduction*; Wiley, 1981; p 513.

cathodic because of its less active (more electropositive) potential. This is usually demonstrated in the laboratory by setting up a partitioned cell in which one-half is aerated and the other half is deaerated with a steel electrode immersed in each half. The measured open-circuit potentials and couple current flow direction indicates the afore-mentioned anode/cathode relationship. However, as noted by Tomashov,[8] actual corrosion rate of the steel electrode in the aerated half of the cell is greater in chloride solutions, despite being the cathode in the differential-aeration macrocell, than in the deaerated half. This apparent contradiction is due to enhanced microcell action in solutions containing chloride that prevents anodic passivation of the steel. Lower corrosion rates of steel piles in the mud zone (low oxygen environment) relative to the submerged zone above it (aerated environment) were cited as a practical manifestation of this apparently anomalous behavior. In other words, in seawater, chloride ions and associated high conductivity generates microcell activity that overrides the differential-aeration macrocell. In environments where steel can 'passivate,' for example, chloride-free waters or soils, the classical differential-aeration mechanism is operative.

In anaerobic seawater environments, the oxygen reduction reaction is negated. However, high corrosion rates observed under anoxic conditions have been attributed to microbiological activity, in particular, sulfate-reducing bacteria (SRB, e.g., *Desulfovibrio*

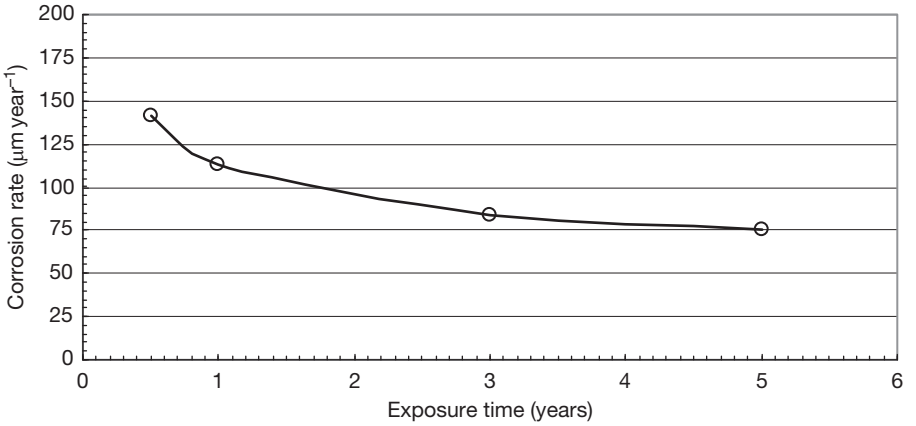


Figure 5 Decrease in general corrosion rate of carbon steel in quiescent seawater associated with development of corrosion products and macrofouling.

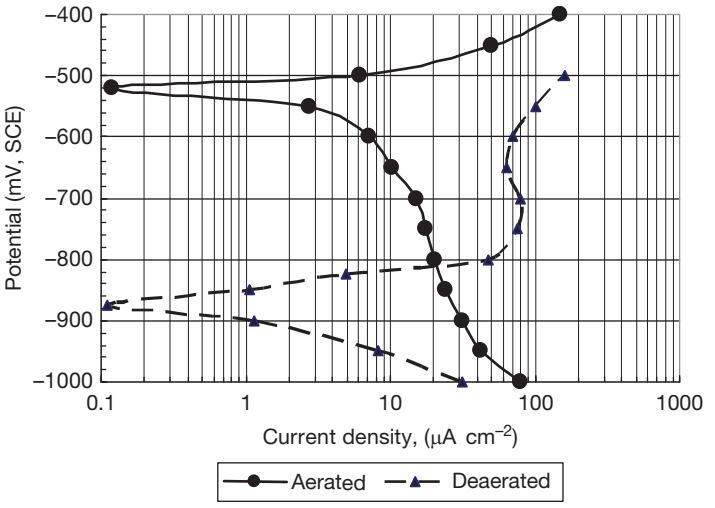


Figure 6 Polarization curves for carbon steel in aerated versus deaerated seawater.

desulfuricans). These bacteria reduce sulfate to sulfide under favorable conditions.[9–11]

While SRB can influence corrosion in anaerobic environments, propagation of corrosion attack is generally most prevalent under conditions of alternate exposure to anaerobic and aerobic environments. In principle, an impervious sulfide film covering the entire metal surface may be protective. However, in practice, such films are rarely impervious and/or continuous. Thus, corrosion at small exposed areas of the substrate is accelerated as the sulfide film serves as a large cathodic surface. The adverse effect of sulfide on corrosion of copper heat exchanger alloys is well known and discussed later in this chapter. Bacteria other than SRB can also influence corrosion behavior in seawater.

As discussed earlier, calcareous deposits form on cathodic surfaces. Although enhanced under cathodic protection conditions, such deposits can also form at cathodic areas associated with corroding surfaces. Hydroxyl ion generation by the oxygen reduction reaction at the cathodes increases the pH at the metal surface. The increase in pH alters the equilibrium, which can result in precipitation of calcium carbonate scale.

$$Ca^{2+} + HCO_3^- + OH^- \rightarrow CaCO_3 + H_2O$$

Under cathodic overprotection conditions, formation of magnesium hydroxide is favored.

$$Mg^{2+} + 2OH^- \rightarrow Mg(OH)_2$$

Adherent calcareous deposits slow the diffusion of oxygen to the metal surface. For immersed surfaces, this implies polarization of the oxygen reduction reaction, and hence a decrease in cathodic protection current demand or a decrease in general corrosion rate when there is no cathodic protection.

Calcareous deposits play a critical role in cathodic protection of bare metal structures in seawater. Early pioneering work by Cox generated several patents, which described the 'best' current density ranges to produce the most protective calcareous deposits[12,13] that were referred to as 'electrocoating.' Over the years, many research efforts have been pursued for a better understanding and optimization of calcareous deposits.[14–24] It has been shown that the deposits are affected by many variables that include current density, potential, seawater velocity, temperature, and chemistry.

For example, **Figure 7** shows current densities associated with calcareous deposits formed in synthetic seawater at several potentiostatically-controlled protection potentials in 50 h. Variation of deposit morphology with protection potential is illustrated in the SEM photomicrographs in **Figure 8**. At −1100 mV (SCE), the deposits consisted predominately of $CaCO_3$ (aragonite) and a trace of $Mg(OH)_2$ (brucite). Although the deposits were thick, they were poorly adherent and not very protective since the current density did not decrease significantly. More adherent and protective deposits, composed primarily of $CaCO_3$ (aragonite), were formed at −800, −900, and −1000 mV (SCE). At −750 mV (SCE), complete cathodic protection was not attained, and the deposits comprised a mixture of $CaCO_3$ and iron corrosion products (Fe_2O_3 and Fe_3O_4). Data are not included in **Figures 7** and **8** for deposits formed at −1200 mV (SCE); however, the current density was $>10^4$ mA m^{-2}, the deposits were thick, gelatinous, mainly $Mg(OH)_2$, and nonprotective. In practice,

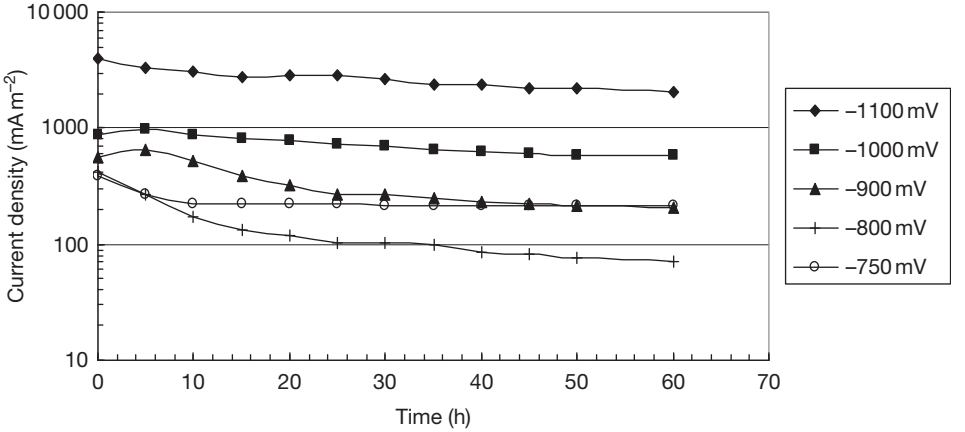


Figure 7 Current density versus time for clean steel at various protection potentials versus SCE.

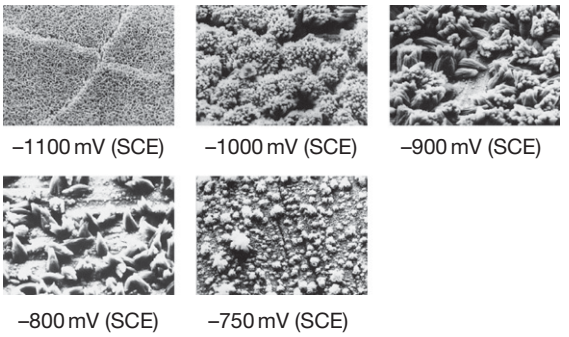


Figure 8 Morphology of calcareous deposits formed at various potentials in synthetic seawater in 50 h.

many impressed current cathodic protection systems operate at constant current. As calcareous deposits form under constant current conditions, the current density increases at local areas, the potential of the protected metals shifts in the negative direction, and the morphology of the calcareous deposits changes. From a research viewpoint, it is easier to study the formation of, and protection afforded by, calcareous deposits under potentiostatic conditions rather than constant current conditions. Moreover, cathodic protection from well-designed sacrificial-anode systems, which are more common in seawater applications, approximates potentiostatic conditions, that is, anode output decreases as calcareous deposits are formed. However, Humble[14] has shown that the Ca/Mg ratio in calcareous deposits decreases with increasing current density as illustrated in **Figure 9**. Surprisingly, it has been found that Mg^{2+} ions in seawater inhibit $CaCO_3$ crystal nucleation and precipitation.[25] However, the greater concentration of Mg^{2+} in open seawater compared to Ca^{2+} assures that $CaCO_3$ will precipitate as aragonite.[26] In other natural waters, precipitation of $CaCO_3$ as calcite is favored because the concentration of Ca^{2+} is usually much greater than that of Mg^{2+}.

Table 4 provides some quantitative and qualitative data on calcareous deposit properties relative to cathodic protection current densities. The solubility of $CaCO_3$ decreases with increasing temperature, while that of $Mg(OH)_2$ increases. Thus, precipitation of $CaCO_3$ is favored in warm seawater, and that of $Mg(OH)_2$ in colder seawater. This has been postulated as one possible explanation of higher current density requirements for cathode protection of steel structures in colder seawater. It is apparent that calcareous deposit formation involves complex phenomena; thus, ideally, characterization and optimization of calcareous deposits for specific applications are best performed locally, *in situ*.

It is generally accepted that protection for materials such as aluminum alloys, titanium alloys, stainless steels, and nickel alloys is conferred by surface passive oxide films. In other words, these materials possess active–passive characteristics. In seawater, active corrosion is suppressed if the material exhibits passive behavior. The air-formed oxide films are thin (typically <10 nm). The composition of the base metal, heat treatment, surface cleaning, surface finish, and the exposure environment are among the primary factors that determine the stability of the passive film. There is some evidence that adsorbed species from the exposure environment, for example, oxygen, contribute to the properties of the passive film.[28]

The most deleterious species toward passive films in seawater is chloride. Film breakdown and rehealing is a dynamic process. However, under certain conditions rehealing may not occur after corrosion has initiated at local areas. This can lead to propagation of attack and high rates of localized corrosion. A classical case of this is crevice corrosion of susceptible stainless steels, for example, under gaskets, under O-rings, at threaded joints, under silt and macrofouling deposits, and so on. Previously, this type of attack was ascribed to oxygen concentration cells, that is, absence of oxygen in the crevice and plentiful supply outside the crevice. However, the modern view is a little more complex. Crevice corrosion has been modeled mathematically and experimentally in a number of ways.[29–33] Deeper and tighter crevices are more susceptible to both the initiation and propagation of crevice corrosion. Metal-to-nonmetal crevices are generally tighter than metal-to-metal crevices because of the easier deformation of the nonmetal. However, metal-to-metal crevices, although less tight, may be equally susceptible to attack because both metal surfaces can contribute metal ions for hydrolysis. In dissimilar alloys, crevice attack on the more susceptible metal occurs first. As the crevice electrolyte becomes more aggressive by the autocatalytic process, crevice corrosion can initiate subsequently on the more resistant alloy.[34]

Electrochemical measurements have been used to demonstrate potential ennoblement and increased cathodic kinetics on passive surfaces covered by biofilms in seawater.[35,36] It has also been shown that biofilms are more prevalent in warm seawater, and hence decrease crevice corrosion initiation time compared to cold seawater. Above a threshold temperature

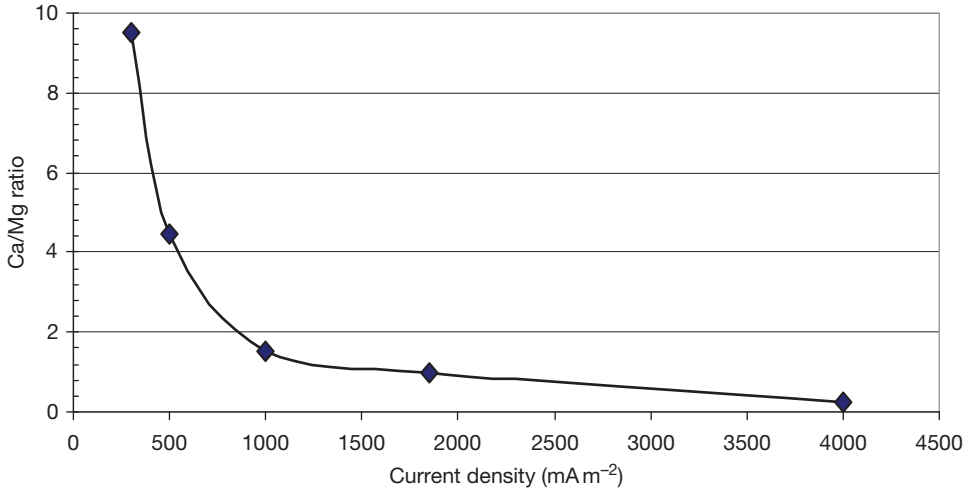


Figure 9 Ca/Mg ratio in calcareous deposits formed in seawater at various cathodic protection current densities over 3 months. Reproduced from Humble, H. A. *Corrosion* **1948**, *4*(7), 358–370.

Table 4 Calcareous deposits properties relative to cathodic protection current densities

Initial current density (mA m^{-2})	*Duration (days)*	*Calcareous film formed*	*Current density for continued protection (mA m^{-2})*
2000	1	Soft with high Mg content	20–30
1000	2–3	Fairly hard	20–30
500	5–6	Hard, mainly calcareous	20–30
200	20–30	Medium thickness, hard	30–40
100	60–90	Medium thickness, hard	40
60	180	Light, hard	40–50

Source: Doremus, E. P.; Doremus, G. L. *Corrosion* **1950**, *6*(7), 216–224.

of ~50 °C, crevice corrosion susceptibility in natural seawater apparently decreases.[32] It has been hypothesized that this behavior is due to reduced microbial activity in biofilms above such temperatures. This is analogous to reduction in crevice corrosion propensity in seawater treated with biocide, for example, 2 ppm residual chlorine, which inhibits biofilms and microbial activity. Crevice corrosion observed at elevated temperatures where biofilms are absent and microbial activity diminished must be explained by other mechanisms. For example, crevice corrosion of titanium is considered to occur at temperatures >70 °C.

Empirical data and experience indicate that some alloys are clearly much more resistant to crevice corrosion than others under the same crevice geometry and seawater conditions. It has been postulated that this is due to more stable passive films formed as a result of certain minimum combinations of alloying additions, primarily Cr, Mo, and N. Interestingly, it has been shown that hydrolysis of Cr^{3+} and Mo^{3+} produces the lowest pH values for the crevice electrolyte.[37] This appears to be the case when insufficient Cr and Mo (and possibly other synergistic elements) are present on the alloy surface to produce a passive film that prevents hydrolysis, acidification, and anodic dissolution. It has been suggested that addition of N to stainless steels can produce ammonium (NH_4^+) ions that raise the pH and effectively neutralize any acidification due to hydrolysis in the crevice.[38]

It is often stated that increasing oxygenation of seawater (e.g., by increasing flow rate) alleviates crevice corrosion. This is based on the widely held view that higher flow rates minimize biofilm attachment and allow easy access of oxygen which is necessary to maintain passivity. That would be the case if oxygen was able to reach all areas of the metal (including the crevice) quickly. In fact, increasing oxygen supply worsens the situation because the oxygen-reduction cathodic reaction is enhanced by depolarization on surfaces outside the crevice, while oxygen diffusion

into the crevice and neutralization of the acidic conditions within the crevice, both necessary for repassivation, remain severely restricted by the crevice geometry. In other words, the trapped acidic electrolyte in the crevice and lack of oxygen ingress keeps the crevice very active, while the area outside the crevice serves as an efficient cathode. Moreover, the rate of attack within the crevice can be very severe since the ratio of the cathode (area outside the crevice) to anode (area inside the crevice) is often quite large in many components in seawater immersion applications. This is illustrated by the example in Figure 10.[39]

Another poorly appreciated fact is the behavior of active–passive metals in deaerated solutions. As stated previously, materials such as stainless steels depend on the formation and maintenance of a passive film on the alloy surface exposed to seawater. A widely held but mistaken view is that deaerated seawater is detrimental because oxygen is unavailable to sustain passivity. However, the fact of the matter is that removal of oxygen from seawater is not normally deleterious toward such alloys because oxygen reduction is the primary cathodic reaction that controls the propagation of pits and/or crevice attack. This has been demonstrated experimentally. Since the seawater pH is near-neutral, the alternative hydrogen-discharge cathodic reaction is negligible. Thus, even if oxygen is unavailable to maintain passivity, corrosion of active–passive alloys in seawater will be imperceptible unless some other cathodic reaction is viable. Even with anaerobic bacteria active under anoxic conditions, corrosion damage occurs predominately when subsequent exposure to aerated conditions occurs.

2.18.3.1 Galvanic Corrosion

Figure 11 is a depiction of the typical galvanic series of metals and alloys in seawater that is widely available to engineers. Unfortunately, the severe limitations of adequately predicting galvanic corrosion from such a series alone are commonly underestimated. The series basically represents a ranking of metals and alloys according to their free corrosion potentials in a given environment (in this case, flowing natural seawater at ambient temperature). For a given metal or alloy, the ranking can be influenced by a number of variables, for example, temperature, flow rate, fouling, chlorination, crevices, oxygen content, and so forth. The series provides only qualitative information on likely galvanic corrosion if two dissimilar materials in the series are electrically coupled in the environment. Corrosion of the anodic material (i.e., one with a relatively more electronegative corrosion potential) is likely to be accelerated, and corrosion of the cathodic (i.e., one with a relatively more electropositive corrosion potential) material is slowed or stopped. However, the degree of galvanic interaction cannot be determined quantitatively from the series alone. If Ohm's law applied, then the amount of galvanic attack would be a function of the potential difference between the two dissimilar materials and the total circuit resistance. However, polarization behavior of the materials and their relative area ratios have been found to be much more significant than the potential difference. The example Figure 12 shows that the potential difference between titanium and steel in seawater is nearly double that of copper versus

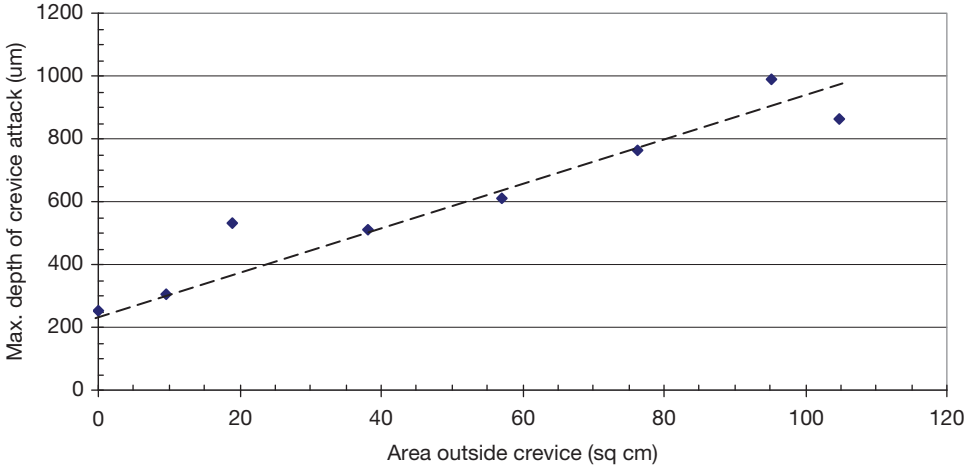


Figure 10 Maximum depth of crevice attack on 17-Cr stainless steel in seawater versus exposed cathodic (passive) surface area outside crevice. Adapted from Ellis, O. B.; LaQue, F. L. *Corrosion* **1951**, *7*(11), 362.

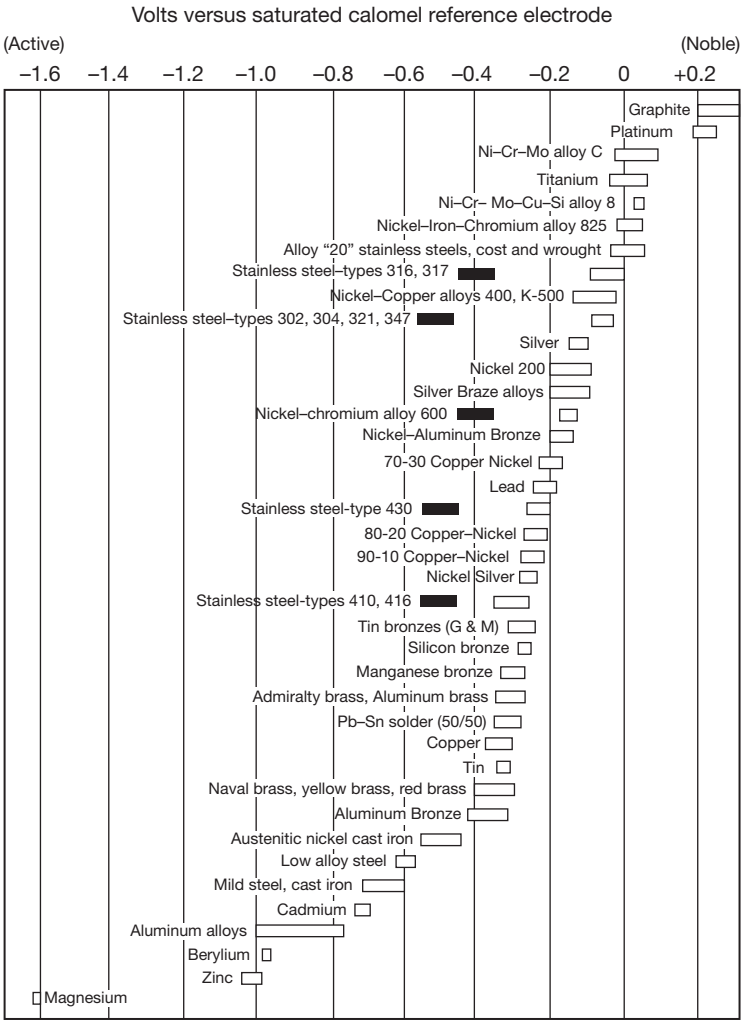


Figure 11 Typical galvanic series of metals and alloys in flowing natural seawater. Solid bars represent active corrosion potentials for otherwise passive alloys.

steel. However, for the same anode/cathode area ratio, the galvanic current density for the titanium/steel couple (and hence expected corrosion rate of steel) is nearly nine times lower compared to the copper/steel couple. This behavior is due to the polarization of the cathode, for example, oxygen reduction on titanium is less efficient than on copper. Polarization is influenced by surface films. For example, biofilms on metal surfaces generally enhance reduction reactions and hence galvanic corrosion in certain couples.[41] Although chlorination is used for biofouling control (e.g., to maintain heat transfer), low levels have also been found to reduce galvanic corrosion, for example, of copper alloys coupled to more noble alloys, because reduction reactions are polarized appreciably.[42,43] If calcareous deposits form on cathodic surfaces, they hinder oxygen diffusion, and hence can reduce galvanic corrosion effects.

It is generally accepted that galvanic attack on the anodic member in a dissimilar metal couple intensifies as the cathode/anode area ratio increases. However, geometry and spatial disposition of the anode and cathode can also play an important role. For example, in a retubed seawater heat exchanger, titanium tubes represent a large surface area with respect to the original copper–nickel tube sheet. This generally results in high galvanic corrosion rate of the latter. In a plate-type heat exchanger, the surface area of the titanium plates is very large compared to say the adjoining copper–nickel piping. However, galvanic corrosion of the latter is not usually

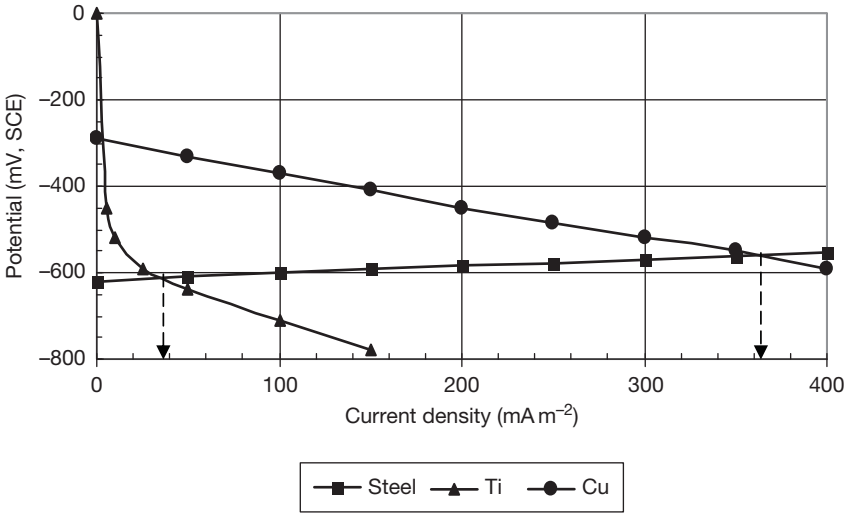


Figure 12 Polarization curves for carbon steel (anode) versus copper (cathode) and titanium (cathode), illustrating galvanic current density depends on polarization behavior rather than galvanic couple potential difference. Adapted from LaQue, F. L. *Marine Corrosion: Causes and Prevention*; Wiley, 1975; p 195.

significant because the tortuous pathways between the titanium plates represent a high electrolytic resistance.

The galvanic series only relates to consideration of two materials at a time in a galvanic couple. In many practical applications, multiple metals may be involved. For example, if aluminum, steel, copper, and titanium components are all coupled in a system, the galvanic series only indicates that titanium would be cathodic and aluminum would be anodic with respect to each other. The galvanic series would not indicate whether steel and copper would behave anodically or cathodically in the overall system. This could only be determined after measuring the mixed potential (when all the subject metals are electrically coupled).

Another widely misunderstood aspect of galvanic corrosion is the interaction of graphite with other metals. Because of its typically most noble position in the galvanic series, it is often misstated that graphite forms adverse galvanic couples with all metals. Graphite (and graphite-containing materials) can certainly accelerate corrosion of metals such as copper, steel, and aluminum, especially for large cathode/anode area ratios. However, graphite is galvanically compatible with titanium, most Ni–Cr–Mo alloys, and many stainless steels. For example, attack observed on stainless steels in contact with graphite gaskets is often incorrectly attributed to galvanic corrosion. However, practical experience indicates that such attack (and of similar magnitude) is in fact due to crevice corrosion which occurs even when the graphite is replaced by a nonconductive gasket material.[44]

2.18.4 Carbon Steels

Carbon steels are widely used in seawater applications because of their ready availability, extensive range of mechanical properties, ease of fabrication, and usually lowest initial cost. However, carbon steels are very susceptible to corrosion in marine environments, and hence require some form of corrosion control. As stated earlier, corrosion rate is affected by a complex interaction of factors that include DO, temperature, depth, flow rate, and microbiological activity.

Figure 13 illustrates the typical corrosion profile of steel piling subjected to various marine zones. The highest corrosion rate in the splash and spray zone is attributed to constant wetting that produces a thin film of seawater through which oxygen diffuses readily. The high corrosion rate, constant wetting conditions, and the mechanical effects of splashing preclude development of protective corrosion products. Even if coatings are applied prior to installation, they are virtually impossible to maintain in this zone. However, successful long-term protection has been achieved with metallic sheathing[46,47] (e.g., using alloys such as NiCu alloy 400 and 90/10 CuNi). Nonmetallic wraps have also been used for protection in the splash zone.[48,49]

In the tidal zone, the structure is exposed to alternate immersion. The lower general corrosion rates in this zone are ascribed to oxygen shielding by macrofouling. However, localized attack can be aggressive if SRB activity under anaerobic macrofouling

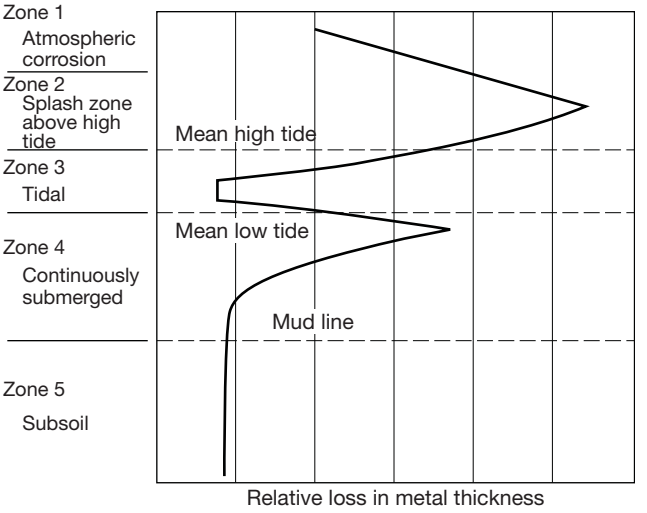


Figure 13 Corrosion profile of steel piling exposed to various zones in a natural marine environment for 5 years. Reproduced from Humble, H. A. *Corrosion* **1948**, 5(9), 23–28.

conditions is high. A particularly severe form of attack known as 'accelerated low water corrosion (ALWC)' afflicts steel sheet piling in some ports and harbors at areas just above the lowest astronomical tide.[50–52] The phenomenon is not fully understood presently. However, empirical evidence suggests a strong influence of microbiological organisms, especially SRB possibly synergized by acid producing bacteria; corrosion penetration can exceed 1 mm year^{-1} on a given steel surface. Apparently, structures affected by ALWC are usually exposed to tidal and brackish waters.

As discussed earlier, reduction of DO is the primary cathodic reaction controlling corrosion of steel in aerated marine environments. Although oxygen solubility decreases with increasing temperature, corrosion of steel immersed in seawater generally increases with temperature and with flow rate. The data in **Table 5** illustrate the increase in corrosion rate of steel in warmer seawater during the summer months despite the lower oxygen content. **Figure 14** depicts increase in corrosion rate with flow velocity in ambient temperature seawater in the absence of significant corrosion products.[53] Adherent corrosion products can diminish the effect of flow. However, above a critical wall shear stress, corrosion products are prevented from forming and existing ones removed by erosion.

The corrosion rate of steel, based on mass loss, in quiescent seawater is typically on the order of ~0.13 mm year^{-1} after several years of exposure. A 5-year worldwide seawater corrosivity test program

Table 5 Seasonal variations in corrosion rate of steel in natural seawater at ~0.2 m s^{-1}

Period	*Ave temp (°C)*	*Ave dissolved oxygen (ppm)*	*pH*	*Ave corrosion rate (μm year^{-1})*
Jan–Mar	10.5	9.2	8.0	100
Apr–Jun	22.4	6.9	8.0	130
Jul–Sep	27.2	6.0	8.0	140
Oct–Dec	14.1	8.2	8.1	110

indicated that local conditions can exert a strong influence on corrosion rate and types of attack.[54] In some cases, pitting rates were much higher than general corrosion rates, as illustrated by the data in **Figure 15**. Complete localized penetration of the 6-mm plate thickness occurred on some, but not all test panels at specific exposure locations. This type of accelerated localized attack resulting in large, round pits has been attributed previously to sea urchins[55] – which apparently 'scrub' corrosion products, and thus expose fresh metal surface to enhanced corrosion rate. Extensive data analysis and modeling on corrosion of steel in seawater have been performed by Melchers[56]; the primary conclusions were that velocity affects corrosion in the early stages until marine fouling and corrosion products develop – thereafter, influence of fouling thickness is significantly reduced. DO and temperature rather than depth *per se* affect corrosion; for a given exposure period, test specimens first exposed during autumn and winter exhibited less corrosion than those first exposed in spring and summer.

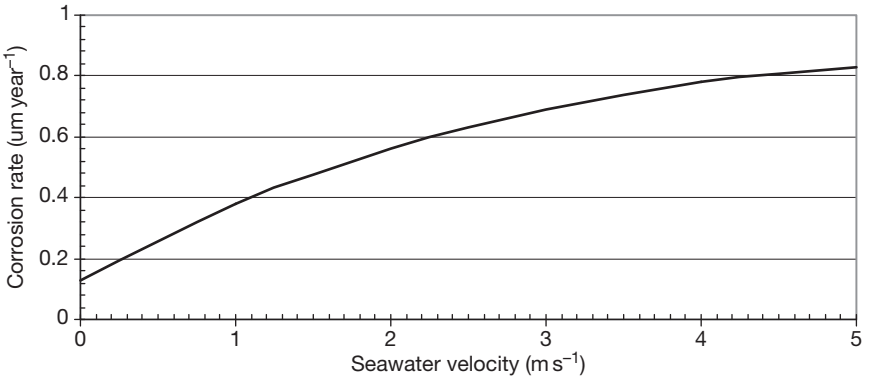


Figure 14 Corrosion rate of clean carbon steel in ambient temperature flowing natural seawater on 38 days. Reproduced from Oubner, R.; Beech, I. In *Corrosion/1999*; NACE International: Houston, TX, 1999; Paper No. 99318.

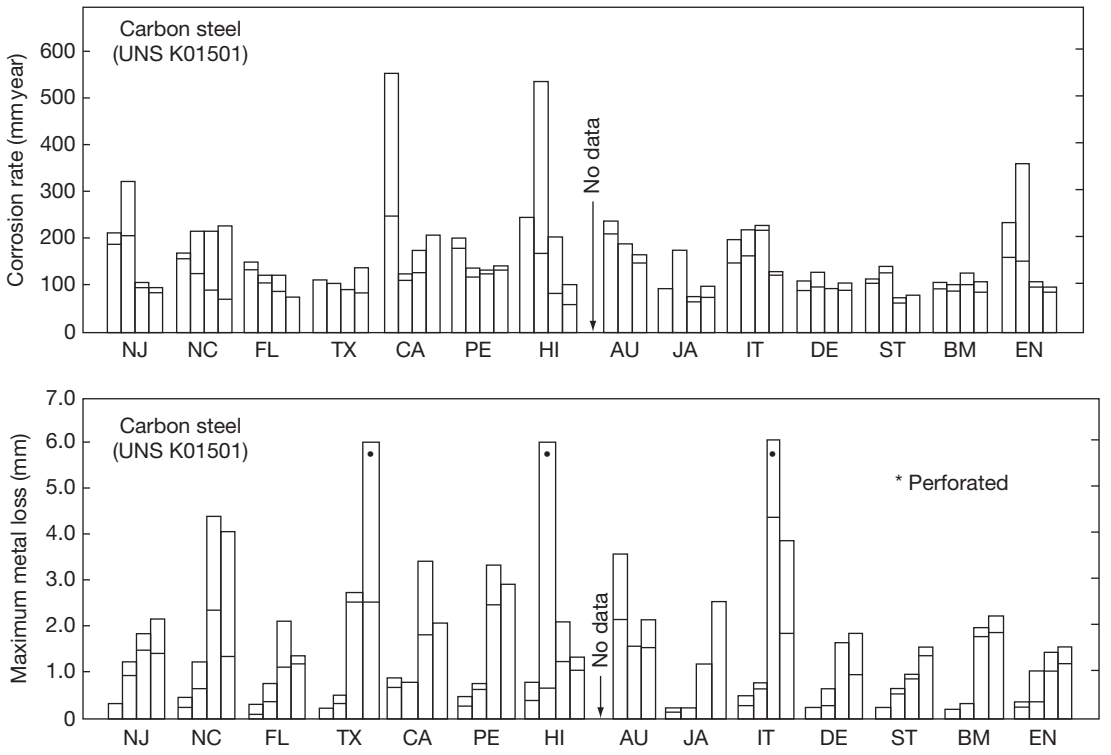


Figure 15 Mass loss corrosion rate and maximum depth of pitting attack on carbon steel in seawater at 14 worldwide test locations. For each location, the bars represent (from the left) 0.5, 1, 3 and 5-year exposures. NJ – New Jersey, NC – North Carolina, FL – Florida, TX – Texas, CA – California, HI – Hawaii, AU – Australia, JA – Japan, IT – Italy, DE – Denmark, ST and BM – Sweden, EN – England. Reproduced from Phull, B. S.; Pikul, S. J.; Kain, R. M. *ASTM STP 1300*; American Society for Testing and Materials: West Conshohocken, PA, 1997; pp 34–73.

It is commonly believed that macrofouling reduces oxygen transport and, consequently, overall corrosion rate of steel in seawater. However, data from tests indicate similar corrosion behavior for carbon steel in filtered and unfiltered seawater – suggesting that microfouling may be more important than macrofouling in controlling corrosion behavior. This appears contradictory to the view that biofilms enhance the oxygen reduction reaction, for example, in microbiologically influenced corrosion of stainless steels.

In shallow immersion, pitting depths on carbon steel can be an order of magnitude higher than general corrosion rates calculated from mass loss as illustrated in **Figure 16**.[57] Determination of pitting rates

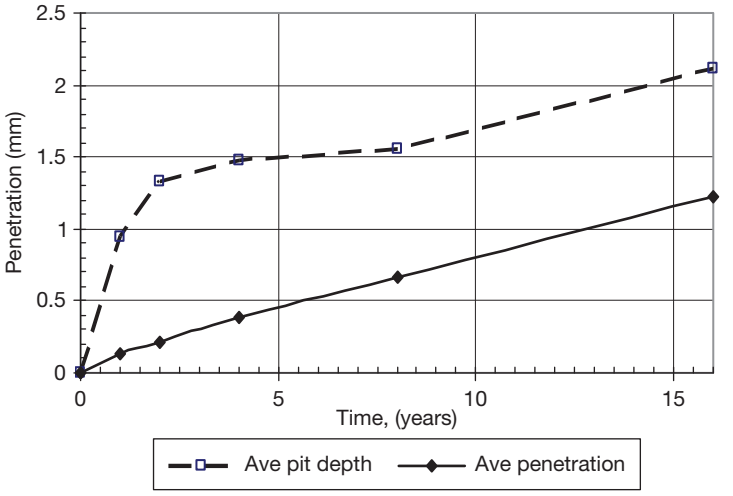


Figure 16 Corrosion of carbon steel in natural seawater. Average penetration calculated from mass loss of six replicate test panels – slope corresponds to ~0.07 mm year^{-1} corrosion rate; average pit depth is the mean of 60 measured pits, five deepest pits on each surface of six test panels. Reproduced from Southwell, C. R.; Alexander, A. L. *Mater. Prot.* **1970**, *9*(1), 14–23.

can be an onerous task. A common but mistaken practice involves measuring the maximum or average pit depth and calculating a pitting rate. There are a number of flaws with this approach. First, pitting is not linear with time, especially for long-term exposures in seawater. Second, the data are usually based on measuring different pits on different test samples, that is, specific pits are not monitored as a function of time owing to the obvious experimental shortcomings of this approach. Pitting is a statistical process; the number, size, and locations of pits on a given metal surface can vary over a wide range. Thus, computing pitting rates from limited data, and assuming linear growth rate with time is fraught with problems.

Figure 17 shows the general corrosion rate of carbon steels exposed for 1 year and DO content as a function of seawater depth.[58] It is apparent that the corrosion profile closely follows that of the oxygen concentration. The following relationship derived from linear regression analysis[59] indicates a greater dependence on oxygen content compared to temperature:

$$\text{Corrosion rate } (\mu\text{m year}^{-1}) = 21.3 + 25.4[\text{O}_2(\text{ml l}^{-1})] + 0.356[T(^{\circ}\text{C})]$$

Other data have indicated more complex corrosion behavior. For example, in unpolluted seawater, corrosion rate was dependent on oxygen content, whereas in polluted seawater, temperature had a greater influence.[60] Polluted seawater typically has the following characteristics (compared to unpolluted seawater): lower oxygen content, lower pH, higher sulfide and sulfate-reducing bacteria concentrations, and higher ammonia content. Corrosion data suggest that at similar temperatures, the effect of lower oxygen content in polluted seawater is offset by the presence of pollutants,[60] especially sulfides.

Under immersion conditions, low alloy steels behave fairly similar to carbon steels. Small amounts of alloying additions do not evidently alter the corrosion rate of steel significantly in seawater. However, as the alloy content is increased further (e.g., >5 wt%, especially Cr addition), a reduction in general corrosion rate is often replaced by deep, localized pits. It has been reported that low-alloy steels can cause galvanic corrosion of plain carbon steels despite a small difference in their corrosion potentials.[61]

Higher strength carbon steels are susceptible to hydrogen embrittlement (HE) in seawater, especially when polarized to excessively negative potentials. Factors that affect the susceptibility of steels in seawater to HE include yield strength (which is influenced by composition and heat treatment), electrochemical potential, and pollutants (especially sulfide). In general, susceptibility increases with increasing yield strength. High-strength steels are generally considered to be those with YS >700 MPa. Steels with YS <700 MPa are generally considered resistant to hydrogen embrittlement. Since alloying additions affect yield strength, it is difficult to delineate their influence on HE behavior. However, limited data[62] on steels of varying compositions but similar yield strengths indicate negligible

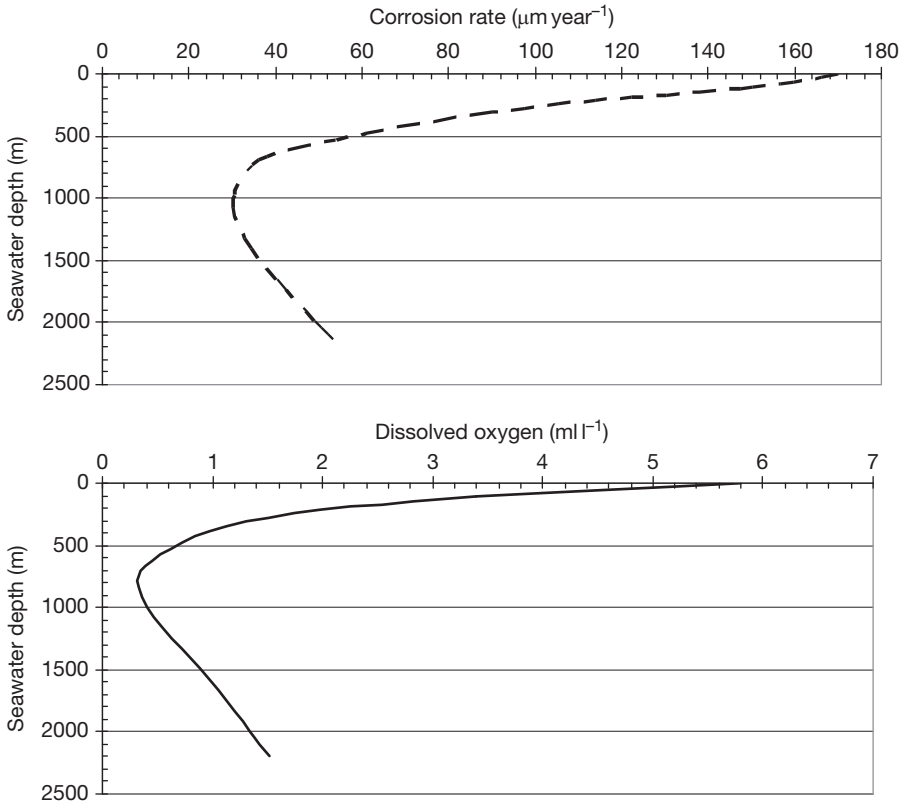


Figure 17 Corrosion rate of carbon steel exposed for 1 year and DO concentration versus seawater depth (DO ml $l^{-1} \times 1.429 = mg\,l^{-1}$); note similarity in profiles versus depth. Reproduced from Reinhart, F. M. Corrosion of Metals and Alloys in the Deep Ocean, Technical Report No. R-834, Civil Engineering Laboratory, Port Hueneme, CA, 1976; p 14.

effects of S or P (each <0.03%), Si or Cr (each <2.1%), Mo (<1.2%), and Co (<3%). Detrimental effects were indicated for C (0.3–0.4%), Mn (0–3%), and slightly deleterious for Ni (>6%). Tempering or aging heat-treatments at higher temperatures lower yield strengths and hence typically increase HE resistance. Ultra high-strength steels (e.g., with YS $\geq\sim$1400 MPa) can suffer HE even in distilled water. Elevated temperatures increase crack growth rates, probably as anodic dissolution rate increases, following Arrhenius behavior.[63]

Polarization of high-strength steel to more noble potentials than the free corrosion potential can markedly increase HE propensity as indicated by significantly shorter time to failure. Cathodic polarization to a narrow potential range (e.g., −0.8 to −0.9 V SCE) is generally beneficial since anodic dissolution processes are stifled. However, cathodic polarization to more active potentials (especially more negative than −1 V, SCE) again lowers cracking resistance. Pollutants, in particular sulfur-containing species, have been found to increase susceptibility to HE.

Carbon steel has been used extensively in the past for construction of flash chambers in thermal desalination plants. In principle, the corrosion rate should be negligible if DO is completely stripped from the seawater used in the process. However, in practice, it is very difficult to deoxygenate the seawater fully. Typically, oxygen concentration is reduced from ppm to ppb levels. However, increasing temperature has a far greater effect on corrosion than low oxygen levels. For example, data from a thermal desalination pilot-plant test revealed that even at a 100 ppb DO level and seawater flow-rate of $\sim$1.7 m s^{-1}, the corrosion rate of carbon steel doubled from 0.3 mm $year^{-1}$ at 81 °C to 0.6 mm $year^{-1}$ at 120 °C.[64] Oldfield and Todd[65] developed the following expression for calculating corrosion rate:

$$\text{Corrosion rate (mm year}^{-1}) = (0.0117 \times C_o \times U_o^{0.9})/Pr^{0.75}$$

where C_o is bulk concentration of oxygen in moles cm^{-3}, U_o is bulk flow rate in cm s^{-1}, and Pr is the

Prandtl number (defined as ratio of kinematic viscosity, v, to diffusion coefficient, D). The predicted corrosion trends were in good agreement with previously published field and test data; lower actual corrosion rates in the field were attributed to presence of corrosion products and scaling.

Millscale on steel surfaces is cathodic to the substrate. While the general corrosion rate of steel, based on mass loss data, may not appear significantly different, the presence of millscale can produce deep pitting by galvanic action due to the adverse cathode (millscale) to anode (exposed steel) area ratio.

2.18.5 Cast Irons

The initial corrosion behavior and corrosion rates of unalloyed gray and ductile cast irons in seawater are similar to those of carbon steels. However, in gray cast irons, as corrosion progresses, dissolution of the ferrite phase exposes increasing the amounts of the flake graphite network. This dealloying phenomenon is known as graphitic corrosion. It is commonly but incorrectly referred to as graphitization – which is really a metallurgical term used to describe decomposition of iron carbide (Fe_3C) to iron (Fe) and graphite (C) at elevated temperatures. At low velocities, the iron corrosion products are often 'trapped' by the flake graphite network. In such cases, visual inspection can greatly underestimate material degradation because reduction in section thickness is not apparent. Ultrasonic inspection and/or destructive examination are usually necessary to monitor graphitic corrosion before catastrophic failure occurs.

Graphitic corrosion may advance to a stage where it can induce galvanic corrosion of other materials coupled to the cast iron. For example, initially steels are galvanically compatible with gray cast iron, while copper alloys are cathodic. However, graphitic corrosion of the cast iron can eventually generate or reverse the galvanic couple where the exposed graphite behaves cathodically to steels and copper alloys and accelerate their dissolution by galvanic action.[66] Graphitic corrosion and its consequences are not a major problem with ductile irons because the graphite is present as discrete nodules instead of a flake graphite network.

Addition of ~20% Ni to cast iron increases toughness and shock resistance. However, the austenitic microstructure is still susceptible to graphitic corrosion, if the graphite is present in the flake form. Depending on alloy composition, small Mg additions (e.g., ~0.1%) during production convert the flake graphite to nodular form. This significantly increases the tensile strength, impact resistance, and graphitic corrosion resistance. Corrosion rates of austenitic cast irons (commonly referred to by a trade name, NiResist®) in seawater are often an order or magnitude lower than gray or ductile irons. The benefit of increasing Ni content in the cast iron is particularly notable for erosion–corrosion resistance in seawater at high velocities as illustrated by the data in **Table 6**. Sporadic stress corrosion cracking (SCC) problems observed in austenitic cast iron seawater pumps have been attributed to improper stress-relief heat treatment.[67]

2.18.6 Stainless Steels

The addition of >11% Cr to Fe generates a Cr- and Fe-containing passive oxide film on the metal surface. There is a very marked reduction in general corrosion rate in seawater compared to carbon steel or low alloy steels. However, the stainless steels are highly susceptible to localized corrosion unless much more Cr and/or other alloying additions such as Mo, N, and Ni are also present. Thus, determination of general corrosion rates of stainless steels in seawater from weight loss results can be misleading because most of the metal dissolution is typically associated with pitting and crevice corrosion.

There are several hundred commercial grades of stainless steels, with widely ranging compositions and properties. Obviously, a large number of grades are unsuitable for seawater applications. Typically, stainless steels are divided into several broad classifications, *viz.* martensitic, ferritic, austenitic, superaustenitic, duplex, superduplex, and precipitation hardening

Table 6 Effect of increasing Ni content in cast irons on erosion–corrosion behavior in seawater at 8 m s^{-1} at 28 °C for 60 days

Alloy	*% Ni content*	*Ave corrosion rate (mm year^{-1})*
Cast iron	–	6.9
2% Ni cast iron	2	6.1
Type 1 Ni-Resist	13.5–17.5	0.74
Type 2 Ni-Resist	18–22	0.79
Type 3 Ni-Resist	28–32	0.53

Source: Shreir, L. L., Ed. *Corrosion*; Newnes Butterworths, 1979; Vol. 1, pp 3–106.

(PH). Even within each classification, there can be a wide range of compositions.

Martensitic stainless steels are susceptible to pitting, crevice corrosion, SCC, and hydrogen embrittlement (especially under cathodic protection) in seawater.

High-Cr ferritic grades, such as Type 29–4C (29Cr–4Mo), are considered to have much greater resistance to localized corrosion compared to more leaner grades such as Type 430 (17Cr). Ferritic grades are resistant to chloride stress corrosion cracking (Cl^- SCC), but susceptible to hydrogen embrittlement, especially when subjected to cathodic protection, unintentional or otherwise.

Despite their popularity, particularly for marine atmospheres, many austenitic grades (such as many of the 300-series) are unsuitable for seawater service; unfortunately, they are often selected because of their attractive mechanical and welding properties, that is, without a complete understanding of their limitations, especially corrosion resistance. They are susceptible to Cl^- SCC; susceptibility trend decreases with increasing Cr or Cr + Ni content. They are resistant to hydrogen embrittlement except in the severely cold-worked condition. Type 304 is highly susceptible to pitting corrosion in seawater; Type 316 is more resistant due to the Mo addition but certainly not immune. The biggest limitation of these grades is their high propensity to crevice corrosion, for example, under gaskets, O-rings, washers, sleeves, macrofouling deposits, silt, and so on which are discussed later. Superaustenitic grades have higher alloy content (~20% Cr, 25% Ni, and ~6% Mo, or more) for increased resistance to crevice corrosion; their higher Ni content also enhances resistance to SCC.

In seawater, the localized corrosion resistance of duplex stainless steels is similar or slightly better than many austenitic grades. They are slightly magnetic and susceptible to hydrogen embrittlement, for example, under cathodic over-protection conditions, or in the presence of hydrogen sulfide. Superduplex stainless steels have higher Cr and nitrogen contents (e.g., alloy 2705; 27Cr–5Ni–3Mo–0.2N), which improve corrosion resistance notably, but does not make them immune under all conditions. They are much more resistant to chloride SCC than austenitic stainless steels.

PH stainless steels are susceptible to pitting, crevice corrosion, SCC, and hydrogen embrittlement (especially under cathodic protection) in seawater.

The highly-oxygenated thin liquid film environment, and absence of biofouling in the splash zone represent conditions under which stainless steels generally perform well at ambient temperatures and if critical crevices are not present. Leaner compositions may develop superficial rust staining. Stainless steel sheathing fabricated from 300-series austenitic stainless steel (e.g., Types 304 and 316) has been successfully used for long-term splash-zone protection at ambient temperatures.

The leaner grades of stainless steel are susceptible to localized attack under biofouling attachments that usually proliferate in the intertidal zone. **Figure 18** shows classical crevice corrosion of Type 304 stainless steel under barnacles.

The data in **Table 7** show localized corrosion penetration for several stainless steels after 16 years exposure in the intertidal and full immersion zones in Panama.[69] There was complete perforation of the 6-mm thick plate material by localized corrosion for all the grades tested under full immersion conditions. In the intertidal zone, the martensitic Type 410 material was also completely perforated, while the deepest pits on the other three grades ranged from ~1 to ~3 mm. Type 304 stainless steel pits in seawater like Type 302 and Type 321 (Type 321 is an analog of Type 304L, but stabilized with Ti for averting sensitization in lieu of lowering the carbon content). Attack typically initiates under biofouling attachments and at cut edges where inclusions in the material are exposed prominently. The acidic environment generated in the pits by metal-ion hydrolysis often leads to subsurface attack in the form of tunneling. On stationary structures, such attack often progresses vertically downward due to gravity. The addition of 2–3% molybdenum in Type 316 stainless steel imparts some resistance, but

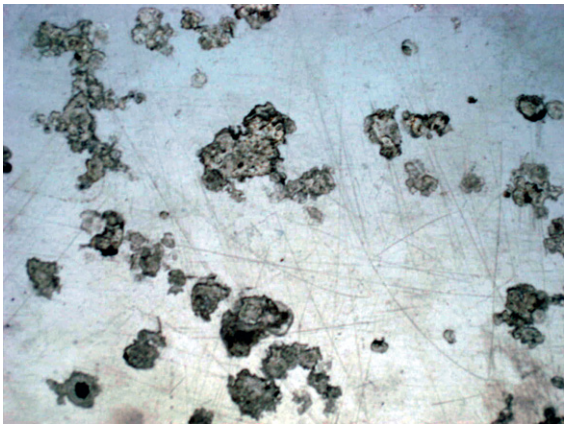


Figure 18 Crevice corrosion of Type 304 stainless steel under barnacles in natural seawater.

Table 7 Depth of localized corrosion attack on various stainless steels in natural seawater intertidal and immersion zones after 16 years' exposure

Alloy	*Exposure zone*	*Ave of 20 deepest pits 1 year*	*Ave of 20 deepest pits 8 years*	*Ave of 20 deepest pits 16 years*	*Deepest pit over 16 years*
410 SS	Tidal	1.17	1.70	2.72	P*
	Immersion	1.55	4.08	5.38	P
302 SS	Tidal	0.15	1.47	1.27	2.79
	Immersion	1.78	3.56	3.84	P
321 SS	Tidal	0.18	1.42	1.40	2.36
	Immersion	1.63	4.90	6.02	P
316 SS	Tidal	0.13	0.41	0.33	0.91
	Immersion	1.14	3.96	2.41	P

Source: Southwell, C. R.; Bultman, J. D.; Alexander, A. L. *Mater. Prot.* **1976**, *15*(7), 9–26.
P – Perforation (of 6-mm plate).

not immunity to pitting in seawater compared to Type 304. Sufficiently high flow rates that prevent macrofouling attachments would be expected to mitigate pit initiation. However, even in the absence of macrofouling, attack can initiate if the inclusion (e.g., MnS) concentration is high. For example, rampant pitting is encountered in Type 303 'free machining' stainless steels because they contain sulfur or selenium alloying additions, which produce copious electrochemically active inclusions in the microstructure.

Localized corrosion experienced with Type 304 stainless steel often leads to selection and specification of Type 316 stainless steel in many seawater applications. This is unfortunate despite a wealth of available information in the published literature. Virtually all 300-series stainless steels are susceptible to crevice corrosion in seawater. It is not widely appreciated that Type 316 stainless steel is just as susceptible to this form of localized corrosion as Type 304 in typical seawater applications containing crevices, for example, at flange faces, under gaskets, under O-rings, under washers, at threaded connections, under lap joints, and so on. Crevice corrosion can also occur under disbonded paint films, under biofouling, under silt, and so forth. There is another mistaken, but again, unfortunately, widely held view that low carbon 'L' grades of 300-series stainless steels are somehow more resistant to pitting and crevice corrosion in seawater. This is simply not the case. The 'L' grades are beneficial for mitigating intergranular attack, which would otherwise occur at sensitized areas of the microstructure, for example, heat-affected zones associated with welds. **Figure 19** shows typical crevice corrosion of a Type 316L stainless steel pump shaft bearing sleeve in seawater service. Thus, if crevices are present, and some form of corrosion-control,

Figure 19 Crevice corrosion of Type 316L stainless steel pump shaft sleeve in natural seawater.

for example, cathodic protection, is not applied, Type 316 stainless steel will crevice corrode in most aerated seawater applications. This fact is used in crevice corrosion testing in seawater where Type 316 specimens are routinely used as susceptible controls. In other words, the test is invalid if the Type 316 stainless steel controls do not crevice corrode.

Crevice corrosion of stainless steels in seawater has been investigated extensively over the past 30 years. The importance of DO transport to the stainless steel surface to maintain passivity is widely accepted. For example, the following minimum velocity relationship was derived from one laboratory study[70] on smooth-wall, Type 316 stainless steel tubes to prevent fouling in seawater:

$$v = 2.74 + 0.56 \log d$$

where v is the velocity ($\mathrm{m\,s^{-1}}$) and d is the tube inside diameter (m).

Another rule-of-thumb $\sim 2\,m\,s^{-1}$ minimum seawater velocity is commonly recommended to prevent settlement of silt and macrofouling under which localized attack can initiate and propagate. This type of approach is fraught with problems in practice because it is virtually impossible to maintain flow without interruptions. Experience indicates that macrofouling which attaches during extended stoppages is not usually removed when flow is resumed. Furthermore, it is very difficult to eliminate 'man-made' crevices completely in most components and seawater applications. The seawater electrolyte in the crevice is stationary even if high flow conditions exist outside the crevice. Thus, if a material is susceptible to crevice corrosion, maintaining sufficient flow to prevent fouling (outside the crevice) will not prevent attack within the crevice, as also discussed earlier.

The data in **Table 8** show that the maximum depth of crevice attack in seawater for three stainless steel alloys at 28 °C seawater temperature is greater than at 12 °C, but decreases markedly at 50 °C.[71] As discussed earlier, this behavior is often attributed to enhanced biofilm activity in warm seawater compared to cold or hot seawater.

As stated earlier, Cr and Mo are the primary alloying elements that influence localized corrosion resistance of stainless steels in seawater. Service experience[72] and tests[73] have shown that significant crevice corrosion resistance in ambient temperature seawater is attained with austenitic stainless steels of the nominal composition 20Cr–25–Ni–6Mo. However, even these 6Mo alloys have been shown to be prone to crevice attack if the seawater temperature exceeds their critical crevice corrosion temperatures (CCT), typically in the 30–45 °C range depending on the exact composition. Crevice attack on these alloys has also been demonstrated in tests conducted on painted panels even with relatively small areas left uncoated.[74]

Table 8 Maximum depth of crevice attack in natural seawater in 30-day test at three temperatures

Alloy	*Maximum depth of crevice attack (mm)*		
	12 °C	28 °C	50 °C
304 SS	0.56	0.81	0.12
316 SS	0.34	0.57	0.05
Alloy 20	0.18	0.20	0

Source: Asphahani, A. I.; Manning, P. E.; Silence, W. L.; Hodge, G. F. In *Corrosion/80*; NACE International: Houston, TX, 1980; Paper No. 29.

Duplex stainless steels are also potentially susceptible to crevice corrosion depending on alloy composition, crevice geometry, and service or test conditions. The higher Cr-containing alloy 2507 is generally more resistant to crevice corrosion than the lower Cr alloy 2205. Successful use of higher Cr duplex stainless steel has been reported in offshore and desalination seawater pumps, valves, and piping systems[75] normally at temperatures below their CCT (~40 °C) and chlorine levels of 0.5 ppm or less.

Critical crevice corrosion temperatures of stainless steels are usually determined using the ASTM G 48 standard test.[76] Test conditions are accelerated by using an aggressive test solution consisting of 6% $FeCl_3 + 1\%$ HCl. Test specimens are fitted with PTFE multiple crevice assembly (MCA) washers torqued to 0.28 N m and exposed to the test solution at various thermostatically controlled temperatures. The CCT of the test alloy is the temperature at which crevice attack ≥ 0.025 mm is detected after 72 h immersion. Critical pitting temperatures (CPT) also commonly reported in the literature refer to pitting corrosion observed in the absence of artificial crevices; any edge attack on the test specimens is discounted. For any given alloy, the CPT is always greater than the CCT. However, the latter is generally considered more meaningful for practical applications because it takes into account the unavoidable crevices that are either man-made (e.g., gaskets, O-rings, sleeves) or naturally occurring (e.g., barnacles, thermal scale, silt).

In addition to CCT, stainless steels are also frequently ranked for likely crevice corrosion behavior in seawater based on their pitting resistance number (PREN). The following is a typical empirical expression for PREN based on alloy (wt%) composition:

$$\text{PREN} = \%\text{Cr} + 3.3\%\text{Mo} + A\%\text{N}$$

where the value of A is typically chosen to be 30 for austenitic stainless steels, and 16 for duplex stainless steels. **Table 9** shows the PREN and CCT values for a number of stainless steels.[77] It is apparent that as the PREN increases, the CCT increases. The following expression was proposed[78] for austenitic stainless steels:

$$\text{CCT} = (2.7 \times \text{PREN}) - 0.81$$

At first sight, this appears to be a very convenient method for ranking alloy performance in terms of expected localized corrosion resistance. However, for reliable materials selection, neither the PREN nor the CCT (as determined by the ASTM G 48

Table 9 Typical PRE numbers versus CCT and CPT values for various stainless steels

Alloy	*UNS No.*	*PRE*	*CCT (°C)*
316L	S31603	26	<5
317L	S31783	28	0
2205	S31803	34	17.5–25
904L	N08904	36	15–25
2507	S32750	43	35–43
254SMO	S31254	46	30–60
AL–6XN	N08367	47	43
654SMO	S32654	63	60+

Source: Davies, M.; Scott, P. J. B. *Guide to Use of Materials in Waters*; NACE International: Houston, TX, 2003; p 90.

test method) provides a consistently accurate prediction of actual crevice corrosion resistance in seawater service. In other words, even alloys with high PREN and CCT values have been shown to crevice corrode, for example, in tests with severe crevice geometries.[74]

In crevice corrosion testing in seawater, crevice geometries can be created in a number of ways. For example, for flat sheet or plate stock, nonmetal or metal crevice-forming washers are commonly used as bolted assemblies; rubber O-rings and vinyl sleeves are used for tubular and round bar, and so on. Some of these devices and methods are described in ASTM standard G 78.[79] For irregular-shaped components or surfaces, for example, protruding weld beads, that do not readily permit the use of conventional crevice-formers, the use of paint coatings as crevice-formers has been proposed.[80] This methodology is based on the observation of crevice attack under coatings (e.g., marking inks, paint overspray, disbonded coatings, etc. on stainless steel surfaces) in actual practical applications.

Surface finish can also affect crevice corrosion resistance of stainless steels. The generally greater resistance of rough-ground surfaces to crevice attack, compared to smooth polished ones, is explained in terms of crevice tightness with respect to the crevice former. A certain degree of surface grinding and/or pickling is considered essential to remove the surface layer that is considered to have suffered some Cr depletion during manufacturing and/or fabrication (e.g., hot rolling, annealing, welding, etc.). The surface treatment is also intended to remove any extraneous particles (e.g., embedded iron) from tools, which can impair corrosion resistance; removal of embedded iron by immersion in acid solution (e.g., HNO_3) is often erroneously referred to as 'passivation.' Chromium carbide ($Cr_{23}C_7$) precipitation occurs at grain boundaries, typically in the 500–800 °C temperature range during welding and heat treatment of austenitic stainless steels. Corrosion resistance of areas adjacent to the carbide particles is lowered because of Cr depletion. However, despite this, no significant incidences of intergranular corrosion in ambient temperature seawater have been reported. On the other hand, precipitation of sigma phase (FeCr intermetallic), which can form in the temperature range 565–925 °C in Mo-containing austenitic steels, is considered more deleterious for seawater corrosion resistance.[81] Although their compositions are not exactly the same, cast stainless steels often exhibit lower resistance to localized corrosion than the corresponding wrought material. In the case of austenitic stainless steels, this is generally due to the residual ferrite phase in castings (necessary to counteract shrinkage during solidification). Welds represent 'miniature' castings, and hence their corrosion resistance can be lower than the parent material of similar nominal composition. Laboratory tests have shown that pickling of welds increased the critical pitting temperature of a 25-Cr duplex stainless steel by ~20 °C[82]; presumably, the reason for this is elimination of heat tint and surface Cr depletion. A 'soft' start up has been recommended in situations where pickling is not possible in piping systems; apparently the procedure consists of initial exposure to cold seawater for at least 2 days, followed by exposure to cold chlorinated seawater for at least 5 days, before turning on the heat exchangers.[75] This method may not be very practical in geographical areas where cold seawater is not readily available.

As stated earlier, the type of crevice geometry and its dimensions can play a critical role in the determining crevice corrosion resistance of stainless steels. Generally, crevice corrosion severity increases with decreasing crevice gap (tightness) and increasing crevice depth. This has been demonstrated by mathematical modeling[83] and by experimental evaluations.[84,85] The critical values of pH and chloride concentration of crevice solutions to initiate crevice attack have also been determined[83]; results for four different stainless steels are depicted in **Table 10**. In crevice corrosion testing, ideally, the actual crevice geometry (gap and depth) and crevice-formers (gaskets, washers, O-rings, etc.) intended for the service environment should be used. Lack of proper simulation could result in incorrect indication of corrosion behavior in service, for example, insufficiently tight and deep crevices could lead to a false prediction of corrosion resistance and vice versa.

Table 10 Critical crevice corrosion solution values

Alloy	*[Cl⁻] (M)*	*pH*
430 SS	1.0	2.90
304 SS	2.5	2.10
316 SS	4.0	1.65
904L	4.0	1.25

Source: Oldfield, J. W.; Sutton, W. H. *Br. Corros. J.* **1980** ,*15*(1), 31–34.

Polluted seawater containing sulfides have not adversely affected the performance of stainless steels in seawater,[86,87] both at moderate ($2.4\,m\,s^{-1}$) as well as high velocities[88] (up to $50\,m\,s^{-1}$). This observed corrosion resistance is based on the results of testing and experience in seawater, which either already contained sulfide or to which sulfide was deliberately added. Lean stainless steel grades can however, be subjected to microbially influenced corrosion (e.g., especially at welds) where the sulfide is generated biogenically as a byproduct, especially under quiescent conditions.

In the absence of solid particulate matter (e.g., entrained sand), stainless steels can generally tolerate seawater velocities in the order of $40\,m\,s^{-1}$ without any erosion–corrosion problems.[88] However, in service application such as pumps and valves, proper hydraulic design and operation are essential to preclude cavitation problems. Stainless steels are not immune to cavitation, but typically exhibit greater resistance than 'softer' materials such as carbon steels, copper-base alloys, and aluminum-based alloys.

In the passive condition, stainless steels are galvanically compatible with each other in seawater despite their somewhat different positions in the galvanic series. In the passive condition, stainless steels are not usually adversely affected by galvanic corrosion when coupled to more noble materials, for example, titanium, gold, graphite, etc. As stated earlier, when localized corrosion of stainless steel is observed at areas in intimate contact with such noble materials, there is a common tendency to attribute attack on the stainless steel to galvanic corrosion. Tests conducted with inert gasket materials, such as PTFE and graphite-impregnated gaskets have shown that attack on 300-series stainless steel flanges is just as severe as, or more so, with the PTFE gaskets compared to graphite-containing ones.[44] The attack on the stainless steel can be explained as crevice corrosion with both types of gasket. Conversely, it can be argued that if it was galvanic corrosion, graphite should produce attack on even those stainless steels that exhibit resistance to crevice attack with PTFE gaskets. It has been shown that high alloy stainless steels are adversely affected when coupled to graphite in chlorinated but not natural seawater.[89] But this is also the case when stainless steels with nonabsorbent inert gasket materials are subjected to chlorine levels >2 ppm.[90] It has been reported that crevice corrosion associated with synthetic fiber gasket materials which absorb water is less severe because of dilution of the crevice electrolyte.[91] This phenomenon has been used to advantage in the design of an electrochemical corrosion cell where crevice corrosion of a flat-stock test specimen is eliminated by flooding the specimen sealing gasket with distilled water.[92]

Crevice corrosion has been investigated electrochemically in a number of ways.[80] For example, in a remote crevice assembly, crevice formers are used to sandwich a small specimen (anode) which is then coupled to a separate, larger cathode specimen made of the same alloy as the anode. In an alternative compartmentalized cell design, the anode and cathode are exposed to two different solutions (anode – acidic, deaerated, high Cl^- solution; cathode – neutral, aerated solution), respectively. In both these designs, the crevice corrosion current associated with the couple is usually measured using a zero-resistance ammeter (ZRA). Although, the anode and cathode members are made from single pieces of the test alloy to eliminate unwanted crevices and immersed connections, unfortunately localized attack is sometimes observed on one or both electrodes at the waterline. Such attack is usually associated with salting out due to evaporation of the test solutions.

Comparisons have been made between pitting corrosion of stainless steels and crevice attack. This stems from the similarity observed between chemistry (low pH, high Cl^-) of the electrolyte trapped in a pit and in a crevice. Pits can be regarded as microcrevices (e.g., initiating at sulfide, inclusions) associated with the microstructure, extraneous embedded particles (e.g., iron), under heat tints, 'peaks' folded over by mechanical processes such as coarse machining, grinding, and so on. **Table 11** shows that pitting corrosion of Type 316 stainless steel is reduced by more than an order of magnitude in 'deaerated' seawater at 105 °C, but not eliminated because 25 ppb O_2 is probably sufficient to initiate and sustain some attack. The degree of localized corrosion attack on susceptible stainless steels shows a decreasing trend with exposure depth in the ocean

Table 11 Maximum pit depths on stainless steel in natural and deaerated seawater[93]

Alloy	*Seawater conditions*	*Exposure (days)*	*Max. pit depth (mm)*
316 SS	Aerated, natural, ambient temperature	486	2.4
316 SS	Deaerated, 105 °C, 25 ppb DO	547	0.12
304 SS	Deaerated, 105 °C, 25 ppb DO	547	0.60

as the dissolved O_2 content decreases in the deeper water; however, this trend is general but not necessarily always one-to-one, indicating that other factors are probably implicated in the corrosion process.

In ambient temperature seawater, nonsensitized austenitic and ferritic stainless steels are considered resistant to stress corrosion cracking (SCC). SCC of Type 316 stainless steel was reported in the vapor zone above boiling seawater, but not on fully immersed specimens.[81] Since oxygen is usually essential for SCC, the likely explanation is that boiling eliminated dissolved O_2 from the seawater. Other data indicated that SCC of Type 316 stainless steel in hot seawater (126 °C) containing 600 ppb O_2 was mitigated when the O_2 level was reduced to 30 ppb.[93]

Martensitic and precipitation-hardening (PH) stainless steels are more prone to SCC. In deep-ocean water tests,[94] SCC of 15–7PH in the RH950 and RH1150 condition, and 17–7PH in the RH1050 condition, was reported. No correlation between SCC susceptibility and tensile strength was observed. These failures occurred at low O_2 levels, leading to the inference that hydrogen embrittlement may be the operative mechanism, for example, due to acidity generation associated with localized corrosion. Extrapolation of results of accelerated SCC tests in boiling NaCl and $MgCl_2$ solutions (e.g., per ASTM standard G 38) to seawater applications should be done very cautiously. Although temperatures >50 °C are often cited as necessary for SCC of stainless steels, experience indicates that failures can occur even at ambient temperatures where localized corrosion (e.g., crevice attack) generates low-pH conditions; of course, the material has to be in a susceptible condition, for example, sensitized, containing martensite, or containing high tensile stresses, and so on. Dynamic tests such as slow strain rate and crack growth rate are now strongly favored over static tests which may otherwise falsely indicate resistance, because the former expose the base metal by almost constantly deforming and cracking the passive film above certain strain rates. Another major attribute of the dynamic SCC tests is the much shorter testing time compared to the static tests.

The greater SCC resistance of cast stainless steels over wrought material is usually ascribed to residual ferrite phase in the former. Polluted or putrid seawater containing sulfides may increase the susceptibility to hydrogen embrittlement of martensitic and PH and stainless steels; this propensity is increased when the material is deliberately or inadvertently under cathodic protection. Duplex stainless steels may also be susceptible under the latter condition. If SCC propensity decreases above a certain threshold temperature, especially in sulfide environments, it is reasonable to infer that the cracking mechanism is hydrogen embrittlement since outward diffusion of atomic hydrogen is temperature dependent. On the other hand, cathodic polarization increases the susceptibility to hydrogen embrittlement.

SCC failure of Type 316 stainless steel in venting systems of MSF desalination plants has been attributed to bromine.[95] Chlorination of seawater for biofouling control converts the bromide ions present in the seawater to bromine, which dissolves to produce hypobromous and bromic acids. If ammonia is present, bromoamines are also produced. Acidification of the seawater downstream for scale control releases hydrogen bromide into the vapor, which can then attack the susceptible stainless steels such as Type 316 in the venting system:

$$Cl_2 + 2Br^- \rightarrow Br_2 + 2Cl^-$$
$$Br_2 + H_2O \rightarrow HBr + HOBr$$

Cavitation susceptibility and ranking are often performed in laboratory tests in which the test specimen is vibrated ultrasonically in seawater by a magnetostrictive transducer. Cavitation resistance is based on mass loss per unit time and visual examination. Results for a number of alloys are discussed in the literature.[96–98] Such laboratory tests have shown decreased cavitation damage due to cathodic protection.[97]

2.18.7 Copper Alloys

Copper alloys rely on the formation of protective corrosion product films for corrosion control in marine environments. These films are sometimes described as 'passive.' In the author's opinion, the term 'passive film' should be reserved for situations

(materials and conditions) where the film is very thin (typically <100 nm), optically transparent, extremely adherent, ennobles the potential of the metal appreciably, and not formed as a 'voluminous' deposit by the incorporation of species from the environment. For example, films on stainless steel, titanium, and aluminum would be considered as passive. The protective corrosion product films on copper alloys are much thicker and also quite complex; in clean seawater, copper hydroxychloride is often a major component.

In the splash and intertidal zones, 90/10 copper–nickel sheathing has been successfully used for corrosion control on offshore platforms. Unlike paint coatings, such sheathing requires little maintenance – which is extremely difficult in the splash zone as discussed earlier. Because of high resistance to macrofouling, copper–nickel has also been used for the construction of small boat hulls and as adhesive-backed foils on ferries and fireboats.[99] Macrofouling resistance is apparently reduced when the copper–nickel is galvanically coupled to a more anodic material such as steel, or if it is subjected to cathodic protection, *albeit* inadvertently. However, the macrofouling attachments are relatively easily removed compared to bare or coated steel.

Copper alloys are very desirable for seawater applications because of high resistance to biofouling. They have been widely used for heat exchanger tubing because of high heat transfer properties. However, they are relatively soft and hence subject to erosion–corrosion in seawater above certain velocities. The high turbulence in seawater pumps restricts the choice of copper alloys largely to nickel–aluminum–bronze. As a general rule of thumb critical velocities for small diameter heat exchanger tubing determined from laboratory testing and field experience are shown in **Table 12**.[100] These values are still widely used in design and operation of copper-alloy heat exchangers. It has been postulated that each copper-based alloy has a critical shear stress associated with it.[101] When this shear stress exceeds the adhesion strength of the protective corrosion product film, attack of the base metal ensues. Critical shear stresses, depicted in **Table 13**,[101] determined using fluid mechanics and evaluated by laboratory testing now support the view that as the piping diameter increases, higher critical velocities than those shown in **Table 12** can be used without increasing the risk of erosion–corrosion for the alloy in question.

One of the most important corrodents toward copper alloys in seawater service is sulfide. Sulfide can be entrained in the seawater (e.g., from bottom sediment or polluted harbors) or generated *in situ* by sulfate-reducing bacteria, as discussed previously. The latter mechanism is very common when initially clean seawater putrifies after a few days of stagnation, for example, in heat exchangers, pumps, and piping systems during outages or shutdowns. The role of sulfides has been extensively investigated.[102–108] The evidence suggests that protective corrosion product films allowed to form on copper alloys in clean seawater for several months can resist short subsequent exposure to sulfides. Conversely, exposure to sulfide conditions from the outset interferes with the formation of desirable protective films. In such cases, sulfides have a deleterious effect on pitting corrosion and erosion–corrosion resistance. Thus, commissioning of copper alloy equipment such as heat exchangers using clean seawater for several months initially appears to be critical. However, this task should not be underestimated as access to large quantities of clean seawater at coastal areas is difficult

Table 12 Rule of thumb maximum velocities for copper alloy heat exchanger tubing in seawater

Alloy	*Suggested max. design velocity ($m\,s^{-1}$)*[a]
Copper	0.9
S-bronze	0.9
Admiralty brass	1.5
Aluminum brass	2.4
90/10 CuNi	3.0
70/30 CuNi	3.7

[a]Velocities may be increased by 0.3–0.6 $m\,s^{-1}$ in deaerated brines encountered in heat recovery exchangers in desalination plants.
Source: LaQue, F. L. *Marine Corrosion: Causes and Prevention*; Wiley, 1975; p 267.

Table 13 Critical velocities and shear stresses for copper alloys in seawater

Alloy	*UNS No.*	*Test temp (°C)*	*Critical velocity ($m\,s^{-1}$)*	*Critical shear stress ($N\,m^{-2}$)*
DHP Copper	C12200	17	1.3	9.6
As-inhibited Al brass	C68700	12	2.2	19.2
90/10 CuNi	C70600	27	4.5	43.1
70/30 CuNi	C71500	12	4.1	47.9
Cr-modified 90/10 CuNi	C72200	27	12.0	296.9

Source: Efird, K. D. *Corrosion* **1977**, *33*(1), 3–8.

at best. Experience indicates that once incorporated in the corrosion product film, sulfides are very difficult to remove completely and problems continue.

A classical laboratory study[102] of the effects of sulfide on 90/10 and 70/30 CuNi in seawater disclosed the following: corrosion rate increased as the sulfide content was increased from 7 to 250 ppb; attack was localized at lower sulfide concentrations, and more general at higher ones; once attack initiated, there was minimal seawater velocity effect over the 0.5–5.3 m s^{-1} range; attack continued even in clean seawater after initial exposure to 1000 ppb sulfide for 1–5 days, or 10 ppb sulfide for 30 days. Addition of 0.2 ppm ferrous ions which is sometimes used to mitigate erosion–corrosion was found to reduce but not eliminate sulfide-related corrosion of the CuNi alloys.[103]

Protective corrosion product films reduce the corrosion rate markedly in clean seawater with time.[109] In the ambient temperature range, the corrosion rate increases with increasing temperature. However, at high temperatures, such as in MSF desalination plants, the corrosion rate decreases with increasing temperature[110] in the 54–121 °C temperature range as illustrated in **Figure 20**. This may be due to the lower oxygen levels and precipitation of scales due to inverse solubility. However, at elevated temperature, corrosion rate increases with increasing dissolved O_2 even in the ppb range as depicted in **Figure 21**.[111] Decrease in corrosion rate has also been reported for deep waters where oxygen levels were lower.[112]

It should be emphasized that copper alloys with >70% copper are resistant to biofouling compared, for example, to carbon steel, stainless steels, aluminum, titanium, but not completely immune. Macrofouling resistance of copper alloys is usually explained in terms of either continuous, but slow release of copper ions, or presence of copper-rich corrosion products on the metal surface – both of which are purportedly 'toxic' toward macroorganisms.[113,114] The situation is confounded by the fact that while increasing alloy content lowers the copper ion release (and hence likelihood of fouling), it also affects the nature of corrosion products. It is not known why, for instance, barnacles attach at some but not most areas of such copper alloys. It has been suggested that attack on the metal under macrofouling occurs after the organisms die.[113] Fouling is more likely in quiescent than flowing seawater. A minimum velocity of 1–2 m s^{-1} was proposed to mitigate macrofouling.[115] Higher velocities to prevent macrofouling may increase the risk of erosion–corrosion.

Pitting corrosion of copper alloys can occur under deposits due to differential aeration, and is often exacerbated by sulfides. Another factor that can increase pitting susceptibility is the presence of residual carbon films on the copper surface.[116] The carbon is believed to derive from oxidation of lubricating oil residues left behind from manufacturing (e.g., tube drawing) and subsequent annealing heat treatment. It has been hypothesized that the carbon film acts as a cathode and promotes localized corrosion of the base metal at exposed areas in the film. Ideally, the carbon film should be removed after

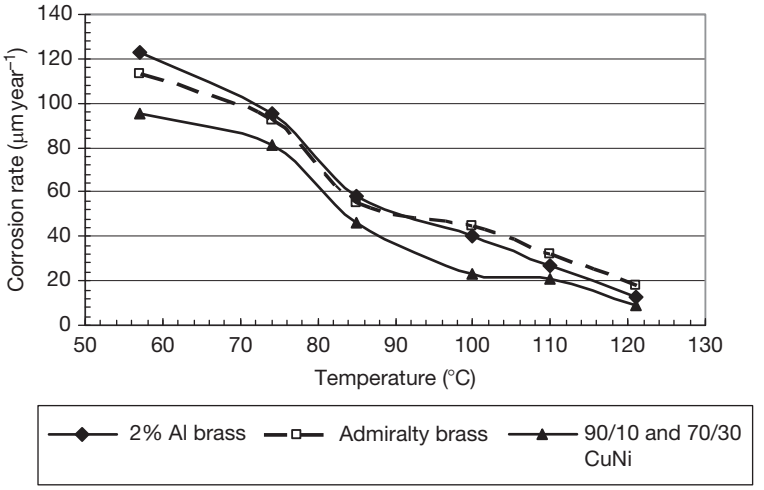


Figure 20 Corrosion rate versus temperature for three copper-base alloys in seawater; DO 40–600 ppb, pH 6.2–7.8, 156 days. Reproduced from Mattson, M. E.; Fuller, R. M. A study of materials of construction in distillation plants, Joint report by Office of Saline Water and the International Nickel Company, Research report No. 163, 1965.

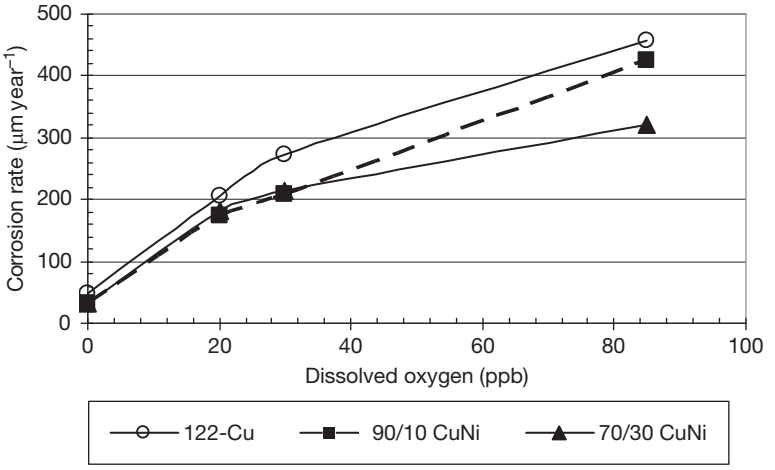


Figure 21 Corrosion rate versus DO in seawater for three copper-base alloys at 107 °C in a once-through desalination system; seawater flow rate ~2 m s^{-1}, pH 7.2–7.5, CO_2 <10 ppm, 15–30 days. Reproduced from Schreiber, C.; Osborn, O.; Coley, F. H. *Mater. Perform.* **1968**, *7*(10), 20–25; NACE International: Houston, TX.

manufacturing. Detection of the carbon film is not always easy. Since the carbon film is insoluble, one method involves placing a drop of nitric acid on the test surface and visually observing any flakes of thin gray/black matter float up to the top. The films are often even more difficult to detect after corrosion failures have occurred. Potential measurement has been proposed as a detection method in such cases.[117]

A number of copper-based alloys tested concurrently in an ancillary seawater condenser operated for over 8 months revealed that Admiralty brass (70/30 Cu/Zn) suffered from impingement corrosion at the inlet and outlet ends as well as adjacent to deposits.[118] Iron and manganese additions to copper–nickel alloys have been shown to enhance erosion–corrosion resistance.[119] However, some other tests indicated iron additions to be deleterious,[120] especially >1%.[121] This was probably related to the distribution of iron in the microstructure. Uniform distribution improves corrosion resistance, while segregation increases susceptibility to localized attack. A simple magnetic permeability test is used to detect iron segregation.[122] Small Cr additions to copper–nickel alloys improves erosion–corrosion resistance; however, localized corrosion resistance in quiescent conditions may be impaired by Cr segregation. Any restrictions on flow caused by macrofouling (e.g., shells) or pebbles lodged in the tube can cause severe local turbulence, and hence erosion–corrosion or impingement attack at associated (usually immediately downstream) areas. Experience shows that less protective scales are formed as the seawater pH is lowered, for example, by pollutants or dissolved CO_2. This is not unexpected since corrosion products and mineral scales are more difficult to form at lower pH values. Absence of protective scales at lower pH will increase erosion–corrosion tendency.

The role of gas bubbles in promoting erosion–corrosion has been discussed.[121,123] In rotating spindle and disc tests, air bubbles in the seawater were stated to be beneficial toward CuNi alloys.[121,124] However, results of impingement attack in flowing loop tests indicated that both air and nitrogen bubbles were harmful.[121] Since collapsing bubbles are thought to be the cause of cavitation damage, copper alloys, being considerably softer are more likely to be affected by this form of damage compared to, say, stainless steels, nickel-based alloys, and titanium.

Investigations on the effect of entrained sand particles on erosion–corrosion resistance of aluminum brass indicated the following: erosion–corrosion increased with increasing sand content (>300 ppm) and particle size[125]; cathodic protection with or without ferrous iron injection was beneficial when the sand particles were 50 μm in size, but not when they were 250 μm[126]; less attack in polluted seawater containing sand was attributed to protective sulfide films.

Galvanic corrosion of copper alloys is problematic in seawater when they are coupled to cathodic materials such as stainless steels, nickel-based alloys, titanium, and graphite, especially for small anode/cathode area ratios. In chlorinated seawater, a reduction in galvanic corrosion of copper alloys coupled to stainless steels has been observed[43,128] and ascribed to reduced biofilm

activity, which affects cathodic polarization kinetics, and hence galvanic corrosion rate.

Copper alloys are cathodic to many materials, and can thus cause unacceptable galvanic corrosion of anodic materials such as carbon steel, cast iron, aluminum, and zinc – particularly for large cathode/anode area ratios. Gray cast irons are initially anodic to copper; however, graphitic corrosion of gray cast iron can eventually expose enough graphite at the exposed surface so as to reverse the galvanic couple and accelerate corrosion of the copper alloy. In principle, the free corrosion potentials of copper-based alloys in seawater are fairly close to one another, and thus one might expect them to be galvanically quite compatible. In principle, this is the case. However, regardless of potentials, an adverse cathode/anode area ratio between two copper alloys can produce unacceptable galvanic attack on the more anodic material. Increasing flow rate and temperature typically accelerates galvanic corrosion of copper alloys. Presence of sulfides (especially 50 ppb) in seawater was found to promote galvanic attack on Naval M-bronze and nickel aluminum bronze fittings coupled to 90/10 or 70/30 CuNi piping.[128] The effect of sulfides appears to be greater at lower seawater temperatures.[129]

Electrically isolated copper components in a system can undergo self-corrosion in seawater environments. The release of even imperceptible amounts of copper ions can pose an indirect problem for anodic materials that come in contact with them. For instance, copper ions can plate out on aluminum, steel, and zinc, leading to deposition or cementation corrosion. In other words, copper ions plate out on the anodic material (as a cathodic reaction) and cause corrosion of the base metal. However, once the ions have plated out, galvanic corrosion (usually as pitting) of the base metal ensues even if the original source of the copper ions is removed. Copper-colored areas surrounding pits in the anodic material are usually a telltale sign of this phenomenon. For example, it has been observed in aluminum fuel tanks on pleasure boats, particularly those operating in seawater environments. Copper corrosion products enter the tank from water condensate in copper piping and/or copper alloy valves in the fuel system.

Copper-based alloys are quite resistant to SCC in clean seawater. SCC of brasses is generally associated with ammonia, amines, or nitrates that are either entrained as pollutants or can form *in situ* under decaying biofouling attachments. Copper–nickel alloys are considered resistant to such environments. Stress relief of cast nickel aluminum bronze alloys is recommended to minimize the risk of SCC failure in critical seawater service applications.

Dealloying is quite common in many copper-base alloys. It is often observed in marine artifacts submerged for long periods in seawater. Reddish coloration due to copper deposition is a telltale sign of dealloying. Alpha brasses (e.g., 85/15 Cu/Zn) can be inhibited against dealloying with Sn or As additions. Unfortunately, such additions are usually not considered to be as effective in beta brasses (e.g., 60/40 Cu/Zn). Unlike in brasses, dealloying in copper–nickels is usually superficial. In nickel aluminum bronzes, dealloying involves the selective dissolution of aluminum from the alloy. Proper heat treatment to minimize the beta phase alleviates this problem.

Copper-base alloys can suffer from another form of localized attack known as metal-ion concentration cell corrosion. Attack usually occurs at the mouth of a crevice, in contrast to localized corrosion that occurs inside the crevice, for example, in susceptible stainless steels and nickel-base alloys.

2.18.8 Nickel-Based Alloys

The primary nickel-based materials used in seawater include nickel–copper alloys (e.g., the 70/30 NiCu alloy 400, or its high strength version, alloy K-500) and Ni–Cr–Mo–Fe alloys (e.g., alloy 625, C-276, alloy 59, alloy 686, etc.). As with stainless steels, these nickel-based alloys are also believed to depend on the formation of passive films for corrosion resistance.

Because of its established record of good performance in the splash zone, alloy 400 has been successfully used as sheathing for protection of carbon steel in the splash zone.[46] Unlike the copper–nickel alloys described in the previous section, the NiCu and other nickel-based alloys are not resistant to macrofouling in the tidal and immersion zones. In fact, nickel-base alloys tend to behave more like stainless steels. General corrosion rates in clean seawater are low to negligible, and when attack occurs, it is usually as pitting and/or crevice corrosion; corrosion initiation is more likely in quiescent or low flow seawater.

For unalloyed nickel and nickel–copper alloys (e.g., alloys 400 and K-500), pitting in low flow seawater is common. Unlike with susceptible stainless steels, the pits tend to be open, hemispherical, and do not grow significantly in depth, for instance, after ~1 year in seawater.[130] Alloying nickel with copper appreciably improves the pitting resistance. Seawater

temperature has been shown to dramatically affect pitting corrosion behavior of alloy 400; maximum pit depth increased in the 18 to 30 °C range, but then decreased markedly at 50 °C.[131] A similar trend was discussed earlier for crevice corrosion of stainless steels and attributed to 'degraded' biofilms at 50 °C. The pitting potential of alloy 400 determined electrochemically decreased with temperature.

Corrosion rate of alloy 400 has been shown to increase as the sulfide concentration in seawater increased from 0 to 5 ppm.[128] Ni–Cr alloys were not affected by sulfides in ambient temperature seawater.

Alloying additions of ~20% Cr and ~8% Mo to nickel-based alloys (e.g., in alloy 625) impart significant resistance to pitting and crevice corrosion in seawater. As with stainless steels, addition of ~20% Cr alone does not ameliorate localized attack. Table 14 provides a comparison of relative corrosion resistance of several nickel-based alloys immersed in quiescent seawater for ~18 months.[132] Apparently, from a corrosion resistance viewpoint, cobalt in Ni-based alloys behaves like nickel. Alloy C-276 was found to be highly-resistant to localized and general corrosion in seawater up to 288 °C[71]; it also exhibited crevice corrosion resistance at 107 °C in deaerated seawater (DO 28 ppb).

Crevice corrosion of alloy 400 is not generally as severe as, say, for 300-series stainless steels. This can be explained by the fact that hydrolysis of Ni and Cu ions from alloy 400 does not produce a highly acidic and aggressive crevice solution compared to Cr and Mo ions from susceptible stainless steels.

The erosion–corrosion behavior of nickel-based alloys is quite similar to that of stainless steels, that is, notably superior to that of copper–nickel alloys. Thus, nickel alloys can tolerate high velocities, turbulence, and impingement which makes them candidates for seawater pumps, valves, piping systems, and so on. Test data summarized in Table 15[133,134] show that, like stainless steels, most nickel-based alloys exhibit acceptably low erosion–corrosion rates up to ~40 m s^{-1}. With the exception of the NiCu alloys, their cavitation resistance is also expected to be similar to that of stainless steels.

Nickel–copper alloys are often selected for seawater service because their higher Ni content provides enhanced resistance to chloride stress corrosion cracking (Cl^- SCC) compared to stainless steels. Deep ocean tests of 2 years duration confirmed the SCC resistance of Alloy 825 and Alloy C.[135] Alloys G-3, 625, and C-276 were resistant to SCC in U-bend tests of several hundred hours duration in synthetic seawater.[81]

Table 14 Localized corrosion resistance of some Ni-based alloys after ~18 months' exposure in quiescent ambient-temperature seawater

Alloy	*Max depth of crevice attack (μm)*
Unalloyed Ni	2.06
400	432
K-500	51
600	457
825	229
625	0
C 276	0

Source: Lennox , T. J., Jr. In *Corrosion/82*; NACE International: Houston, TX, 1982; Paper No. 64.

Table 15 Corrosion resistance of Ni-base alloys at ambient-temperature high velocity seawater

Alloy	*Seawater velocity ($m\,s^{-1}$)*	*Corrosion rate ($\mu m\,year^{-1}$)*
C	39	50
825	43	75
718	41	50
X-750	41	50
600	41	75
400	43	100
K-500	41	100

Source: Boyd, W. K.; Fink, F. W. Corrosion of Metals in Marine Environments, Battelle Columbus Labs, OH, Report No. MCIC-78-37, 1978; p 103.
Moller, G. E. *Soc. Pet. Eng. J.* **1977**, *17*(2), 101–110.

Because of their austenitic microstructure, nickel-based alloys are generally resistant to hydrogen embrittlement in the annealed condition. However, risk of hydrogen embrittlement (HE) increases as the strength level increases due to cold work or heat treatment. At similar strength levels, increasing aging temperature is apparently favorable in lowering HE risk. HE of 70/30 NiCu alloy K-500 bolts has been reported for an offshore platform application under cathodic protection conditions.[136] Susceptibility to HE was traced to high hardness and subsequent cathodic protection. Annealing after rolling the threads to alleviate high hardness before precipitation hardening was recommended to mitigate the problem.

Accelerated attack would be predicted for active materials coupled to Ni-based alloys because of the noble positions of the latter in the galvanic series. However, the successful use of alloy 400 sheathing on carbon steel in the splash and tidal zones, even in the

absence of cathodic protection, suggests galvanic compatibility which has been attributed to easy cathodic polarization of the nickel alloy.[137] It should be emphasized that this would not be the case if the steel/nickel-alloy area ratio was small. Pitting and crevice corrosion of NiCu alloy trim in pumps and valves is mitigated by cast iron bodies. Although nickel-based alloys are anodic to titanium, they are galvanically compatible because the increased corrosion rates are still very small from a practical viewpoint.[138]

2.18.9 Titanium Alloys

Titanium alloys are among the most corrosion resistant materials in seawater service. The general corrosion rate in seawater is extremely low, for example, 0.8 nm year^{-1},[140] not surprisingly, it is often reported as 'nil.' Titanium alloys are immune to pitting and crevice corrosion in all marine environments at ambient temperatures. Thus, no localized corrosion is encountered in the atmospheric, splash, tidal, immersion, or mud zones. Titanium is not resistant to biofouling. Therefore, macrofouling attachments similar to those on carbon steel, stainless steel, and aluminum alloys are also commonly observed on titanium alloys. However, unlike the other alloys, there is no localized corrosion of titanium under the biofouling or any silt deposits. No corrosion allowance is necessary for titanium because of its very high corrosion resistance in ambient-temperature seawater. Thus, titanium is widely used for heat exchanger applications, especially retubing in shell-and-tube heat exchangers. Grade-1 titanium is the premier material of choice for plate-and-frame heat exchangers because of high corrosion resistance and easy fabrication properties. However, because of extensive demand, there is currently a worldwide shortage of grade-1 titanium. Other materials such as Ni–Cr–Mo alloys and superaustenitic stainless steels are being evaluated as alternatives.

Crevice corrosion of titanium in seawater has been reported in tight crevices at temperatures >70 °C.[141,142] The crevice electrolyte was also found to be acidic, suggesting that a similar oxygen-depletion and metal-ion hydrolysis mechanism (as discussed earlier for stainless steels) may be operative. However, the actual mechanism is probably more complex since crevice corrosion resistance is increased by small additions of Pd, Mo, or Ni to the titanium. Iron, carbon steel or low-alloy steel particles embedded accidentally into a titanium surface can lead to localized attack at temperatures >80 °C. This has been proposed as a special case of crevice corrosion,[142] that is, occlusion by titanium metal smears or laps may create micro crevices that allow the generation of acidic conditions, preventing repassivation and localized attack, and possibly hydrogen adsorption. Thus, the mechanism is probably more complex than the presence of just micro crevices, since the attack has not apparently been observed if austenitic stainless steel, nickel, or copper particles are embedded in the titanium surface. Evidently, the Pd-containing grades exhibit significant resistance to attack associated with smeared iron particles.

Titanium is quite noble in the galvanic series for seawater. It can sometimes exhibit potentials anodic to graphite, platinum, zirconium, Ni–Cr–Mo–Fe alloy C, and NiCu alloy 400, and even some 300-series stainless steels in the passive condition. However, titanium is galvanically compatible with these materials. Any attack on the titanium is more likely to be crevice corrosion in tight crevices at temperatures >70 °C, as discussed earlier. In fact, despite being slightly more noble than titanium, attack is much more likely on the 300-series stainless steels, alloy 400, and even alloy C, at crevices regardless of galvanic coupling to titanium. Corrosion of the more active materials such as copper-base alloys, carbon steels, cast irons, aluminum alloys, zinc, and magnesium will be accelerated if galvanically coupled to titanium. The degree of attack acceleration will depend on the relative anode/cathode area ratio, system geometry, and cathodic reduction reaction kinetics on the titanium, which are influenced by seawater temperature, velocity, fouling, and so on. **Figure 12** shows that in seawater, titanium is more easily polarized than copper due to its higher overvoltage for oxygen reduction. Consequently, galvanic attack on carbon steel coupled to titanium is less severe compared to corrosion of steel coupled to copper, in each case for the same anode/cathode area ratio. The potential difference between titanium and steel is nearly double that between steel and copper – affirming that potential difference is not reliable for predicting galvanic attack on the anodic material. The free corrosion potential of titanium in aerated and deaerated seawater is fairly steady in the ~20–135 °C temperature range. However, as temperature increases further, the potential shifts in the electronegative direction by ~500 mV at ~200 °C. Potentials of other alloys also exhibit electronegative shifts.[143]

Unfortunately, titanium is susceptible to hydriding when its potential shifts more electronegative

than −0.7 V (SCE) at elevated temperature.[144] This can happen when titanium is cathodically polarized by coupling to anodic materials such as steel, aluminum, or zinc, or by impressed current cathodic protection. In ambient temperature seawater applications, a potential shift to no more than −1 V (SCE) has been suggested.[144] Hydrogen is generated by a two-step process: (1) discharge of H^+ to atomic H by electron consumption, and (2) H recombination to generate molecular H_2 gas bubbles. If excessive hydrogen is generated, there is an increasing uptake of atomic hydrogen, which diffuses into the titanium, reacts with the α phase, and precipitates as brittle, acicular, titanium hydride (TiH_2) needles at grain boundaries. Hydriding is exacerbated at temperatures >80 °C. About 500–600 ppm, hydrogen can cause excessive hydriding which embrittles the titanium, causing potentially significant reduction in ductility and fracture toughness. Anodizing reduces hydrogen uptake. The beta titanium alloys do not hydride easily because of considerably higher solubility of hydrogen (several 1000 ppm) in the beta phase.

It was prematurely concluded from early tests that unwelded titanium was highly resistant to stress corrosion cracking (SCC) in seawater[145] even when stressed to 75% of the yield strength. However, those tests used smooth test specimens. Subsequent tests disclosed that while alloy composition and heat treatment influenced SCC resistance, presence of notches simulating stress risers was essential. Environmental variables did not greatly affect the SCC behavior, this is not surprising because of the very high resistance of titanium to corrosion in seawater. Generally, all the α alloys (e.g., Ti–8Al–1Mo–1V and Ti–7Al–3Nb–2Sn), and near α alloys that contain small amounts of β phase stabilizers such as Al, Zr, Sn, and O (e.g., Ti–6Al–5Zr–0.5Mo–0.2Si and Ti–8Al–1V–1Mo) are susceptible to SCC in seawater. SCC resistance is markedly reduced when Al >5% and oxygen >0.4% in α alloys. For the α + β alloys, such as grade 5, oxygen concentration <0.1% is necessary to preclude SCC; commercial alloys contain typically 0.15% oxygen. In α + β alloys, cracking occurs primarily in the α phase. Contradictory behavior has been indicated for the effect of β stabilizers on SCC resistance in seawater. In some cases, V and Nb additions were detrimental, in others, >2% V was advantageous when Al in the 7–8% range was present in the alloy. In α + β alloys, Si and Mn were apparently deleterious.[146]

The SCC resistance of titanium alloys in seawater is greatly affected by heat treatment. Generally, small grain size is attained by aging just below the β transus (above the β transus, α or α + β phases transform to β) which enhances SCC resistance. For the susceptible alloys, the SCC behavior is also highly influenced by stress concentration; for example, there is a significant decrease in fracture stress in seawater at a notch radius of 125 μm compared to 250 μm. In slow strain rate tests in simulated seawater, SCC behavior of a number of titanium alloys was found to be highly influenced by the strain rate[147]; typically, they were most susceptible at strain rates of $\sim 7 \times 10^{-5}\,s^{-1}$. Other laboratory tests revealed that crack velocity decreased linearly with cathodic polarization of titanium alloys over the −400 to −900 mV (SCE) potential range[148]; the opposite behavior would be expected if there is sufficient uptake of hydrogen, for example, due to extended precharging time and/or polarization to more negative potentials in test. Crack velocity was found to increase with temperature, exhibiting Arrhenius behavior.[149] In tests performed in surface seawater and at depths of ~800 and ~2000 m, only butt-welded, nonstress relieved, Ti–13V–11Cr–3Al alloy stressed to 75% of yield strength exhibited SCC.[150]

The passive oxide film, its ability to reheal, and high inherent hardness of titanium confer high resistance to erosion–corrosion and cavitation in seawater. Velocities as high as $\sim 37\,m\,s^{-1}$ can be tolerated in clean seawater. Minimum velocities of the order of $>2\,m\,s^{-1}$ are necessary to prevent attachment of macrofouling in seawater heat exchanger tubing and piping. Provided they are galvanically compatible, titanium inserts can be used in the inlet ends of heat exchanger tubing experiencing erosion–corrosion at those locations. Ordinary silt in seawater does not lower erosion–corrosion appreciably. Higher-strength titanium alloys, such as grade 5 (Ti–6Al–4V), can be substituted for enhanced resistance if large, angular, abrasive particles produce erosion of the commercial purity (CP) grades.

The anodic breakdown potential of titanium in ambient temperature seawater is ~10 V. This allows titanium to be used as a substrate for platinum in impressed current anodes. If higher breakdown voltages are necessary in specific applications, niobium is substituted for titanium.

2.18.10 Aluminum Alloys

The low density of aluminum alloys ($\sim 2700\,kg\,m^{-3}$) makes them attractive for marine engineering

applications such as seacraft. However, the corrosion resistance is quite variable depending on alloy composition and heat treatment. Table 16 summarizes a commonly used designation system for aluminum alloys. For some applications, 'Alclad' products are specified. Alclad refers to a composite wrought (e.g., sheet) material which has an aluminum-alloy core with a metallurgically bonded thin 'skin' of either pure aluminum or another aluminum alloy on its surface for corrosion protection, for example, high-purity (more corrosion resistant) aluminum over alloy 2024, or alloy 7072 applied over alloy 6061 as sacrificial coating.

Aluminum relies on the formation of a passive aluminum oxide (Al_2O_3) film for corrosion resistance. However, the film properties are affected by alloy composition, metallurgical condition, and environmental factors, and widely ranging corrosion behavior is exhibited in seawater. In seawater immersion, the major corrosion concerns with aluminum alloys are galvanic corrosion when electrically coupled to cathodic materials, crevice corrosion and pitting under macrofouling and silt deposits, exfoliation, stress corrosion cracking, and metal-ion (especially Cu) deposition corrosion. The high strength 2xxx series and the 7xxx series aluminum alloys are widely used in aerospace applications. However, the alloying additions of Cu (in 2xxx) and Zn (in 7xxx) also make them more susceptible to corrosion in seawater.

Unalloyed aluminum has very low general corrosion rate in the splash zone because DO diffuses easily through the thin seawater layer to sustain passivity associated with the Al_2O_3 film. However, pure aluminum is considered too 'soft' for many splash-zone applications. The stronger 5xxx and 6xxx alloys are more suitable. In the tidal zone, aluminum alloys are subject to biofouling and localized attack (pitting and/or crevice corrosion) underneath macrofouling attachments. Long-term exposure tests[133] showed that the average penetration for alloys 1100 and 6061 in the tidal zone after 16 years exposure was very low, <15 μm; however, the maximum pit depths were ~1–1.5 mm. Under immersion conditions, the average penetration increased but was still very low from a practical viewpoint, <30 μm; however, maximum pit depths were in the range ~1–2 mm.

The effect of alloying additions is complex. For example, depending on specific alloy, <0.2% Cu and ~0.5–1.2% Mn were beneficial, but 0.4–0.8% Fe was detrimental toward corrosion resistance of aluminum in synthetic seawater.[151] Heat treatment affects the distribution and size of the intermetallic particles that influence mechanical properties by precipitation hardening. The desired mechanical properties of 5xxx alloys cannot be attained by conventional heat treatment alone; instead, they are strain hardened by mechanical working and, depending on alloy, may be subjected to some annealing treatment. The selected data in Table 17 show that the general corrosion rate of alloy 5052 in the H22 temper was an order of magnitude lower than in the annealed 'O' condition; the H32 temper reduced the corrosion rates of alloys 5086 and 5456 by half, but the H113 temper did not significantly affect the corrosion rate of alloy 5085.[152]

Table 16 Aluminum wrought alloy designations

Series	*UNS No.*	*Major alloying element(s)*	*Heat treatable*
1xxx	A91xxx	None	No
2xxx	A92xxx	Cu	Yes
3xxx	A93xxx	Mn	No
4xxx	A94xxx	Si	No
5xxx	A95xxx	Mg	No
6xxx	A96xxx	Mg, Si	Yes
7xxx	A97xxx	Zn, Mg	Yes
8xxx	A98xxx	Ti, Li	No
Temper (heat treatment) designations			
O		Annealed	
H (e.g., H1, H2, etc.)		Strain-hardened only	
T (e.g., T4, T6, etc.)		Solution heat treated; aged	

Table 17 Corrosion rates of 5xxx aluminum alloys in seawater

Alloy	*Exposure duration (months)*	*Exposure depth (m)*	*Corrosion rate (μm year^{-1})*
5052-O	35	1700	78
5052-H22	35	1700	10
5052-H34	13	770	5
5454-H32	6.5	760	18
5454	13	770	15
5456-H321	13	770	28
5456-H32	13	770	15
5456-H343	6.5	770	10
5085	13	770	13
5085-H113	13	770	15
5086	13	770	20
5086-H32	13	770	10
5086-H34	13	770	15
5086-H112	6	2	28

Source: Reinhart, F. M. Corrosion of Metals and Alloys in the Deep Ocean, Technical Report No. R-834, Civil Engineering Laboratory, Port Hueneme, CA, 1976; pp 204–209.

The increase in corrosion rate of aluminum alloys with seawater depth is often attributed to lower oxygen content of seawater, which presumably prevents rehealing of the aluminum oxide film. Lower seawater pH (and hence less likelihood of calcareous deposit formation) in deep water has been proposed as an alternative explanation.[153] However, this rationalization appears to be contradicted by data[154] from artificial seawater tests indicating that corrosion rate of Al–3Mg alloy decreased with decreasing pH in the 6–9 range. The mechanism is probably much more complex since alloy composition appears to affect corrosion behavior even at the same seawater depth. For example, maximum depth of pitting was <50 μm for alloys 5086-H34 and 5083-H113 exposed for 1 year at a depth of ~250 m, whereas the maximum pit depths for alloys 2219-T81 and 6061-T6 at the same depth were almost 2 mm.

Tests on alloys 3004, Alclad 3004, Aclad 3003, and 5052 exhibited less mass loss, but more pitting in cold seawater (~10 °C); in warm seawater (~27 °C), overall mass loss was greater but corrosion was uniform with no discernible pitting.[155] It was shown electrochemically that the pitting potential was not significantly altered by temperature,[153] however, ennoblement of the free corrosion potential indicated greater propensity to pitting in colder seawater. Data for Al–3Mg alloy indicate that corrosion rate increased steadily with temperature in the ~20 to ~70 °C range, it then decreased somewhat with a minimum at ~90 °C, and finally increased rapidly above boiling to ~125 °C.[154] Increased corrosion at the lower temperatures was attributed to DO. A significant decrease in pitting corrosion of alloy 5052 was reported for hot, flowing seawater environment associated with desalination when the temperature increased from 82 to 108 °C.[156]

Alloy composition and heat treatment appear to affect pitting corrosion behavior as illustrated by the data in **Table 18** which show maximum depths of attack on selected aluminum alloy plate specimens after 10 years exposure in seawater. Electrochemical studies have indicated that alloying additions of Zn lower the pitting potential of aluminum; Cu additions raise it, while Mg additions exhibit a minor effect. However, heat treatment to attain maximum hardness lowered the pitting potential of Al–Cu alloys; this is probably related to grain boundary depletion of Cu associated with precipitation of $CuAl_2$ intermetallics particles at grain boundaries. **Figure 22** shows an interesting relationship between pitting propensity and free corrosion potential for a number of aluminum alloys.[175] There is no general agreement on the growth rate of pits with time. Some data suggest that pit depth decreases with time; other data imply opposite behavior. Despite wide variations in pitting corrosion behavior, aluminum alloys are highly susceptible to crevice corrosion, and therefore require protection, especially when tight crevices are present. It is not known with certainty whether the crevice corrosion mechanism in aluminum alloys is similar to that in stainless steels because it has not been studied as extensively.

Aluminum alloys have been reported to be generally resistant to seawater up to velocities of ~6 m s^{-1}.[157,158] However, critical velocities vary with alloy composition, heat treatment, and exposure time[159,160]; and attack morphology can change from pitting to increased general metal loss as velocity increases. Aluminum alloys being relatively soft are more susceptible to cavitation damage under turbulent flow conditions than, for example, stainless steels, nickel-based alloys, and titanium alloys. This means

Table 18 Maximum depths of attack on some aluminum alloy plate specimens after 10 years immersion in natural seawater at ambient temperature

Alloy	*Maximum pit depth (mm)*
1100-H14	1.02
3003-H14	0.53
6061-T4	0.36
6061-T6	2.41
7072	1.42 (perforated)
7075-T6	1.68

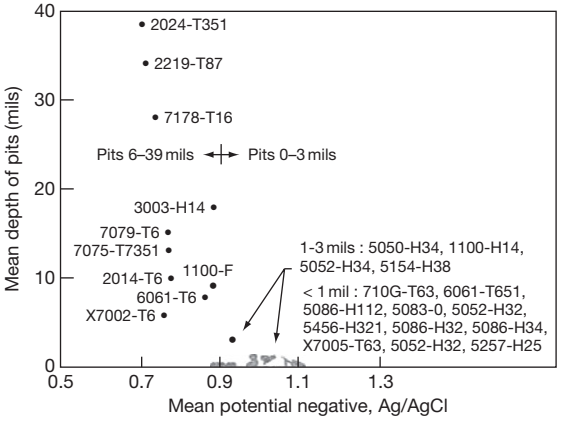


Figure 22 Depth of pitting on Al alloys versus free corrosion potential in seawater. Reproduced from Schumaher, M., Ed. *Seawater Corrosion Handbook*; Noyes Data Corporation: NJ, USA, 1979; p 61.

that aluminum alloys can be used, for example, in the construction of seacraft hulls operating at moderately high velocities, but not for seawater pump applications.

Aluminum alloys have high thermal conductivities, which should make them very attractive for heat transfer surfaces, for example, in seawater heat exchanger applications. However, several factors limit their widespread utility. For example, aluminum alloys are highly susceptible to galvanic corrosion when coupled to components made of cathodic materials such as copper-based alloys, stainless steels, nickel-based alloys, and titanium alloys. Even if they are electrically isolated, only tens of ppb levels of copper ions (e.g., generated by self-corrosion of copper-based alloys upstream) are sufficient to cause severe corrosion by the cementation mechanism[161] (also known as deposition corrosion). Aluminum alloys are prone to biofouling such that sufficient velocities would be required to prevent crevice corrosion under macrofouling attachments.

An extensive study was performed in simulated seawater on galvanic corrosion rates determined from current density measurements for a number of aluminum alloys coupled to dissimilar materials as well as between the aluminum alloys themselves.[162] The data showed that all of the aluminum alloys tested were incompatible with Cu, Ag, and Ni. In addition, alloys 1100, 6061, and 7075 were not compatible with all the stainless steels, nickel alloy 718, low alloy steel 4130, and Ti–6Al–4V. However, surprisingly, alloy 2024 and 2219 were compatible with these materials (low alloy 4130 excepted). The aluminum alloys were compatible with each other except alloy 1100 with alloy 7075. In another test, alloy 5086-H32 incurred accelerated attack when coupled to carbon steel, 90/10 CuNi, Type 304 stainless steel, and 60/40 brass.[163] The degree of galvanic attack is influenced by the anode/cathode area ratio. Laboratory tests[164] have confirmed that increasing seawater velocity will increase galvanic attack on aluminum alloys in seawater because the most common cathodic reaction is oxygen reduction, which is under diffusion control.

Pure Al is generally considered resistant to stress corrosion cracking (SCC) in seawater. Aluminum 5xxx alloys are prone to intergranular SCC failure when the Mg addition is >5% which results in supersaturation and allows precipitation of a continuous Mg_2Al_3 intermetallic phase at the grain boundaries; this phase is anodic to the matrix. Restricting the Mg addition to <4% results in discontinuous Mg_2Al_3 phase and SCC resistance. This was confirmed by immersion tests in seawater.[165] In Mg additions >5%, addition of small amounts of Bi, Cr, Zr, B, and Be have been found to enhance SCC resistance. Alloy 6061 in the fully aged T6 temper is resistant to SCC in seawater.[165]

Several of the highest strength 7xxx aluminum alloys were found to be susceptible to SCC in natural seawater,[166] especially in the T6 temper to attain maximum strength.

Lithium additions lower the density of the aluminum alloy but increase strength and elastic modulus. However, fracture toughness and SCC resistance are lowered.

Alloy 6061 is susceptible to SCC in the T4 temper if slow quench follows high heat treat temperature. However, the T6 temper (fully aged) imparts high SCC resistance.

Exfoliation is a form of corrosion that occurs along grain boundaries in rolled aluminum alloy sheet and plate products. Corrosion often initiates at exposed grain boundaries (end-grain attack) or at pit sites that have exposed the elongated grain microstructure to the environment.The voluminous corrosion products of aluminum, $Al(OH)_3$, introduce internal stresses that are sufficiently to cause exfoliation, that is, lifting off of material in layers that comprise corrosion products and unaffected metal. Exfoliation resistance is derived from the absence of second phase precipitates, especially at grain boundaries, and where the free corrosion potentials of the precipitates and matrix are closely matched. The 5xxx aluminum alloys are resistant to exfoliation in seawater at Mg additions <3%. 6xxx series are resistant to intergranular corrosion (IGC) except when excess Si is present. Al alloys 5456-H116 and-H117 have been used for boat hull plates to resist IGC in H321 temper.

As-quenched T3 or T-4 tempers, or naturally aged 2xxx alloys are prone to SCC and exfoliation, but only slightly to IGC, whereas 7xxx alloys are more susceptible to IGC. Overaging to T6 and T8 tempers improves SCC and exfoliation resistance but not IGC. As for 2xxx alloys, rapid quenching and overaging increases resistance significantly, for example, 7075-T63 and T-73, 7178-T76.

Cathodic protection can mitigate SCC of aluminum alloys when they are polarized to a potential in the −1.2 to −1.4 V (SCE) range. However, overprotection, for example, with Mg anodes, and especially quiescent conditions raise the surface pH significantly as OH^- ions accumulate at the aluminum alloy surface. The high-pH conditions can result in 'cathodic corrosion' of the aluminum alloy by the

formation of an aluminate complex:

$$Al + 2OH^- \rightarrow AlO_2^- + H_2$$

Seawater flowing at moderate velocities can prevent build-up of OH^-, and hence reduce risk of cathodic corrosion. Even in quiescent conditions, the cathodic corrosion issue for aluminum alloys can be almost entirely eliminated by the use of zinc anodes for cathodic protection.

Aluminum is an active metal in the galvanic series. Therefore, in principle, it could be considered for cathodic protection of carbon steel and other engineering alloys. However, the application is not straightforward. As stated earlier, a passive Al_2O_3 film usually forms on aluminum and its structural alloys. This is contrary to the requirements of a sacrificial anode material which must corrode efficiently to provide protective current for cathodic protection. Thus aluminum is alloyed with small amounts of Zn or Sn and trace levels of In or Sn (although Hg was used previously) to activate the aluminum surface. Aluminum alloys find widespread usage as sacrificial anodes in seawater, for example, on offshore platforms and pipelines because of their significant weight advantage and high current capacity ($A\,h\,kg^{-1}$) compared to zinc. The development and performance of aluminum anodes have been reviewed extensively in the literature.

2.18.11 Magnesium

Magnesium and its alloys are not normally utilized in seawater because they are highly reactive. Reduction of DO is not normally the primary cathodic reaction, but rather the reduction of water to evolve hydrogen.

This reaction occurs particularly easily on cathodic impurities in Mg such as Fe, Cu, Cd, Zn, Ni, and so on. Being anodic and very reactive, corrosion of Mg is greatly accelerated by galvanic coupling to most other engineering materials, for example, carbon steel, stainless steels, nickel-based alloys, copper-based alloys, titanium alloys, and so on. In fact, Mg sacrificial anodes are widely used for cathodic protection of steel used in soil environments. Use of Mg for cathodic protection in seawater environments is rare because of the risk of overprotection of steel, which can cause hydrogen embrittlement in high strength steels, cathodic blistering and disbondment of coatings, and cathodic corrosion of amphoteric materials such as aluminum alloys, and high rate of self corrosion of the magnesium.

Mg alloys have been used in a very limited capacity for deep-sea diving suits – the main attraction being very light weight due to the low density of Mg. However, the Mg has to be properly coated to prevent direct contact with seawater which would otherwise result in severe corrosion. This could be especially catastrophic at any holidays in the coating if the Mg was inadvertently coupled to a cathodic material.

The very high corrosion rate of Mg can be used to advantage in special applications, for example, 'time release' links on oceanographic instruments deployed in deep ocean environments.

2.18.12 Zinc

Although zinc is also an active metal in the galvanic series, it is much less reactive than magnesium. While the self corrosion of zinc is surprisingly low, its other major limitations are similar, that is, susceptibility to galvanic attack when coupled to steel, stainless steels, copper alloys, and so on. Historically, zinc is the most common sacrificial material used for cathodic protection in seawater.

2.18.13 Nonmetallic Materials

Nonmetallic materials do not generally corrode in the same manner as metals and alloys in seawater, that is, by thinning. Instead, they may suffer marine borer attack which is usually not very deep, except in wood. In addition, nonmetallics may undergo swelling and some reduction in mechanical strength – which is usually less for synthetic materials than those derived from nature (e.g., to make ropes). Deterioration is generally greater in warm seawater and close to the seabed. Other hazards are fish bites, for example, on cables. Results of long-term exposures of polymeric materials have been reviewed.[167,168] Nonmetallic materials are often reinforced to increase strength, for example, glass fibers in epoxy resins, carbon fibers in epoxies and aluminum, steel in concrete, and so on. Reinforcements can affect the performance of the composite material in seawater, for instance, carbon fibers can accelerate galvanic corrosion of aluminum and corrosion of steel reinforcement can lead to spalling of concrete. Rubber linings and reinforced polymer (e.g., flake glass polyester) coatings applied to metal substrates can provide corrosion protection. However, elastomers exhibit reduction in elongation with prolonged exposures; silicone–rubber and ester-based polyurethanes deteriorate more extensively. Thus, meticulous attention to selection, application, inspection, and maintenance is essential for successful

usage. While they may provide adequate corrosion protection, the major limitations of polymeric materials include elevated service temperatures, high pressure containment, and fatigue life.

For seawater service, concrete with sulfate resistance (Type V cement) must be specified. The steel reinforcement requires corrosion protection for long-term performance, for example, by cathodic protection. Ceramics are generally considered resistant to seawater.

2.18.14 Corrosion Protection Methods

In principle, most of the methods used for corrosion control in other aqueous environments are applicable to seawater. However, the major factors contributing to the corrosivity of seawater that must be taken into consideration include high chloride and sulfate content, presence of microorganisms, biofouling, and putrefaction.

Carbon steel corrodes in aerated seawater at rates that require either a corrosion allowance or other means of corrosion control if a long service life is required. Corrosion allowance requires careful monitoring to avoid unexpected failures. It is probably most useful for applications where some areas of localized perforation caused by seawater can be tolerated without loss or mixing of fluids, or reduction in structural (load-bearing) properties, for example, nontubular structural support members.

Coatings are the most common method of corrosion control. Apart from suitability for seawater service, the single most important requirement for a successful coating is proper surface preparation and application conditions – which are often best achieved under shop conditions. Once a coating system is applied, even under initially optimal conditions, repairing or replacing it in service can be extremely challenging. For applications such as the underwater areas of ship hulls, it is fairly easy to perform coating inspections, repairs, and replacements during periodic dry docking. In the case of fixed structures, such as offshore platforms and pipelines, similar access to the underwater areas is almost impossible. Building cofferdams around the structure and evacuating them to carry out repairs and replacements can only be done more economically in very shallow waters. Although some coatings that can be applied underwater (e.g., by divers) are available, they are fraught with problems, for example, surfaces cannot usually be adequately cleaned and prepared, and coating application and inspection are very difficult. Hence, such coatings are typically only considered for emergency situations. Antifouling coatings are utilized on ships for controlling macrofouling. Presently, most antifouling coatings contain either copper or cuprous oxide as the antifouling agent. Other coatings based on low surface energy are also available to minimize adhesion of macrofouling, but they are more expensive. Both types of coatings are consumed with time and lose their effectiveness, and thus need replacement periodically.

As mentioned previously, it is possible to protect simple structural steel shapes by sheathing. For example, splash-zone areas on offshore platforms have been successfully protected by alloy 400 (nickel–copper) or 90/10 CuNi sheathing. Hot risers are not usually protected by sheathing because of periodic visual inspection requirements that would require removal of the sheathing. Metallurgically cladding steel with a more corrosion resistant alloy can also be used but is again applicable to only relatively simple geometries such as piping.

The most widely used method for corrosion control of steel structures immersed in seawater is cathodic protection using sacrificial zinc or aluminum anodes. Cathodic protection is only effective below the waterline, regardless of tide level. Cathodic protection and coatings complement each other if they can be used together for corrosion control. The coating reduces the area of metal requiring protection; conversely, cathodic protection can extend the coating life. However, it is essential to select a coating that is compatible with cathodic protection to obviate cathodic disbondment (blistering) of the coating. The anode material must meet tight alloy composition and performance requirements before it is specified. Sacrificial anodes generally require no maintenance, except replacement when consumed. However, it must be recognized that replacement of anodes on stationary structures is difficult and very costly; thus, in such cases, the original anodes are typically designed to last the life of the structure (e.g., 30 years). Impressed current cathodic protection systems serve the same function but, contrary to popular belief, require more attention, for example, in monitoring proper operation of rectifiers and reference electrodes, especially for systems with automatic control, stray current effects, diver safety concerns, and so on. Impressed current anodes based on platinized niobium are now preferred in seawater over platinized titanium because of the much higher breakdown voltage of niobium ($\sim$70 V) compared to titanium ($\sim$9 V). Cathodic protection of structures

and components in seawater is extensively covered in the published literature.

Calcareous deposits (discussed previously), formed in seawater on cathodically protected metal surfaces are essential for polarization of the structure, and to achieve significant reduction in current demand by the structure. Although used extensively for carbon steels, cathodic protection can also be used to mitigate corrosion of nonferrous materials in seawater, for example, for pitting and crevice corrosion of stainless steels. In fact, carbon steel can be used as sacrificial anode material. Cathodic protection can also be effective for controlling galvanic corrosion between dissimilar materials, for example, between waterbox, tube sheet, and tubes; and erosion–corrosion of copper alloys, for example, inlet-end impingement in heat exchangers.

Cathodic protection for corrosion control in the interior of pipes has been investigated but is not commonly used for a variety of reasons. The biggest problem is limited 'throwing power' of the protective current. Thus, depending on pipe size, an anode may be required, say, every few meters for protecting bare steel. Continuous anodes along the length of the pipe for internal cathodic protection have also been investigated but present many limitations including economics.

Modification of the seawater by deoxygenation is an effective method of corrosion control. As discussed previously, it is used in a number of applications, for example, MSF desalination and seawater injection during secondary recovery. For steels, while corrosion rate decreases with oxygen content, it increases markedly at elevated temperatures even when only a few hundred parts per billion oxygen is present. Thus, effective control and monitoring of oxygen is essential for this method of corrosion control, especially at high temperatures. Deaeration has not been found to be harmful for passive materials such as stainless steels that often rely on oxygen for film healing.

Conflicting requirements often preclude use of a single material in most engineering applications. Since dissimilar materials are often unavoidable, the issue of galvanic corrosion must be addressed. In principle, electrical isolation between anode and cathode members can be effected by using nonconductive (insulating) hardware to mitigate galvanic corrosion. However, in many applications, this is difficult to achieve because alternative electrical grounding paths exist, for example, hangers or pedestal supports in piping systems. In many systems, however, even if electrical isolation is possible, it is not permitted for personnel safety reasons, for example, electrical safety codes may preclude galvanic isolation in systems where ground fault currents would otherwise pose potential electrical shock hazards. In design, a small cathode/anode area ratio is usually selected to minimize galvanic corrosion. Coating either the cathode or both the cathode and anode can be a practical means of controlling galvanic corrosion. The anode alone should not be coated since inevitable holidays and flaws in the coating can severely accelerate corrosion attack at those locations because of the resulting very adverse anode/cathode area ratio. However, it should be recognized that coatings for galvanic corrosion control have limitations where they are difficult to apply and maintain, and where they can incur damage rather easily, for example, in pumps.

Mitigation of galvanic corrosion in piping systems at dissimilar-metal flanged-joints, without electrical isolation, has been addressed in a number of ways. Since the galvanic corrosion damage of the anodic material is restricted over a relatively short distance that is either known from experience or can be determined by mathematical modeling, one approach is to insert a heavier-wall waster spool between the original anode and cathode. The length of the spool piece is a function of the pipe diameter and polarization characteristics of the cathode.[169] Consumption of the waster spool piece can be monitored, for example, by ultrasonic thickness gauging from the outside, so that it can be replaced in a timely manner. A novel technique known as the 'bi-electrode' method[170,171] has been proposed to control galvanic corrosion. Two ring or sleeve electrodes with inert metal surfaces are inserted in the piping system on either side of the joint. The sleeves which are electrically isolated from the piping system are polarized with respect to each other by an external DC power supply, such that the electrical field created by current flow annuls the galvanic current. Since the galvanic current can vary with operating conditions such as flow, temperature, oxygen content, and so on, the externally applied counter current between the bielectrodes has to adjust automatically – a problem that has been found to be challenging to date.

For many components, coatings and/or cathodic protection are not practical methods of corrosion control, for example, pumps, valves, entire heat exchanger tubing in shell-and-tube heat exchangers, plates in plate-and-frame heat exchangers, internal piping, and so on. In such cases, corrosion resistant materials must be selected. The selection is based on

practical experience and test data. For example, as discussed earlier, Type 304 and 316 stainless steels are unsuitable materials for seawater service because of extreme susceptibility to crevice corrosion in aerated seawater; in other words, they should not be used for immersion service in seawater without protection (e.g., cathodic protection) if crevices are present. The use of a resistor-controlled cathodic protection system has been proposed for controlling internal corrosion in chlorinated stainless steel piping systems.[172,173]

Unfortunately, selection of more highly alloyed grades such as Type 317 stainless steel or alloy 825 that contain greater amounts of Cr and Mo does not bestow sufficient resistance to crevice corrosion. Crevice corrosion initiation time is usually longer in cold seawater. Conversely, it is known that localized corrosion susceptibility increases with increasing temperature. Critical pitting and critical crevice corrosion temperature data are used to provide guidelines for the upper limits of alloys. For example, it is generally recommended that superaustenitic 20Cr–25Ni–6Mo stainless steels should not be subjected to seawater temperatures $>\sim 35\,^{\circ}C$ for prolonged periods. Pitting resistance equivalent numbers (PREN) based on alloy composition (primarily Cr, Mo, and N) have been advocated as indicators of localized corrosion resistance of stainless steels and nickel-based alloys in seawater and other chloride environments. However, crevice corrosion resistance does not depend just on alloy content; it is also affected by crevice geometry, tightness, depth, crevice-forming materials, their deformability, and surface finish of both the crevice formers and the alloy in question.[174] Thus, although many materials, such as Ni–Cr–Mo–Fe alloy C-276, have very high resistance to crevice corrosion in seawater, they are by no means completely immune under all crevice conditions. This is also the case with titanium, which is very resistant to localized attack in seawater at ambient temperatures, but is susceptible to crevice corrosion at $>70\,^{\circ}C$. Titanium alloys containing noble metals additions, such as palladium in grade 7, are more resistant to crevice corrosion.

Copper–nickel alloys have been successfully used for heat exchanger piping, waterbox and sheathing applications in clean seawater. In the freely corroding condition, they are resistant to macrofouling but not entirely immune. Their major limitations are high susceptibility to accelerated corrosion by sulfides, for example, in polluted seawater and bottom silt; or generated *in situ* as a result of putrefaction of stagnant seawater by sulfate reducing bacteria. Avoiding contact with polluted seawater is difficult in certain situations, for example, when shipboard piping and heat exchangers are commissioned in harbors. Guidelines based on practical experience suggest that the risk of sulfide production and consequent initiation of corrosion of copper–nickel alloys increases when they are exposed to stagnant seawater in closed systems for more than $\sim$2 days. Thus, it is imperative that such conditions be avoided by draining the system, flushing with fresh water and drying, if possible. Alternatively, the water can be treated with suitable biocides that are compatible with the alloy. Nickel aluminum bronze pumps have been used successfully in unpolluted seawater. This is not a good choice of material for auxiliary pumps because of the risk of seawater putrefaction during protracted standby periods and hence sulfide attack. However, heat treatment is necessary to preclude dealloying, especially in the heat-affected zones associated with welds, and the consequent increased risk of fatigue cracking. Such areas are often beyond the reach of cathodic protection from bare steel piping that might be beneficial in other situations where the current 'throwing power' is more effective.

Other factors that need to be considered in the selection of copper-base alloys include sensitivity to erosion–corrosion at high flow rates and galvanic corrosion when coupled to stainless steels, titanium, and Ni–Cr–Mo alloys. For instance, copper–nickel heat exchanger tubes that have suffered from sulfide and/or erosion–corrosion attack are often retubed with titanium. This results in unacceptable galvanic corrosion of the copper–nickel tube sheet. This situation is usually alleviated by the application of cathodic protection which must be controlled and monitored carefully to preclude hydriding of the titanium tubes that can be caused by overprotection. Tube support plate spacing also has to be addressed to preclude vibration of the low-modulus, thin-walled titanium tubes, and subsequent mechanical fatigue failure.

Chlorination is widely used for controlling biofilms and macrofouling; the latter interferes with seawater flow. Both biofouling and macrofouling can initiate corrosion of engineering materials such as many stainless steels, nickel–copper alloys, and even some copper-based alloys. At low levels, $\sim$0.5 ppm residual, chlorine does not accelerate corrosion of these materials or carbon steel significantly. However, as the chlorine concentration increases to $\sim$1–2 ppm, risk of general corrosion on the active metals and localized corrosion on certain passive metals

increases. Reduction of galvanic corrosion of certain materials, for example, copper-alloys coupled to stainless steels and nickel alloys, has been observed especially in laboratory tests and explained on the basis of biofilm suppression and attendant reduction reaction kinetics on the cathode. As the chlorine content increases to ~5 ppm or more, corrosion propensity of many stainless steels and nickel–copper alloys increases noticeably. Titanium is unaffected by even very high levels of chlorine in seawater – in fact, titanium electrodes are commonly used in the electrolytic cells to generate chlorine from seawater.

2.18.15 Test Methods

Test methods for assessing materials performance in seawater are extensively described in the literature and standards. In principle, most of these test methods are not applicable exclusively to seawater. However, it is important to recognize the important differences between natural seawater that contains micro and macroorganisms that can influence corrosion behavior of materials, and synthetic seawater that does not contain representative species which mimic biological effects. The equipment used for testing in seawater must also be corrosion resistant to prevent test interruptions or confounding of results by extraneous corrosion products introduced into the system.

References

1. Lyman, J.; Abel, R. B. *J. Chem. Educ.* **1958**, *35*(3), 113–115.
2. Reinhart, F. M. Corrosion of Metals and Alloys in the Deep Ocean, Technical Report No. R-834; Civil Engineering Laboratory: Port Hueneme, CA, 1976; p 2.
3. Dexter, S. C.; Culberson, C. *Mater. Perform.* **1980**, *19*(9), 23.
4. Compton, K. G., *Corrosion* **1970**, *26*, 448.
5. Korbin, G., Ed. *Microbiologically Influenced Corrosion*; NACE International: Houston, TX, 1993; p 5.
6. Korbin, G., Ed. *Microbiologically Influenced Corrosion*; NACE International: Houston, TX, 1993; p 48.
7. Whitfield, M.; Jagner, D., Eds. *Marine Electrochemistry, A Practical Introduction*; Wiley, 1981; p 513.
8. Tomashov, N. D. *Theory of Corrosion and Protection of Metals*; Macmillan: New York, 1966; p 470.
9. *ASM Handbook, Vol. 13A: Corrosion, Fundamentals, Testing, and Protection*; ASM International: Materials Park, OH, 2003; pp 398–416.
10. Uhlig, H. E.; Revie, R. W. *Corrosion and Corrosion Control*, 3rd ed.; Wiley: New York, 1985; p 94.
11. *ASM Handbook, Vol. 13A: Corrosion, Fundamentals, Testing, and Protection*; ASM International: Materials Park, OH, 2003; p 648.
12. Cox, G. C. US Patent 2,200,469, May 14, 1941.
13. Cox, G. C. US Patent 2,417,009, March 11, 1947.
14. Humble, H. A. *Corrosion* **1948**, *4*(7), 358–370.
15. Guillen, M. A.; Feliu, S. *Rev. Metall.* **1966**, *2*(6), 519–532.
16. Shigeno, H.; Umino, T.; Fukazawa, H. In *Proceedings of the 5th International Congress on Metallic Corrosion*; NACE International: Houston, TX, 1975; pp 619–623.
17. Wolfson, S. L.; Hartt, W. H. *Corrosion* **1981**, *37*, 70–76.
18. Phull, B. S. PhD Thesis, University of Manchester, UK, 1981.
19. Hartt, W. H.; Culberson, C. H.; Smith, S. W. In *Corrosion/83*; NACE International: Houston, TX, 1983; Paper No. 59.
20. Ambrose, J. R.; Yaniv, A. E.; Lee, U. R. In *Corrosion/83*; NACE International: Houston, TX, 1983; Paper No. 60.
21. Culberson, C. H. In *Corrosion/83*; NACE International: Houston, TX, 1983; Paper No. 61.
22. Fisher, K. P.; Finnegan, J. E. In *Corrosion/89*; NACE International: Houston, TX, 1989; Paper No. 581.
23. Fisher, K. P.; Finnegan, J. E. *Corrosion/89*; NACE International: Houston, TX, 1989; Paper No. 582.
24. de Oliveira, R.; de Souza Pimenta, G. In *12th International Corrosion Congress*; NACE International: Houston, TX, 1993; pp 2278–2284.
25. Pytkowich, R. M. *J. Geology* **1965**, *73*(1), 196.
26. Philipponeau, G. In *8th International Congress on Metallic Corrosion*; DECHEMA: Frankfurt, Germany, 1981; pp 196–199.
27. Doremus, E. P.; Doremus, G. L. *Corrosion* **1950**, *6*(7), 216–224.
28. Uhlig, H. E.; Revie, R. W. *Corrosion and Corrosion Control*, 3rd ed.; Wiley: New York, 1985; pp 69–71.
29. Oldfield, J. W.; Sutton, W. H. *Br. Corros. J.* **1978**, *13*(1), 13.
30. Oldfield, J. W.; Sutton, W. H. *Br. Corros. J.* **1978**, *13*(3), 104.
31. Kain, R. M. *Mater. Perform.* **1984**, *23*(2), 24–30.
32. Kain, R. M.; Lee, T. S. In *ASTM STP 866*; Haynes, G. S., Baboian, R., Eds.; American Society for Testing and Materials: West Conshohocken: PA, 1985; pp 299–323.
33. Kain, R. M. Seawater crevice corrosion testing of stainless steel, Ni-base and Cu-Ni alloys: Perspectives on methodologies and interpretation of results, EUROCORR/2006.
34. Kain, R. M. Private communication.
35. Dexter, S. C.; Gao, G. Y. *Corrosion* **1988**, *44*(10), 717–723.
36. Johnsen, R.; Bardal, E. *Corrosion* **1985**, *41*(5), 296.
37. Suzuki, T.; Yamake, M.; Kitamura, Y. *Corrosion* **1973**, *29*, 18.
38. Lei, M. K.; Zhu, X. M. *J. Electrochem. Soc.* **2005**, *152*(8), B291–B295.
39. Ellis, O. B.; LaQue, F. L. *Corrosion* **1951**, *7*(11), 362.
40. LaQue, F. L. *Marine Corrosion: Causes and Prevention*; Wiley, 1975; p 195.
41. Dexter, S. C.; LaFontaine, J. P. *Corrosion* **1998**, *54*(11), 851–861.
42. Ferrara, R. J.; Taschenberg, L. E.; Moran, P. J. *Corrosion/85* Paper No. 211: NACE International: Houston, TX, 1985.
43. Wallen, B.; Henrikson, S. *Werkst. Korros.* **2004**, *40*(10), 602–615.
44. Kain, R. M. In *Corrosion/98*; NACE International: Houston, TX, 1998; Paper No. 702.
45. Humble, H. A. *Corrosion* **1948**, *5*(9), 292–302.
46. Kirk, W. W. *Mater. Perform* **1987**, *26*(9), 23–28.
47. Phull, B. S.; Kain, R. M. In *Applications and Materials Performance*; Proceedings of the Nickel-Cobalt/97 International Symposium; Smith, F. N., McGurn, J. F., Lai, G. Y., Sastri, V. S., Eds.; Canadian Institute of Mining, Metallurgy and Petroleum, 1997; Vol. IV, pp 141–150.
48. Chaker, V.; Lindemuth, D. In *Corrosion/90*; NACE International: Houston, TX, 1990; Paper No. 376.

49. Smith, M.; Bowley, C.; Williams, L. In *Corrosion/2002*; NACE International: Houston, TX, 2002; Paper No. 02214.
50. Linder, B. Cathodic protection of sheet steel piling in seawater, containing sulfate-reducing bacteria, particularly with respect to accelerated low water corrosion (ALWC), Swedish Corrosion Institute Report No. 55467, 1999.
51. Oubner, R.; Beech, I. In *Corrosion/1999*; NACE International: Houston, TX, 1999; Paper No. 99318.
52. Breakell, J. E.; Foster, K.; Siegwart, M. *Management of Accelerated Low Water Corrosion in Steel Maritime Structures*; CIRIA: London, 2005; publ. no. C634.
53. LaQue, F. L. *Marine Corrosion: Causes and Prevention*; Wiley, 1975; p 143.
54. Phull, B. S.; Pikul, S. J.; Kain, R. M. *ASTM STP 1300*; American Society for Testing and Materials: West Conshohocken, PA, 1997; pp 34–73.
55. LaQue, F. L. *Marine Corrosion: Causes and Prevention*; Wiley, 1975; p 14.
56. Melchers, R. E. *ASTM STP 1300*; American Society for Testing and Materials: West Conshohocken, PA, 1997; pp 20–33.
57. Southwell, C. R.; Alexander, A. L. *Mater. Prot.* **1970**, *9*(1), 14–23.
58. Reinhart, F. M. *Corrosion of Metals and Alloys in the Deep Ocean, Technical Report No. R-834*; Civil Engineering Laboratory: Port Hueneme, CA, 1976; p 14.
59. Beavers, J. A. Koch, G. H.; Berry, W. E., Eds. *Corrosion of Metals in Marine Environments;* MCIC Report MCIC-86-50, 1986, pp 6–33.
60. Beavers, J. A., Koch, G. H., Berry, W. E., Eds. *Corrosion of Metals in Marine Environments*; MCIC Report MCIC-86-50; pp 6-42–6-45.
61. Preiser, H. S. In *Handbook of Corrosion Protection for Steel Pile Structures in Marine Environments*; American Iron and Steel Institute, 1981; pp 67–100.
62. Beavers, J. A. Koch, G. H.; Berry, W. E. Eds. *Corrosion of Metals in Marine Environments;* MCIC Report MCIC-86-50, 1986, pp 6–49.
63. Van Der Sluys, W. A. *Eng. Fract. Mech.* **1969**, *1*, 447.
64. Schrieber, C. F.; Coley, F. H. *Mater. Perform.* **1976**, *15*(7), 47–54.
65. Oldfield, J. W.; Todd, B. Corrosion considerations in selecting metals for flash chambers, Eurocorr/77, 1977.
66. LaQue, F. L. *Marine Corrosion: Causes and Prevention*; Wiley, 1975; p 8.
67. Shreir, L. L., Ed. *Corrosion;* Newnes Butterworths, 1979; Vol. 1, p 3.106.
68. Miyasaka, M.; Ogure, N. In *Corrosion/86*; NACE International: Houston, TX, 1986; Paper No. 324.
69. Southwell, C. R.; Bultman, J. D.; Alexander, A. L. *Mater. Prot.* **1976**, *15*(7), 9–26.
70. Mollica, A.; Trevis, A. In Proceedings of the 4th International Congress on Marine Corrosion and Fouling; Juan-les-Pins: France, 1976.
71. Asphahani, A. I.; Manning, P. E.; Silence, W. L.; Hodge, G. F. In *Corrosion/80*; NACE International: Houston, TX, 1980; Paper No. 29.
72. Deverell, H. E.; Maurer, J. R. *Mater. Perform.* **1978**, *17*(3), 15–20.
73. Bond, A. P.; Dundas, H. J. *Mater. Perform.* **1984**, *23*(7), 39–43.
74. Kain, R. M. In *ASTM STP1399*; Dean, S. W., Delgadillo, G. H., Bushman, J. B., Eds.; American Society for Testing and Materials: West Conshohocken, PA, 2000; pp 284–299.
75. Francis, R.; Byrne, G. In *Corrosion/2003*; NACE International: Houston, TX, 2003; Paper No. 03255.
76. ASTM Standard G 48.
77. Davies, M.; Scott, P. J. B. *Guide to Use of Materials in Waters*; NACE International: Houston, TX, 2003; p 90.
78. Kovach, C. W.; Redmond, J. D. In *Corrosion/93*; NACE International: Houston, TX, 1993; Paper No. 267.
79. ASTM Standard G 78.
80. *ASM Handbook Vol. 13A Corrosion: Fundamentals, Testing, and Protection*; ASM International, 2003; pp 549–561.
81. Sedriks, A. J. *Int. Met. Rev.* **1982**, *27*(6), 321–353.
82. Francis, R.; Burton, G. R. In *Corrosion/2000*; NACE International: Houston, TX, 2000; Paper No. 630.
83. Oldfield, J. W.; Sutton, W. H. *Br. Corros. J.* **1980**, *15*(1), 31–34.
84. Kain, R. M. In *Corrosion/81*; NACE International: Houston, TX, 1981; Paper No. 200.
85. Kain, R. M. In *Corrosion/82*; NACE International: Houston, TX, 1982; Paper No. 66.
86. Deverell, H. E.; Davis, J. A. In *Corrosion/78*; NACE International: Houston, TX, 1978; Paper No. 27.
87. Kain, R. M.; Zeuthen, A. W.; Maurer, J. R. In *Corrosion/97*; NACE International: Houston, TX, 1997; Paper No. 97423.
88. Weber, J. In *Materials to Supply Demand*; Proceedings of ASM International Conference; Harrison, B. C., Ed.; 1981; pp 49–80.
89. Turnbull, A. *Corrosion* **1999**, *55*(2), 206.
90. Wallen, B. *Some Factors Affecting Stainless Steel Corrosion in Seawater;* ACOM: Avesta Sheffield, 1990; Vol. 4, pp 1–8.
91. Rogne, T.; Drugli, J. M. In *Stainless Steel World'99 Conference*; KCI, 1999; p 527.
92. ASTM Standard G 160.
93. Oldfield, J. W.; Todd, B. *Desalination* **1979**, *31*(3), 365–383.
94. Reinhart, F. M. *Corrosion of Metals and Alloys in the Deep Ocean,* Technical Report No. R-834, Civil Engineering Laboratory, Port Hueneme, CA, 1976; p 174.
95. Lee, W. S. W.; Oldfield, J. W.; Todd, B. *Desalination* **1983**, *44*(5), 209–221.
96. Plesset, M. S. *Corrosion* **1962**, *18*(5), 181t–188t.
97. Simoneau, R.; Fihey, R.; Roberge, R. In *Cavitation Erosion in Fluid Systems Conference Proceedings,* June 1981; pp 71–81.
98. Al Hashem, A. H.; Tarish, H.; Akbar, A. A. In *Corrosion/2007*; NACE International: Houston, TX, 2007; Paper No. 07253.
99. Powell, C. A.; Michels, H. T. In *Corrosion/2000*; NACE International: Houston, TX, 2000; Paper No. 00627.
100. LaQue, F. L. *Marine Corrosion: Causes and Prevention*; Wiley, 1975; p 267.
101. Efird, K. D. *Corrosion* **1977**, *33*(1), 3–8.
102. Gudas, J. P.; Hack, H. P. *Corrosion* **1979**, *35*(6), 259–264.
103. Gudas, J. P.; Hack, H. P. *Mater. Performance* **1980**, *19*(4), 49–54.
104. Syrett, B. C. *Corros. Sci.* **1981**, *21*(3), 187–209.
105. Sanchez, S. R.; Schiffrin, D. J. *Corros. Sci.* **1981**, *22*(6), 245–248.
106. Eiselstein, L. E.; Syrett, B. C.; Wing, S. S.; Caliguiri, R. D. *Corros. Sci.* **1981**, *23*(3), 223–239.
107. Francis, R. *The effect of sulfide and chlorine on the corrosion of copper alloy heat exchanger tubes,* INCRA Project, 289A/3, final report 1984.
108. Al Hajji, J. N.; Reda, M. R. *Corros. Sci.* **1992**, *34*(1), 163–177.
109. Efird, K. D.; Anderson, D. B. *Mater. Perform* **1975**, *14*(11), 37–40.
110. Mattson, M. E.; Fuller, R. M. A study of materials of construction in distillation plants, Joint report by Office of

Saline Water and the International Nickel Company, Research report No. 163, 1965.
111. Schreiber, C.; Osborn, O.; Coley, F. H. *Mater. Perform* **1968**, *7*(10), 20–25.
112. Reinhart, F. M. Corrosion of Metals and Alloys in the Deep Ocean, Technical Report No. R-834, Civil Engineering Laboratory, Port Hueneme, CA, 1976; p 56.
113. LaQue, F. L.; Clapp, W. F. *Trans. Electrochem. Soc.* **1945**, *87*, 103–125.
114. Efird, K. D. In *Corrosion/75*; NACE International: Houston, TX, 1975; Paper No. 124.
115. Clapp, W. F. In *Corrosion Handbook*; Uhlig, H. H., Ed.; Wiley, 1948; p 443.
116. Retief, R. *Br. Corros. J.* **1973**, *8*(4), 264–269.
117. Shone, E. B.; Grim, B. C. *Trans. Inst. Mar. Eng. UK* **1985,** 98, Paper No. 11.
118. Otsu, T.; Sato, S. *Comparative corrosion test of condenser tube materials by model condenser at Meiko power station,* Sumitomo Light Metal Technical reports, Oct 1961, *2*(4), pp 299–322.
119. Tracy, A. W.; Hungerford, R. L. *Proc. ASTM* **1945**, *45*, 591.
120. Stewart, W. C.; LaQue, F. L. *Corrosion* **1977**, *8*(8), 259–277.
121. Syrett, B. C. *Corrosion* **1976**, *32*(6), 242–252.
122. ASTM Standard A 342.
123. Anderson, D. B. *Mater. Perform.* **1971**, *10*(11), 26.
124. Danek, G. J., Jr. *Naval Eng. J.* **1966**, *78*(5), 763–769.
125. Sato, S.; Nagata, K. Factors affecting corrosion and fouling of condenser tubes of copper alloys and titanium, Sumitomo Light Metal Technical reports, July 1978, *19*(3–4), pp 83–94.
126. Sato, S.; Nosetani, T.; Yamaguchi, Y.; Onda, K. *Factors affecting the sand erosion of aluminum brass condenser tubes*; Sumitomo Light Metal Technical reports, Jan 1975, *16*(1–2), pp 23–37.
127. Maligas, M.; Vicic, J. In *Corrosion/97*; NACE International: Houston, TX, 1997; Paper No. 426.
128. Hack, H. *J. Test. Eval.* **1980**, *8*(2), 74–79.
129. Gehring, G. A., Jr. Effects of sulfide, sand and cathodic protection on condensers, EPRI Report No. CS-4562, Electric Power Research Institute, Palo Alto, May, 1986.
130. Southwell, C. R.; Alexander, A. L. *Mater. Prot.* **1969**, *8*(3), 39–44.
131. Tipton, D. G.; Kain, R. M. In *Corrosion/80*; NACE International: Houston, TX, 1980; Paper No. 36.
132. Lennox, T. J., Jr. In *Corrosion/82*; NACE International: Houston, TX, 1982; Paper No. 64.
133. Boyd, W. K.; Fink, F. W. Corrosion of Metals in Marine Environments, Battelle Columbus Labs, OH, Report No. MCIC-78-37, 1978; p 103.
134. Moller, G. E. *Soc. Pet. Eng. J.* **1977**, *17*(2), 101–110.
135. Reinhart, F. M. Corrosion of Metals and Alloys in the Deep Ocean, Technical Report No. R-834, Civil Engineering Laboratory, Port Hueneme, CA, 1976; p 126.
136. Efird, K. D. *Mater. Performance* **1985**, *24*(4), 37–40.
137. LaQue, F. L. *Marine Corrosion: Causes and Prevention*; Wiley, 1975; p 115.
138. Bomberger, H. B. *J. Electrochem. Soc.* **1954**, *101*(9), 442–447.
139. Ross, R. W., Jr.; Tuthill, A. H. *Mater. Performance* **1990**, *29*(4), 65–69.
140. Feige, N. G.; Murphy, T. J. *Corrosion* **1966**, *22*(11), 320–324.
141. Griess, J. C., Jr. *Corrosion* **1968**, *24*(4), 96–109.
142. Covington, L. C. *Corrosion* **1979**, *35*(8), 378–382.
143. Smith, C. A.; Compton, K. G. *Corrosion* **1975**, *31*(9), 320–326.
144. *ASM Metals Handbook Vol. 13: Corrosion*; ASM International, 1987; p 673.
145. Brown, B. F. *Mater. Res. Stand.* **1966**, *6*(3), 29–133.
146. Beavers, J. A.; Koch, G. H.; Berry, W. E. Eds. *Corrosion of Metals in Marine Environments;* MCIC Report MCIC-86-50, 1986, pp 3–27.
147. Powell, D. T.; Scully, J. C. *Corrosion* **1968**, *24*(6), 151.
148. Beck, T. R. *J. Electrochem. Soc.* **1967**, *114*, 551–556.
149. Beck, T. R. In *Theory of Stress Corrosion Cracking in Alloys*; Scully, J. C., Ed.; NATO Scientific Affairs Division: Brussels, 1971; pp 64–85.
150. Reinhart, F. M. Corrosion of Metals and Alloys in the Deep Ocean, Technical Report No. R-834, Civil Engineering Laboratory, Port Hueneme, CA, 1976; pp 228–238.
151. Beavers, J. A. Koch, G. H.; Berry, W. E. Eds. *Corrosion of Metals in Marine Environments;* MCIC Report MCIC-86-50, 1986; pp 2–31.
152. Reinhart, F. M. Corrosion of Metals and Alloys in the Deep Ocean, Technical Report No. R-834, Civil Engineering Laboratory: Port Hueneme, CA, 1976; pp 204–209.
153. Dexter, S. C. *Corrosion* **1980**, *36*(8), 423–432.
154. Konstantinova, E. V.; Nazarova, N. P. *Prot. Met.* **1976**, *12*(3), 279–281.
155. Larsen-Basse, J. *Mater. Perform.* **1984**, *23*(7), 16–21.
156. Ahmed, Z. *Anti-Corros. Methods Mater.* **1981**, *28*(7), 4–10.
157. Davis, J. A.; Gehring, G. A., Jr. *Mater. Performance* **1975**, *14*(4), 32–39.
158. Herrigel, H. R. *Mater. Perform.* **1976**, *15*(3), 43–45.
159. Davis, J. A.; Gehring, G. A., Jr. In *Corrosion/75*; NACE International: Houston, TX, 1975; Paper No. 123.
160. Gehring, G. A., Jr.; Peterson, M. H. *Corrosion* **1981**, *37*(4), 232–242.
161. Dexter, S. C. *J. Ocean Sci. Eng.* **1981**, *6*(1), 109–148.
162. Mansfeld, F.; Hengstenberg, D. H.; Kenkel, J. V. *Corrosion* **1974**, *30*(10), 343–353.
163. Lennox, T. J., Jr.; Peterson, M. H.; Smith, J. A.; Groover, R. E. *Mater. Perform.* **1974**, *13*(2), 31–36.
164. Mansfeld, F.; Kenkel, J. V. *Corrosion* **1977**, *33*(7), 236–240.
165. Beavers, J. A., Koch, G. H.; Berry, W. E., Eds. *Corrosion of Metals in Marine Environments;* MCIC Report MCIC-86-50, 1986, pp 2–64.
166. Reinhart, F. M. Corrosion of Metals and Alloys in the Deep Ocean, Technical Report No. R-834, Civil Engineering Laboratory, Port Hueneme, CA, 1976; pp 221–223.
167. Connolly, R. A.; DeCoste, J. B.; Gaupp, H. L. *J. Mater.* **1970**, *5*, 339.
168. Munitz, A. J. *Undersea Technol.* **1966**, *7*, 45.
169. Francis, R. A. *Galvanic Corrosion: A Practical Guide for Engineers*; NACE International, 2001; pp 120–124.
170. Shifler, D. A.; Hack, H. P.; Melton, D. G. In *Corrosion/98*; NACE International: Houston, TX, 1998; Paper No. 706.
171. Tossey, B. M.; Beavers, J. A. In *Corrosion/2007*; NACE International: Houston, TX, 2007; Paper No. 07244.
172. Bardal, E.; Johnsen, R.; Gartland, P. O. In *Corrosion/92*; NACE International: Houston, TX, 1992; Paper No. 408.
173. Johnsen, R.; Gartland, P. O.; Valen, S.; Drugli, J. *Mater. Performance* **1996**, *35*(7), 17.
174. Kain, R. M. In *Corrosion/91*; NACE International: Houston, TX, 1991; Paper No. 508.
175. Schumaher, M., Ed. *Seawater Corrosion Handbook*; Noyes Data Corporation: NJ, USA, 1979; p 61.

2.19 Corrosion in Soils

J. F. D. Stott and G. John
Intertek-CAPCIS Ltd, Bainbridge House, 86-90 London Road, Manchester M1 2PW, UK

Glossary

Aerobic The presence of oxygen in an environment (including oxygen from the air).

Anaerobic The absence of oxygen in an environment (including oxygen from the air).

Backfill Material that is applied around the component that is buried.

Bacteria Unicellular microorganisms, typically a few micrometers in length and having many shapes including spheres, rods, and spirals.

Biogenic Arising from living processes.

Brownfield site Former commercial or industrial site, earmarked for commercial development or industrial projects. Potentially contaminated land.

Cathodic protection A decrease of corrosion rate of a metal by cathodic polarization such that the net anodic current is decreased.

Clay Natural constituent of soil. Comprises particles of less than 0.005 mm mean diameter ($<5\,\mu m$).

Conductivity The specific ability of a material to conduct ionic or electronic charge. In SI units, the conductivity is defined as the conductance (in Siemens or $Ohms^{-1}$) between the opposing faces of a cube of side $1\,m^2$ (SI units: Sm^{-1}).

Corrosion product A species formed as a result of corrosion; most commonly a solid (e.g., rust) but may be liquid, gaseous, or ionic.

Corrosion protection A process of modification to a material or its corrosive environment such as to decrease its corrosion rate.

Corrosion rate The rate at which corrosion proceeds. It may be expressed as loss of thickness or cross-section (e.g., $\mu m\,year^{-1}$), as a rate of mass loss per unit area (e.g., $g\,m^{-2}\,d^{-1}$), or via Faraday's Law, as a current density (e.g., $A\,m^{-2}$). Although the term implies that corrosion rate is constant, this is not always the case.

Corrosivity assessment or aggressivity assessment Estimation of likely corrosion rates that will occur on metal/alloys in a particular soil environment on the basis of consideration of soil type, resistivity, soil and groundwater chemistry, etc. Also referred to as aggressivity assessment.

Disturbed soils Soils that are mechanically disturbed, for example, by excavation.

Electrical resistance (ER) Method for measuring corrosion rate where metal loss of a probe element is determined by measurement of electrical resistance of the element.

General corrosion or uniform corrosion A form of corrosion that results in a more-or-less uniform thickness loss from the surface of a material.

Graphitic corrosion A form of corrosion of grey cast iron in which the metallic constituents are selectively removed leaving the graphite (often called 'graphitization,' although the latter term is deprecated because of its alternate use to describe the formation of graphite, for example, by the decomposition of carbides in steel).

Greenfield site Agricultural or forest land or undeveloped (virgin) site earmarked for commercial development or industrial projects.

Groundwater Natural water contained within soil.

Hot-dip coating or galvanizing A method of application of a coating by immersion of a material in a bath of molten metal: for example, hotdip galvanizing, for the formation of a zinc coating on steel.

Linear polarization (1) The approximate linear dependence between current density and potential at low values of electrode polarization (typically less than $\sim\pm 10\,mV$). (2) An electrochemical test method that uses the principle of linear polarization to derive an estimate of the electrode reaction (e.g., the corrosion rate). Not the same as Ohm's law.

Long-line corrosion Form of differential concentration corrosion, occurring on pipelines crossing different soil conditions or for piles crossing different soil layers. The large surface area in benign/aerated conditions causes increased corrosion in more active/deaerated conditions. See also macro-cell corrosion.

Macro-cell corrosion Form of differential concentration corrosion.

Metabolism The complete set of chemical reactions that occur in living cells. These processes are the basis of life, allowing cells to grow and reproduce, maintain their structures, and respond to their environments.

Microbially influenced corrosion (MIC) A form of corrosion that is associated with the presence of microorganisms, including bacteria, fungi, and algae. Corrosion is commonly influenced by the metabolites of the organism and/or by the production of a biofilm.

Microorganism or microbe An organism that is microscopic (too small to be seen by the human eye). The study of microorganisms is called microbiology. Microorganisms can be bacteria, fungi, archaea, or protists, but not viruses and prions, which are generally classified as nonliving.

Moisture-holding capacity of a soil A term applied to the ability of a soil to hold water present in the form of capillary water.

Nondisturbed soils Soils in the natural state that have not been affected by excavation.

pH The negative logarithm of the hydrogen ion activity defined as $pH = -\log_{10} aH^+$, where aH^+ is the activity of the hydrogen ion in an electrolyte.

Pit A cavity in the surface of a material that extends into its interior.

Pitting (corrosion) A form of localized corrosion that results in the formation of pits, often associated with a cover of corrosion product or perforated metal over the pit.

Polarization resistance The slope (dE/di) at the corrosion potential of the current density versus electrode potential under Linear Polarization. The polarization resistance is inversely proportional to the corrosion rate provided that charge transfer is the controlling electrode reaction. See also Charge Transfer Resistance.

Redox An abbreviation for reduction–oxidation.

Redox potential The equilibrium potential of an inert electrode (such as platinum) in a solution; a measure of the oxidizing or reducing tendency of the solution.

Reference electrode An equilibrium half-cell electrode of reproducible potential against which an unknown electrode potential can be measured. Commonly used examples include the standard hydrogen electrode (Pt, H_2/H^+), the calomel electrode ($Hg/Hg_2Cl_2/Cl^-$), and the silver/silver chloride electrode ($AgCl/Cl^-$).

Resistivity The inverse of the conductivity of a material or solution.

Sabkha A flat area between a desert and an ocean, characterized by a crusty surface consisting of evaporite deposits (including salt, gypsum, and calcium carbonate), windblown sediments, and tidal deposits.

Sand Natural constituent of soil. Comprises particles of mean diameter 0.07–~2 mm in size (70–2000 μm).

Selective backfill Backfill material that has been selected to exhibit specific properties, particularly material of low corrosivity.

Silt Natural constituent of soil. Comprises particles of mean diameter 0.005–0.07 mm (5–70 μm).

Sulfate-reducing bacteria (SRB) Term describing several groups of anaerobic bacteria that use sulfate as an oxidizing agent for their energy-generating metabolism, reducing it to sulfide, which appears as H_2S, sulfide ions, metal sulfides, or a combination of these forms according to the conditions.

Standard hydrogen electrode (SHE) A redox electrode which forms the basis of the thermodynamic scale of oxidation–reduction potentials. The potential of platinum electrode in an acid solution with $aH^+ = 1\ mol\ kg^{-1}$, purged with hydrogen at a pressure of 101 325 Pa, which is declared to be zero at all temperatures.

Stray-current corrosion A corrosion process caused by unwanted currents flowing through paths other than the intended circuit and usually arising from an adjacent direct current source (e.g., traction supply, welding station, cathodic protection supply) or from an induced alternating current source (e.g., power transmission conductors).

Substrate (1) The base material on which a coating is present. (2) Any substance (other than materials required in trace quantities only) required for the growth or metabolism of a microorganism.

Thermal spraying A process for the application of a coating to a substrate by forming an atomized spray of small particles of molten material and directing onto a substrate with a suitable profile.

Underground corrosion A process involving corrosion of a buried material where the earth's soil forms a component of the corrosive environment.

Abbreviations

ASTM American Society for Testing and Materials
AWWA American Water Works Association
BC Before Christ
BS British Standards
DC Direct current
DIN Deutsches Institut für Normung e.V. (German Institute for Standardization)
DIPRA Ductile Iron Pipe Research Association
DMM Design decision model
EN EuroNorm (European Standard)
ER Electrical resistance
HDPE Hugh density polyethylene
LPR Linear polarization resistance
MIC Microbiologically influenced corrosion
MTBF Mean time before failure
NACE National Association of Corrosion Engineers
NBS National Bureau of Standards
PREN Pitting resistance equivalent numbers
PVC Polyvinyl chloride
SHE Standard hydrogen electrode
SRB Sulfate reducing bacteria
TS Tank size
UNS Unified Numbering System

Symbols

k_W, k_P Scale constants (mm, mil, g m^{-2}, or oz ft^{-2})
m, n Shape constants, with $n, m \leq 1$ (typically $n, m \approx 0.33–0.67$)
pH Soil pH
t Period of exposure (years)
y Typical variable
Cl Chloride ion content
***E*(age)** Unconditional predicated age to corrosion failure
M Moisture content
P Pit depth (mm or mil)
P Tank structure-to-soil potential
R Resistivity
S Sulfide content
SC Stray current magnitude
W Uniform metal loss (measured as mm, mil, g m^{-2}, or oz ft^{-2})
ϵ Standard error

2.19.1 Introduction

Corrosion in soils (also referred to as 'underground corrosion') is applicable to a wide range of different components (e.g., pipelines, underground storage tanks, tubular and sheet steel piles, ground anchors, power and telecommunication cables, etc.), which are either directly buried, driven, or otherwise positioned underground (sometimes referred to as 'below grade') and which can, therefore, be exposed to a wide range of different conditions (ranging from dry aerated sand to water-logged anaerobic clay). Any consideration of the subject of corrosion in soils must address both natural and contaminated ('green field' and 'brown field') sites and must consider ferrous alloys (such as cast iron, ductile iron, carbon/low-alloy steels, various stainless steels), as well as alloys of copper, aluminum, zinc, and lead. It must also include the performance of metallic coatings (e.g., galvanizing and thermally sprayed coatings). However, in all cases, the basic corrosion mechanisms are the same as for corrosion in natural waters. Examples of corrosion in soils for different components are shown in Figures 1–4.

Degradation of buried nonmetallic components (such as polyvinyl chloride (PVC) or high density polyethylene (HDPE) pipe) and concrete (including corrosion of reinforcing steel in buried concrete foundations) are not considered here.

Figure 1 Corrosion of cast iron pipe (due to external stray current).

Figure 2 Burst failure of carbon steel oil pipeline (due to lack of cathodic protection).

Figure 3 Corrosion of buried copper pipe exposed to H_2S containing ground water. Reproduced from *Corrosion Atlas*, 3rd ed.; compiled by During, E. D. D.; Elsevier Science Publishers: London, 1997; (Item 06.05.05.01).

Figure 4 Bronze spearheads (Knossos, Crete, *ca.* 1500 BC).

The main aspect that differentiates corrosion in soils from corrosion in other natural environments is the contribution that the soil (and associated ground water) plays in modifying the corrosion mechanism and in particular the corrosion rate. In practice, cast iron, ductile iron, and carbon/low-alloy steels behave essentially in the same manner for a given soil type and condition; their corrosion rates can range from negligible ($<0.005\,\text{mm year}^{-1}$) to extremely rapid (with localized corrosion of $>1\,\text{mm year}^{-1}$). The differences are related to soil type, soil resistivity, ground water composition, soil condition, and the activity of bacteria, especially sulfate-reducing bacteria (SRB).

It is not the intention here to provide full details of soil genesis, for which the reader is directed to specialist publications[1] for a detailed description. However, an outline description of different soil types and methods of classification is included, in order to assist in the understanding of the variations in soil that directly relate to the corrosion processes.

Corrosion in soils has been studied worldwide over many years; to date, the most extensive and comprehensive study was conducted by Romanoff, for the National Bureau of Standards (NBS) in the US in the 1950s and 1960s. This work was published as a series of NBS circulars and presentations,[2,3] covering the results of over 37 000 different samples, comprising numerous different metals/alloys and 95 different soils, exposed for up to 17 years. The work is still relevant today as the benchmark for describing performance of various metals and alloys exposed to a wide range of soils and is the ultimate source for much of the published data relating to the subject.

Other studies have been carried out since the work by Romanoff up to the present day in North America,[4,5] the United Kingdom,[6,7] Sweden,[8] Germany,[9] and Russia[10] and elsewhere. Whilst these different studies provide further information on specific local soils and on alternative/new alloys, the overall results are complementary to the original findings of Romanoff.

The original studies of corrosion in soil mainly concerned pipelines; this was later expanded to include driven piles. However, since the 1980s significant interest has also been focused on underground storage tanks, because of increased concern about environmental damage caused by leaking tanks, especially for small scale (retail) storage facilities at petrol (gasoline) stations, domestic fuel storage facilities, etc. The long-term performance of ground anchorage systems (high stressed anchorages, soil nails, tie-bars, etc.) is another area where there is an increasing requirement to demonstrate long-term performance, especially for systems that provide stabilization for major critical civil structures such as flood defense, dams, bridges, tunnels, etc.

In archaeology, the extent of deterioration of metals over millennia of exposure is of interest. The condition of bronze and iron artifacts can vary greatly depending on local ground conditions.

2.19.2 The Nature of Soils

Encarta Dictionary defines soil as "The top layer of most of the Earth's land surface, consisting of the unconsolidated products of rock erosion and organic decay, along with bacteria and fungi." While demonstrating the natural variability in soils, such a definition does not fully convey an impression of the countless variations in soil that occur. The type and condition of the soil are directly affected by moisture, principally entering through rainfall, but also from flooding, tidal action, and human interaction (e.g., from irrigation schemes, dewatering, or land reclamation schemes), as well as by contamination from industrial processes. It is the combination of the genesis of the soil and the source of moisture within the soil that leads to variations in soil chemistry.

Often locations are classed as 'green field' or 'brown field' sites; green field sites are those where there has been no previous industrial use, while brown field sites are locations of former industrial use (factories, storage yards, waste tips, etc.). For brown field sites there is, therefore, also the possibility that contaminated soil may be present, the soil being contaminated by the presence of man-made chemicals or

other human alteration in the natural soil environment. In addition to the direct discharge of industrial wastes to the soil, contamination often occurs from the rupture of underground storage tanks, application of pesticides, percolation of contaminated surface water, oil dumping or leaching of wastes from landfill sites. The most common chemicals involved are petroleum hydrocarbons, pesticides, and heavy metals.

An important consideration is that statutory or other routine surveys of contaminated land for data acquisition are carried out primarily for public health and environmental reasons and are rarely, if ever, intended for corrosivity assessment. Most of the data, thereby, generated are of little or no consequence for soil corrosivity evaluation, whereas other data of vital importance for appraisal of soil corrosivity are often not generated by such surveys. A soil corrosivity survey is, therefore, a distinct exercise from a general contaminated land survey.

2.19.2.1 General Soil Texture and Structure

In any assessment of potential corrosivity, it is important that the nature of the soil itself is considered; however, this is a vast topic.[1] The more important aspects on soil texture and structure that follow are based on the corresponding chapter, by Harris and Eyre, in the previous edition.[11]

Soils are commonly named and classified according to the general size range of their particulate matter. Thus, sandy, silt, and clay types derive their names from the predominant size range of inorganic constituents. Particles between 0.07 mm and ~2 mm are classed as sands. Silt particles range from 0.005 to 0.07 mm, and clay particle size ranges from 0.005 mm mean diameter down to that of colloidal matter.

The proportion of the three size groups will determine many of the properties of the soil. Although a number of systems have been used to classify soils according to texture, the one shown in **Figure 5** represents the most commonly used terminology for various proportions of sand, silt, and clay.

As soils contain organic matter, moisture, gases, and living organisms as well as mineral particles, it is apparent that the relative size range does not determine the whole nature of the soil structure. In fact, most soils consist of aggregates of particles within a matrix of organic and inorganic colloidal matter rather than of separate individual particles. This aggregation gives a crumb-like structure to the soil and leads to friability, more ready penetration of moisture, greater aeration, less erosion by water and wind, and generally greater biological activity. The loss of the aggregated structure can occur as the result of mechanical action, or by chemical alteration such as excess alkali accumulation. Destruction of the structure or 'puddling' greatly alters the physical nature of the soil.

Mention should be made of the soil profile (section through soil showing various layers) because it is important to recognize that the soil's surface gives a very poor indication of the underlying strata. Pipelines are buried a meter or so and foundations/anchorages typically more than 10 m below surface soils and corrosion surveys based on surface observations give little information as to the actual environs of the pipe when buried.

2.19.2.2 The Clay Fraction

Clays make up the most important inorganic constituents of soil. They consist of various minerals depending on the mineral composition of the parent material, and on the type and degree of weathering. Often clays may be grouped in a family series, depending upon the weathered condition, as, for example, montmorillonite → illite → kaolinite. Weathering of montmorillonite causes loss of potassium and magnesium, which alters the crystalline structure, and eventually kaolinite results. In this example (and also for other clay mineral groups) marked changes occur in the physical properties of a soil as clay minerals undergo the weathering process.

Montmorillonite clays absorb water readily, swell greatly, and confer highly plastic properties to a soil. Thus soil stress occurs most frequently in these soils and less commonly in predominantly kaolinitic types. Similarly, a soil high in bentonite will show more aggressive corrosion than a soil with a comparable percentage of kaolinite. A chalky soil usually shows low corrosion rates. Clay mineralogy and the relation of clays to corrosion deserve attention from corrosion engineers.

2.19.2.3 Aeration and Oxygen Diffusion

The pore space of a soil may contain either water or a gaseous atmosphere. Therefore, the aeration of a soil is directly related to the amount of pore space present and to the water content. Soils of fine texture due to high clay content contain more closely packed particles and have less pore capacity for gaseous diffusion than an open-type soil such as sand.

Oxygen content of soil atmosphere is of special interest in corrosion. It is generally assumed that the

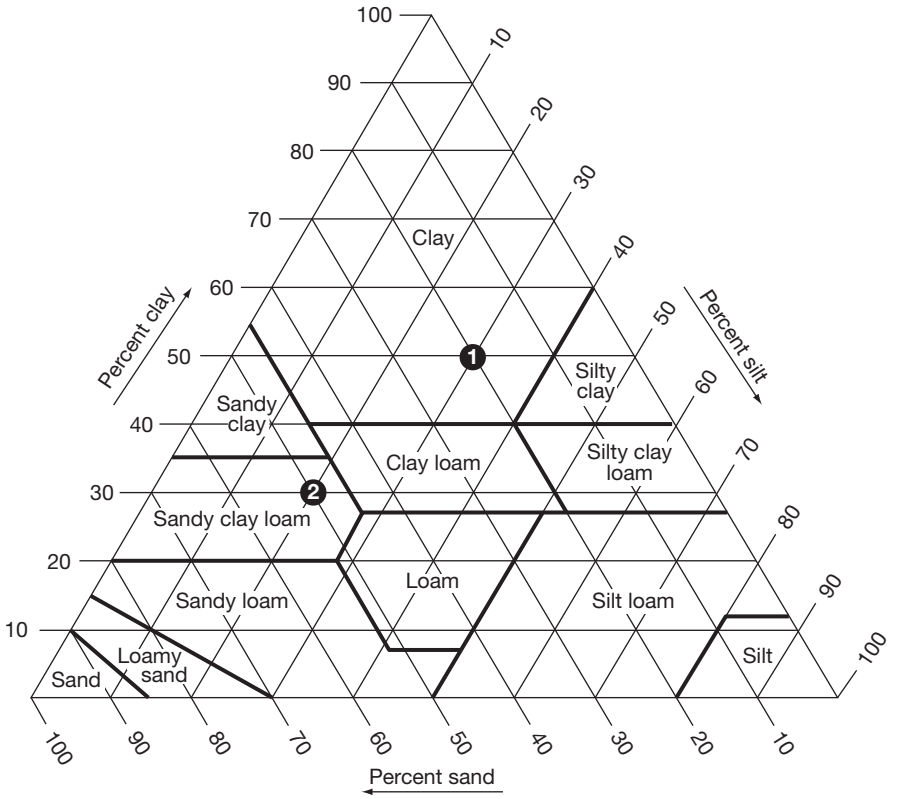


Figure 5 Proportions of sand, silt and clay making up the various groups of soils classified on the basis of particle size [(*)]. Reproduced from *Soil Survey Manual* 1993. Soil Conservation Service. U.S. Department of Agriculture Handbook 18. United States Government Printing Office, Washington, DC. (ISBN-10: 1410204170) (http://soils.usda.gov/technical/manual).
[(*)] The corresponding percentages of clay, silt, and sand present in the soil are located on relevant axis of the triangle. The values are then projected inwards (along the lines as indicated on the axis). The name of the area in which the lines intersect is the class name for the particular soil. For example, a soil with 50% clay, 30% silt, and 20% sand (location ❶) is a 'clay,' while a soil with 30% clay, 20% silt, and 50% sand (location ❷) is a 'sandy clay loam,' etc.

gases of the upper layers of soil are similar in composition to the atmosphere above the soil, except for a higher carbon dioxide content. However, biological activity within the soil tends to decrease the oxygen content and replace the oxygen with gases from metabolic activity, such as carbon dioxide. Most of the biological activity occurs in the upper 150 mm of soil, and it is in this region that diffusion would be most rapid. Factors which tend to increase microbial respiration, such as the addition of large amounts of readily decomposed organic matter, or factors which decrease diffusion rates (water saturation) will lead to development of anaerobic conditions within the soil. Anaerobic conditions favor the proliferation of SRB, which produce hydrogen sulfide, sulfide ions, or inorganic sulfides according to the conditions and may give rise to very high rates of localized corrosion under conditions of near neutral pH and low oxygen concentration, which would normally be considered benign. Such situations are discussed in **Section 2.19.2.6** below.

The main difference between 'buried' and 'driven' components relates to the availability or otherwise of oxygen. Where the components are exposed to 'backfill,' the soil will normally be aerated (at least initially), whereas for driven components (i.e., where the item is placed in 'undisturbed soil') the oxygen content can be negligible and in these cases corrosion of the bulk of the item will also be negligible (see **Section 2.19.3.5** below).

2.19.2.4 Water Relations

No soil corrosion occurs in a completely dry environment. Water is needed for ionization of the oxidized state at the metal surface. Water is also needed for ionization of soil electrolytes, thus completing the

circuit for flow of a current maintaining corrosive activity. Apart from its participation in the fundamental corrosion process, water markedly influences most of the other factors relating to corrosion in soils. Its role in weathering and soil genesis has already been mentioned.

2.19.2.4.1 Free ground water

At some depth below the surface, water is constantly present. This distance to the water table may vary from a <1 m to >100 m, depending upon the geological formations present. Only a small amount of the metal used in underground service is present in the ground water zone. Such structures, as well as casings and under-river pipelines, are surrounded by ground water. The corrosive conditions in such a situation are essentially those of an aqueous environment.

2.19.2.4.2 Gravitational water

Water entering soil at the surface from rainfall or some other source will tend to move downwards. This gravitational water will flow at a rate governed largely by the physical structure regulating the pore space at various zones in the soil profile. An impervious layer of clay, a 'puddled' soil, or other layers of material resistant to water passage may act as an effective barrier to the gravitational water and cause zones of water accumulation and saturation. This is often the situation in highland swamp and bog formation. Usually, gravitational water percolates rapidly to the level of the permanent ground water.

2.19.2.4.3 Capillary water

Most soils contain considerable amounts of water held in the capillary spaces of the silt and clay particles. The actual amount present depends on the soil type and weather conditions. Capillary moisture represents the important reservoir of water in soil, which supplies the needs of organisms living in or on the soil. Only a portion of capillary water is available to plants. The moisture-holding capacity of clay is much greater than that of a sandy type soil. Likewise, the degree of corrosion that occurs in soil will be related to its moisture-holding capacity, although the complexities of the relationships do not allow any direct and simple quantitative predictive application.

2.19.2.5 Variable Conditions

An important aspect of corrosion in soil is the effect that varying soil type and composition have on corrosion, giving rise to different conditions on different sections of a buried structure; this could be due to changes in moisture content, soil structure (clay to sand), or soil chemistry (pH, chloride concentration, oxygen content, etc.), resulting in galvanic corrosion.

For horizontal systems (i.e., pipelines, metallic cable sheaths, etc.), such processes are often referred to as 'long-line corrosion' where the pipeline crosses from one soil condition to another. Vertical changes in conditions are equally important, principally for driven components (such as piles or soil nails). The effect can be extensive, with the section of a structure in the higher pH, low chloride, or more aerated ground acting as a large cathode over tens or even hundreds of meters in length, leading to accelerated corrosion in the lower pH, deaerated, or higher chloride areas. A classic example is that of a pipeline running in normally high resistance, well aerated sand crossing an area of *sabkha* (i.e., salt rich desert area). Sabkha, being very high in salt content and generally low resistivity ground, is naturally more corrosive than the surrounding sand and hence would result in naturally higher corrosion rates. However, for a continuous pipeline, the line sections in the adjacent, normal sand also act as a good cathode, increasing the rate of attack in the shabka area.

2.19.2.6 Effect of Bacterial Activity on Corrosion in Soils

Microorganisms or 'microbes' are commonly defined as living entities that are too small to be visible to the human eye. Soils that contain more than a few percent moisture abound with an incalculable variety of microbial activity. However, fundamentally, the vast majority of those microbes perform a single overall metabolic function to mineralize organic matter of plant and animal origin and ultimately to convert it to carbon dioxide, inorganic nitrates, inorganic phosphates, etc. Such microbes, termed 'heterotrophs,' have little direct relevance to corrosion of buried metal.

A more detailed account of microbially-influenced corrosion (MIC) is presented in **Chapter 2.20, Corrosion in Microbial Environments** of this work, to which the reader is referred. In the context of soils, the most important group of microorganisms associated with corrosion belongs to the specialized group of bacteria known as sulfate reducing bacteria (SRB); the vast majority of MIC failures are related to their activities.

SRB thrive in oxygen-free soils and other anaerobic environments, where they obtain their required carbon from organic nutrients and their energy from the reduction of sulfate ions to sulfide. Waterlogged

soils often provide ideal environments for SRB proliferation, especially in brackish or estuarine water areas, which contain abundant sulfate ions. Sulfide appears as hydrogen sulfide (H_2S), sulfide ions, metal sulfides, or a combination of these according to conditions. Sulfides are highly corrosive to many metals, mainly because they form complex semiprotective layers that are also conductive and cathodic to the parent metal, leading to intense localized (shallow pitting) attack in areas of film breakdown. In the case of ferrous metal corrosion, the characteristic black iron sulfide (FeS) corrosion products liberate H_2S on acid treatment, distinguishing it from black iron oxide (magnetite). The corrosion products are often loose and when dislodged exhibit pits lined with bright metal corresponding to areas of anodic dissolution, but are sometimes also hard and adherent. Typical pictures of the morphology of this type of attack are shown in **Chapter 2.20, Corrosion in Microbial Environments**.

The effect of high SRB activity to corrosion rates in soils and corrosivity assessments are described in the relevant sections of this chapter (below). Suffice to say at this point that the most severe conditions with respect to MIC typically occur under conditions of low oxygen content and near-neutral pH, which would be otherwise regarded as benign. In this respect, low oxidation/reduction (redox) potential is an important indictor of intense SRB activity.

2.19.3 Corrosion Rates in Soils

2.19.3.1 Carbon and Low Alloy Steels

The monumental work of Romanoff,[2] which still forms the basis of much of our understanding of corrosion in soils, demonstrated that, in practice, all ferrous alloys (cast iron, wrought iron, mild steel, low carbon steel, etc.) are attacked at essentially the same rate in a particular soil. However, the rate of attack is not normally linear with time and can generally be modeled by an exponential relationship, as shown in eqn [1]:

$$W = k_w \cdot t^m \text{ or } P = k_p \cdot t^n \qquad [1]$$

where t is period of exposure (years); W, uniform metal loss (measured as mm, mil, $g\,m^{-2}$ or $oz\,ft^{-2}$); P, pit depth (mm or mil); k_W, k_P, scale constants (mm, mil, $g\,m^{-2}$ or $oz\,ft^{-2}$); m, n, shape constants, with n, $m \leq 1$ (typically n, $m \approx 0.33$–0.67).

The scale and shape constants are dependent on the soil. For example, data for different ferrous materials exposed to two different soil types (sand and clay respectively) over a 17 year period are shown in **Figure 6**, which presents both the raw data and the calculated regression lines for the two conditions. The shape constants are 0.702 in clay and 0.809 in sand (independent of units), with scale constants of 0.636 and 0.215 mm for clay and sand, respectively. The error bands shown in **Figure 6** are the 95% prediction range from linear regression analysis of the data which gives factors of $\times 10^{\pm 0.094}$ for clay and of $\times 10^{\pm 0.236}$ for sand. Because the data follow an exponential relationship, a plot of $\log(y)$ versus time takes the form of a straight line, which allows the standard error (ε) of the predicted value of $\log(y)$ to be obtained. The 95% prediction limit (for $\log(y)$) is given as $\pm 1.96 \cdot \varepsilon$; the corresponding limit is for y is then given by $10^{\pm 1.96 \cdot \varepsilon}$.

However, while Romanoff found that, in general, all ferrous material alloys behaved similarly in the same soil, some low alloy steels did show variable performances in specific soil types. Generally steels with chrome content of ~5% (with or without molybdenum, at ~0.5%) exhibited corrosion rates (weight loss or pitting) of about half of that plain low alloy/low carbon steels, except in very poorly aerated conditions, where they exhibited greater pitting, as shown in **Figure 7** (steels *D*, *E*, and *H*).

In the presence of high SRB activity, see **Section 2.19.2.6**, special considerations apply and localized, shallow pitting corrosion may be much higher than would otherwise be expected. In some instances, cast

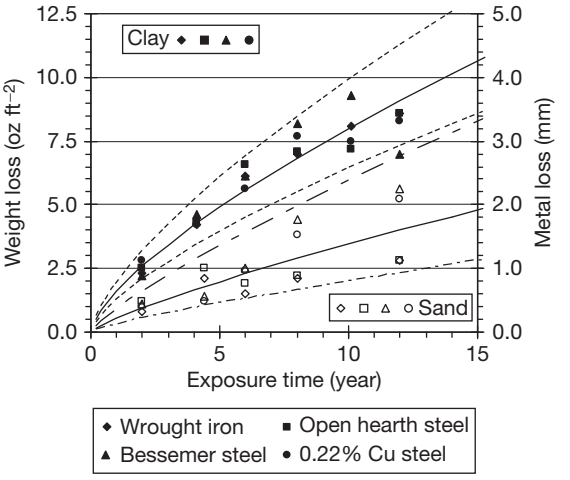


Figure 6 Comparison of measured metal loss for a range of steels in different soil types (clay and sand). Reproduced from Romanoff, M. *Underground Corrosion*; National Bureau of Standards, Circular 579 (April 1957). [Republished by NACE International, Houston, Texas, USA, 1989 ISBN 0–915567–47–4].

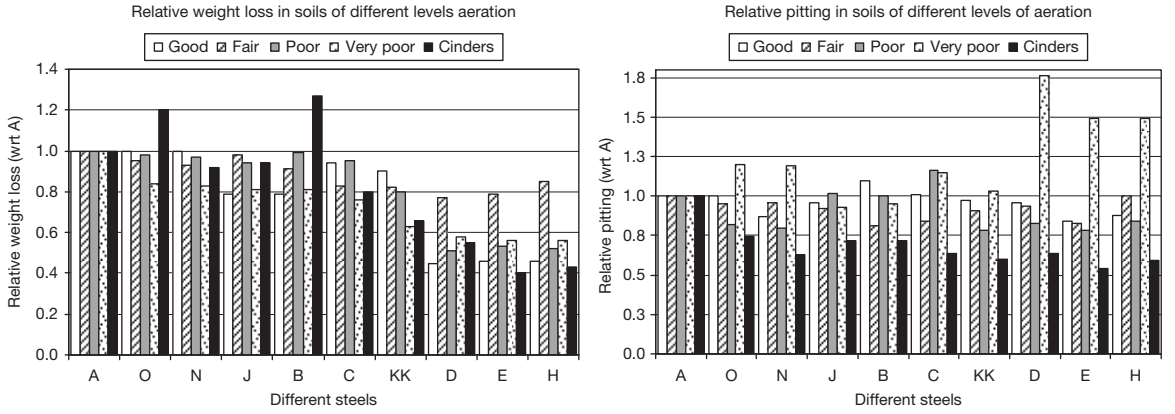


Figure 7 Relative performance of different low alloy steels in different soil types. Reproduced from Romanoff, M. *Underground Corrosion*; National Bureau of Standards, Circular 579 (April 1957). [Republished by NACE International, Houston, Texas, USA, 1989 ISBN 0–915567–47–4].

iron pipes of 8 mm thickness have become perforated within a year under such conditions, while perforation rates of 1–2 mm year^{-1} at ambient temperature are not uncommon. Under anaerobic/aerobic cycling conditions, where SRB mediated sulfide corrosion may be exacerbated by the production of elemental sulfur, and/or at elevated temperatures corrosion rates of 3–5 mm year^{-1} may be encountered.

2.19.3.2 Other Metals and Alloys

Romanoff obtained only limited experimental data for stainless steels in soils, as summarized in **Table 1**, which shows improved performance with regards to both general metal loss and pitting except for ferritic steels (grade UNS S41000/EN 1.4006 and UNS S43000/EN 1.4016), which were still relatively high at 0.75 and 0.31 mm year^{-1}, respectively.

Table 1 Corrosion performance of some stainless steels

Stainless steel grade		*General metal loss (μm year)*	*Pitting rate (μm year)*
UNS system	*EN system*		
S41000	1.4006	750 ± 650	68 ± 44
S43000	1.4016	310 ± 280	65 ± 31
S30400	1.4301	7.8 ± 7.2	15 ± 8.0
S31600	1.4401	1.3 ± 0.88	0.42 ± 0.82
S30200	1.4319	0.37 ± 0.30	0.91 ± 1.8
S30900	1.4822	0.15 ± 0.11	0.91 ± 1.8

Source: Modified from Romanoff, M. *Underground Corrosion*; National Bureau of Standards, Circular 579 (April 1957); [Republished by NACE International, Houston, Texas, USA, 1989 ISBN 0–915567–47–4].

Since the work carried out by Romanoff, a wider range of stainless steels are now available for pipework, tanks, and anchorage systems. Typically, the higher the pitting resistance equivalent number (PREN) the better the corrosion resistance for stainless steels in soils and with the main factor being consideration of chloride content and resistivity, SRB can also attack stainless steels.[12]

The corrosion behaviors of other metals, including copper, lead, and zinc obtained by Romanoff[2] are summarized in **Figure 8** for exposure in a combination of several different soil types and for tidal marsh, clay, and sandy loam.

2.19.3.3 Influence of Soil Parameters on Corrosion Rate

Romanoff concluded that the main parameter that influenced the corrosion rate of ferrous materials was the extent of aeration, with the highest corrosion rates occurring in soils with poor or very poor aeration (poor or very poor drainage) compared to lower corrosion in soils possessing good aeration (good drainage). Pooled data for steel samples from 47 test sites[2] grouped by degree of aeration as defined by Romanoff are shown in **Figure 9**. However, it is clear that aeration alone is insufficient to account for the variability among different soils.

The same pooled data from Romanoff, if grouped by resistivity, soil type, pH, or salt content, show similar levels of correlation. Consequently, these factors have been incorporated into the development of various methods for assessing corrosivity, as described later.

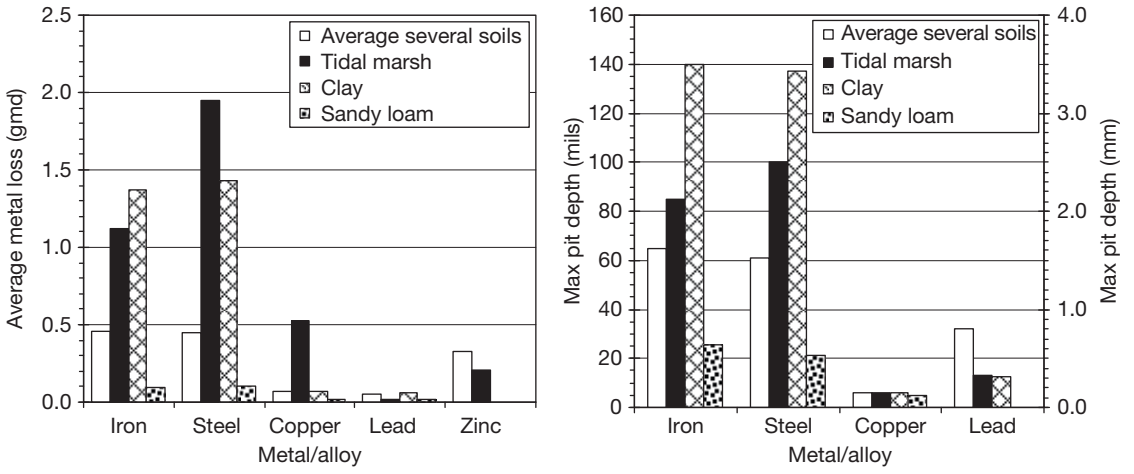


Figure 8 Average metal loss rate ($g\,m^{-2}\,day^{-1}$) and maximum pit depth (normalized to 12 years exposure) for iron, steel, copper, lead and zinc. Reproduced from Romanoff, M. *Underground Corrosion*; National Bureau of Standards, Circular 579 (April 1957). [Republished by NACE International, Houston, Texas, USA, 1989 ISBN 0–915567–47–4].

2.19.3.4 Long-Term Corrosion Rates in the Soil Environment

For most engineering applications, the rate of deterioration is considered over periods ranging from 20 to 100 years; however, corrosion in soils is a rare instance where exposure periods of more than 1000 years may also be relevant, both for strategic storage facilities (nuclear fuel storage, national archives, etc.) and with regard to archaeological artifacts.

In fact, ancient historical periods are defined by the metallurgy that was available at the time, from the Copper Age (4000–3500 BC), various Bronze Ages (Early 3500–2000 BC, Middle 2000–1600 BC, and Late 1600–1100 BC), and Iron Age (from 1100 BC to 500 AD). Lead was also widely used from 3000 BC onward. The different natural corrosion rates for the various metals and alloys, together with the soil conditions in which the metallic items were left, lead to a wide variation in the quantity and condition of artifacts found in archaeological sites, with some older artifacts fabricated from bronze being preserved in better condition than newer iron artifacts.

Studies over recent years have examined the impact of different soil parameters on bronze artifacts[13] and have shown that aeration and pH are the main factors, with highly aerated and/or acidic soils leading to the most severe deterioration, while low aeration (i.e., clay soils) leads to best preservation. These findings have wider implications as environmental changes, and in particular impact of acid rain on soils, could lead to deterioration of artifacts that have been reasonably preserved for thousands of years but are located in soils of poor buffering capacity.

For iron artifacts, while acidic conditions will also lead to severe attack, poorly aerated/clay soils tend to lead to the converse condition,[14] which is consistent with the experience on modern ferrous materials as described above. Another factor that is important is the presence of soot in the soil, formed for example, from a funeral pyre or from destruction of a dwelling by fire, which can also lead to more extensive corrosion of the artifact.

2.19.3.5 Disturbed and Undisturbed Soils

One of the main factors relating to corrosion rate in soils is whether the component is placed in disturbed or undisturbed condition. Disturbed soil is defined as one that has been mechanically moved and replaced (e.g., backfill used in laying a pipeline, burying a tank, etc.), whereas in the case of undisturbed soils the component has been driven directly into soil, for example, driven sheet steel piles, H-piles, tubular steel piles, or soil nails.

The process of mechanically moving the soil affects it by breaking up the soil (which can affect the cohesiveness of the soil) and also aerates the soil. Conversely, for driven components, the mechanical property and cohesive form of the soil are unaffected. Separate work by Romanoff in the 1960s[3] found that a soil which, if disturbed, would be considered corrosive, resulted in negligible corrosion if undisturbed.

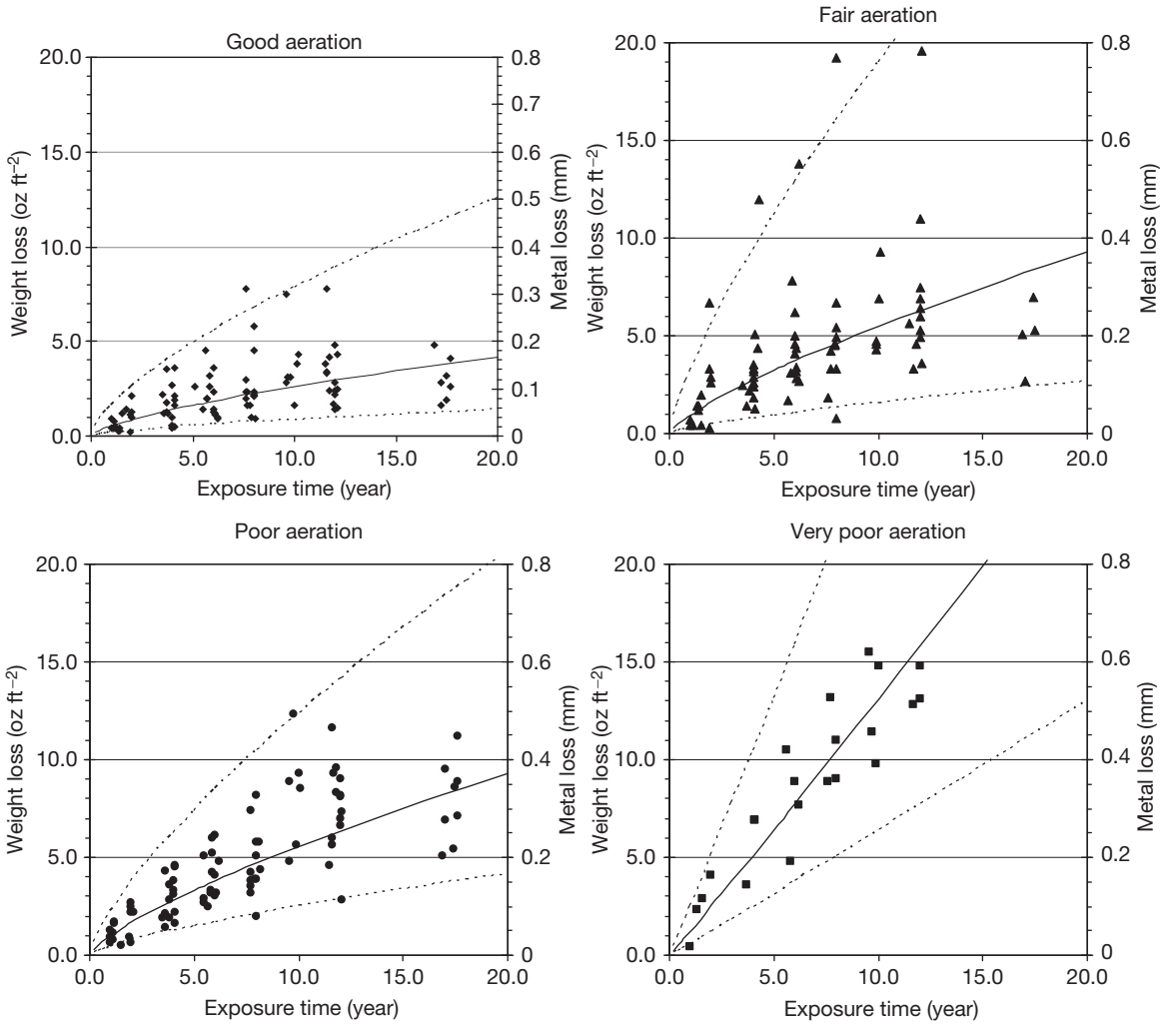


Figure 9 Pooled weight loss data for Bessemer steel (3'' Pipe) from 47 different test sites, group by degree of aeration. Reproduced from Romanoff, M. *Underground Corrosion*; National Bureau of Standards, Circular 579 (April 1957). [Republished by NACE International, Houston, Texas, USA, 1989 ISBN 0–915567–47–4] – with regression lines and 95% prediction limits.

This work combined specific field trials plus observations of steel piling that were removed after 50 years exposure.

In many cases, the extent of metal loss was minimal and only limited isolated pitting observed, which for most types of piling is not structurally significant. In some cases, it was reported that the original mill scale was still intact when the pile was removed.

Examination of piling removed from permafrost[15] also showed that no corrosion occurred on the section of pile that was permanently in frozen conditions, with little corrosion occurring in the area above the frozen ground.

A survey carried out in 2000[4] confirmed Romanoff's overall findings relating to steel piling for most soils. However, cases of significant corrosion of steel piling (single side metal loss rates of 0.06–0.11 mm year^{-1}) were found where the ground was either contaminated by cinders/coal, or had low resistivity ($<20\,\Omega$ m) or where acidic (pH <4) conditions occurred. In particular, this study identified problems at the interface between disturbed and nondisturbed areas, where differential concentration cells (macro-cells) are formed, as described above and as shown in **Figure 10**. Here the corrosion is concentrated on the section of pile in the nondisturbed soil at the interface.

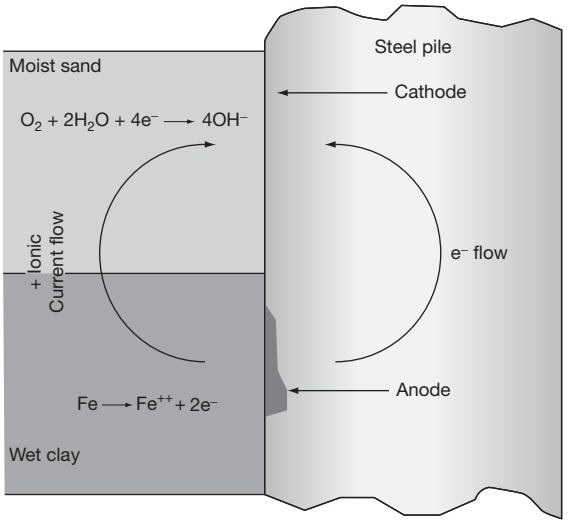


Figure 10 Creation of differential concentration (differential aeration) cells between disturbed and nondisturbed soils. NACE Technical report 05101 *State-of-the-Art Survey on corrosion of steel piling in soils* (Dec-2001); published by NACE International (item no 24216): Houston, TX, USA.

2.19.4 Corrosivity (Aggressivity) Assessment

2.19.4.1 General

Various different methods for assessing the likely corrosivity of soils have been developed. In most cases, the different procedures assign ratings (values) to the various influencing factors and the sum of the ratings is used to assess aggressiveness. Schemes have also been proposed that incorporate a statistical approach, whilst others provide general guidance based on generic conditions or rely on direct measurement of corrosion rate in the soil.

The methodology used most extensively in the UK and Europe is on the basis of original work carried out in Germany[9] and first published in English in Beakman and Swenk.[16] In Germany, the system has evolved into DIN 50 929 Part 3.[17] The Eyre & Lewis system,[18] which is a derivative of the original German system, has been adopted by the UK Highways Authority for assessment of new bridge and road projects.

2.19.4.2 Resistivity

The simplest method of assessing soil corrosivity is solely on the basis of resistivity measurements. There are several different resistivity criteria; the most widely used method is summarized in Table 2.

Table 2 Assessment of soil corrosivity based on soil resistivity

Soil resistivity (Ω m)	*Corrosivity rating*
>200	Essentially noncorrosive
100–200	Mildly corrosive
50–100	Moderately corrosive
30–50	Corrosive
10–30	Highly corrosive
<10	Extremely corrosive

Source: Reproduced from NACE Resource Centre (http://events.nace.org/library/corrosion/SoilCorrosion/Severity.asp).

Sandy soils are high up on the resistivity scale and therefore considered the least corrosive on the basis of this simplistic criterion. Clay soils, especially those contaminated with saline water, are at the opposite end of the spectrum. The soil resistivity parameter is widely used in practice and generally considered to be the dominant variable in the absence of microbial activity. However, it is widely acknowledged that resistivity alone does not determine overall soil corrosivity and hence most systems include assessment of other parameters.

2.19.4.3 Influence of SRB

A large scale study of many factors at 59 sites in the United Kingdom[6] lead to the finding that aggressive sites were characterized by soil resistivity of less than 20 Ω m or a mean redox potential more negative than +400 mV (the redox potential is defined as the potential difference measured between a platinum electrode and a suitable reference electrode in contact with the soil, normalized to standard hydrogen potential (SHE) and a pH of 7[19]). Low redox potential is attributed to reduced sulfur compounds, particularly sulfides, produced by the activities of SRB. 'Borderline' cases were classified according to water content – those containing more than 20% water being deemed aggressive soils. Similarly, all soils with a mean soluble iron content of over 120 mg kg^{-1} were found to be aggressive.

It is the authors' experience that, in addition to redox measurements (which should, be performed on undisturbed soil *in situ*), it is also useful to test small samples of soil for the direct presence of SRB. A pragmatic procedure, developed by the authors, involves two different sized samples inoculated (in triplicate) into vials of SRB culture medium

(see **Chapter 2.20, Corrosion in Microbial Environments**), which are then immediately resealed. One set of samples (milli-sample) would comprise 100–200 mg of soil material, the other set (micro-sample) less than 10 mg of material (i.e., literally only a few grains). If, after 14 days incubation at ~30 °C, all the milli- and micro-sample test vials turn black, the soil is designated as moderate to severely infected with SRB, and likely to present a high threat of MIC. If the micro-samples only turn black, the soil may be designated as low to moderately infected with SRB and hence, as a lower threat of MIC.

2.19.4.4 DIN 50 929 Part 3 Criteria

DIN 50 929 Part 3[17] provides by far the most comprehensive methodology and has been validated on many different projects, following comparison of predicted metal loss to actual metal loss.

In DIN 50 929 Part 3, the aggressiveness of soil is determined by consideration of soil type, extent of contamination, resistivity, pH, moisture content, soil condition (i.e., disturbed/undisturbed), total alkalinity/acidity (buffer capacity), combined chloride/sulfate content (aqueous extract), sulfate content (acid extract), sulfide content, presence of ground water, vertical/horizontal homogeneity, and external stray current.

DIN 52 929 Part 3 also assesses ground water in a similar manner to soil by considering water mobility, position (fully submerged, at water/air interface or in splash zone), chloride and sulfate content, buffer capacity, calcium content, pH, and external stray current.

The individual parameters are assigned values ranging from +4 (nonaggressive) to −12 (strongly aggressive), on the basis of defined criteria. The sum of the various items can range from a best case of +4 to a worst case of −47, and from this the overall corrosivity (aggressivity) is assessed. This also allows an estimate of the expected corrosion rate and the practical range (essentially a factor of × 2 on the median corrosion rate), as shown in **Table 3**.

2.19.4.5 Eyre and Lewis System

The Eyre and Lewis system[18] is a development of the original German assessment method[9] and has been adopted by the UK Highway Agency[20] for determining corrosion protection requirements for all highway structures.

The main differences between the two systems is that the Eyre and Lewis system of assessment includes the redox potential of the soil and assesses the buffering capacity (i.e., the extent to which the pH could change if acid or alkali is added) by measuring the bicarbonate content of soil as opposed to the analysis required in DIN 50 929 part 3.

2.19.4.6 AWWA/DIPRA

DIPRA (Ductile Iron Pipe Research Association, formally CIPRA – Cast Iron Pipe Research Association) developed a classification for assessing soils for determining whether protective measures are required for ductile iron pipes for use in water and waste water applications, which was first published in 1968.[21] This method was adopted by AWWA (American Water Works Association) and is now more commonly referred to as the 'AWWA 10-point system.'[22] It is used widely in North America, where it defines whether or not the minimal level of protection as required by AWWA, that is either no protection or application of plastic liner, for cast iron water and sewer pipes. It has also been adopted as part of other assessment schemes.[7]

The system assigns scores dependent on resistivity, pH, redox, sulfides, and moisture content, with a total score greater than 10 considered corrosive and requiring application of plastic liner. While it would appear to be similar to other systems as described above,[9,17,18] and there is evidence of good correlation between the predicted score and actual performance,[5] there is also evidence that it can underestimate the soil corrosivity.[23,24] This is on the basis of the observed corrosion performance being at

Table 3 Assignment of aggressiveness and expected corrosion rates for soils

Sum of ratings	*Aggressiveness*		*General corrosion rate (mm year^{-1})*	*Localized attack (mm year^{-1})*
≥0	Ia	Virtually nonaggressive	0.005 (range 0.0025–0.01)	0.03 (range 0.015–0.06)
−1 to −4	Ib	Weakly aggressive	0.01 (range 0.005–0.02)	0.05 (range 0.025–0.1)
−5 to −10	II	Aggressive	0.02 (range 0.01–0.04)	0.2 (range 0.1–0.4)
<−10	III	Strongly aggressive	0.06 (range 0.03–0.12)	0.4 (range 0.2–0.8)

Source: Reproduced from DIN 50 929 Part 3. Probability of Corrosion of Metallic Materials when Subject to Corrosion from the Outside. Buried and Underwater Pipelines and Structural Components, Sep-1985.

variance to the predicted corrosion condition. Similar variances between predicted corrosivity by the DIPRA method and actual experience, where other systems (e.g., DIN 50 929 part 3) showed good correlation, have also been observed by the authors. This is, in part, because most assessments only take into account the point scores, while the full procedure (as detailed in AWWA C105) actually requires the soil type to be taken into consideration as well. It is this omission that, in the authors' opinion, is one of the reasons for the variable results of the system.

DIPRA has recognized the limitations in its original system and published an extended methodology in 2004, referred to as Design Decision Model™ (or DMM™).[23,24] This extends the general hazard assessment to include the effect of chlorides, bimetallic effects, and known corrosive environments (i.e., soil type) and also now includes a risk assessment as shown in Figure 11 (i.e., considering the combination of both likelihood of failure and consequence of failure, used in corrosion risk assessments as described in other chapters of this publication).

As the risk of corrosion increases the required level of protection (as outlined in Section 2.19.6 below) is then increased, as follows:

1. Negligible – install the pipe as-manufactured with its protective standard shop coating/annealing oxide system.
2. Low – encase the pipe in polyethylene.
3. Medium – encase the pipe in polyethylene or encase the pipe and provide bonded joints.
4. High – encase the pipe in polyethylene and provide bonded joints or provide life-extension cathodic protection currents, with or without encasement.
5. Severe – apply cathodic protection.

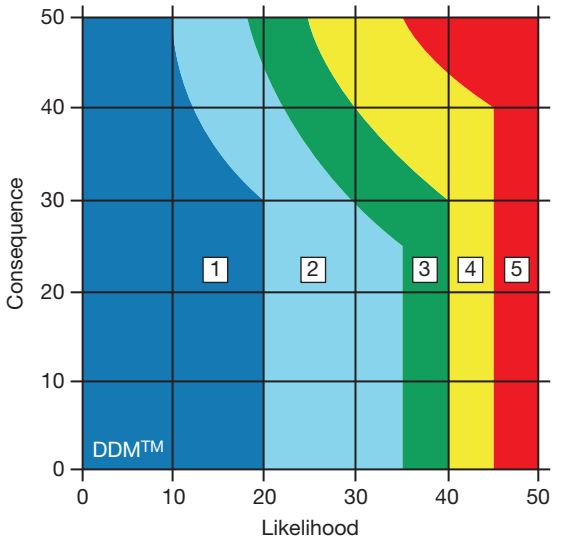


Figure 11 Design decision method (DDM™) corrosion risk assessment. Reproduced from Kroon, D. H.; Lindermuth, D. L. *Corrosion Protection of Ductile Iron Pipe*, Paper no 04046, NACE Corrosion 2003, New Orleans, March 2004; NACE International; Houston, Texas.

2.19.4.7 Mean Time Before Failure (MTBF)

As noted previously, there has been increasing concern relating to environmental contamination due to leaking underground storage tanks caused by corrosion, especially in North America. This led to the development of ASTM standard G 158[25] for assessing the condition of buried tanks and applying remedial measures to ensure integrity, which is now mandated in the United States.

One of the methods described in the standard is the so called MTBF method, which determines the likely time to failure (through wall pitting in this case) of buried storage tanks, on the basis of a statistical assessment of various soil parameters and was developed from previous published work.[26] The parameters used in the assessment are obtained either directly on site or from samples analyzed in a laboratory and compared to a data set of corrosion rates obtained from a number of other sites, so as to be able to demonstrate a statistical confidence level of 99%.

The standard indicates that a multivariate nonlinear regression analysis that contains the following minimum essential variables may be used, for example, as shown in eqn [2]:

$$E(age) = f(R, M, Cl, pH, S, SC, TS, P) \qquad [2]$$

where $E(age)$ = unconditional predicated age to corrosion failure; R = resistivity; M = moisture content; Cl = chloride ion content; pH = soil pH; S = sulfide content; SC = stray current magnitude; TS = tank size; and P = tank structure-to-soil potential.

However, the standard also allows other parameters such as redox potential, sulfate content and others tests as may be considered appropriate to be included.

It is clear that while the overall procedure as described in ASTM G 158 is generally sound, it first requires the user to develop a sufficiently large data set of soil types, environmental conditions, and associated measured corrosion rates of actual structures before any assessment for further tanks can be carried out. Several commercial organizations have

developed data sets and are therefore able to carry out such assessments. The requirement for a 99% confidence level would normally be sufficient to compensate for any limitation of data for a particular parameter. In the absence of a specific database to allow an assessment, the extensive data from the original work by Romanoff[2] could be utilized.

Other statistical approaches to the development of predicting corrosion rates and time to failure have also been proposed by other workers.[27,28]

2.19.4.8 Direct Measurement Approach

As an alternative (or in support of) assessment of soil corrosivity by consideration of relevant parameters, several workers have carried out direct measurement of corrosion in soils, either in the field or on soil samples brought back to the laboratory. Such measurements have used a wide range of different procedures including, for example, corrosion coupons, linear polarization resistance (LPR), electrical resistance (ER), and other electrochemical techniques.

As would by now be realized the most extensive use of exposure coupons to date was carried out by Romanoff,[2,3] who also described some of the earliest electrochemical measurements using what was essentially a LPR measurement, which was reported as giving very good correlation to weight loss measurements.

More recently, the use of ER and LPR probes has been applied to corrosion testing in soils, either directly[29,30] or as part of assessing performance of cathodic protection systems.[31,32]

In any system using buried probes, it should be self-evident that for meaningful information to be obtained the probes must be buried in the same soil and under the same conditions as the structure (pipeline/foundation) they are intended to represent; this is particularly important if they are being used to assess corrosivity for driven components (such as soil nails, driven piles, etc.). It should also be self-evident that an isolated probe/exposure coupon will provide only information on the inherent corrosivity of the soil. If there is a concern relating to either long-line (macro-cell) corrosion or external stray current corrosion, the probe must be electrically connected to the structure of concern. However, there must also be a method of temporarily disconnecting the probe from the structure[31] to allow measurements to take place (when using ER probes) or some other means of isolation (when using LPR probes).

2.19.4.9 Summary

The key features associated with any of the different methods of assessment as applied to soils are the following:

- All soils contaminated with coal, coke, industrial waste, and so on are strongly aggressive.
- The lower the resistivity the more aggressive the conditions, with very low resistivity soil ($<10\,\Omega\,m$) normally associated with strongly aggressive conditions.
- While soil/groundwater pH can range from <4.5 to >9.5, its effect can be less important than the buffering capacity of the soil. Soils which are strongly alkaline buffered (measured as alkalinity to $K_{S\ 4.3}$ of $>1000\ mmol\ kg^{-1}$ or carbonate content $>1500\ mmol\ kg^{-1}$) have sufficient residual alkalinity to be beneficial. Soils which are strongly acidic buffered (as measured by acidity to $K_{B\ 7.0} > 20\ mmol\ kg^{-1}$) are very corrosive, while a low buffering capacity (e.g., $K_{B\ 7.0} < 5\ mmol\ kg^{-1}$) has little impact on overall aggressiveness.
- Where a metal component passes through different soils (e.g., a pipeline running from one soil to another or a pile/anchorage driven vertically through different soil layers) the main concern relates to the formation of differential aeration and/or concentration cells, where corrosion will be located at or just below the interfacial layer.
- Buried structures are also at risk from stray current sources (e.g., from neighboring cathodic protection systems, dc welding, dc railway systems, etc.). This risk is assessed by the change in potential on site and hence measurements of structure-to-soil potential (as required in DIN 50 929 part 3 and others systems) are only directly applicable to existing facilities. Other methods are available to assess stray current risks on site or during the design stage, using computer modeling.[33]
- Where the soil contains active SRB, which can be determined by the presence of sulfides, measurement on site of redox potential or bacterial assay, it can lead to microbiologically influenced corrosion (MIC) which can be severe.

It should also be noted that while some assessment methods (in particular DIN 50 929 Part 3) state that they cover both disturbed and undisturbed soil conditions, in practice the rate of corrosion of components in undisturbed soil is significantly lower. As noted previously, Romanoff[3] concluded that for driven piles the

portion in the undisturbed soil did not suffer significant attack. This concept has been adopted in the recent Eurocode 3 Part 5[34] where projected corrosion rates for a particular soil show essentially a doubling of corrosion rate depending on whether the soil conditions are compacted, driven or noncompacted.

In the case of ground anchorages,[35] the appropriate method of determining aggressiveness will be dependent on the type of anchor. For high strength prestressed ground anchors, the aggressiveness of soils (disturbed or undisturbed) is directly relevant, for example, in ground anchorage systems where the steel component is in direct contact with the soil (e.g., where the tendon has lost its cement grout cover) or the protective system fails (e.g., tie bars with damaged protective tape wrapping). For other anchor types, for example, soil nails or drilled piles, the steel component is inserted into undisturbed soil. For permanent anchorages, where the components are protected by a corrosion protection system, in accordance with most national standards, the steel tendon is not in direct contact with the ground; hence, it is the aggressiveness of the ground water that is most relevant, as it is the ingress of ground water through any breaks in the protective system that is of concern.

There are also many more published soil corrosivity assessment methods than just those described here, with varying degrees of acceptance and success.

Where stainless steels are being considered, their use should be on the basis of chloride content and soil resistivity. In the absence of external stray currents and for soils with groundwater pH > 4.5, suitable stainless steel grades that can be used without coatings or cathodic protection are shown in Table 4.

Table 4 Use of stainless steels in soils

Chloride level (mg kg^{-1})	<500	500–1500	1500–6000
Resistivity (Ω m)	>10	>10	5–10
Suitable grades (UNS/EN)	S30400/ 1.4301 S31600/ 1.4401 S44400/ 1.4521	S31600/ 1.4401 S44400/ 1.4521	S32750/ 1.4410 S31254/ 1.4547

Source: from Pierrie-Jean Cunat, *Corrosion Resistance of Stainless Steels in Soils and in Concrete*; paper presented at CEOCOR, Biarritz, Oct-2001.

2.19.5 Metallic Coatings in Soils

Several different metallic coatings have been used to provide protection to ferrous components in soils, principally galvanizing but also thermal spray coatings of zinc, aluminum, and zinc–aluminum alloys.

The performance of zinc coatings is interesting in that the mechanism of protection in soils differs from that provided in other environments (in particular natural waters and the atmosphere). In soils, initially the zinc acts as a barrier with a generally lower natural corrosion rate (compared to iron or steel) and also provides galvanic protection at any damage site where substrate is exposed. However, the longer term protection from the zinc coating is also provided by the *de facto* formation of insoluble corrosion product comprising zinc salts (mainly sulfates and carbonates), which forms as a semi protective film *in situ*. This extends the level of protection, even when the bulk of the zinc has been consumed.[36]

The overall corrosion rate of zinc is dependent on soil pH and resistivity, the performance of the metal being less efficient in acidic soils. The method for determining the rate of deterioration of zinc in galvanized steel sheets used, for example for culverts, was developed by California Division of Highways[37] (based on the work by Romanoff), as shown in Figure 12. Note 16 Gage galvanized steel sheet is nominally 0.06 inch (1.5 mm) thick with 3.5 mil (88 μm) thick zinc coating. Gage 14, 12, 10, and 8 galvanized sheets are respectively 0.075, 0.105, 0.135, and 0.165 inch (1.9, 2.7, 3.4, and 4.2 mm) all with nominal 3.5 mil (88 μm) zinc coating.

Aluminum coatings act only as a barrier, with no significant long term protection once the aluminum has been consumed (which can occur rapidly in alkaline soils).

The use of thermal spray zinc–aluminum alloy, in particular Zn85–Al15, has been shown to provide improved performance, compared to pure zinc coatings, in many soil environments, particularly when used with a finishing layer (typically an thin, semiporous, epoxy coating). The research indicates that the presence of the aluminum reduces the self corrosion of the coating, while still allowing the formation of *in situ* insoluble zinc salts. In fact, by slowing the rate of formation of the zinc salts the protective nature of the film is improved.[38,39] This allows a zinc/aluminum coating to be used in an extended range of soil types compared to a zinc

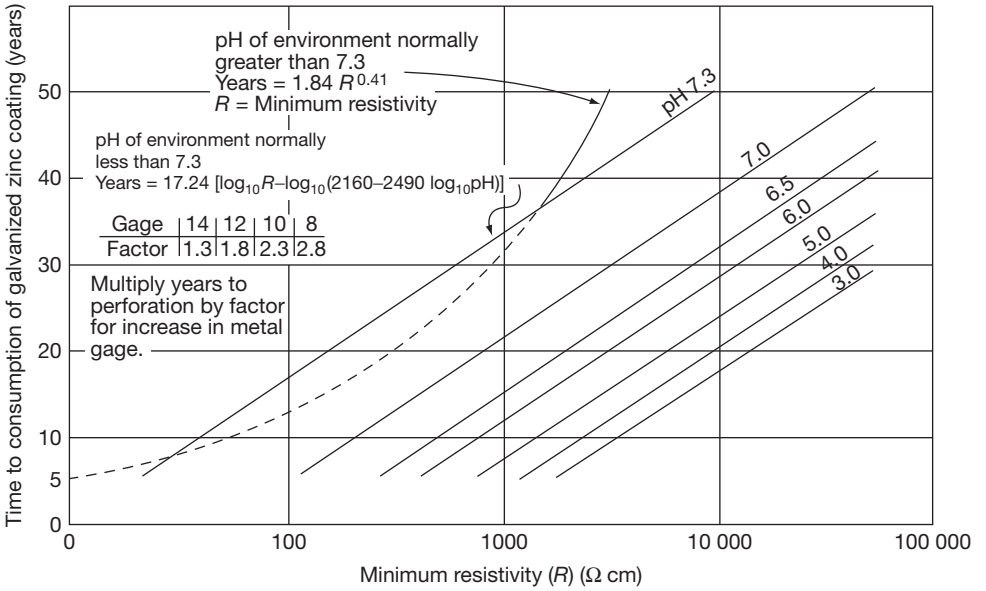


Figure 12 Estimating service life of galvanized steel in soils. Reproduced from *NACE Corrosion Engineers Reference Book*, 3rd ed.; Baboian, R. Ed., ISBN 1–57590–127–7.

coating. However, it is not suitable for very aggressive soils, for example, acidic peaty soils (pH < 5), polluted/contaminated soils (with cinders, slag, etc.), or marine soils below the water table (with resistivity <5 Ω m).

2.19.6 Corrosion Control Options

Practical corrosion control options for buried components are dependent on the structure/component type, soil environment, and impact of any resultant corrosion, and will normally comprise one or more of the following:

- do nothing
- application of a corrosion allowance
- use of selected backfill
- use of protective sheeting (for ductile iron pipe)
- application of organic protective coating (bitumen, epoxy, polyurethane, etc.)
- application of metallic coating (galvanizing, flame sprayed zinc/aluminum alloy)
- cathodic protection (galvanic or impressed current).

As can be seen, the option of do-nothing and/or inclusion of a corrosion allowance can be a practical consideration for some structures/components, where the soil aggressivity may be classed as negligible or low (i.e., general metal loss rates of <0.01 mm year^{-1} or localized pitting rates of <0.05 mm year^{-1}) and where the consequence of any corrosion is acceptable.

The use of a corrosion allowance is normally applied to piles where protective coatings are not suitable (see below). The corrosion allowance is an additional thickness over that necessary for structural strength to allow for metal loss over the design life (as detailed in the new Eurocode 3[34]).

The use of selective backfill, for pipelines or buried tanks, is another method of reducing the corrosion threat for buried components. Any backfill should be sifted to remove stones, which could damage the protective coating. Where the local soil is aggressive, for example, clay based, contaminated with cinders, low pH, or high salt content, imported material should always be used, although it has to be recognized that this is not always practical in remote sites. When selected backfill is used it is also important to ensure that the same material is applied across all components. This is because if backfills of significantly different compositions are used in different locations long-line (macro-cell) corrosion could develop. For example, in 1980 in Ortuella, near Bilbao, Spain, a butane gas tank, which was buried in predominantly limestone fill and connected to a poorly protected small bore pipe, was buried in nonselective backfill. This led to severe corrosion of the pipe causing a leak. This in turn resulted in an explosion in a school

basement and the deaths of forty-seven children and three adults.

Protective plastic sheeting has been used extensively for ductile iron pipes and, where the soil is moderately aggressive, it can provide benefit. It should be noted however, that there is a major variance between European and North American practice; in Europe plastic sheeting is applied over galvanized or thermally sprayed zinc coating plus bitumen, while in North America it has been applied over otherwise uncoated pipe.[40]

Protective coatings are widely applied as the principle method of corrosion control for buried components especially pipelines and buried tanks. They can also be applied to driven components (sheet steel piles, H-piles, soil nails, etc.); however, because the driving process is likely to cause significant mechanical damage to any applied coating, it cannot, in such circumstances, be considered as a primary protection measure. Nonetheless, the use of protective coating can still be beneficial, especially where cathodic protection will also be applied, as it will significantly reduce the current demand. There are a wide range of protective coatings that can be used for buried components (especially pipelines). These include tape wrap systems, bitumen, liquid applied epoxies and polyurethanes, fusion bonded epoxy, and multilayer epoxy/polyurethane coatings. Selection of an appropriate coating depends on the specific soil condition (in particular, whether the component will be below or above the water table). Mechanical damage of coatings from the soil (so called soil stress) can also be a major concern, especially for tape wrap and some epoxy systems).

For small components and cast iron, increasing use is made of metallic coatings (as summarized in **Section 2.19.5**) comprising galvanizing and flame sprayed zinc or zinc–aluminum alloys. However, it should be recognized that metallic coatings are more sensitive to soil pH than organic coatings; in particular, zinc coatings are not suitable for soils with $pH < 5$ and aluminum coatings are unsuitable for $pH > 9$).

Successful cathodic protection can be defined as where the effective corrosion rate is <0.01 mm year^{-1}; consequently use of cathodic protection may not be considered necessary for soils which are classed as 'noncorrosive.' However, for high risk assets (e.g., high pressure oil and gas pipelines, buried hydrocarbon storage tanks, etc.) with unacceptable consequences of failure, cathodic protection is almost universally applied, even in areas where soil aggressivity is negligible; its application to other assets is variable across the world. In some areas, in particular in the tropics, the Middle East, North Africa and elsewhere, cathodic protection is applied to the majority (if not all) of buried facilities, for the simple reason that soils in these areas are often very aggressive.

Cathodic protection can be applied by either galvanic or impressed current systems, the decision being dependent on ground resistivity (zinc galvanic anodes are normally only used in soils with resistivity $<30\,\Omega$ m; magnesium anodes $<150\,\Omega$ m and impressed current otherwise); but selection of the most appropriate system will be influenced by the availability of power supplies, access (for maintenance), required design life, and whether there are any concerns relating to over protection (which can lead to stress corrosion cracking).

References

1. *Soil Survey Manual* 1993. Soil Conservation Service. U.S. Department of Agriculture Handbook 18. United States Government Printing Office, Washington, DC. (ISBN-10: 1410204170) (http://soils.usda.gov/technical/manual)
2. Romanoff, M. *Underground Corrosion*; National Bureau of Standards, Circular 579 (April 1957); Republished by NACE International, Houston, Texas, USA, 1989 ISBN 0–915567–47–4
3. Romanoff, M. *Performance of steel piling in soils*; Paper no 82; presented at NACE Corrosion/69: Houston, TX, 1969.
4. NACE Technical report 05101 *State-of-the-Art Survey on corrosion of steel piling in soils* (Dec-2001); published by NACE International (item no 24216): Houston, TX, USA.
5. Bonds, R. W.; Barnard, L. M.; Horton, A. M.; Oliver, G. L. *J. Amer. Water Works Assoc.* **2005**, *97*(6), 88–98.
6. Booth, G. H.; Cooper, A. W.; Cooper, P. M.; Tiller, A. K.; Wakerley, D. S. *Brit. Corr. J.* **1967**, *2*, 104–118.
7. Jarvis, M. G.; Hedges, M. R. *Water Environ. J.* **1994**, *8*(1), 68–75.
8. Camitz, G.; Vinka, T. G. Corrosion of Steel and Metal Coated Steel in Swedish Soils – Effects of Soil Parameters In *ASTM STP 1013, Effects of Soil Characteristics on Corrosion*; Chaker, V., Palmer, J. D., Eds.; ASTM: Philadelphia, 1989; pp 37–53. [ISBN 0-80321-1189-4].
9. DVGW Arbeitsblatt GW9 *Merkblatt für die Beuteilung der Korrosionsgefährdung von Eisen und Stahl im Erdboden (Instruction sheet for the assessment of corrosion risk of steel and iron in contact with soil)* Aug-1971.
10. Krasnoyarskii, V. V.; Larionov, A. K. *Podzemnaya korroziya metallov i metody bor'by s nei (underground corrosion of metals and corrosion protection)*, Moscow, Ministry of Municipal Economy of the Russian Federation, 1962.
11. Harris, J. O.; Erye, D. Chapter 2.6 *Shrier Corrosion*, 3rd ed.; Butterworth-Heinemann: Oxford, 1994.
12. Pierrie-Jean Cunat *Corrosion Resistance of Stainless Steels in Soils and in Concrete*; paper presented at CEOCOR, Biarritz, Oct-2001.
13. Tylecote, R. F. *J. Archaeol. Sci.* **1979**, *6*, 345–368.
14. Neffa, D.; Dillmannc, P.; Bellot-Gurletd, L.; Berangere, G. *Corros. Sci.* **2005**, *47*(2), 515–535.
15. Romanoff, M. Corrosion Evaluation of Steel Test Piles Exposed to Permafrost Soils, Paper no 81; presented at NACE Corrosion/69: Houston, March 1969.

16. Von Baeckmann, W.; Schwenk, W. *Handbook of Cathodic Protection*, 1st ed.; Portcullis Press: England, 1975.
17. DIN 50 929 Part 3. *Probability of Corrosion of Metallic Materials when Subject to Corrosion from the Outside. Buried and Underwater Pipelines and Structural Components*, Sep-1985.
18. Eyre, D.; Lewis, D. A. *Soil Corrosivity Assessment* (Contractor Report 54); Transport and Road Research Laboratory: Crowthorne, UK, 1987.
19. BS 1377–3: 1990 *Method of test for soils for civil engineering purposes – Part 3: Chemical and Electrochemical tests.*
20. DMRB BD 42/00 *Design Manual for Roads and Bridges*, Vol. 2 *Highway Structures: Design (Substructures, Special Structures and Materials)*, Section 1 *Substructures*, Part 2 *Design of embedded retaining walls and bridge abutments*, UK Highways Agency (2000) http://www.standardsforhighways.co.uk.
21. Smith, W. H. *J. Am. Wat. Wks. Ass.* **1968**, *60*, 221–227.
22. ANSI/AWWA C105/A21.5–05 *American National Standard for Polyethylene Encasement for Ductile-Iron Pipe Systems* Dec-2005 (Appendix 1).
23. *The Design Decision Model™ for corrosion control of ductile iron pipelines* ‘‘DIPRA DDM/7–06/5 M, July 2006. http://www.dipra.net.
24. Kroon, D. H.; Lindermuth, D. L. *Corrosion Protection of Ductile Iron Pipe*, Paper no 04046, NACE Corrosion 2004, New Orleans, March 2004; NACE International; Houston, Texas.
25. ASTM G 158 – 98 (2004) *Standard Guide for Three Methods of Assessing Buried Steel Tanks*.
26. Bushman, J. B.; Mehalick, T. E. Statistical Analysis of Soul Characteristics to Predict Mean Time to Corrosion Failure of Underground Metallic Structures. In *Effects on Soil Characteristics on Corrosion*; ASTM STP 1013; Chakar, V., Palmer, J. D., Eds.; ASTM: Philadelphia, 1989; pp 107–118 [ISBN 0–80321–1189–4].
27. Song, G. L.; Cao, C. N.; Lin, H. C. *Corrosion* **1995**, *51*(7), 491–497.
28. Kajiyama, F.; Koyama, Y. *Corrosion* **1997**, *53*(2), 156–162.
29. Ansuini, F.; Yaffe, M.; Chaker, V. *Corrosion Rate Sensors for Soil, Water and Concrete*; Paper 14; presented at NACE Corrosion/95 conference: Houston, TX, 1995.
30. Li, M. C.; Han, Z.; Lin, H. C.; Cao, C. N. *Corrosion* **2001**, *57*(10), pp 913–917.
31. NACE Technical Report 05017 *Report on Corrosion Probes in Soil or Concrete* Aug-2007; NACE International: Houston, TX, USA.
32. Sun, X. *Real-Time Monitoring of Corrosion in Soil Utilizing Coupled Multi-Electrode Array Sensors,* Paper no 05381; presented at NACE Corrosion/2005: Houston, TX, Apr- 2005.
33. Aylott, P.; Cotton, I.; Hassanein, A. *The Application of Modelling Systems at the Design Stage to the Mitigation of Stray Current Interference,* paper no 03706; NACE Corrosion/2003: San Diego, CA, USA, 2003.
34. EN 1993–5: 2007, *Eurocode 3: Design of steel structures – Part 5: Piling.*
35. John, G.; Littlejohn, G. S. Durability of Buried and Encased Ground Anchorages and Structural Components; Requirements for Corrosivity Assessment and Protection to Ensure Service Life, In *Ground Anchorages and Anchored Structures in Service*; Littlejohn, G. S., Ed.; Thomas Telford: London, 2008, pp 263–272.
36. A Comprehensive Ductile Iron Pipeline System; Pont-a-Mousson SA: France, 1992.
37. *NACE Corrosion Engineers Reference Book*, 3rd ed.; Baboian, R. Ed., ISBN 1–57590–127–7.
38. Nouail, G. *A New Coating for Ductlie Iron Pipes based on Zinc-Aluminium 85–15 alloy;* 3R International: Rohre Rohrleitungsbau Rohrleitungstransport, 2001; Vol. 40, Chapter 2, pp 120–124.
39. Nouail, G.; Tournier, A. *Utilisation des alliages de zinc-aluminium comme protection anticorrosion - application au revetement des tuyaux en fonte*, paper presented at CEOCOR Colloque Biarritz, 2001.
40. Rajani, B.; Kleiner, Y. *J. Am. Wat. Wks Assoc.* **2003**, *95*(11), 110–125.
41. *Corrosion Atlas*, 3rd ed.; Compiled by During, E. D. D.; Elsevier Science: London, 1997; (Item 06.05.05.01).

Further Reading

Chaker, V.; Palmer, J. D., Eds.; *Effects of Soil Characteristics on Corrosion;* ASTM STP 1013 1989, ISBN 0–8031–1189–4.
Escalante, E., Ed.; *Underground Corrosion*, ASTM STP 741, 1981, ISBN: 0–8031–0703-X.
Handbook of Corrosion Engineering, Roberge, P. R.; McGraw-Hill Professional, 1999. ISBN 0070765162.
Littlejohn, G. S., Ed. *Ground Anchorages and Anchored Structures in Service*, Institute of Civil Engineers; Thomas Telford Publishing: London, 2008, ISBN 978–07277–3561–4.
Practical Handbook of Corrosion Control in Soils, 1st ed.; Bradford, S. A., Ed.; CASTI Publishing Inc.: Edmonton, Alberta, 2006. ISBN 1-894038-48-7.
Romanoff, M. *Underground Corrosion*, National Bureau of Standards, Circular 579 (April 1957); Republished by NACE International: Houston, TX, USA, 1989. ISBN 0–915567–47–4.
Wilmott, M. J.; Jack, T. R. Corrosion by Soils. In *Uhlig's Corrosion Handbook*, 2nd ed.; Winston, R. R. Ed.; John Wiley & Sons, 2000, ISBN 978–0–471–15777–9.

2.20 Corrosion in Microbial Environments

J. F. D. Stott
CAPCIS Ltd., CAPCIS House, 1 Echo Street, Manchester M1 7DP, UK

Glossary

Anaerobic In the absence of oxygen, cf. aerobic.

Autotroph An organism that produces complex organic compounds from inorganic carbon dioxide and an external source of energy, such as sunlight or chemical reactions of inorganic compounds. All green plants and certain specialized bacteria are autotrophs.

Bacteria Unicellular microorganisms, typically a few micrometers in length and having many shapes, including spheres, rods, and spirals.

Biocide Usually refers to a chemical agent that is intended to kill microbiological life, as distinct from sterilization by heat or radiation.

Biofilm A complex aggregation of microorganisms marked by the excretion of a protective and adhesive matrix and also often characterized by surface attachment, structural heterogeneity, genetic diversity, and complex community interactions.

Biogenic Arising from living processes.

Biostud One of a set of small, replaceable metal coupons mounted inside a side stream, used to monitor biofilm.

Chlorination The process of adding elemental chlorine, hypochlorous acid, or hypochlorite ions in solution to water as a biocide to reduce the prevalence of microorganisms.

Fungi Microorganisms with cells much larger than bacteria that digest their food externally, absorbing nutrient molecules into their cells. They were originally classified as simple plants; however, they have been separated as heterotrophs.

Heterotroph An organism that requires organic substrates to get its carbon for growth and development in contrast with autotrophs, which use inorganic carbon dioxide as their sole carbon source. All animals are heterotrophic, as well as fungi and many but not all bacteria.

Mesophilic Microorganisms that grow optimally within the temperature range 20–40 °C.

Metabolism The complete set of chemical reactions that occurs in living cells. These processes are the basis of life, allowing cells to grow and reproduce, maintain their structures, and respond to their environments.

Microbially-influenced corrosion (MIC) The deterioration of the essential properties of a material due to reactions caused or promoted by the activity of microorganisms.

Microorganism Or *microbe* is an organism that is microscopic (too small to be seen by the human eye). The study of microorganisms is called microbiology. Microorganisms can be bacteria, fungi, archaea, or protists, but not viruses and prions, which are generally classified as nonliving.

mg l^{-1} A concentration term to define the amount of a substance which is the practical numerical equivalent to the derived SI unit of milligrams per cubic decimeter (mg dm^{-3}).

Nitrate-reducing bacteria (NRB) Nitrate-reducing bacteria, comprising many groups of bacteria that use nitrate as an oxidizing agent for their energy-generating metabolism, reducing it to nitrogen gas, sometimes with the generation of nitrite or other nitrogen oxides as by-products.

Planktonic Bacteria that are suspended or free floating in a liquid environment.

Redox Reduction/oxidation potential. High redox potentials (E_h positive) reflect oxidizing conditions, whereas low redox potentials (E_h negative) reflect reducing conditions.

Sessile Stationary bacteria that are permanently attached to a surface as a biofilm.

Side stream A small test pipe, generally equipped with a series of isolation valves, through which water flows and designed to hold multiple flush mounted, replacable metal studs for the measurement, analysis, or testing of biofilms in industrial water systems.

Souring The appearance of hydrogen sulfide (H_2S) in production fluids as a consequence of the introduction and activity of sulfate-reducing bacteria, which generate H_2S within a waterflooded zone of (particularly) a subsurface petroleum reservoir.

Sulfate-reducing bacteria (SRB) Sulfate-reducing bacteria, comprising several groups of anaerobic bacteria that use sulfate as an oxidizing agent for their energy-generating metabolism, reducing it to sulfide, which appears as H_2S, sufide ions, metal sulfides, or a combination of these forms according to the conditions.

Substrate Any substance required for the growth or metabolism of a microorganism, other than materials required in trace quantities only.

Thermophilic Microorganisms that grow optimally at temperatures above the mesophilic range and most typically at greater than 55 °C.

Water injection A method used in crude oil production where water is injected back into the subsurface petroleum reservoir to increase pressure and thereby stimulate production.

2.20.1 Introduction

'Microbiologically-induced corrosion,' nowadays more commonly termed 'microbially-influenced corrosion' (MIC), is a rather special case of the widespread occurrence of biodeterioration of materials. It is important to recognize that MIC of metals is not fundamentally different from other types of aqueous electrochemical corrosion. Microorganisms do not directly consume metals in the way that they do various types of organic substances; it is rather that aggressive environments are created by certain types of microorganisms as by-products of their energy-obtaining metabolism.

Microorganisms can be viewed as 'catalytic' entities, bringing about chemical reactions that would otherwise not occur because of their high activation energy. Such processes may involve the production of aggressive chemical agents, such as acids, or the consumption of substances that are important in corrosion reactions, such as oxygen or nitrite corrosion inhibitors. Alternatively, the physical presence of microorganisms may form a slime or poultice, leading to differential aeration corrosion or crevice corrosion. Some microorganisms also chemically break down the desirable physical properties of lubricating oils and protective coatings.[1–3]

Microorganisms, including the corrosion-inducing microorganisms, are present almost everywhere in soils, freshwater, seawater, and air. Therefore, the mere presence of microorganisms in an environment does not necessarily indicate a corrosion problem. What is important is the number of microorganisms of the relevant types. The salient characteristics of these microorganisms from the engineer's viewpoint are that they are small, ubiquitous, potentially very rapid-growing, and subject to certain common restraints such as temperature, pH, and nutrient availability. A corrosion problem exists only when conditions become favorable for a specific microbial population to 'explode,' giving rise to millions of cells per gram of environmental material. Even then, most microorganisms are harmless from the standpoint of corrosion; therefore, it is important that microbiological assessment be type specific.

The role of microorganisms in the corrosion of metals was the subject of a number of classic review articles in the 1970s and 1980s[4–7]; though now somewhat dated, they are still worth reading for their clarity and objectivity. For more detailed recent reviews of the subject, the reader is referred to European Federation of Corrosion Report, number 29, published in 2000,[8] the '*Practical Manual on Microbiologically Influenced Corrosion*,' published by NACE in 2001,[9] the very practical and engineering-orientated report titled '*Microbiologically Influenced Corrosion and Biofouling in Oilfield Equipment*,'[10] also published by NACE, and the '*Microbiologically Influenced Corrosion Handbook*' published by the Industrial Press, USA.[11] For a detailed review of emerging tools and modern trends in MIC technology, the reader is referred to the research review article by Videla and Herrera.[12]

2.20.2 Microorganisms

2.20.2.1 General

Microorganisms or 'microbes' are commonly defined as living entities that are too small to be visible to the human eye. The main differentiating feature, other than size alone, between microorganisms and larger, more complex organisms is that there is little or no degree of cell differentiation, one cell being able to emulate the functions of the whole organism. The microorganisms traditionally comprise bacteria, fungi, algae, and small unicellular animal-like protists, but not the minute and very simple (in biological terms) viruses and prions, which are now generally classified as 'nonliving.' Microorganisms are often described as single-celled (unicellular) organisms; however, some unicellular protists are visible to the human eye, and some multicellular species are microscopic. It seems to be a general biological principal that every criterion we devise for defining and classifying things breaks down at some point. We need to remember that classification – the process of putting objects into 'pigeon holes' – is something fundamental to human beings rather than nature. However, such classification is useful and necessary in reducing chaos to some semblance of order.

2.20.2.2 Bacteria

The true bacteria are unicellular microorganisms, neither animals nor plants, but belonging to a separate and more primitive domain of organisms called

prokaryotes. They are a few micrometers in size and have many different shapes, including spheres, rods, and spirals. Though often occurring in pairs, clusters, or chains owing to incomplete separation after cell division, such aggregates are not essential to their existence. Bacteria are ubiquitous, living in almost every possible habitat on the planet, including soil, underwater, and even deep inside the earth's crust. Some types are aerobic, requiring oxygen, whereas other types are anaerobic, growing in the absence of oxygen.

The defining characteristic of bacterial cells, apart from their very small size, is that they do not contain a nucleus or other membrane-bound discrete structures having specialized functions (organelles). Bacteria are also far more metabolically varied than higher organisms, and the various diverse species are able to obtain energy by carrying out a wide range of chemical reactions. Most types of bacteria are heterotrophs, meaning that they do not fix their own carbon through photosynthesis, but use the oxidation of organic carbon fixed by other organisms. However, a few specialized bacteria carry out inorganic reactions involving oxidation or reduction of nitrogen, sulfur, or iron. It is the activities of these latter groups, the so-called chemolithotrophic bacteria, that most concern the corrosion engineer, even though they represent a tiny minority of the bacteria in the biosphere.

2.20.2.3 Fungi

Fungi were originally classified as plants; however, they have more recently been separated, as they are heterotrophs and are now thought to be more closely related to animals than to plants. The fungi are now placed in their own kingdom. Fungal cells are much larger than those of bacteria and always contain nuclei. The simplest fungi are yeasts, which are completely unicellular, as are bacteria, and reproduce by 'budding.' Most fungi are filamentous organisms, often referred to as moulds, comprising a tanged mass of filaments (hyphae), known collectively as a mycelium. The higher fungi include most common mushrooms, which are clearly large and multicellular, at least in one phase of their life cycle, but are biochemically and genetically closely related to the microscopic fungi. The fungi of most interest to the corrosion engineer are those that attack some types of polymers and paints and the 'kerosene fungus,' which secretes organic acids and is implicated in the corrosion of aircraft integral fuel tanks.

2.20.2.4 Algae

Algae are pigmented organisms (most commonly green, brown, or red) possessing the same photosynthetic mechanism as higher plants but having no true tissues, despite the enormous size of some marine algae (seaweeds). Their main interest to the engineer is the development of fouling masses in environments where light can penetrate, giving rise to microenvironments beneath which harmful bacteria may proliferate.

2.20.2.5 Biofilms

Though bacteria and other microorganisms can and do exist in the free floating state, they all have a propensity to attach to surfaces, where they proliferate, giving rise to slime and deposit formation. Microorganisms that float or swim in water are called *planktonic* bacteria. Microorganisms that attach to a substrate (e.g., metal) surface are called *sessile* bacteria. Biofilms comprise accumulations of living and dead microorganisms and their secretions together with entrained organic and inorganic debris. Microbial corrosion generally occurs beneath such biofilms. In addition to corrosion, biological fouling in industrial water systems can, for example, result in decreased efficiency of heat exchangers and plugging of tubes; slimes in cooling towers can cause rot of timber slats. Biofilm that detaches from surfaces can plug pumps and pipe work in cooling systems.

Most of the microorganisms present in industrial systems are present in biofilms and the most common types of biofilms are largely bacterial in nature. As noted earlier, various types of bacteria can grow with or without oxygen. Bacteria that need oxygen to grow are called *aerobes* or aerobic bacteria. Bacteria that need little or no oxygen to grow are called *anaerobes* or anaerobic bacteria. All these broad types of bacteria have a role in a typical biofilm. **Figure 1** shows the steps in biofilm.

Biofilm formation begins when (1) small amounts of organic material attach to the metal surface; (2) microbes then attach onto the organic layer; (3) a thick biofilm develops; (4) the biofilm becomes thick enough to exclude oxygen from the metal surface and forms an anaerobic zone near the surface; (5) aerobic and anaerobic bacterial species develop where conditions are most favorable for them; the biofilm thickens until parts of it are torn away by flowing water. The process begins again when the detached biofilm adsorbs onto another part of the

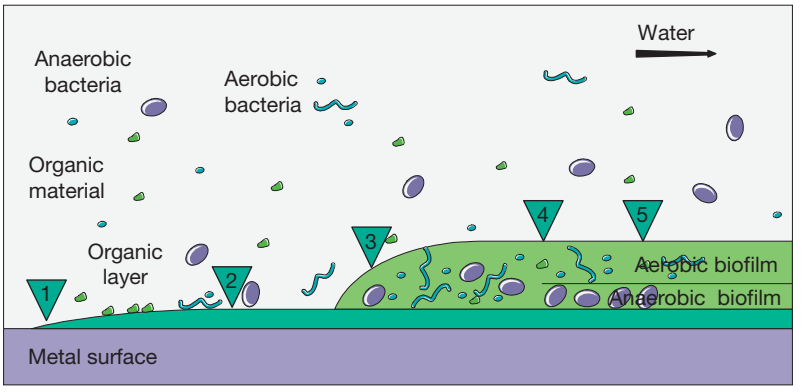


Figure 1 Steps in biofilm formation.

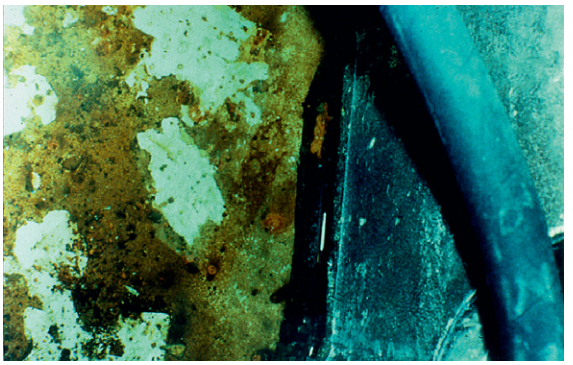


Figure 2 Dense biofilm on a metallic surface.

metal surface. The time required for this process is very much dependent on temperature and nutrient availability, but may be as short as 10–14 days under favorable conditions. The biofilm can thus provide very favorable conditions for both anaerobic and aerobic bacteria, growing in consort. Localized corrosion can develop on the metal surface and can cause corrosion pitting. An example of dense biofilm formation in an industrial water system is shown in Figure 2.

Although it is true that corrosion-inducing microorganisms do not exist in isolation, but rather they form part of the complex microbiological consortia of biofilms, there has nevertheless been a tendency by recent reviewers to approach the subject of MIC from the point of view of the microbiologist rather than the engineer and to emphazise the complexity and diversity of such communities. It is not really necessary to either identify or to understand in detail these multifarious microbial populations in order to comprehend the presence and activity of specific problem microbes.

2.20.3 Types of MIC

As emphasized in Section 2.20.1, the immediate cause of corrosion is electrochemical in all cases we are about to consider. The ways in which metabolic activities of certain microorganisms are believed to bring about or contribute to electrochemical corrosion may be broadly categorized as follows:

(a) Direct chemical action of metabolic products such as organic acids, sulfuric acid, or reduced sulfur compounds, especially sulfides or chelating agents.
(b) Generation of an electrochemical cell by deposition of cathodic and conducting metal sulfides, often in conjunction with mechanism (a).
(c) Cathodic depolarization by removal of hydrogen associated with anaerobic growth, frequently suggested to occur in conjunction with mechanisms (a) or (b) or both. This proposed mechanism is contentious and is now less widely believed than it formerly was.
(d) Localized electrochemical effects because of changes in oxygen concentration, pH, etc., which establish local differential cells often in conjunction with mechanisms (a) and (b). The physical presence of the biomass in itself contributes to this mechanism.
(e) Removal of inhibitory substances, most notably nitrite corrosion inhibitors but also possibly amines.

(f) Direct degradation of protective coatings by utilization of one or more of their constituents as a source of carbon, thereby reducing the efficiency of the protective system.

The most important of these mechanisms will now be considered in turn, with brief reference to illustrative case histories and with a particular emphasis on the monitoring of trends and the identification of remedial measures.

2.20.4 Corrosion by the Sulfate-Reducing Bacteria

2.20.4.1 Situations Effected

The most important group of microorganisms associated with corrosion belongs to the rather diverse group of bacteria known as the sulfate-reducing bacteria (SRB); the vast majority of MIC failures are related to their activities.

SRB live in oxygen-free environments (i.e., SRB are anaerobic bacteria), where they obtain their required carbon from organic nutrients and their energy from the reduction of sulfate ions to sulfide. SRB will only flourish and cause damage if they can obtain sufficient sulfate, which is abundant in many freshwaters, seawater, and soils. Sulfide appears as dissolved or gaseous hydrogen sulfide (H_2S), HS^- ions, S^{2-} ions or metal sulfides, or a combination of these according to conditions. Sulfides are highly corrosive to many metals. In the case of iron and mild steel, the characteristic black iron sulfide (FeS) corrosion products liberate H_2S on acid treatment, distinguishing it from black iron oxide (magnetite). The corrosion products are often loose and when dislodged exhibit pits lined with bright metal corresponding to areas of anodic dissolution, but are sometimes also hard and adherent.

The most commonly encountered SRB type belongs to the genus *Desulfovibrio* (**Figure 3**). In natural conditions, SRB grow in association with other microorganisms and use a range of carboxylic acids and fatty acids, which are the common by-products of other microorganisms. They couple the oxidation of these carbon sources to sulfate reduction. For example:

$$CH_3COO^-(\text{acetate}) + SO_4^{2-} \rightarrow S^{2-} + CO_2 + H_2O + HCO_3^- \quad [1]$$

Bacteria of this type can also use molecular hydrogen instead of organic carbon, to sulfate reduction:

$$4H_2 + SO_4^{2-} \rightarrow S^{2-} + 4H_2O \quad [2]$$

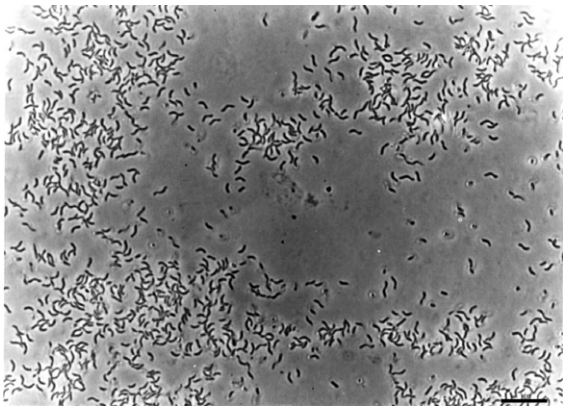


Figure 3 *Desulfovibrio*, a common genus of SRB as viewed under a phase contrast microscope. Bar = 10 μm in length.

Conditions at the base of even thin biofilms can be ideal for growth of SRB, with high organic nutrient status, no oxygen, low oxidation–reduction (redox) potential, and protection from biocidal agents. SRB activity can thereby produce active sulfide-mediated corrosion even in systems where the bulk liquid phase has a low nutrient status, a high oxygen concentration, and will not support growth of anaerobic bacteria.

Most types of SRB will only grow in the pH range 5–10 and the temperature range 5–45 °C, up to a pressure limit of ~500 atm,[13] although some less common groups of SRB have the ability to grow at elevated temperatures (thermophiles) up to at least 70 °C. The thermophiles are most commonly encountered in produced water from hot petroleum reservoirs; however, Postgate,[14] in his definitive monograph on the microbiology of the SRB, cites two cases of corrosion by thermophilic SRB activity, one case being in a tank of hot molasses at a sugar refinery, the second being a transformer tank buried in London clay, where the temperature was 60–80 °C for long periods.

The muddy bottoms of any relatively stagnant bodies of water with a high biological oxygen demand often support massive growth of SRB, as may waterlogged soils. Any metallic installations buried or immersed in such environments can be expected to suffer badly from SRB-mediated external MIC. The most serious economic problem is to pipelines, though sheet piles, hulls of ships, piers, etc., are also frequently attacked. In some instances, cast iron pipes of 0.25 in. thick have become perforated within a year under such conditions, while perforation rates 1–2 mm $year^{-1}$ at ambient temperature are not uncommon. Under anaerobic/aerobic cycling

conditions, where SRB-mediated sulfide corrosion may be exacerbated by the production of elemental sulfur, and/or at elevated temperatures, corrosion rates (CRs) of several millimeters per year may be encountered. This type of corrosion typically takes the form of shallow pitting attack, which may be either widespread or highly localized. A classic example of such morphology is shown in **Figure 4**; this is a case of unprotected sheet steel piling attack in highly SRB-infested estuarine mud. The metal surface was grit blasted to remove the copious black iron sulfide corrosion product.

Oil field water handling systems, especially wet crude oil production pipelines and water injection pipe work, are beset by internal corrosion problems due to SRB-mediated internal MIC, most commonly at the bottom of line locations of water and particulate drop out. Low-velocity lines and storage tanks, aircraft jet fuel tanks and lines, hydrotest waters, cooling waters, and facilities handling wastewaters suffer similarly.

The petroleum production industry has been particularly plagued by the activities of SRB, because it handles large volumes of deaerated water. These waters can become very 'sour' with H_2S if infection with SRB occurs. The phenomenon of unexpected increase in H_2S concentrations in produced fluids from petroleum reservoirs (reservoir souring) has been observed over a period of many years in different areas of the world.[15–18] It is a very serious problem for that industry and considerable money has been invested in biocide-treatment programs for water injection, with the principal aim of killing or controlling this group of microorganisms. For a more detailed modern appraisal of the problems of microbiological control in oilfield water systems, the reader is referred to the book by Davies and Scott.[19]

A large-scale study of external MIC of buried steel structures at 59 sites in the United Kingdom[20,21] lead to the suggestion that aggressive sites were characterized by soil resistivity of less than 2000 Ω cm or a mean redox potential more negative than +400 mV (on the hydrogen scale) at pH 7. Borderline cases were classified according to the water content – those containing more than 20% water being deemed aggressive soils. Similarly, all soils with a mean soluble iron content of over 120 $\mu g\, g^{-1}$ were found to be aggressive. A particularly corrosive situation occurs when H_2S diffuses upwards into the aerated zone, where it becomes oxidized by air to elemental sulfur.[22] An example of external MIC of a cast iron water main is shown in **Figure 5**. This was one of the most rapid incidences of its kind that has been observed and was accompanied by deposition of elemental sulfur. The site was made-up ground on a new housing estate in the south of England, where the soil type should have been 'nonaggressive' by the foregoing criteria. In fact, the lower horizons were waterlogged with sulfate-rich spring water and the SRB were highly active. H_2S diffused upwards into the aerated zone, where it became partly oxidized by the air to elemental sulfur.

Numerous examples of corrosion of stainless steel and higher active/passive alloys involving sulfide production by SRB have been published.[23,24] The mechanism is discussed in **Section 2.20.4.3**.

Copper alloys are highly susceptible to rapid and severe biogenic sulfide corrosion if conditions allow the growth of SRB (see **Section 2.20.4.4**).

Figure 4 Typical morphology of SRB-mediated MIC of carbon steel, an example from sheet steel piling.

Figure 5 Massive 'black water' corrosion of pipe work under wet lay up because of SRB activity.

2.20.4.2 Mechanism of SRB Corrosion of Iron and Low-Alloy Steel

The mechanism of SRB-mediated sulfide corrosion of ferrous metals[25] involves both H_2S and FeS. CRs tend to be slow at first and then accelerate with time.[26] Initially, semiprotective iron sulfide layers form, which typically reduce the generalized CR compared with sulfide-free environments. Later, sometimes after an exposure period of months, the iron sulfide layer spalls locally and never reforms; in the case of petroleum production pipelines, this localized spalling may be initiated by sand erosion. The remaining iron sulfide film is both electrically conducting and cathodic to carbon steel, thus galvanic attack initiates at the site of the iron sulfide film breakdown, driven by the huge surface area of the FeS cathode. Costello[20] has argued convincingly that the principal cathodic reactions are as follows rather than the mechanism of hydrogen ion reduction:

$$H_2S + e^- \rightarrow HS^- + \tfrac{1}{2}H_2 \quad [3]$$

$$HS^- + e^- \rightarrow S^{2-} + \tfrac{1}{2}H_2 \quad [4]$$

This postulate is supported by the relative pH independence of SRB-mediated sulfide corrosion, except where the pH of the environment is too high or too low to permit SRB growth.

This mechanism is thus a combination of 'Type (a)' MIC, direct chemical action of a metabolic product, in this case H_2S, and 'Type (b)' MIC (see **Section 2.20.3**), the generation of an electrochemical cell, in this case an iron or steel anode with an iron sulfide cathode.

Furthermore, the hydrogen evolved from the cathodic part of the reaction can be used as a source of reducing power by many strains of SRB and coupled to sulfate reduction to generate even more sulfide, a process that is catalyzed by the hydrogenase enzyme. The overall process thus becomes a type of feedback loop: the more the hydrogen available, the more the sulfide generation; the more the sulfide, the more the corrosion; the more the corrosion, the more the hydrogen generated; and so forth. A classic example of this is the so-called 'black water corrosion' of carbon steel pipe work under long-term wet lay up (most especially, following hydrostatic testing), in which more iron sulfide may be generated compared with the stoichiometric amount of organic carbon that is available to 'fuel' sulfate reduction (**Figure 6**). A schematic representation of the mechanism of SRB-mediated MIC of iron or low-alloy steels is presented in **Figure 7**.

Figure 6 SRB-mediated external MIC of a cast iron water pipe accentuated by elemental sulfur.

Some workers have suggested that hydrogenase can even directly depolarize steel surfaces by removal of bound atomic hydrogen. This suggested mechanism, the 'direct cathodic depolarization theory' ('Type (c)' MIC – see **Section 2.20.3**), was very prominent in the early literature on SRB-related corrosion and is still cited by some workers, though it has lost widespread acceptance in the light of work showing that hydrogenase extracts are only active on combined, molecular hydrogen (H_2).

Figure 8 shows a longitudinal section of oilfield water gathering line, which suffered a through wall internal perforation because of SRB-mediated MIC. Note the thick layer of black, iron sulfide corrosion product, which has broken down locally to allow massively concentrated localized corrosion at the resulting anodic site surrounded by an overwhelmingly large volume of cathodic and conductive FeS. The surface oxidation of the thick FeS layer to give a layer of brown rust (with included elemental sulfur) is due to air exposure of the specimen and is typical.

2.20.4.3 SRB-Mediated Corrosion of Stainless Steel Alloys

The phenomenon of severe SRB-mediated MIC of stainless steels and higher active/passive alloys is best termed as microbially assisted, chloride-induced pitting attack. It should not be confused with the related but separate mechanism of pitting and crevice corrosion of these materials promoted by aerobic biofilms in seawater, which is discussed in **Section 2.20.6**. SRB-mediated effects are generally more rapid and severe.

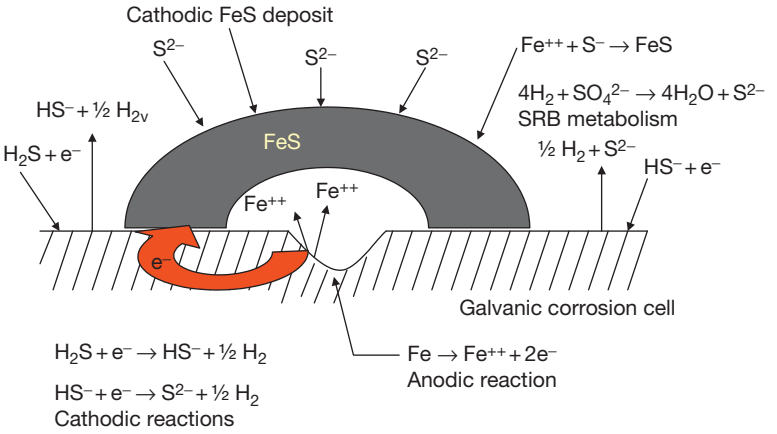


Figure 7 SRB-mediated galvanic corrosion cell.

Figure 8 SRB-mediated internal MIC of an oilfield water gathering line.

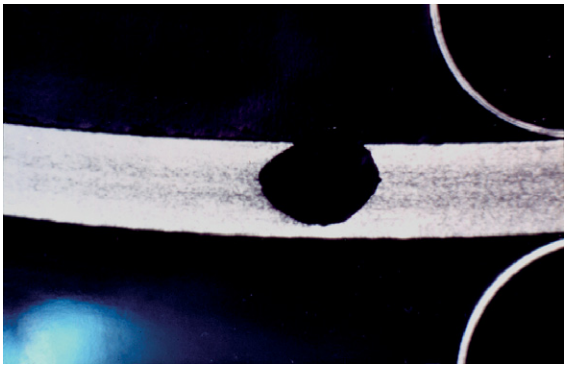


Figure 9 Microbially assisted chloride-induced internal pitting attack on AISI 304 stainless steel pipe work showing classic flask-shaped pit.

Active/passive alloys are susceptible to localized pitting by chloride ions. Reduced sulfur compounds produced by SRB, primarily sulfide but also traces of other substances such as thiosulfate, are believed to bind to the metal surface and catalyze the initiation of the stable pitting process at much lower chloride concentrations and at lower temperatures than would otherwise be the case. Once a pit is established, the anodic reaction generates acidic conditions, accelerating the CR and causing flask-shaped cavities to form. This is totally different from the shallow pitting attack of low-alloy materials exemplified in **Figures 4, 5** and **8**. The pits may be concealed beneath ferric hydroxide and the process is often associated with copious amounts of brown colored biological slime, containing iron-oxidizing bacteria known as *Gallionella*. Bacterial metabolism is only involved in the pit initiation stage. Once the pit has formed and acidic chloride conditions develop within, the pit propagation process is the same as with any other case of chloride-induced pitting. Claims that MIC of stainless steels is characterized by a particular pit morphology that is different from conventional chloride-induced pitting attack should therefore be viewed with scepticism.

An example of this type of corrosion occurred when AISI 304 grade stainless steel pipe work was installed at a newly built chemical plant in the tropics (**Figure 9**). A hydrostatic pressure test was made using local water, which was subsequently left *in situ* in a static condition for several weeks. The 5 mm thick wall pipe work began to leak in hundreds of places at the 6 o'clock position. Initially, futile attempts were made to clamp the pits but much of

the pipe work eventually had to be replaced. The pipes were found to contain slime masses of iron-oxidizing bacteria of the genus *Gallionella* (**Figure 10**). The aerobic iron bacteria had depleted all the dissolved oxygen, allowing SRB to become active, giving rise to sulfide formation. The water was warm, stagnant, but relatively low with respect to chloride ion concentration (20–30 mg l^{-1}); the presence of the extensive slime had allowed an inorganic chemical mechanism to initiate and run rife at chloride levels that would conventionally be regarded as relatively safe. This case history stands as a good example of the interaction of 'microbiological' and 'conventional' corrosion processes.

2.20.4.4 MIC of Copper and Copper Alloys

There is a widespread belief that copper and its alloys are toxic to microorganisms and are therefore not susceptible to MIC. This notion is false; copper alloys are much less toxic to bacteria than they are to macrofouling organisms such as seaweed or shellfish; they suffer severe attack by sulfide if conditions allow the growth of SRB. 'Under-deposit' or 'under-fouling' pitting corrosion of copper–nickel heat exchanger tubes due to biogenic sulfide, produced by SRB, has frequently been observed.

Figure 11 shows a gunmetal impeller located in a polluted harbor, which failed in a matter of weeks after the pump was switched off because of growth of SRB in the stagnant pump. The component had previously given 2 years of trouble-free operation in aerated seawater.

Numerous cases of mysterious 'under-deposit' or 'under-fouling' pitting corrosion of copper–nickel heat exchanger tubes have been observed, especially in marine environments. Many of these failures are demonstrably due to biogenic sulfide produced by SRB. Copper–nickel alloys are much more susceptible to the initiation of MIC by SRB-generated sufides during the first few weeks of exposure to the operating environment, at which time the protective oxide film is incompletely formed. MIC of aged copper–nickel alloys, with a well-formed oxide film, is also known, but is much less common than for newly exposed material and seems to require considerably more severe conditions, that is exposure of higher concentration of sulfide for a longer duration. The promotion of protective oxide films and avoidance of exposure to stagnant contaminated water during early exposure is critical in order to avoid such problems. If wet lay up with contaminated water is unavoidable, the use of an organic biocide treatment is highly recommended.

Although the most severe corrosive effects on copper alloys, including cupronickel, have typically been associated with microbially generated sulfides, attention has also been drawn to other activities of surface-associated microbial populations, which in some cases can lead to localized corrosion. These are the production of acidic metabolites, the binding of copper by microbial cells and their extracellular slimes, and the alteration of the nature and porosity of the oxide layer because of biofilm penetration. In particular, it should be recognized that the level of acidity at the biofilm metal interface may be much higher than that of the bulk fluid. It has been reported that a biofilm-forming marine bacterial species increased the corrosion potential of copper by 145 mV and the CR fivefold in a seawater medium and that electrochemical impedance spectroscopy data has shown CRs under biofilms in water of pH

Figure 10 *Gallionella* slime mass from corroded AISI 304 stainless steel pipe work.

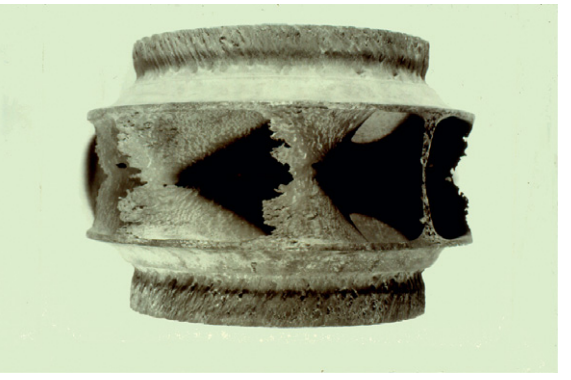


Figure 11 SRB-mediated sulfide corrosion of a gunmetal pump impellor.

8 to be equivalent to those observed under abiotic conditions at pH 6.8.

It is worth noting that when copper alloys are electrically linked into a cathodic protection system they tend to lose their antifouling properties and become highly susceptible to attack by biogenic sulfide produced by SRB colonies underneath the consequent fouling 'mat.'

From the mid 1980s onward, a distinctive type of localized corrosion has been seen in copper pipes carrying potable water, which Campbell and coworkers[27] termed 'Type 1½' pitting. Type 1 and Type 2 are relatively well-known forms of pitting corrosion in copper water pipes; Type 1 occurs in cold water and is associated with the presence of residual carbon films, whereas Type 2 occurs only at temperatures above 60 °C, in copper pipes carrying soft water with a sulfate content exceeding the bicarbonate content. The newly identified type of pitting was found to be associated with relatively large mounds of corrosion product and occurred at relatively low temperatures, as in the case of Type 1, but the products consisted principally of basic copper sulfates and black cupric oxide, as in Type 2. The phenomenon was initially reported from Britain, Germany, Sweden, and the Middle East, and in all cases, a gelatinous film of microbial origin and consisting largely of polysaccharides was present.

This occurrence has since been reported from numerous other locations[28] and is commonly associated with blue water, in which the copper concentration in solution can reach tens of milligrams per liter. This should not be confused with 'Type 3' pitting, associated with soft waters and elevated pH (above 8.0), which is also known as blue water corrosion.

The occurrence of this type of copper corrosion has been convincingly demonstrated to be associated with the presence of biofilms comprising extracellular polymeric products of various types of aerobic heterotrophic bacteria, particularly members of the genera *Pseudomonas*, S*phingomonas*, or *Acidovorax*.[28,29] When the corroded copper surfaces are treated with 25% nitric acid in a similar manner to a carbon film test, a clear gelatinous membrane of biofilm floats on the surface of the acid. This corrosion mechanism is also critically dependent on water composition and operating conditions. Soft waters with less than 40 mg l^{-1} hardness expressed as $CaCO_3$, low bicarbonate alkalinity (< 20 mg l^{-1}), high organic content, mean temperatures of between 20 and 45 °C, frequent stagnation, and lack of chlorination or chlorine depletion are factors that favor this mechanism. One distinctive characteristic is that this type of copper tube corrosion does not have the likelihood of failure at maximum between 1 and 3 years of installation, unlike other common mechanisms.

The mechanism of copper tube failure associated with non-SRB biofilms is still uncertain. Certainly, the uronic acid carboxyl residues of the extracellular biopolymers are rather acidic (typical pK_a values are in the range 4.2–5.6)[28] and their uronic acid or pyruvyl substitutions are known to possess high affinity binding sites for copper.[30] Furthermore, the heterogeneous nature of aerobic biofilm distribution is certain to create oxygen concentration cells. Siedlarek et al.[31] called attention to the property that oxygen and carbon dioxide, but not chloride ions, can readily pass through the biofilm layer. This allows the formation of copper oxides or hydroxides beneath the biofilm coating but not the deposition of copper (I) chloride. Therefore, it is argued, amorphous copper (I) oxide, which inhibits anodic metal dissolution, cannot readily be formed from hydrolysis of copper (I) chloride in the presence of chloride impermeable biofilm. It is likely that all the foregoing mechanisms play a role to a greater or lesser extent and that they are interactive with one another.

Optimization of the anion ratio in the water with regard to chloride, sulfate, and bicarbonate, minimizing stagnation, ensuring adequate chlorination, and avoiding warming of cold water pipes from proximity to hot water pipe runs are considered to be appropriate measures to minimize the risk of biofilm-related corrosion of copper pipework in potable water systems.[28]

2.20.5 Microbially Influenced Acid Corrosion

2.20.5.1 Sulfuric Acid Corrosion by Sulfur-Oxidizing Bacteria

Massive and rapid general corrosion of metal, concrete, or limestone under aerobic conditions may be caused by the action of sulfuric acid formed by the oxidation of sulfur or sulfide by members of the bacterial genus *Thiobacillus*. The sulfur-oxidizing bacteria belong to an unusual group of microorganisms called chemolithotrophs. These microorganisms obtain energy not by oxidation of organic compounds but by oxidation of inorganic sulfur compounds (including sulfides) to sulfuric acid; they build up their cell material by fixation of carbon dioxide, and hence they are classed as autotrophs. Purkiss[32] stated that the following interlinked reactions are

performed by mixed cultures of *Thiobacillus* acting on elemental sulfur or sulfides:

$$2H_2S + 2O_2 \rightarrow H_2S_2O_3 + H_2O \quad [5]$$

$$5S_2O_3^{2-} + 4O_2 + 4H_2O \rightarrow 5SO_4^{2-} + H_2SO_4 + 4S \quad [6]$$

$$4S + 3O_2 + 2H_2O \rightarrow 2H_2SO_4 \quad [7]$$

Certain of the *Thiobacillus* species will also leach sulfide mineral ores according to the following reaction[33]:

$$4FeS_2 + 15O_2 + 2H_2O \rightarrow 2Fe_2(SO_4)_3 + 2H_2SO_4 \quad [8]$$

The important points to note are that these bacteria require oxygen (the opposite of the situation with SRB) and a source of reduced sulfur. The end product is sulfuric acid; one species (*Thiobacillus thioxidans*) has been claimed to remain active at a pH value as low as 0.7, corresponding to more than 5% sulfuric acid.[34] In the author's experience, such an observation would be very unusual, a pH limit in the order of 1.5–2.0 being more common.

Thiobacilli are fairly commonly encountered in nature; it is believed that they obtain their energy from simple sulfur-containing breakdown products of sulfide minerals or from sulfide produced in an adjacent locality (or in the same locality under anaerobic conditions) by SRB. Normally, they are not found in large numbers, but if a high concentration of a suitable nutrient substrate occurs locally, their activity may be great and result in a zone of intense acidity. Such conditions can most commonly arise in sewage, in which sulfide concentration may be high owing to the action of putrefactive bacteria, unless aeration is efficient. Suitable conditions can also arise locally in paper mill effluents, damp or wet sulfur cargos, and 'made-up' ground into which industrial waste material sometimes finds its way. For this reason, sulfuric acid corrosion of underground pipelines is sporadic and unpredictable. Though *Thiobacillus*-mediated corrosion is occasionally encountered in other industrial situations, such as cooling water systems, such instances are surprisingly uncommon.

At first sight, it might seem inconsistent that various microorganisms can obtain energy both from the oxidation of sulfide to sulfate and from the oxidation of sulfate to sulfide. However, the answer is that one chemical reaction is not simply the reverse of the other: sulfur oxidizers consume oxygen, whereas SRB consume organic carbon. The two reactions can be summarized as follows:

$$\text{Thiobacillus sulfide} + \text{Oxygen} \rightarrow \text{Sulfate} + H^+\text{ions} \quad [9]$$

$$\text{SRB sulfate} + \text{Organic carbon} \rightarrow \text{Sulfide} + \text{Carbon dioxide} + \text{Water} \quad [10]$$

Combining reactions [9] and [10] one derives the following overall reaction [11], a variant of the familiar air oxidation of organic carbon to CO_2.

$$\text{Overall oxygen} + \text{Organic carbon} \rightarrow \text{Carbon dioxide} + \text{Water} + H^+\text{ions} \quad [11]$$

Figure 12 shows a severe example of internal corrosion of a mortar-lined ductile iron sewage water pipe, which failed by sulfuric acid corrosion due to the activity of *Thiobacillus* sulfur-oxidizing bacteria. As it is typical, the attack has occurred in the air gap at the top of the pipe where there is no actual sewage contact. **Figure 13** explains how this occurs. The rate of microbial activity in the sewage effluent is sufficiently high that the rate of consumption of oxygen exceeds its rate of replenishment by diffusion from the air space. SRB then become active and produce H_2S – some of which diffuses into the air space at the top of the line where it redissolves in a surface film of condensed moisture. The sulfur oxidizers convert the H_2S containing solution to sulfuric acid, which then attacks the metal. Corrosion takes the form of generalized thinning around the '12 o'clock' position.

2.20.5.2 Acid Attack on Concrete

Attack on concrete by sulfuric acid produced by sulfur-oxidizing bacteria of the genus *Thiobacillus* can be a severe problem. The most widely studied problem is the collapse of concrete sewers. As in the case of attack

Figure 12 Corrosion of a mortar-lined ductile iron wastewater pipe because of activity of *Thiobacillus*.

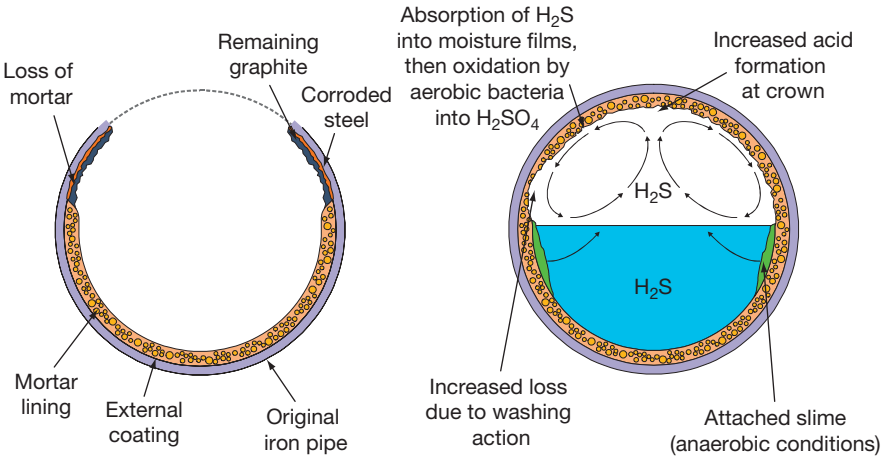


Figure 13 MIC processes leading to collapse of iron or steel sewer pipes.

on metallic sewer pipes, this occurs due to growth of SRB in the anaerobic sewage, giving rise to H_2S. The gaseous H_2S migrates to the air space at the top of the pipes where it is converted to sulfuric acid by the sulfur oxidizers, leading to low pH, rapid dissolution of the concrete, and collapse of the sewer from the top. Under favorable conditions for the bacteria, average concrete dissolution can be greater than 12.5 mm in 50 years, with higher local rates at the air/sewage interface.[36]

When steel reinforcing bars are exposed, metal corrosion occurs, leading to spalling and cracking of the concrete.[35] Corrosion of reinforced concrete on offshore installations is a more recent problem. In seawater displaced crude oil storage tanks, crude oil stimulates SRB, with sulfur oxidizers again producing sulfuric acid in aerobic zones.[37]

2.20.5.3 Acid Corrosion by Fungi

A completely different type of microorganism to the various groups of bacteria is the fungi. The cells are much larger and grow in dense mats of material. Filamentous fungi are aerobic in nature and are involved in a wide range of biodeterioration problems; timber, paper, fuel oils, cloth, etc., can all be attacked by fungi. Fungal activity generates organic acids, leading to a low pH in the surrounding water and under the microbial mat. In addition to direct acid corrosion, oxygen concentration cells are set up between zones of metal covered with oxygen-depleted fungal mats and those areas where no fungi are present. This electrochemical cell drives the metal dissolution reaction beneath the fungal mat. A further complication is that SRB can sometimes proliferate in the oxygen-depleted niches beneath the fungal mats, giving rise to a combination of acid corrosion enhanced by sulfide corrosion and differential aeration effects.

By far the most troublesome fungus in engineering systems is a type regularly called *Cladosporium resinae*, though now formally reclassified as *Hormoconis resinae*. *H. resinae* has the ability to thrive in the presence of kerosene and other hydrocarbons, which it uses as a carbon source for oxidation. It also produces spores that can survive extremes of temperature, only to germinate when more moderate conditions prevail. *H. resinae* is a continuing problem in fuel storage tanks and in aluminum integral fuel tanks of aircraft.

Brown slimy mats of *H. resinae* may cover large areas of aluminum alloy, causing pitting, exfoliation, and intergranular attack because of the organic acids produced by the microbes and differential aeration cells. This type of problem is largely confined to tropical, humid locations and is particularly severe in short-haul aircraft and those that experience a lot of idle time where condensation builds up.

The above problem was unknown until the 1950s, when it became prevalent after the introduction of gas-turbine-engined aircraft that use kerosene rather than gasoline as a fuel and the adoption of integral or 'wet wing' fuel tanks to replace the earlier rubberized fabric 'bay tanks.'[38] The problem of fungal growth in the fuel tanks of jet aircraft has generally diminished as the design of fuel tanks has improved to facilitate better drainage of condensed water and as biocides such as organoboranes have gained acceptance as fuel additives. However, this type of corrosion has reappeared in recent years, due to increasingly stringent environmental restrictions of biocides in fuel, combined with certain aircraft designs in which the lessons that were learned during the 1950s and 1960s

about the importance of efficient water drainage in integral fuel tanks have seemingly been forgotten.

2.20.6 MIC of Higher Alloys by General Aerobic Biofilms in Natural Seawaters

While the most rapid and severe MIC attack of stainless steels and higher active/passive alloys is commonly associated with SRB (see **Section 2.20.4.3**), several investigators[39–42] have reported positive shifts in the open circuit potentials for metals exposed to natural seawater. This positive shift has been attributed to the development of general mixed bacterial slimes on the exposed metal surface. A wide range of bacteria (e.g., *Pseudomonads* and *Flavobacteria*) can secrete large amounts of slime-like organic material under both aerobic and anaerobic conditions. Although the exact mechanism by which the bacterial film enhances the oxygen reduction kinetics is not fully understood, it has been postulated that the oxygen reduction is catalyzed by enzyme action, leading to either an increase in the exchange current density or a decrease in the cathodic Tafel constant.

In the absence of corrosion, the free corrosion potential of high molybdenum austenitic and duplex stainless steels in artificial seawater was found to rise from ~ -100 to 0 mV and reach a stable value of $\sim$130 mV (vs saturated calomel electrode) after $\sim$10 days.[42] In microbially active seawater, however, there appears to be fairly common agreement that the potential continues to rise and reaches stable values of 300 ± 50 mV after 20 days. There are indications, however, that the final potential is relatively independent of composition but is dependent on factors such as temperature and flow rate.

The length of time to reach the steady potentials has been shown to be closely related to the time necessary for microorganisms to settle and grow on the metal surface. Close examination of some of the potential versus time curves reported in the literature reveals that there is an initial rise in potential in the first 2–3 days, followed by a more rapid rise until the potential stabilizes after $\sim$20 days. This is consistent with the time found by Gunderson *et al.* for the bacterial count to reach a level of $\sim 10^6$ cells cm^{-2}.

The effect of the potential rise because of bacterial slime is to increase the susceptibility of these alloys to crevice corrosion, but this is accompanied by an increase in the cathodic reaction rate, and therefore, once crevice initiation occurs, there can be a marked increase in the crevice propagation rate.

Temperature has a number of important effects. First, the rate at which a critical crevice solution (CCS) is achieved is believed to be a function of the passive current density within the crevice, which may be increased by 2–3 orders of magnitude because of the positive potential shift caused by bacterial slimes compared with sterile seawater. It may also increase by 1–2 orders of magnitude because of an increase in temperature from 10 to 35 °C. Second, it is reported that there is a maximum temperature for biofilm growth of $\sim$30–32 °C, although this may be higher in warm climates such as the Mediterranean because of greater natural tolerance.

This phenomenon explains why the crevice CRs for 254 SMO and SAF 2507 have been found to be 2–3 decades higher at 25 °C than at 40 °C. At lower temperature, the biofilm activity increases the potential, and therefore increases the risk of crevice initiation, but also increases the cathodic reaction efficiency and consequently increases the rate of propagation. At 40 °C, the higher temperature also increases the risk of crevice initiation but the loss of the biofilm activity results in lower cathodic efficiencies and lower propagation rates. The anodic current densities were found to be 2–3 decades higher at the lower temperature. Mathematical modeling was used to show that, in the absence of biofilms, a much larger cathode-to-anode area ratio was required to sustain high CRs, and that improved mass transport of oxygen because of higher flow rates had a greater effect on the propagation rate than in the presence of the biofilm.

2.20.7 Other Microorganisms Associated with MIC

As already discussed, biological slimes are commonly found in the water phases of industrial process plant. Conditions at the base of even thin slimes can be anaerobic and therefore ideal for the growth of SRB, with high organic nutrient status, oxygen depletion, low redox potential, and protection from biocidal agents. In such situations, SRB-mediated sulfide corrosion can be exacerbated by fermentative acid production. An illustration of how easy it is for anaerobic conditions to develop even beneath thin microbiological slime layers is provided by the well-known process of tooth decay; the secretion of organic acids, which leads to such dental decay, is a completely anaerobic process – in the presence of oxygen, sugars

are oxidized to carbon dioxide instead of being fermented to produce acids. Given that even a thin layer of tooth plaque can give rise to anaerobic conditions, it can be appreciated how colonies of the strictly anaerobic SRB can similarly thrive beneath slimes in the nominally fully aerated environments of industrial water systems.

A particularly common type of biological slime is that produced by iron bacteria, which are found in freshwaters with a high concentration of reduced iron compounds. The bacteria require aerobic conditions and form colonies that are impregnated with ferric oxides. *Sphaerotilus* forms dense, cotton wool like growths in organically polluted rivers, the bacteria forming filaments, which are surrounded by a sheath of slime and iron oxide. The base of the colony may be highly anaerobic and SRB can develop under certain conditions. Another genus of iron bacteria, *Gallionella*, is able to use iron sulfides as a source of reduced iron, and is therefore commonly found in association with SRB. The consortium, in the form of tubercle, is one of the classic examples of growth of organisms inside ferrous water pipes.

Some components of slimes also use nitrogen-containing compounds as an energy source. They are involved in the cycling of nitrogen in the environment. Ammonia and amines are produced by microbial decomposition of organic matter under both aerobic and anaerobic conditions (ammonification). These compounds are oxidized to nitrite and nitrate by aerobic bacteria such as *Nitrosomonas* or *Nitrobacter* species. *Nitrobacter* is very efficient at destroying the corrosion-inhibition properties of nitrite-based inhibitors by oxidation of nitrite to nitrate, unless a biocidal agent is included in the inhibitor formulation.[43] Whereas under anaerobic conditions, a wide range of bacteria can reduce nitrite to nitrogen gas, sometimes with the production of ammonia as a by-product. The detrimental effects on brass by the release of ammonia at the surfaces of heat exchanger tubes have also been highlighted.[44]

2.20.8 Testing for MIC Activity

Assessment of a microbial population in industrial fluids is normally carried out by the use of well-established culture techniques involving specific growth media; this test format was developed for oilfield water systems, but is suitable for other industrial situations as well. Enumeration of SRB and total viable bacteria is most commonly by a 'serial extinction dilution technique' in liquid media, using vials containing specially formulated SRB growth medium.[45] The test comprises initially a *serial dilution step*, followed by an *incubation step*. Ideally, the serial dilution part of the test should be carried out within a few hours of obtaining the subject samples (typically done on site); however, this is frequently not possible. The results become less reliable with increasing sample age. At the end of the incubation period, the SRB population density in the original sample is determined to the nearest order of magnitude by the number of vials in each dilution series that turn black because of bacterial sulfide production. According to current standards, the prescribed incubation period for SRB is a minimum 28 days, but a very good indication of the SRB count can usually be obtained after 10–14 days of incubation.

An example of an SRB serial extinction dilution set is shown in **Figure 14**. Similar serial dilution tests, using a different type of growth medium, are used to enumerate total bacteria; the vials that show growth turn cloudy rather than black, often accompanied by a color change of an indicator dye to indicate organic acid production.

For the most accurate assessment, threefold or fivefold replicate serial dilution enumeration is carried out; this allows the mean probable number (MPN) of bacteria to be derived from standard MPN tables. Test kits of this type are commercially available, sometimes in a simplified form. There are also some rapid assay kits based on antibody reactions to SRB enzymes, which do not require an incubation period to obtain a result. All simplified and rapid kits compromise accuracy and reliability to a greater or lesser degree. They should be regarded as no more than crude indicators, unless calibrated against standard serial dilution techniques for a specific environment.

Most of the SRB and other bacteria present in industrial systems are sessile. Planktonic bacterial

Figure 14 Serial dilution test for SRB, blacking (positive SRB count) to 5 orders of magnitude demonstrates heavy infestation.

counts are often misleadingly low, whereas much larger populations may be concealed in sediments, deposits, and biofilms on metal surfaces. Increasingly, there is a trend to monitor SRB, and microorganisms generally, using metal coupons similar to corrosion test coupons. Such coupons are sometimes termed 'bioprobes' or 'biostuds,' and for ease of recovery, are often placed in side-stream devices in flowing pipelines (**Figure 15**).

If surface deposits or coupons are available, they should be sonicated or scraped into vials of culture media, serially diluted on site, then incubated in the manner of water samples. In order of priority, samples from industrial systems should be taken from bottom, metal surfaces, dead ends, and sumps. Sampling should be on a regular basis, weekly, monthly, or quarterly, depending on the extent of concern.

Low sessile population counts of below ~1000 SRB colony forming units (cfu) per gram of deposit or per square centimeter of surface are not levels about which much concern need be expressed. However, periodic testing should be undertaken to check trends. Populations will oscillate, but if they show a general upward drift, then action will be necessary sooner or later.

Planktonic SRB counts provide a less reliable indicator than do sessiles. However, recurrent counts of greater than ~10 cfu ml^{-1}, any general upward trend, or any positive planktonic SRB counts in a nominally biocide-treated system are a cause at least for further investigation of sessiles. Any planktonic SRB counts above ~1000 cfu ml^{-1} is indicative of a heavy proliferation somewhere in the system, probably in deposits on surfaces, especially in stagnant or low flow areas.

Total viable bacteria counts (TVCs) are usually much higher than SRB counts. Planktonic TVCs of up to 100 000 per milliliter are common in many natural waters and industrial systems. Higher population densities, greater than 100 000 per milliliter probably, indicate a significant build-up of slime somewhere. In biocide-treated systems, TVCs should be several orders of magnitude lower than in the untreated water. The incubation period for TVCs is only ~3 days, considerably shorter than for SRB. Therefore, these counts can give a useful early warning of any problem with biocide-treatment efficiency.

SRB counts and TVCs are relatively straightforward, routine analytical techniques. Tests for nitrite-utilizing bacteria or sulfur oxidizers are more complex, specialist techniques. However, problems associated with these organisms are usually clearly evident from the undesirable chemical changes they bring about.

As a supplement to microbiological testing, it is useful to carry out some chemical analyses. An important test is for sulfide in corrosion deposits. In the case of iron sulfide, a simple test is to warm a small sample of the solid with a few milliliters of 50% hydrochloric acid in a glass test tube. If sulfide is present in the deposit, a slightly moistened strip of lead acetate paper in the mouth of the tube will turn brown, indicating H_2S evolution. If a lot of sulfide is present, the characteristic 'rotten egg' odor will also be apparent.

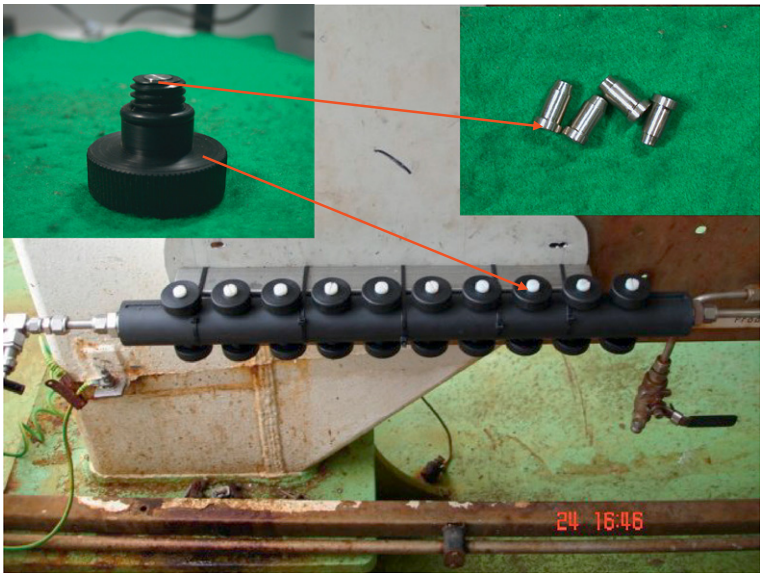


Figure 15 A typical side stream device incorporating 'biostuds' for monitoring of biofilms.

This test needs to be undertaken within a short time of removing samples of deposit from the corrosion site, because wet or damp iron sulfide will eventually oxidize in air to brown rust. The color change of corrosion deposits from black to brown is, in itself, a good indication of sulfide corrosion product. Any such deposits that are taken back to the laboratory for more sophisticated chemical analysis should be preserved in anaerobic bags to prevent the oxidation of FeS, which would otherwise occur.

Tests for sulfate, total organic carbon (TOC), pH, and oxygen concentration are also useful indicators of the potential for SRB growth. If sulfate concentration is low, then there is little hazard of SRB activity. The pH range most suited to SRB growth is 5–10. Outside these limits, bulk phase growth of SRB is very limited. Oxygen concentration will also affect SRB growth in bulk phases but a high value does not mean that SRB are not present in pockets or under debris. A high oxygen concentration may allow other organisms to develop. The interpretation of those tests, together with other operating parameters, is discussed in Section 2.20.9.

2.20.9 MIC Risk Assessment Based on Operating Conditions

If is frequently the case that the engineer is required to make an appraisal of MIC risk with limited information on the corrosion history and little or no historic microbiological test data. The conclusions that may be drawn about the risk of MIC in such cases are largely based on consideration of operating parameters. This is often the situation for the assessment of internal corrosion risk of pipelines, water handling systems, or process plant. The discussion that follows refers specifically to SRB-related problems, which are by far the most common MIC issue.

A group of researchers at Shell Petroleum Co. have developed an approach to risk assessment of carbon steel pipelines, based on the details of water chemistry and operation parameters.[46] Their MIC CR calculation is based on the following equation:

$$\mathrm{CR}(\mathrm{mm\ year}^{-1}) = C \times F^{p} \qquad [12]$$

with

$$F = f_1 \times f_2 \times \cdots \times f_n \qquad [13]$$

where C is a constant ($C = 2\ \mathrm{mm\ year}^{-1}$), the fs are factors for the various influencing parameters and p is a power law index (0.57).

Although such an approach does provide a valuable empirical recognition of the relative contribution of the factors that encourage or inhibit MIC, the calculation of actual CRs on such a basis should be viewed with extreme caution, given our present limited understanding of the effects of the various factors, the most important of which are discussed below.

2.20.9.1 Presence of Water

All bacteria require free water in order to proliferate. In the case of oil pipelines, MIC is to be considered only under separated flow conditions; if all water is entrained in the hydrocarbon, there is no risk of MIC.[46] In the author's experience, it is unwise to rely on dispersions to provide fully oil-wetted surfaces when the bottom solids and water content (BS and W) is in the order of 2% or greater, except under very turbulent hydrodynamic conditions. It is best to take an empirical and conservative view as to whether lines may be wet enough to sustain SRB growth at the present time and during their future operating life.

2.20.9.2 Salinity

According to results presented by Postgate,[14] if the salinity of water, expressed as w/w percentage NaCl, is more than ~7%, there is unlikely to be any SRB growth. That observation broadly agrees with the author's own experience and similar statement is reiterated in the more recent Shell Petroleum Co. work.[39] However, it should be noted that other workers have suggested that some SRB strains may tolerate considerably higher salt concentrations.

2.20.9.3 Temperature

Most types of bacteria grow optimally at temperatures between 10 and 45 °C, which is the highest risk temperature range. However, there are thermophilic strains of SRB and some other microorganisms that grow at higher temperatures up to at least 70 °C and their activity cannot be discounted.

2.20.9.4 pH

Generally, there is unlikely to be SRB proliferation below pH 5 or above pH ~10, though the pH range for SRB that are protected in heavy deposits may be somewhat wider.

2.20.9.5 Anaerobic Conditions

SRB will only proliferate and produce sulfide in the absence of oxygen. Clearly, deaerated environments are at greatest risk; however, there are many cases of problems in slimes and deposits where the bulk water phase is essentially aerobic, for example, in cooling waters. This is especially the case for static or slow flow conditions.

2.20.9.6 Nutrient Status of Water Phase

SRB growth is much less likely if the aqueous sulfate concentration or the organic carbon content is below $\sim 10\,mg\,l^{-1}$ or the ammoniacal nitrogen (not nitrate or nitrite) is below $\sim 1\,mg\,l^{-1}$. An exception, in the author's experience, is the case of plant under long-term lay up with static, high-sulfate water (e.g., seawater), where a corrosive 'black water' condition may develop over a period of months even with very low organic carbon content. This is believed to be due to reduction of sulfate by SRB using cathodic hydrogen formed by corrosion on steel surfaces.

2.20.9.7 Flow Rate

Flow rate influences the nature of biofilm formation and the rate of nutrient delivery. As flow rate increases, biofilms become less bulky and only the more adherent films remain on the metal surface. However, above a certain threshold, the initiation of biofilm formation is severely limited. The transition zone is considered to be between 2 and $3\,m\,s^{-1}$. At the other extreme, stagnation is often associated with the severest MIC incidents.

2.20.9.8 Cleaning Frequency

As discussed earlier, in order to initiate the MIC process, the free-floating ('planktonic') SRB cells need to attach to a metal surface (i.e., become 'sessile') and proliferate to form a biofilm with its own entrained microenvironment. It is that process that is disturbed by regular mechanical cleaning, such as brush 'pigging' of pipelines. The incipient biofilm is disrupted and largely removed and this *is the most* effective means of minimizing SRB growth and obviating MIC. MIC is uncommon in pipelines that are frequently pigged, even without biocide treatment. The more frequent the pig runs, the less time the biofilm has to recover.

2.20.9.9 Use of Antimicrobial Chemical Treatments

Any operation that disrupts the biofilm and thereby limits its proliferation will act to alleviate MIC. The application of periodic or continuous biocide or other antimicrobial chemical treatment is one such factor, which is discussed in more detail in Section 2.20.10.

2.20.10 Control of MIC

2.20.10.1 General Principles

For situations where chemical treatment of the environment by biocides is not possible (e.g., soils, harbors, marine fouling, etc.), the remedies available reflect measures taken to obviate other forms of corrosion and may be summarized as follows.

2.20.10.1.1 Provision of a nonaggressive surround

For buried steel, provision of a backfill of sand or chalk around the material to ensure good drainage and aeration is helpful. Incorporation of a very sparingly soluble biocide into the backfill material has been tried with some success. Another method is to incorporate biocide into protective coatings, the best-known example probably being the use of strontium chromate primers, which are very effective in controlling the kerosene fungus in aircraft wing integral fuel tanks, until such time as the chromate becomes deleted by leaching into the condensed water phase. However, the use of such treatments is both expensive and becoming increasingly less acceptable for environmental reasons.

2.20.10.1.2 Cathodic protection

This may be in the form of sacrificial anodes or impressed voltage, sometimes in conjunction with a coating. There is a long-standing precept, current in the industry, that structures need to be held at potentials more negative than $\sim$–0.9 V (vs Ag/AgCl) in order to protect against anaerobic corrosion by SRB. Horvath and Novak[47] presented the most reasoned argument for this view, which has appeared numerous times in print, though the evidence is not very convincing.

2.20.10.1.3 Protective coatings

The coating must be resistant to biodegradation and to chemical attack by the metabolic products of microbial activity, principally H_2S and sulfuric

acid. Biodegradation resistance of the more long-established wrappings and coatings forms a separate study and was the subject of a site study by British Gas in the 1970s.[48]

2.20.10.1.4 Biocides

For internal protection of pipelines, cooling systems, vessels under hydrotest, etc., biocide treatment is the obvious remedial measure, in combination with cleaning techniques. Their selection use is explained in more detail below.

2.20.10.2 Biocide Treatments

A microbiocide must first be effective in greatly inhibiting microbial activity; second, it must be economical in a treatment program; and third, it must meet the constraints imposed by increasingly stringent environmental legislation. In Europe, the availability of biocides is now restricted to a limited range of substances that are registered under the European Union's 'Biocidal Substances Directive,' which was primarily intended for control of agricultural pesticides but has caught industrial biocides within its net.

Biocides can be broadly divided into two types: oxidizing and nonoxidizing. Commonly accepted oxidizing biocides are 'chlorine' (actually mainly hypochlorous acid plus hypochlorite ion), chlorine dioxide, chlorine donors, bromine, ozone, and hydrogen peroxide. Chlorine is by far the single most prevalent industrial biocide in use today; it is popular because of its efficiency, low cost, and relative environmental acceptability. The disadvantages of chlorine are its high reactivity with many other chemicals, short half-life, necessitating continuous dosing of flowing systems, and difficulties with dosage control – there is a narrow window between the effective dose and an overdose, which can give rise to serious corrosion problems for a wide range of materials. Achieving a continuous chlorine dose with that 'window,' typically 0.2–1.0 mg l^{-1}, is complicated by the 'chlorine demand' of natural waters. During continuous chlorination practice, sufficient chlorine must be generated to overcome the chemical demand and leave a free chlorine residual. This is variously expressed as free residual chlorine (FRC), free available chlorine (FAC), total residual chlorine (TRC), or total available chlorine (TAC). In each case, it is a measure of the residual oxidizing power rather than the true concentration of dissolved chlorine, since a number of different 'free chlorine species' are produced during water chlorination.

Chlorination can be achieved by dosing with chlorine gas or sodium hypochlorite solution and in the case of treating large volumes of saline water by electrolytic chlorine generation. When water is chlorinated, whether by chlorine gas, addition of sodium hypochlorite ('bleach') solution, or by electrolytic generation, it very quickly reacts to establish the equilibrium:

$$Cl_2 + H_2O \Leftrightarrow HOCl + HCl \quad [14]$$

$$HOCl \Leftrightarrow H^+ + OCl^- \quad [15]$$

One of the by-products is H^+ and therefore the concentration of OCl^-, HOCl, and Cl_2 are pH dependent. At a pH of greater than 8, chlorine is present predominantly as OCl^- (hypochlorite ion), whereas at pH 7.5 there is a mixture of hypochlorite and hypochlorous acid (HOCl) and at lower pH values hypochlorous acid predominates. HOCl is ~80 times more effective as a biocide than OCl^-, which means that chlorination is more effective as a biocidal treatment for neutral or slightly acidic 'soft' waters than it is for alkaline 'hard' waters. The exception is seawater, which has an alkaline pH of ~8.2 but can, nevertheless, be effectively chlorinated due to the presence of bromide ions, which react with hypochlorous acid to form hypobromous acid (HOBr); 'chlorination' of seawater is really therefore bromination. HOBr has a different pH-dependent dissociation profile to HOCl, making it much less dissociated into OBr^- ions under mildly alkaline conditions and therefore effective as a biocide in seawater.

Continuous chlorination of industrial water systems might seem prudent for control of MIC as well as slime control on heat exchangers and for public health reasons, particularly the control of the waterborne pathogen *Legionella*, the causative organisms of legionnaire's disease. However, in some high demand systems with process leaks maintaining a continuous 'FRC' level is not economically feasible. The alternative is to maintain a 'FRC' concentration from 0.3 to 0.8 mg l^{-1} for up to 2 h per treatment period in order to kill as much as possible of the sessile biofilm. The rate of recontamination will suggest the frequency of treatment.

The shortcomings of chlorination for certain systems with high pH and/or high chlorine demand have led to an increase in the popularity of chlorine dioxide (ClO_2) as an alternative oxidizing biocide. Chlorine dioxide has more oxidizing power than 'chlorine,' yet is less affected by the chemical composition of the water than is chlorine and it forms a true nonionized solution of a gas, which is not pH dependent. ClO_2 is usually generated at plant site by

the reaction of chlorite and hypochlorite, with the addition of acid to ensure the optimal pH for the reaction. For this reason, it is somewhat more complex and expensive than chlorination. Other oxidizing biocides are less extensively used in industrial water systems, but their mode of action is similar to chlorine and chlorine dioxide.

The alternative approach to the use of oxidizing biocides is to use much higher concentrations of organic, nonoxidizing biocides,[49] often blended with surfactants to increase their penetrating power into biofilms and other deposits. Some of the most common organic biocides are THPS (tetrakis hydroxy phosphonium sulfate),[50] aldehydes (glutaraldehyde and formaldehyde),[51] quaternary ammonium compounds,[52] biguanides,[53] isothiazolones,[54] chlorinated phenols, methylene-bis-(thiocyanate), dithiocarbamates, sulfones, and thiones. All the various types of organic biocides have advantages and disadvantages relative to one another and to oxidizing biocides, which are beyond the scope of this chapter. However, it is worthwhile to elucidate some general principles about biocide selection.

First, where some form of biocide treatment is in use, it should not necessarily be assumed that it must be effective. Biocides have very limited ability to penetrate existing deposits and are therefore much more effective when used in conjunction with frequent mechanical cleaning processes such as brush cleaning or 'pigging' of pipelines.

The first stage of biocide selection is a laboratory evaluation of candidate biocides. For plant under wet lay-up, recirculating water systems, or other batch processes, there are only two parameters that need to be optimized: type of biocide and concentration. However, for oilfield water systems, such as secondary recovery water injection systems, continuous treatment with biocides at an effective dose is not feasible. Continuous low-level injection of organic biocide is of no value; indeed continuous low-level exposure may possibly be detrimental in the long term, because there is some possibility that it may encourage the development of more biocide tolerant strains of bacteria. Periodic batch treatment is therefore required to treat biofilms and this means that there is a substantial matrix of parameters: type of biocide, concentration, duration of batch, and frequency of treatment. The first three factors may be optimized by means of a laboratory 'time kill test'[45] using real or simulated system water at system temperature, bacteria such as mixed cultures of SRB, isolated from the system for which the treatment is being selected, and most importantly, sessile rather than planktonic populations. More realistic, though more complex and expensive, tests can be performed on side stream units at site. As with batch chlorination, the rate of recontamination will suggest the frequency of treatment and this parameter can only be optimized on the basis of site experience.

Given the relative complexity of optimizing biocide dosages properly, it is not surprising that organic biocides are commonly either underdosed or overdosed. Underdosing ultimately results in lack of control of MIC, possibly involving biocide-resistant strains of microorganisms and negates the whole purpose of the treatment strategy. Conversely, overdosing involves not only unnecessary expenditure over a prolonged period but also the introduction of unnecessarily large amounts of toxic substances into the environment.

2.20.10.3 Alternatives to Biocide Treatments

Given the shortcomings of biocide treatments, specific antimicrobial treatments based on manipulation of microbial ecology have long been mooted; one such alternative, continuous nitrate treatment, has recently become commonplace for selective control of SRB in oilfield water systems. Though mainly employed as an agent against H_2S generation in deep petroleum reservoirs ('souring'), nitrate (or, less commonly, nitrite) based treatments have also been used in some locations as anti-MIC agents in topside facilities, either alone or in combination with biocide.

The objective of these treatments is the suppression of sulfide generation by the selective manipulation of indigenous bacteria through nutrient addition. There are several possible mechanisms by which nitrate can suppress the accumulation of sulfide.

Nitrate-reducing bacteria (NRB) generally 'outcompete' SRB for common carbon and energy sources such as acetate and longer chained fatty acids. This shifts the flow of electrons in energy-generating metabolism of bacteria away from sulfate reduction toward nitrate reduction. Ideally, this process gives rise to nitrogen as the end product, but sometimes it also results in the production of nitrite.

SRB themselves may preferentially use nitrate instead of sulfate as the electron acceptor in their energy-generating metabolism.

NRB may produce traces of compounds (nitrogen oxides and/or nitrites) that raise the oxidation–reduction (redox) potential of the environment to a

level that is inhibitory to the growth of SRB or these substances may act as toxins to SRB.

Finally, specialized nitrate-reducing sulfide oxidizing bacteria may use the sulfide produced by SRB as the electron donor for nitrate or nitrite reduction. In this case, the production of sulfide is not completely inhibited, but the consumption of sulfide by NRB prevents its accumulation.

Depending on the microbial populations present, any of the above mechanisms or combinations of mechanisms may be operative. The addition of nitrate does not directly kill SRB and, in some cases, will not inhibit their activity, but will still achieve the desired effect – the reduction or removal of sulfide. Several field tests have also demonstrated the potential usefulness of nitrate for controlling SRB. The Institute of Petroleum (now the Energy Institute) in the United Kingdom published a useful introductory review about use of nitrate treatments to control SRB, MIC, and reservoir souring.[55] Certainly, in laboratory tests, these treatments have proven very effective for suppression of SRB activity compared with conventional biocide treatments.

Unfortunately, in practice the use of these agents is by no means always problem free; in particular, the production of nitrites as a by-product of the treatment may itself induce pitting corrosion in a minority of applications; furthermore, nitrate is only active against SRB rather than the wider range of other MIC microorganisms. In summary, it continues to be the case that no single treatment can provide a definitive solution to MIC; the choice of most appropriate agent is very situation specific and will remain, like most decisions relating to living organisms, a challenging compromise.

References

1. Ross, R. T.; Sladen, J. R.; Wienart, L. A. In *Biodeterioration of Materials*; Walters, A. H., Elphick, J. J. Eds.; Elsevier: London and New York, 1975; pp 127–136.
2. Hill, E. C. In *Biodeterioration of Materials*; Walters, A. H., Elphick, J. J. Eds.; Elsevier: London and New York, 1975; pp 157–172.
3. Hill, E. C. In *Microbial Aspects of the Deterioration of Materials*; Lovelock, D. W., Gilbert, R. J. Eds.; Academic Press: London and New York, 1975; pp 127–136.
4. Miller, J. D. A.; Tiller, K. A. In *Microbial Aspects of Metallurgy*; Miller, J. D. A. Ed.; Medical & Technical Publishing: Aylesbury, 1971; pp 61–105.
5. Iverson, W. P. In *Microbial Iron Metabolism*; Niedlands, J. B. Ed.; Academic Press: New York, 1974.
6. Kucera, V. *Microbial Corrosion – A Literature Survey*; Swedish Corrosion Institute: Stockholm, 1980.
7. Pope, D. H.; Duquette, D.; Wayner, P. C.; Arland, H. J. Microbially Influenced Corrosion, a State-of-the-Art Review, Materials Technology Institute, Columbus, OH, Publication Number 13, 1984.
8. Sequeira, C. A. C. Ed.; *Microbial Corrosion*; European Federation of Corrosion, Report Number 29; Institute of Materials: London, 2000.
9. Stoeckler, J. G. II, Ed.; *A Practical Manual on Microbiologically Influenced Corrosion;* NACE International: Houston, TX, 2001; Vol. 2.
10. Anon. *Microbiologically Influenced Corrosion and Biofouling in Oilfield Equipment*; NACE International: Houston, TX, 1990; TCP 3.
11. Borenstein, S. W. *Microbiologically Influenced Corrosion Handbook*; Industrial Press: USA, 1994.
12. Videla, H. A.; Herrera, L. K. *Int. Microbiol.* **2005**, *8*, 169–180.
13. Stott, J. F. D.; Herbert, B. N. *J. Appl. Bacteriol.* **1986**, *60*, 57–66.
14. Postgate, J. R. *The Sulphate-Reducing Bacteria*; Cambridge University Press, 1979.
15. Ginter, R. L. *Bull. Am. Assoc. Petrol. Geol.* **1930**, *14*, 139.
16. Gates, G. L.; Parent, C. F. *Oil Gas J.* **1976**, 115.
17. Al Sawaf, F. D. S. *Econ. Geol.* **1977**, *72*, 608.
18. Ege, S. L.; Houghton, C. J.; Tucker, P. T. Second Annual Workshop on Secondary Recovery Technology, June 1985, London.
19. Davies, M.; Scott, P. J. B. *Oilfield Water Technology*; NACE International: Houston, TX, 2006.
20. Booth, G. H.; Cooper, A. W.; Cooper, P. M. *Br. Corros. J.* **1967**, *2*, 104.
21. Booth, G. H.; Cooper, A. W.; Wakerley, D. S. *Br. Corros. J.* **1967**, *2*, 109.
22. Schaschl, E. *Mater. Perform.* **1980**, *19*, 9.
23. Kobrin, G. *Mater. Perform.* **1976**, *15*, 38–43.
24. Tatnall, R. E. *Mater. Perform.* **1980**, *19*, 88.
25. Costello, J. A. Ph.D. thesis, University of Cape Town, South Africa, 1975.
26. King, R. A.; Miller, J. D. A. *Nature* **1972**, *233*, 491.
27. Campbell, H. S.; Chamberlain, A. H. L.; Angel, P. J. In *Corrosion and Related Aspects of Material for Potable water Supplies*; McIntyre, P., Mercer, A. D. Eds.; Institute of Materials: London, 1993; pp 222–231.
28. Wagner, D.; Chamberlain, A. H. L. *Biodegradation* **1997**, *8*, 177–187.
29. Critchlay, M. M.; Pasetto, R.; O'Halloran, R. J. *J. Appl. Bacteriol.* **2004**, *97*, 590–597.
30. Geesey, G. G.; Beamer, P. J. *Mar. Technol. Soc.* **1990**, *24*, 36–43.
31. Siedlarek, H.; Wagner, D.; Fischer, W. R.; Paradies, H. H. *Corros. Sci.* **1994**, *34*(1), 50–54.
32. Purkiss, B. E. In *Microbial Aspects of Metallurgy*; Miller, J. D. A. Ed.; Medical & Technical Publishing: Aylesbury, 1971; pp 107–128.
33. Kempner, E. S. *J. Bacteriol.* **1966**, *92*, 1842.
34. Le Roux, N. W.; North, A. A.; Wilson, J. C. Tenth International Mineral Processing Congress, London, Institute of Mining and Metallurgy **1974**, 1051.
35. White, G. C. *Handbook of Chlorination and Alternative Disinfectants*, 4th ed.; Wiley: New York, 1999.
36. Thistlethwayte, D. K. B. Ed.; *The Control of Sulphides in Sewerage Systems*; Butterworth: Sydney, Australia, 1972.
37. Morgan, T. O. B.; Steele, A. D.; Gilbert, P. D. Concrete Durability in Acidic Stagnant Water, Proceedings of Congress on Microbial Corrosion, Teddington, March 1983; The Metals Society: London, 1983.
38. Elphink, J. J. In *Microbial Aspects of Metallurgy*; Miller, J. D. A. Ed.; Medical & Technical Publishing: Aylesbury, 1971; pp 157–172.

39. Jurinski, M. P.; Scully, J. R.; Lillard, R. S. Corrosion 1993 Conference, National Association of Corrosion Engineers (NACE), Paper Number 501.
40. Kain, R. D. Corrosion 1987 Conference, National Association of Corrosion Engineers (NACE), Paper Number 348.
41. Kain, R. D. Corrosion 1993 Conference, National Association of Corrosion Engineers (NACE), Paper Number 67.
42. Gallagher, P.; Malpas, R. E. Corrosion 1989 Conference, National Association of Corrosion Engineers (NACE), Paper Number 113.
43. Tarbuck, L. A.; Wyborn, C. H. E. Proceedings of Condenser Biofouling Control Symposium., Electric Power Research Institute, Atlanta, U.S.A., March, 1979.
44. Hoar, T. P. *Corrosion* **1957**, *14*, 1036–1046.
45. Standard Test Method, TM0-194-94. Field Monitoring of Bacterial Growth in Oilfield Systems; NACE International: Houston, TX, 1994.
46. Pots, B. F. M.; John, R. C.; Rippon, I. J.; Maarten, J. J.; Thomas, S.; Kapusta, S. D.; Girgis, M. M.; Whitham, T. Corrosion 2002 Conference, National Association of Corrosion Engineers (NACE), Paper Number 2235.
47. Horvath, J.; Novak, M. *Corros. Sci.* **1976**, *4*, 159.
48. Pankhurst, E. C.; Nicolle, H. G. M.; Allan, K. In *Microbial Aspects of the Biodeterioration of Materials*; *Society for Applied Bacteriology, Technical Series 9*; Academic Press: London, 1975; pp 137–152.
49. Romero, J. M.; Amaya, M.; Hernandez, J. A. *Mater. Peform.* **2004**, 38–40.
50. Larsen, J.; Sanders, P. F.; Talbot, R. E. Corrosion 2000 Conference, National Association of Corrosion Engineers (NACE), Paper Number 123.
51. Gorman, S. P.; Scott, E. M.; Russell, A. D. *J. Appl. Bacteriol.* **1980**, *48*, 161–190.
52. Bessems, E.; Clemmit, A. F. Chemicals in the Oil Industry, Proceedings of a Symposium Organised by Royal Society of Chemistry, Manchester, United Kingdom; Ogden, P. H. Ed.; March 1983; pp 159–170.
53. Gilbert, P.; Pemberton, D.; Wilkinson, D. E. *J. Appl. Bacteriol.* **1990**, *69*, 593–598.
54. Collier, P. J.; Ramsay, A. J.; Austin, P.; Gilbert, P. *J. Appl. Bacteriol.* **1990**, *69*, 569–577.
55. The Energy Institute. The Stimulation of Nitrate-Reducing Bacteria (NRB) in Oilfield Systems to Control Sulfate-Reducing Bacteria (SRB), Microbiologically Influenced Corrosion and Reservoir Souring – An Introductory Review. The Energy Institute (formerly The Institute of Petroleum): London, 2003.

2.21 Corrosion in Alkalis

J. A. Richardson
Anticorrosion Consulting, 5 Redhills Lane, Durham DH1 4AL, UK

Abbreviations

ASTM American Society for Testing Materials
CPVC Chlorinated polyvinyl chloride
ECTFE Ethylene chlorotrifluoroethylene
ETFE Ethylene tetrafluoroethylene
FEP Fluorinated ethylene propylene
FRP Fiber reinforced plastic
HDPE High density polyethylene
IGSCC Intergranular stress corrosion cracking
LME Liquid metal embrittlement
PE Polyethylene
PFA Perfluoroalkoxy
PP Polypropylene
ppm Parts per million
PTFE Polytetrafluoroethylene
PVC Polyvinyl chloride
PVDF Polyvinylidene fluoride
PWHT Postweld heat treatment
PWR Pressurized water reactor
SCC Stress corrosion cracking
SCE Saturated calomel electrode
SHE Standard hydrogen electrode
XPS X-ray photoelectron spectroscopy

Symbols

ΔT Temperature difference (°C)

2.21.1 Introduction

Sodium hydroxide (caustic soda) is one of the world's most commercially important chemicals. It is produced in huge quantities worldwide, mostly as a by-product of the electrolysis of brine to produce chlorine, historically in mercury cells but increasingly in diaphragm or membrane cells. Diaphragm cells are used to produce the majority of the world's caustic soda which leaves the cells at ~10–12% (w/w) concentration, and is subsequently evaporated to produce commercial grades, typically at 50% and 73% (w/w) concentration, or anhydrous. Potassium hydroxide (caustic potash) also has some commercial significance but it is more expensive than and, in volume terms, well-behind caustic soda.

Both sodium and potassium hydroxides are strong bases that are highly dissociated in water at all concentrations to release the metal cation and the hydroxide anion as follows:

$$Na/KOH = Na/K^{+} + OH^{-}$$

Their aqueous solutions have pHs above 7, and pHs 10, 11, and 12 are equivalent to concentrations of 4, 40, and 400 ppm (w/w) sodium hydroxide respectively at 25 °C.

Depending mostly on the method of manufacture, sodium hydroxide may contain contaminants that are potentially significant from the corrosion standpoint, in particular:

1. Chlorides, particularly in the case of the sodium hydroxide produced in diaphragm cells that may contain up to 1% (w/w) sodium chloride, in contrast to mercury or membrane cells that deliver sodium hydroxide with much lower levels of sodium chloride, typically 10 s of ppm.
2. Chlorates, ClO_3^-, are produced at the anode during electrolysis and are normally removed before or during the evaporation stage in the manufacturing process, so are of significance mostly to the performance of manufacturing equipment.
3. Chlorine and hypochlorite that is formed when chlorine, a coproduct of electrolysis, dissolves in sodium hydroxide:

$$NaOH + Cl_2 = NaCl + NaOCl$$

4. Mercury, entrained in the sodium hydroxide produced in mercury cells.

Alkalis differ fundamentally from acids in the mechanism of the metal dissolution process. In acids, soluble metal cations are produced anodically, supported by hydrogen evolution as the primary cathodic process. In alkalis, soluble oxyanions containing the metal are produced by reaction between the metal and hydroxide anions. In the case of the so-called amphoteric metals such as aluminum, zinc, tin and lead, and their alloys, these processes are so strongly favored as to render the metals unsuitable for service under alkaline conditions because of the formation of soluble aluminates, zincates, stannates, and plumbates respectively. For example, a simplified Pourbaix diagram for the aluminum–water system at 25 °C is shown in **Figure 1**, and confirms that the metal will be potentially vulnerable to corrosion above a pH of ~9 as well as below a pH of ~4.

The reactions that give rise to this alkaline vulnerability in aluminum are thought to consist of the initial formation of aluminum hydroxide in a series of three single electron transfers[1]:

$$Al + OH^- = Al(OH)_{ads} + e^-$$

$$Al(OH)_{ads} + OH^- = Al(OH)_{2ads} + e^-$$

$$Al(OH)_{2ads} + OH^- = Al(OH)_{3ads} + e^-$$

The final step is believed to be the chemical dissolution of aluminum hydroxide to form soluble aluminate[1]:

$$Al(OH)_3 + 2OH^- = AlO_2^- + 2H_2O$$

The supporting cathodic process is the direct reduction of water molecules on the aluminum surface with the liberation of hydrogen:

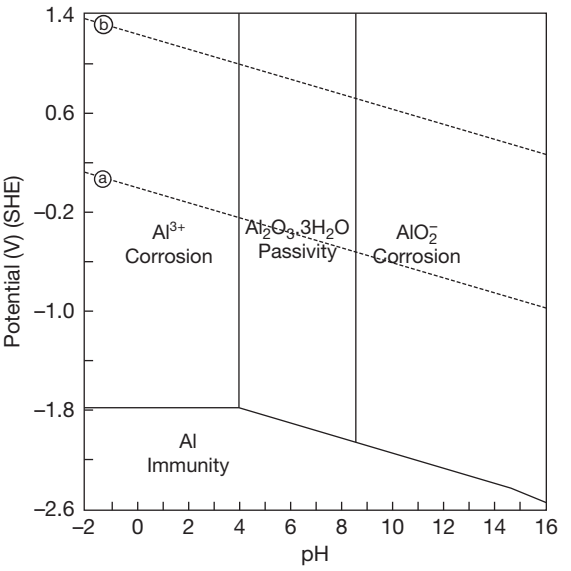


Figure 1 Simplified potential/pH diagram for the Al–H_2O system at 25 °C, based on an equilibrium activity of 10^{-6} g ion per litre.

$$2H_2O + 2e^- = H_2 + 2OH^-$$

Thus, and somewhat counter-intuitively, hydrogen is produced in the corrosion processes occurring adjacent to the metal surface despite the tiny concentration of protons in the bulk fluid.

The corrosion performances of the various classes of material in alkalis have been reviewed extensively elsewhere.[2–5] Much of the data relates to sodium hydroxide, reflecting its commercial significance. There is relatively little data for potassium hydroxide in the public domain. Readers whose main interest is to identify 'what works where' are referred to these sources and the relevant chapters on specific materials in this book. In this chapter, the corrosion performances of materials are reviewed with an emphasis on the principles and mechanisms that underpin their corrosion performances in alkalis.

2.21.2 Steels and Cast Irons

The Pourbaix diagram for the iron–water system at 25 °C shown in **Figure 2**[6] suggests that steels should form protective oxide films, probably composed of magnetite (Fe_3O_4), at pH up to ~14, approximately equivalent to ~4% (w/w), or 1 M, sodium hydroxide.

When the pH is above 14 the diagram indicates that there is a potential corrosion zone arising from

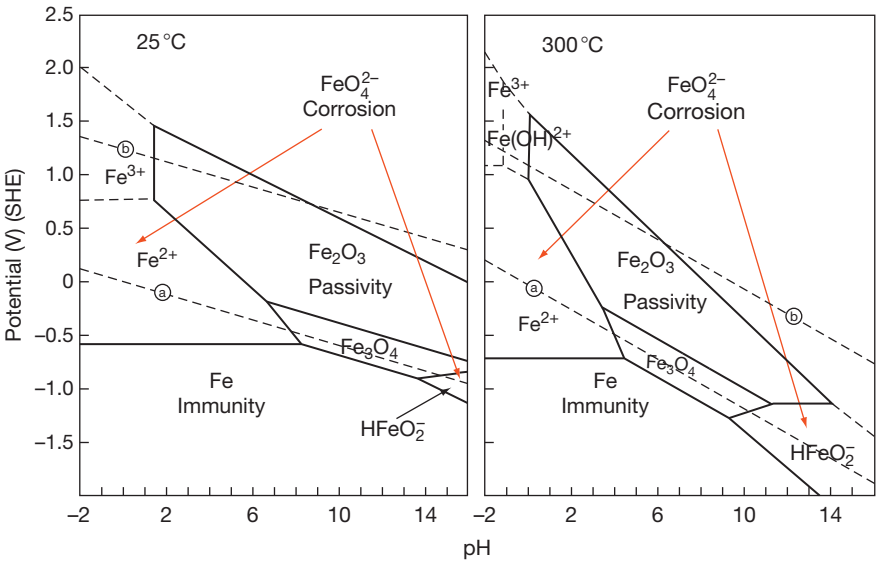


Figure 2 Simplified potential/pH diagrams for the Fe–H_2O system at 25 °C and 300 °C, based on an equilibrium activity of 10^{-6} g ion per litre. Adapted from Chen, C. M.; Aral, K.; Theus, G. J. Computer Calculated Potential-pH Diagrams to 300 °C, Report NP-3137, Electric Power Research Institute, 1983.

the stability of the dihypoferrite anion, $HFeO_2^-$, that forms by the anodic dissolution of the metal in a sequence of steps similar to those for aluminum above, culminating in the dissolution of ferrous hydroxide as the dihypoferrite anion, rather than its further oxidation to solid magnetite on the surface. The overall process can be written as follows:

$$Fe + 2H_2O = HFeO_2^- + 3H^+ + 2e^-$$

As in the case of aluminum, the supporting cathodic process is the reduction of water to hydrogen. The dihypoferrite anion is in equilibrium with the hypoferrite (ferroate) anion as follows:

$$HFeO_2^- + OH^- = FeO_2^{2-} + H_2O$$

This equilibrium moves to the right as the alkalinity increases. Dissolution of the magnetite film also contributes to metal dissolution under alkaline conditions:

$$Fe_3O_4 + 2H_2O + 2e^- = 3HFeO_{2^-} + H^+$$

In practice, carbon steel is widely used for the storage, handling, and transport of sodium hydroxide solutions at concentrations and temperatures well beyond the limits suggested in the Pourbaix diagram, indicative of a kinetic rather than thermodynamic control over the stability of passivity. Although measured corrosion rates vary with purity, velocity, and other factors, the general experience is that corrosion rates are acceptably low in concentrations up to ~75% (w/w) at temperatures up to ~80 °C and perhaps even beyond,[5] assuming that some degree of iron contamination is acceptable. Of the potential contaminants, only chlorates present a significant threat, having the potential to accelerate the corrosion of steel by up to an order of magnitude, depending on the concentration.[5]

The major concern across these concentrations and temperatures is not general corrosion but the risk of intergranular stress corrosion cracking (SCC) to which carbon steels are vulnerable in the presence of tensile residual stresses arising from welding and cold working operations during fabrication. These risks can be mitigated by thermal stress relief following fabrication and some well-established guidelines for determining the need, or otherwise, for postweld heat treatment (PWHT) are available and are summarized in **Figure 3**.[7]

The presence of carbon/carbides in the steel and the segregation of phosphorus at grain boundaries increase the risks of cracking.[8] Low alloy steels vary in their vulnerability to caustic-induced SCC. Tests have suggested that nickel-containing steels are more resistant than carbon steels, molybdenum additions are detrimental and the effects of chromium additions vary with temperature.[9]

At higher temperatures and concentrations, the zone of dihypoferrite stability increases considerably, extending to lower pH values, as shown in the Pourbaix diagram for the iron–water system at 300 °C in

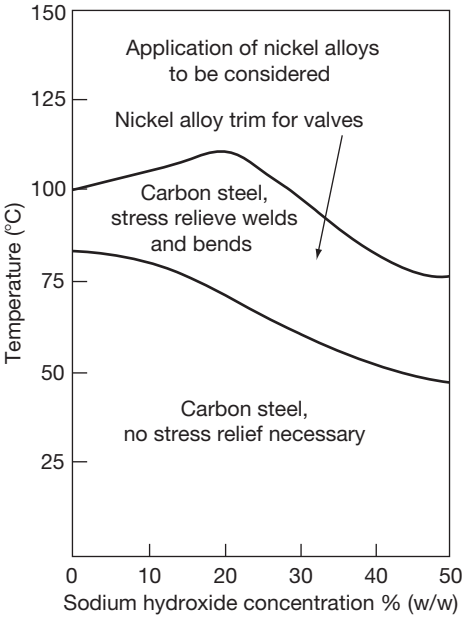


Figure 3 Temperature and concentration limits for stress corrosion cracking susceptibility of carbon steels in sodium hydroxide solutions. Adapted from NACE Standard Practice SP0403–2008. Avoiding Caustic Stress Corrosion Cracking of Carbon Steel Refinery Equipment and Piping; NACE International: Houston, TX, 2008.

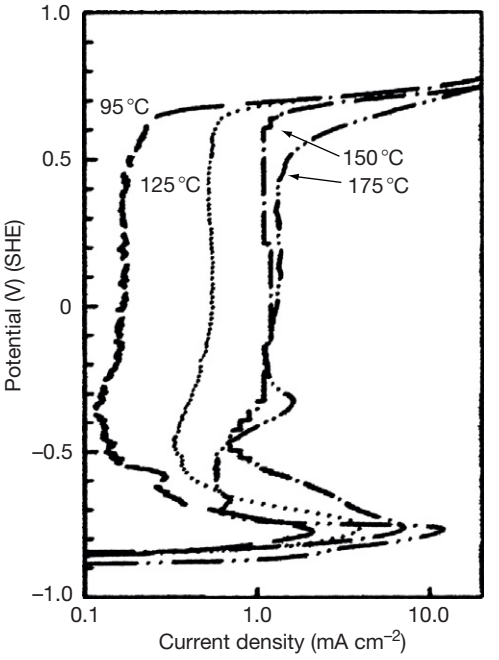


Figure 4 Potentiodynamic polarization curves for AISI 1020 steel in stagnant 2.75 M (~11% w/w) sodium hydroxide at temperatures in the range 95–175 °C. Adapted from Giddey, S.; Cherry, B.; Lawson, F.; Forsyth, M. *Corros. Sci.* **2001**, *43*, 1497–1517.

Figure 2, which suggests that the alkaline corrosion of steels is possible at a pH down to ~9, equivalent to lightly alkalized water. Kinetic studies suggest that the magnetite film on carbon steels begins to lose its protective properties as temperatures increase above the atmospheric boiling point, depending on concentration. Anodic polarization curves for AISI 1020 steel in stagnant 2.75 M (~11% w/w) sodium hydroxide at temperatures in the range 95–175 °C are shown in Figure 4.

The active regions of the curves are thought to be attributable to active iron dissolution as ferroate rather than dihypoferrite, because of the greater solubility of the former in alkaline solutions. Although passivation behavior is evident at all four test temperatures, the passive current increases by around an order of magnitude between the lowest and highest temperatures. Additional studies on rotating electrodes suggest that this is due to magnetite film thinning due to increased solubility as the temperature rises. Film formation competes with film dissolution, and turbulence accelerates film dissolution and corresponding corrosion rates.[10]

Although the corrosion risks that are presented at high temperatures by increased concentrations of sodium hydroxide solutions are well understood, problems can arise in practice in situations where nominally benign concentrations can increase locally under the influence of high heat fluxes and/or adverse geometries, particularly in steam raising equipment. Corrosion control in boilers is based on maintaining the pH of boiler waters, typically in the pH range 9–10, by the addition, at the ppm level, of either free sodium hydroxide or combinations of sodium phosphates that buffer the pH in the required range but have the potential to release free sodium hydroxide, depending on the Na:P ratio. The concentration of these alkaline species can be increased by several orders of magnitude from the ppm to the percent level in films of boiler water on heat transfer surfaces when the metal skin temperature exceeds the boiling temperature of the water at the prevailing pressure. The driving force for concentration is related to the temperature difference between the two, ΔT, as illustrated in Figure 5.[11]

Heat transfer into a crevice, as might exist at a tube/tubesheet joint in a boiler, can exacerbate the driving force for concentration.[12] These concentration processes at local sites in steam raising equipment can result in the on-load, caustic corrosion, sometimes described as 'caustic gouging,' of utility[13] and process plant[14] steel boiler tubes, and also give

rise to caustic-induced SCC of carbon, low alloy and stainless steels, and nickel alloy components, exemplified by the caustic cracking of a carbon steel stub attached to an alloy 825 (N08825) bellows that is illustrated in Figure 6. The root cause of the problem was a malfunctioning desuperheater upstream of the bellows that injected periodic slugs, rather than the intended spray, of water containing a few ppm of sodium hydroxide into the pipe. The water collected in the crevice between the stub and the bellows and successive wetting/drying events resulted in concentration of the caustic in the crevice, leading ultimately to SCC of the stub, driven by the residual stresses around the attachment weld to the bellows.

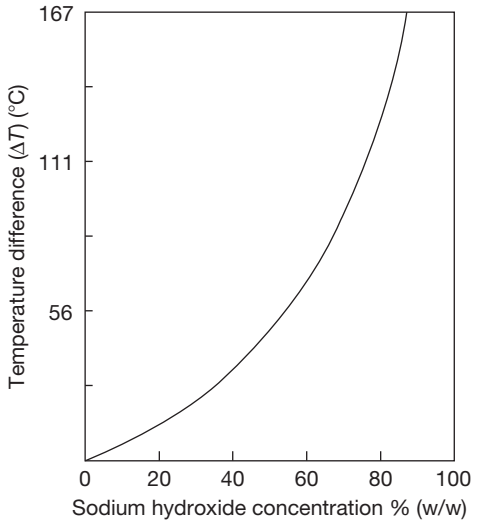


Figure 5 Concentration of sodium hydroxide attainable in a film of boiler water on a heat transfer surface as a function of the temperature difference, ΔT, between the surface temperature and the boiling point of the water. Adapted from *Betz Handbook of Industrial Water Conditioning*, 9th ed.; Betz Laboratories, 1991.

Grey and ductile cast irons have corrosion resistances broadly similar to steels in sodium hydroxide. Ductile irons, in contrast to grey irons, are vulnerable to SCC. A comparative study of pure iron and white cast iron in 14 M (~56% w/w) acid in the temperature range 20–60 °C suggested that iron carbide is detrimental to corrosion performance in relatively concentrated sodium hydroxide.[15] Nickel has the most significant influence on the corrosion resistance of cast irons, as shown by the corrosion rate data in Figure 7 for cast irons with varying nickel contents in boiling 50–65% (w/w) sodium hydroxide.[16]

In particular, the austenitic irons containing >15% nickel (such as F41000, F41002, and F43000) have very good resistance to sodium hydroxide solutions of up to ~70% (w/w) concentration, but can be vulnerable to SCC, particularly if chlorides are present, for which stress relief provides appropriate mitigation.[5] High silicon cast irons, containing typically 14–16% Si (such as F 47003), are not resistant to relatively concentrated sodium hydroxide at elevated temperatures because the siliceous films that provide their resistance to erosion–corrosion are soluble in the alkali.

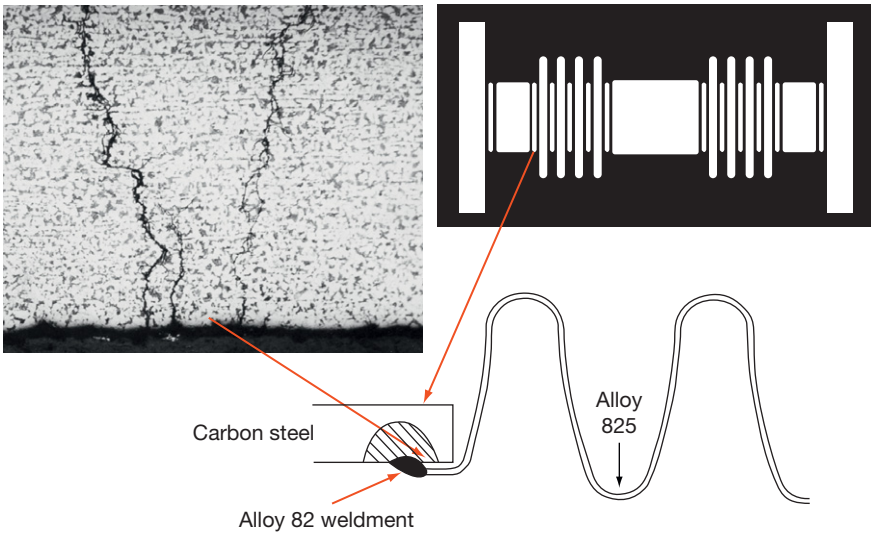


Figure 6 Section through carbon steel bellows stub downstream of a desuperheater in a steam main operating at ~16 bar and 250 °C that has suffered caustic SCC due to sequential wetting/drying in the crevice between the stub and the alloy 825 convolutions.

Although there is very little specific data in the public arena, the performances of steels and cast irons in potassium hydroxide are very similar to their equivalent performances in sodium hydroxide.

2.21.3 Stainless Steels

Potentiodynamic polarization curves in stagnant 15% (w/w) sodium hydroxide solution at 90 °C for the key constituent elements of stainless steels, pure iron, nickel, chromium, and molybdenum, are shown in **Figure 8**.[17] Iron has the most active corrosion potential, and in the early stages of anodic polarization, it dissolves at progressively increasing rates. At a relatively high, critical current density, it exhibits classic passivation behavior due to the formation of magnetite, Fe_3O_4 which shows gradually increasing rates of dissolution as the potential is further increased to a secondary passivation peak, probably associated with the oxidation of Fe_3O_4 to Fe_2O_3. The current density shows little tendency to increase in response to further polarization up to a potential >300 mV at which transpassive breakdown as soluble ferrate, FeO_4^{2-}, occurs.

Molybdenum has a corrosion potential more noble than iron but more active than chromium or nickel. However, under anodic polarization, molybdenum dissolves at progressively faster rates as soluble molybdate, MoO_4^{2-}, and shows no tendency to passivate. Chromium, on the other hand, has the most noble corrosion potential, and its initial anodic polarization behavior is consistent with the presence of a protective passive film of chromium oxide, Cr_2O_3, from the outset with no evidence of an active dissolution peak prior to the onset of passivity. When present, the film is more protective than the corresponding films on iron and nickel, having the lowest current density requirement for maintenance in its passive potential range. However, transpassive breakdown to soluble chromate, CrO_4^{2-}, occurs at a potential

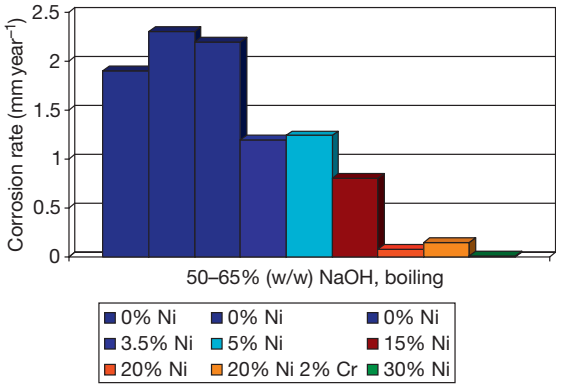


Figure 7 Corrosion rates for cast irons containing varying quantities of nickel tested in boiling 50–65% (w/w) sodium hydroxide solution. Adapted from Corrosion Resistance of Nickel and Nickel-Containing Alloys in Caustic Soda and Other Alkalis; CEB-2, International Nickel Company, 1973.

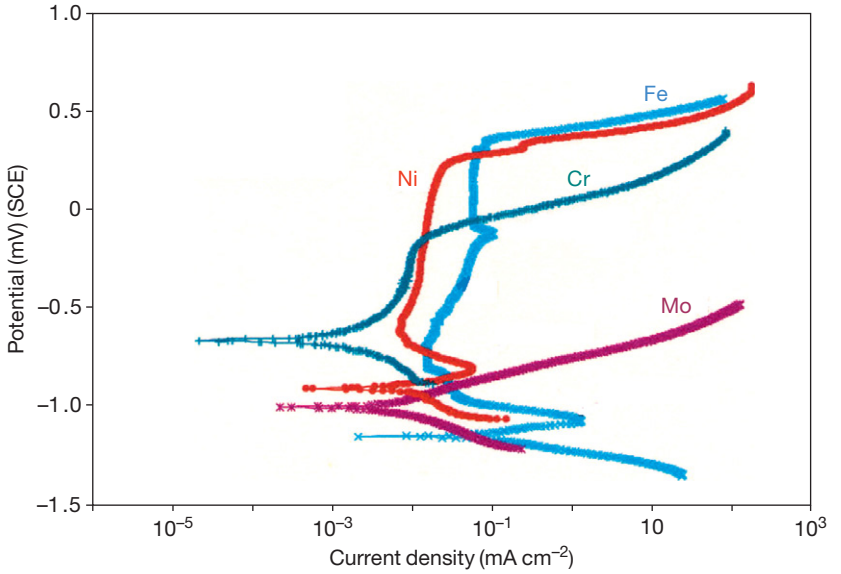


Figure 8 Potentiodynamic polarization curves for pure iron, nickel, chromium, and, molybdenum in stagnant 15% (w/w) sodium hydroxide at 90 °C. Adapted from Bhattacharya, A.; Preet, M. S. In *Corrosion 2008, Corrosion of Duplex Stainless Steels in High pH Caustic Solution*, New Orleans, LA; NACE International: Houston, TX, 2008; Paper 08194.

well below the corresponding breakdown potentials for iron and nickel.

The corrosion potential of nickel is more noble than either iron or molybdenum but less noble than chromium. Nickel's anodic polarization behavior attests to the importance of the element in providing resistance to strongly alkaline environments. The critical current density for passivation is an order of magnitude lower than that for iron. The passive zone extends over a wide potential range within which nickel is protected by nickel oxide, NiO. As the potential increases in the passive range, the current density also increases gradually but the current densities for the maintenance of passivity are only marginally higher than for chromium. Ultimately, transpassive breakdown does occur at a similar potential in iron due to the formation of a soluble species, probably $HNiO_2^-.H_2O$, well beyond the transpassive breakdown potential of chromium.

The polarization curves for the individual alloying elements explain their effects on the performances of stainless steels. The compositions of the commercial grades of stainless steels that are relevant to alkali applications are summarized in Table 1.

Nickel delivers the most significant benefits, providing additional resistance to anodic dissolution at lower potentials and robust passivity across a potential range wide enough to limit the possible promotion of transpassivity by oxidizing agents. Chromium also delivers significant benefits by inhibiting dissolution and promoting passivity at lower potentials but introduces a vulnerability to transpassive dissolution in the presence of oxidizing agents. Molybdenum's benefits are confined to additional resistance to anodic dissolution in the active potential range.

These roles for the individual alloying elements are apparent in the formation of the passive films that form spontaneously on stainless steels in alkalis. Film formation was studied recently[18] in an X-ray photoelectron spectroscopic (XPS) study of the progressive formation of the naturally formed passive films on types 304 (S30400) and 2205 (S32205) stainless steel in 0.4% (w/w) sodium hydroxide solution at ambient temperatures, in which the natural corrosion potential of both increases with time. On first immersion, iron dissolves preferentially and is present in the film ultimately as Fe^{3+} oxy-hydroxide, the content of which decreases with time. Chromium accumulates with time in the film as Cr^{3+} oxy-hydroxide, and the combined effect of iron and chromium depletion is that the metal layer beneath the film becomes enriched in nickel, which accumulates in the film as nickel hydroxide, to a greater extent on

Table 1 Compositions of some wrought stainless steels that are relevant to alkali applications

UNS no	*Common name*	*Fe*	*Ni*	*Cr*	*Mo*	*Cu*	*Si*	*Other*
S43000	430	Balance		16–18				
S43035	439	Balance		17–19				
S44627	E-brite	Balance		26	1			
S44735	29–4C	Balance		29	4			
S32304	2304	Balance	4–5	23				
S31803 S32205	2205	Balance	4.5–6.5	21–23	2.5–3.5			N
S32750 S32760	2507	Balance	6–8	24–26	3–4.5 3–4	<0.5 0.5–1.0		N W, N
S32900		Balance	4.5	28	1.5			
S30400	304	Balance	8–10.5	18–20				<0.08C
S30403	304L	Balance	8–12	18–20				
S30500	305	Balance	10.5–13	17–19				<0.12C
S31600	316	Balance	10–14	16–18	2–3			<0.08C
S31603	316L	Balance	10–14	16–18	2–3			
S31254	254SMO	Balance	17.5–18.5	19.5–20.5	6–6.5	0.5–1.0		N
S30900	309	Balance	12–15	22–24				Mn
S31000	310	Balance	19–22	24–26				Mn
S32654	654SMO	Balance	22	24	7.3			Mn
N08904	904L	Balance	23–28	19–23	4–5	1–2		
N08926	25–6MO	Balance	24–26	19–21	6–7	0.5–1.5		N
S31277	27–7MO	Balance	26–28	20.5–23.0	6.5–8.0	0.5–1.5		Mn, N
N08020	20	Balance	32–35	19–21	2–3	3–4		
N08800	800	Balance	30–35	19–23				
N08028	28	Balance	31	27	3.5	1		
R20033	33	Balance	30–33	31–35	0.5–2.0	0.3–1.2		N

type 304 (S30400) than the 2205 (S32205) stainless steel. The molybdenum in the 2205 (S32205) stainless steel is present as an oxy-hydroxide in the film and in the metallic form at the metal/film interface at the same level as in the parent material. The dealloying effects are exacerbated in more aggressive conditions at higher temperatures and in more concentrated solutions to the extent that 'nickel' layers have been reported on austenitic stainless steel surfaces. It has been shown that, in reality, the dealloying of iron from the surface of type 316 stainless steel proceeds as far as a solid solution with a nickel/iron atomic ratio of ~1.3, equivalent to ~56% nickel, in 50% (w/w) sodium hydroxide solution at 140 °C.[19]

In practice, the relative contribution of chromium and nickel to the corrosion resistances of stainless steels in alkalis depends on and varies with the grade. The standard ferritic grades of stainless steel such as types 430 (S43000) and 439 (S43035), containing up to ~19% chromium, do not find significant application in alkalis. However, the so-called super-ferritic grades containing typically 26–29% chromium with controlled carbon + nitrogen levels or stabilized with titanium or niobium to control vulnerability to intergranular corrosion due to sensitization, typified by the proprietary grades E-brite 26–1 (S44627) and AL 29–4C (S44735), have significant, potentially useful resistance. The corrosion rate data in **Figure 9** demonstrate the superior performance of the super-ferritic grades relative to conventional austenitic grades of stainless steel in tests in 50% (w/w) sodium hydroxide at 140 °C in which a conventional ferritic grade performed relatively badly.[20] Superferritic stainless steels have found commercial application as tubing materials in caustic evaporators, and are not adversely affected to any significant extent by the presence of contaminants. Indeed, they require a sustained supply of oxidizing contaminants such as chlorates to maintain passivity at the prevailing high temperatures and concentrations.[5]

Isocorrosion diagrams for the basic types 304 (S30400) and 316 (S31600) austenitic stainless steels in sodium hydroxide are shown in **Figure 10**.[21] There is little to choose between the grades and their low carbon equivalents, all of which are resistant up to ~20% (w/w) boiling sodium hydroxide solutions. Corrosion rates for some duplex stainless steels in boiling sodium hydroxide solutions are shown in **Figure 11**[22] and suggest that duplex alloys are resistant up to ~30% (w/w) boiling sodium hydroxide solutions, depending on the chromium and nickel contents of the specific grade.

The corrosion rates of various grades of duplex stainless steels in 30–70% (w/w) sodium hydroxide solutions are in the order S32906 <2507 (S32750) <2205 (S32205) <2304 (S32304), confirming that their resistances are determined largely, but not exclusively, by their chromium contents.[23] However, the vulnerability of chromium to transpassive dissolution was confirmed in studies of austenitic and duplex alloys in 30% (w/w) sodium hydroxide at 150 °C. Under deaerated conditions, the strong, passivating effects of both chromium and nickel were confirmed. However, in more strongly oxidizing conditions, chromium was dissolved resulting in thicker, less protective oxide films containing only nickel and iron.[24]

The overall effects of additions of the major alloying elements nickel, chromium, and molybdenum on the corrosion of austenitic and duplex stainless steels

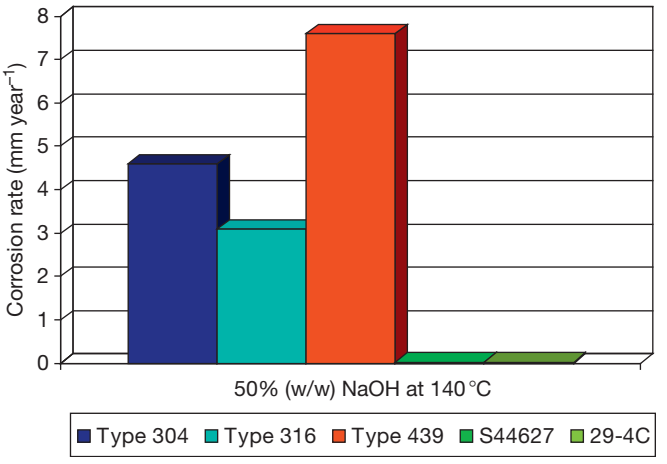


Figure 9 Corrosion rates for various stainless steels in 50% (w/w) sodium hydroxide at 140 °C. Adapted from Deverell, H. E.; Franson, I. A. *Mater. Perform* **1989**, *28*, 52–57.

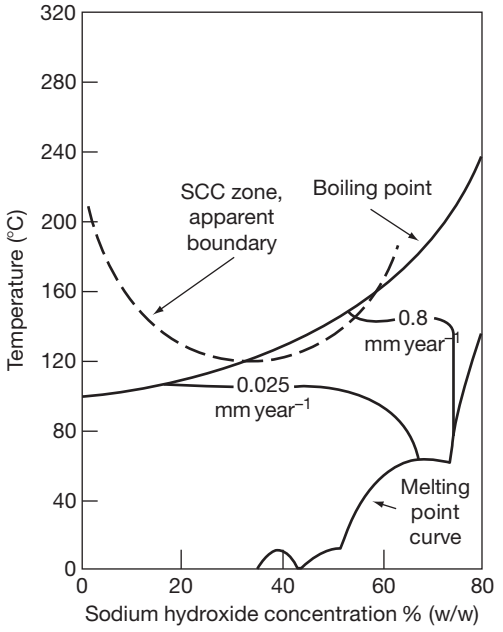


Figure 10 Isocorrosion diagram for types 304 and 316 stainless steels in sodium hydroxide. Adapted from Corrosion Resistance of Nickel and Nickel-Containing Alloys in Caustic Soda and Other Alkalis; CEB-2, International Nickel Company, 1973.

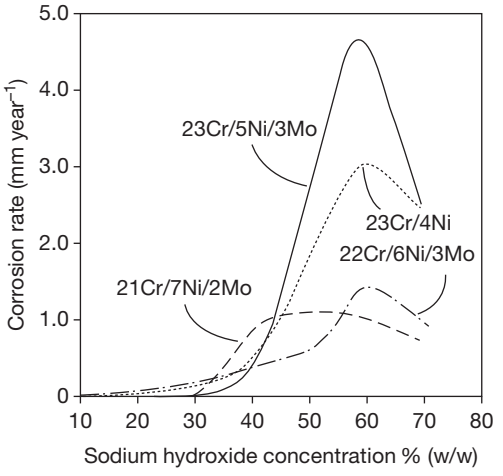


Figure 11 Corrosion rates of duplex stainless steels as a function of concentration of boiling NaOH solutions. Adapted from Horn, E. M. *Corrosion and Environmental Degradation*; Schutze, M., Ed.; Wiley-VCH: London, 2000; Vol. 2, pp 69–111.

in sodium hydroxide are summarized in Figure 12, based on an analysis of the threshold temperature in the range 85–146 °C at which the corrosion rate in 50% (w/w) sodium hydroxide exceeds 0.13 mm

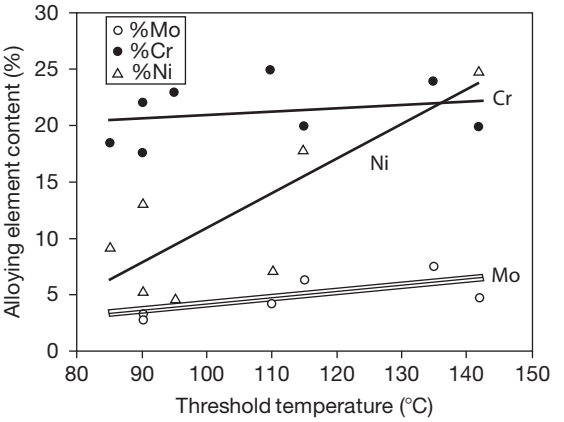


Figure 12 Effects of nickel, molybdenum, and chromium contents of stainless steels on the threshold temperature at which the corrosion rate in 50% (w/w) sodium hydroxide exceeds 0.13 mm $year^{-1}$. Adapted from Davies, M. Corrosion by Alkalies. In *ASM Handbook: Corrosion: Environments and Industries*; ASM International: Materials Park, OH, 2006; Vol. 13C, pp 710–726.

$year^{-1}$ for the grades 304 (S30400), 316L (S31603), 2205 (S32205), 2304 (S32304), 2507 (S32750), 254SMO (S31254), 654SMO (S32654), and 904L (N08904).[5]

The most significant alloying element is clearly nickel, with both chromium and molybdenum demonstrating only minor benefits within the relevant composition ranges. The beneficial effects of nickel additions were demonstrated in a study of stainless steels and nickel alloys in 50% (w/w) sodium hydroxide at 300 °C in the presence and absence of oxygen, the results of which are shown in Figure 13.[25] Clearly, corrosion rates are very sensitive to the presence or otherwise of oxygen but in all conditions, the resistance increases with nickel addition up to a content of ~20%, but not beyond under these relatively aggressive conditions.

At lower temperatures, alloys with >20% nickel such as 904L (N08904) and alloys 20 (N08020), 800 (N08800), and 33 (R20033) provide much improved resistance to sodium hydroxide relative to the basic austenitic grades and cheaper alternatives to nickel based alloys. For example, Figure 14 shows corrosion rates for alloy 33 relative to nickel alloys 600 (N06600) and 625 (N06625) in 50% (w/w) sodium hydroxide solution at 180 °C. Although not as resistant as alloy 600 (N06600), alloy 33 (R20033) corrodes at a rate of <0.1 mm $year^{-1}$ that is acceptable for most applications and is much more resistant than the molybdenum-containing alloy 625 (N06625). Additionally, alloy 33 (R20033) shows no evidence of dealloying after 153 days exposure, in contrast to

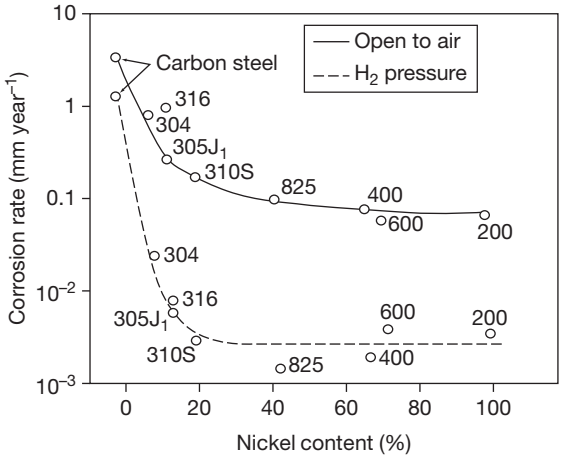


Figure 13 Effect of nickel content on the corrosion rates of stainless steels and nickel alloys in 50% (w/w) sodium hydroxide at 300 °C in the presence and absence of oxygen. Adapted from Yasuda, M.; Fukumoto, K.; Koizumi, H.; Ogata, Y.; Hine, F. *Corrosion* **1987**, *43*, 492–498.

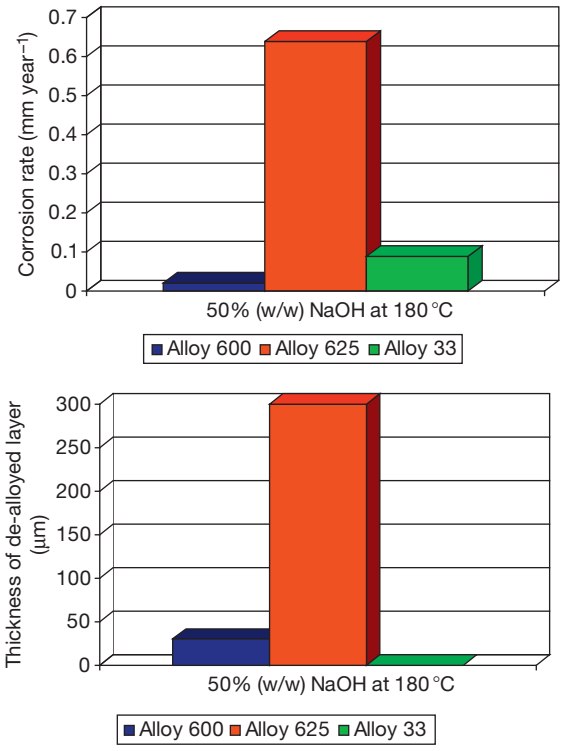


Figure 14 Corrosion rates and dealloyed layer thicknesses for alloys 600, 625 and 33 tested in 50% (w/w) sodium hydroxide solution at 180 °C. Adapted from Paul, L.; Alves, H. *Corrosion 2004, Alloy 33: A Versatile Alloy for Mineral Acid and Other Applications*, New Orleans, LA; NACE International: Houston, TX, 2004; Paper 04226.

both of the nickel alloys that are significantly dealloyed after 35 days exposure.[26]

Stainless steels are vulnerable to SCC in alkaline solutions. It has been proposed that the vulnerability is due to the iron and chromium dealloying processes that occur on stainless steel surfaces in alkaline exposures[19] and this is described in more detail in **Chapter 2.05, Dealloying** in this volume. The conditions of concentration and temperature that present risks to the basic types 304 (S30400) and 316 (S31600) austenitic stainless steels are shown in **Figure 9** which confirms that SCC is possible at temperatures above ~100 °C, depending on concentration. Tests on duplex stainless steels have revealed no evidence of SCC in boiling 20–70% (w/w) solutions, suggesting that they are at least as resistant as the basic austenitic grades.[21,23] Various studies have indicated that chromium is more important than nickel in determining the resistance to SCC in alkaline solutions. For example, in 30% (w/w) sodium hydroxide at 200 °C, the resistances to cracking in separate studies were in the order 2205 (S32205) >316L (S31603) >904L (N08904) and 33 (R20033) >28 (N08028) >825 (N08825) >904L (N08904).[27,28]

Overall, contaminants do not appear to have any significant adverse effects on either the corrosion rates or the vulnerability to SCC of stainless steels in sodium hydroxide solutions. The oxidizing chlorates at the normal levels of contamination appear to be beneficial toward stainless steels in helping to promote and sustain passivity. There is no evidence that SCC risks are exacerbated in sodium hydroxide solutions by the presence of chlorides up to their maximum contaminant levels. However, if the sodium hydroxide is neutralized, the risks of chloride-induced SCC are reintroduced.

Although there is very little specific data in the public arena, the performances of stainless steels in potassium hydroxide are broadly similar to their equivalent performances in sodium hydroxide. A specific study of type 304L (S30403) austenitic stainless steel showed potassium hydroxide to be less aggressive than sodium hydroxide in promoting caustic-induced SCC.[29]

2.21.4 Nickel and Nickel Alloys

The compositions of the commercial nickel alloys that are relevant to alkali applications are summarized in **Table 2**. Nickel is second only to silver in its resistance to concentrated alkalis at elevated

Table 2 Compositions of some wrought nickel alloys that are relevant to alkali applications

UNS no	*Common name*	*Fe*	*Ni*	*Cr*	*Mo*	*Cu*	*Si*	*Other*
N02200	200		>99					<0.15C
N02201	201		>99					<0.02C
N04400	400	<2.5	Balance			28–34		Mn
N05500	500	<2.0	Balance			27–33		Al, Mn
N08825	825	>22	Balance	19.5–23.5	2.5–3.5	1.5–3.0		Ti
N06600	600	6–10	Balance	14–17				
N06690	690	7–11	Balance	27–31				
N10276	C-276	4–7	Balance	14.5–16.5	15–17			W, Co
N06022	C-22	2–3	Balance	20.5–22	13–14			W, Co
N06059	59		Balance	23	16			
N06686	686	<5	Balance	19–23	15–17			W
N06625	625	<5	Balance	20–23	8–10			Nb
N10665	B-2	<2	Balance		28			
N10675	B-3	1.5	Balance	1.5	28.5			W, Co

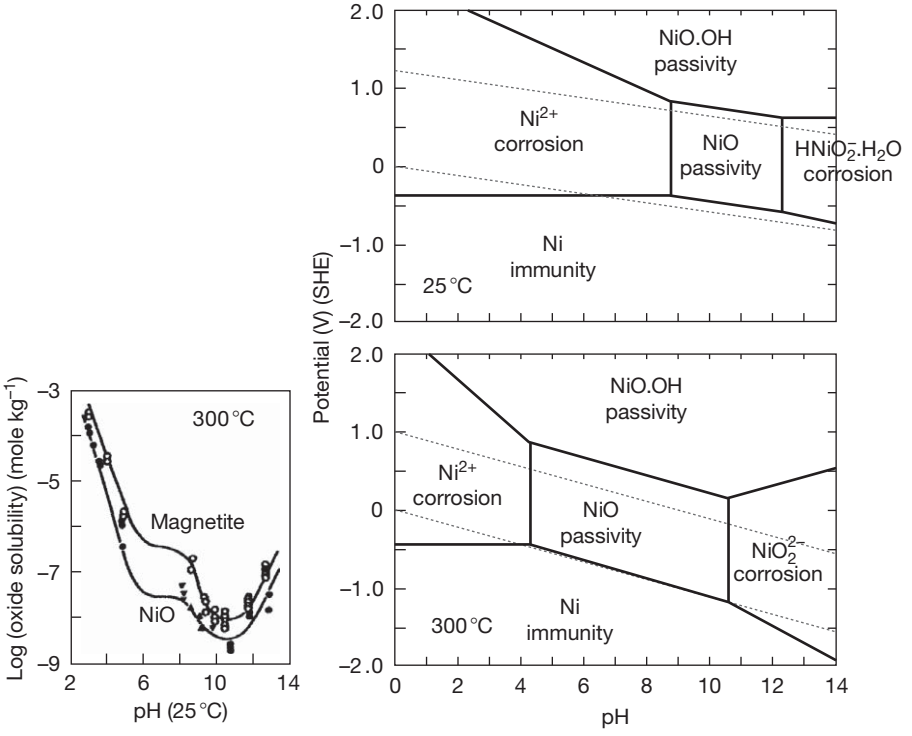


Figure 15 Simplified potential/pH diagrams for the Ni–H_2O system at 25 °C and 300 °C, based on an equilibrium activity of 10^{-5} molal, calculations performed using HSC Chemistry V6.12, Outotec Research, Finland. Oxide solubility data adapted from Turnbull, A.; Psaila-Dombrowski, M. *Corros. Sci* **1992**, *33*, 1925–1966.

temperatures where nickel and its alloys find significant commercial application. The temperature range up to ~300 °C is of particular significance because evaporator tubes that are used for the commercial manufacture of 73% (w/w) sodium hydroxide and steam generator tubes in pressurized water reactors (PWR) in the nuclear industry both operate at ~300 °C. The basis of the good performance of nickel and its alloys under these conditions is the formation and maintenance of highly insoluble, nickel oxide, NiO, films. The Pourbaix diagrams for the nickel–water system shown in Figure 15 confirm the existence of the NiO passivation zone and define its thermodynamic boundaries at 25 °C and 300 °C.

Despite the apparent threat to passivity due to the encroachment of corrosion zones based on the formations of soluble $HNiO_2^- \cdot H_2O$ and NiO_2^{2-} at 25 °C and 300 °C respectively, in practice a kinetic passivation zone extends much further into the alkaline pH region. The passive film forms much more rapidly than it dissolves because of the low solubility of NiO, exemplified by the comparative data for magnetite and NiO at 300 °C[30] that is also shown in **Figure 15**. In practice, both of the commercial purity grades of nickel, 200 (N02200) and 201 (N02201), have excellent resistance to sodium hydroxide across wide ranges of concentration and temperature, as shown in **Figure 16**, and find significant application in caustic evaporation equipment. The low carbon grade 201 is generally preferred for service above 300 °C to avoid problems associated with loss of ductility and intergranular attack due to the precipitation of graphite in the alloy.

However, the corrosion rates of commercially pure nickel are very sensitive to the presence and concentration of oxidizing chlorates and hypochlorites that have a significant influence on the life of caustic evaporation equipment. The test results in **Figure 17** are corrosion rates for a duplex stainless steel (S32906) and nickel 200 (N02200) in boiling 50% (w/w) sodium hydroxide solution with different levels of chloride and chlorate contaminants. The differing effects of oxidizing contaminants on stainless steels and nickel are apparent.

The chromium-free nickel alloys that find significant applications in sodium hydroxide, particularly in evaporation equipment, are the nickel copper alloys 400 (N04400) and 500 (N05500) that have equivalent corrosion resistance to sodium hydroxide up to ~73% (w/w) concentration. However, in keeping with nickel 200 (N02200) and 201 (N02201), their corrosion resistance is very sensitive to the presence of oxidizing agents such as chlorates and hypochlorites that have to be controlled below threshold levels to avoid excessive corrosion and contamination of the product sodium hydroxide with nickel and copper corrosion products.[5,31] The alloys 400 (N04400) and 500 (N05500) are also vulnerable to caustic-induced SCC and liquid metal embrittlement (LME) by mercury in the case of components that handle mercury cell caustic prior to final separation of mercury and its salts from the product caustic.[21] The chromium-free B family of alloys also has good resistance to sodium hydroxide in the absence of oxidizing agents but the advantages they offer over cheaper alternatives are marginal and they have found little application.

The preferred chromium-bearing alloy for sodium hydroxide applications at higher temperatures is alloy 600 (N06600) that has comparable, arguably better, resistance to corrosion than the commercially pure nickel alloys exemplified by the corrosion rates in boiling solutions for the two materials that are shown in **Figure 18**.[5]

Alloy 600 (N06600) has two additional advantages over commercially pure nickel. It is more resistant to

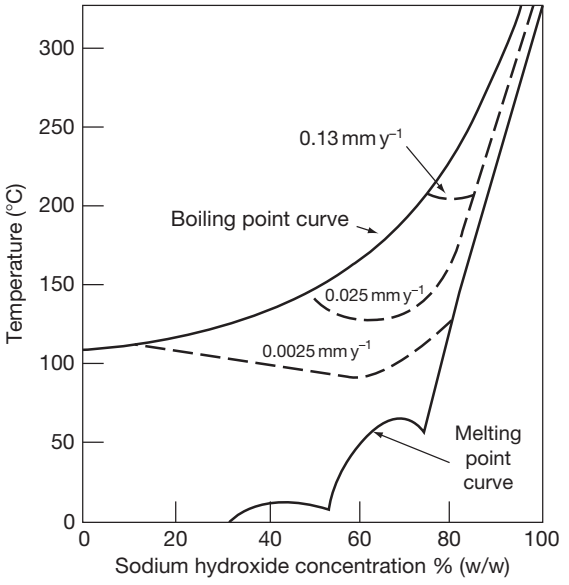


Figure 16 Isocorrosion diagram for nickel 200 and low carbon nickel 201 in sodium hydroxide. Adapted from Corrosion Resistance of Nickel and Nickel-Containing Alloys in Caustic Soda and Other Alkalis; CEB-2, International Nickel Company, 1973.

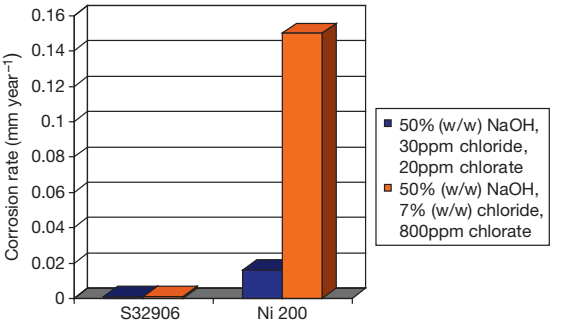


Figure 17 Corrosion rates for a duplex stainless steel and nickel 200 in boiling 50% (w/w) sodium hydroxide solution with different levels of chloride and chlorate contaminants. Adapted from Leander, D. Corrosion Characteristics of Different Stainless Steels, Austenitic and Duplex, in NaOH Environments Stainless Steel World Conference Maastricht, Stainless Steel World, 2003.

the presence of sulfur-containing contaminants in the sodium hydroxide that can promote intergranular penetration of nickel due to the formation of a low melting point nickel/nickel sulfide eutectic. It is also more tolerant to the presence of oxidizing agents such as chlorate and hypochlorite due to the passivation reinforcing effects of the chromium in the base metal. However, on the downside, the alloy 600 (N06600) is vulnerable to caustic-induced SCC at temperatures above ~190 °C in concentrated sodium hydroxide. In principle, the risk of SCC can be mitigated by appropriate thermal stress relief but this requires exposure to relatively high temperatures for significant time periods (for example, ~900 °C for 1 h) and is not always practicable for complex equipment items.

Few corrosion problems have been studied as intensively as the SCC vulnerability of alloy 600 (N06600) that has been used extensively as a steam generator tube material in PWRs in the nuclear industry. This is a complex topic that is described in more detail elsewhere[32] and in **Chapter 2.09, Stress Corrosion Cracking** in this volume. The research has included detailed investigations of the intergranular cracking (IGSCC) of alloy 600 (N06600) in concentrations of ~10% (w/w) sodium hydroxide at ~300 °C. The IGSCC occurs in a relatively narrow band of corrosion potentials close to the Ni/NiO equilibrium that are adopted by the alloy in the deaerated boiler water. Carbon content and carbide morphology play an important role. Intergranular carbides improve IGSCC resistance, and heat treatments for alloy 600 that develop preferred microstructures have been proposed and adopted. Operating experience has established that alloys 800 (N08800) and 690 (N06690) are much more resistant to IGSCC than alloy 600 (N06600). Recently, it has been proposed that internal, intergranular oxidation of chromium is the initiating process for IGSCC in alloy 600 (N06600). In contrast, alloy 690 (N06690) resists IGSCC because its much higher chromium content allows it to form a protective external oxide, and alloy 800 (N08800) has sufficient nickel to resist the dealloying processes that might render it vulnerable to the mechanism for caustic-induced SCC in stainless steels and sufficient chromium to form an external oxide and resist the internal oxidation route to crack initiation.[33,34]

The nickel–chromium–molybdenum alloys such as alloys 625 (N06625), 59 (N06059), and the C family do not find significant application in alkalis because, at the margin, they are less resistant than other nickel based alloys and in conditions where they have useful resistance, there are cheaper alternatives.

The corrosion rates shown in **Figure 19** for various nickel alloys in 70% (w/w) sodium hydroxide solution at 170 °C[35] confirm that the two nickel–chromium–molybdenum alloys are outperformed by the nickel–copper and nickel–chromium–iron alloys. The nickel–chromium–molybdenum alloys are also vulnerable to caustic-induced SCC.[36] Arguably, the nickel–chromium–molybdenum alloys might be more tolerant to the presence of oxidizing agents such as chlorates and hypochlorites, but this has to be offset against the vulnerability of molybdenum to transpassive dissolution in alkaline media. In practice,

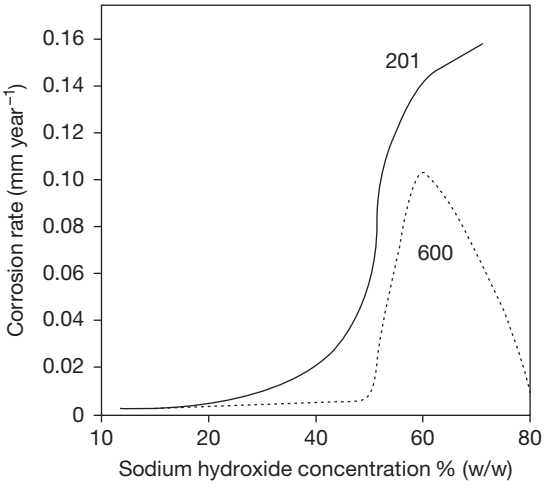


Figure 18 Corrosion rates of alloys 201 and 600 in boiling 50% (w/w) sodium hydroxide solution. Adapted from Davies, M. Corrosion by Alkalies. In *ASM Handbook: Corrosion: Environments and Industries*; ASM International: Materials Park, OH, 2006; Vol. 13C, pp 710–726.

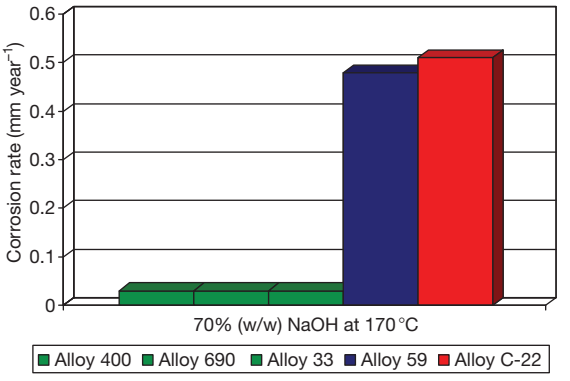


Figure 19 Corrosion rates for various nickel alloys in 70% (w/w) sodium hydroxide solution at 170 °C. Adapted from Heubner, U.; Kohler, M. *High Alloy Materials for Aggressive Environments*; ThyssenKrupp VDM GmbH 1998; VDM Report No 26.

the nickel–chromium–molybdenum alloys are specified only when their corrosion resistance to other fluids are required simultaneously, such as in heat exchange duties for alkaline fluids where the heating or cooling fluid is sufficiently corrosive as to require a nickel–chromium–molybdenum alloy.

Although there is very little specific data in the public arena, the performances of nickel alloys stainless steels in potassium hydroxide are broadly similar to their equivalent performances in sodium hydroxide.

2.21.5 Other Metals and Alloys

Commercial purity zirconium (Zr702, R60702) is protected by a surface oxide film, composed mainly of ZrO_2, that is extremely resistant to alkaline attack, as a result of which zirconium has excellent resistance to alkalis, both aqueous and fused. Thus, it resists 73% (w/w) sodium hydroxide up to 138 °C and molten sodium hydroxide at temperatures above 1000 °C. It is resistant to the presence of oxidizing contaminants unless significant amounts of chlorides are present. It is resistant to SCC in boiling solutions but is very vulnerable to LME if contaminant mercury is present.[5] More detailed information is given in **Chapter 3.14, Corrosion of Zirconium and its Alloys**.

Titanium and its alloys are also dependent on the formation of protective oxide films consisting mostly of titanium dioxide, TiO_2, for their resistances to corrosion. However, the TiO_2 film on titanium is more soluble in alkaline conditions than the equivalent ZrO_2 film on zirconium. In practice, commercially pure titanium (ASTM grade 2, R50400) is resistant over a wide range of concentrations at ambient temperatures but increases in temperature increase the film dissolution rate. Titanium's performance limits are determined as much by vulnerability to the hydriding caused by the hydrogen released in the corrosion process as by the corrosion rate *per se* that might otherwise be acceptable. Hydrogen absorption is particularly prevalent in the 20–40% (w/w) concentration range and at 50% (w/w) sodium hydroxide, the temperature has to be limited to ~80 °C to avoid excessive hydriding and embrittlement.[37]

Copper piping has been used for handling sodium hydroxide solutions in the absence of oxidizing agents and tolerates temperatures up to ~100 °C at 73% (w/w) concentration. The cupronickels (grades 90–10, C70600 and 70–30, C71500) have improved resistance and have found application in caustic evaporation up to 50% (w/w) concentration if some copper contamination is acceptable.[5] However, copper contamination is not acceptable in several major uses of the sodium hydroxide product and this inevitably limits the usage of copper and its alloys in alkaline applications.

Silver has exceptional resistance to sodium hydroxide at temperatures well beyond boiling but its cost precludes other than niche use. The amphoteric metals are clearly precluded from significant application, and tantalum has very poor resistance to caustic solutions and finds no application.

2.21.6 Nonmetallic Materials

2.21.6.1 Thermoplastic and Reinforced Thermosetting Materials

In practice, chemical compatibility is, but, one factor that determines the suitability or otherwise of a thermoplastic material for a specific application. Other factors such as mechanical and fabrication properties, thermal expansion characteristics, permeation properties, etc. are significant design considerations. In practice, other than for small scale equipment, thermoplastics are used more commonly as linings on stronger substrates rather than in solid form, because of reliability and integrity concerns. The application of these materials is covered in more detail in the relevant chapters in this book but the more significant materials are as follows:

1. Polyethylene (PE) and polypropopylene (PP) have good resistance. High density polyethylene (HDPE) and PP resist concentrations up to ~70% (w/w) at temperatures up to ~60 °C and ~100 °C respectively.
2. Polyvinyl chloride (PVC) resists concentrations up to ~70% (w/w) at temperatures up to ~40 °C, although PVC-lined fiber reinforced plastic (FRP) construction is generally preferred to solid PVC on safety grounds. Chlorinated polyvinyl chloride (CPVC) tolerates higher temperatures.
3. Fluoroplastics are very resistant to alkaline solutions, depending upon their degree of fluorination. The highly fluorinated materials polytetrafluoroethylene (PTFE), fluorinated ethylene propylene (FEP), and perfluoroalkoxy (PFA) resist concentrations up to ~70% (w/w) at temperatures up to 150 °C and beyond. The less highly fluorinated materials ethylene chlorotrifluoroethylene (ECTFE), ethylene tetrafluoroethylene (ETFE), and polyvinylidene fluoride (PVDF) are also very resistant but have lower temperature limits for service than the fully fluorinated grades.

FRP constructions can be used, but the common isophthalate and phenolic resins are not resistant to, and glass is attacked by, caustic solutions. FRP construction is therefore based on the use of specific resins such as epoxy, bisphenol-A fumarate, and vinyl ester resins and special surfacing veils of materials such as polyesters and acrylics. Dual laminate construction is also common, involving the use of resistant, thermoplastic liners supported by FRP. Materials selection needs care and experience, and suppliers should always be consulted in the selection of thermoplastic or thermosetting construction for caustic applications.

2.21.6.2 Elastomers

Natural rubber is resistant to concentrations up to 70% (w/w) at temperatures up to ~60 °C, and in the case of hard natural rubber up to temperatures as high as ~90 °C. Several synthetic rubbers, including chlorosulfonated PE (such as Hypalon), butyl, nitrile butadiene, and butadiene rubbers also have good resistance up to ~60 °C and beyond. Silicone and polysulfide rubbers are restricted to temperatures close to ambient.

2.21.6.3 Inorganic Materials

Impervious graphite finds application depending on the impregnant. Of the common impregnants, carbon tolerates up to 80% (w/w) concentration at temperatures up to ~80 °C. The PTFE impregnant is also tolerant to relatively demanding conditions, but phenolic resins are limited to much lower concentrations and are substituted with furane resins for more demanding applications.

Ceramics find only limited application. Borosilicate glass is very restricted in the concentrations and temperatures that it can handle, and glass linings are restricted essentially to dilute solutions at low temperatures.

References

1. Doche, M. L.; Rameau, J. J.; Duran, R.; Novel-Cattin, F. *Corros. Sci* **1999**, *41*, 805–826.
2. Materials Selector for Hazardous Chemicals. Ammonia and Caustic Soda; MTI Publication MS-6, Materials Technology Institute of the Process Industries Inc.: St Louis, MO, 2004; Vol. 6.
3. DECHEMA *Corrosion Handbook: Sodium Hydroxide, Mixed Acids*; Wiley-VCH: London, 2004; Vol. 1.
4. DECHEMA *Corrosion Handbook: Potassium Hydroxide, Ammonia, Ammonium Hydroxide*; Wiley-VCH: London, 2007; Vol. 9.
5. Davies, M. Corrosion by Alkalies. In *ASM Handbook: Corrosion: Environments and Industries*; ASM International: Materials Park, OH, 2006; Vol. 13C, pp 710–726.
6. Chen, C. M.; Aral, K.; Theus, G. J. Computer Calculated Potential-pH Diagrams to 300°C, Report NP-3137, Electric Power Research Institute, 1983.
7. NACE Standard Practice SP0403–2008. Avoiding Caustic Stress Corrosion Cracking of Carbon Steel Refinery Equipment and Piping; NACE International: Houston, TX, 2008.
8. Bandyopadhyay, N.; Briant, C. L. *Corrosion* **1985**, *41*, 274–280.
9. Poulson, B. S. Effect of Alloying Elements on the SCC of Ferritic Steels; Ph.D. Thesis, University of Newcastle-on Tyne, 1972.
10. Giddey, S.; Cherry, B.; Lawson, F.; Forsyth, M. *Corros. Sci* **2001**, *43*, 1497–1517.
11. *Betz Handbook of Industrial Water Conditioning*, 9th ed.; Betz Laboratories, 1991.
12. Lumsden, J. B.; Stocker, J. B.; Pollock, G. A.; Millet, P. J. *Corrosion'96, Heated Crevice Investigation of the Effects of Feedwater Chemistry on Crevice Chemistry and Potential*, Denver, CO; NACE International: Houston, TX, 1996; Paper 111.
13. Staehle, R. W.; Gorman, J. A. *Corrosion'96, Status and Issues in Corrosion on the Secondary Side of Steam Generators*, Denver, CO; NACE International: Houston, TX, 1996; Paper 122.
14. Wood, S. P.; Algeo, W. J.; Call, R. A.; Laronge, T. M. In *Corrosion 2007, Caustic Gouging in a Refinery Boiler – Root Cause and Remediation*, Nashville, TN; NACE International: Houston: TX, 2007; Paper 07451.
15. Bouzek, K.; Bergmann, H. *Corros. Sci.* **1999**, *41*, 2113–2128.
16. Corrosion Resistance of Nickel and Nickel-Containing Alloys in Caustic Soda and Other Alkalis; CEB-2, International Nickel Company, 1973.
17. Bhattacharya, A.; Preet, M. S. In *Corrosion 2008, Corrosion of Duplex Stainless Steels in High pH Caustic Solution*, New Orleans, LA; NACE International: Houston, TX, 2008; Paper 08194.
18. Addari, D.; Elsener, B.; Rossi, A. *Electrochim. Acta* **2008**, *53*, 8078–8086.
19. Deacon, J.; Dong, Z.; Lynch, B.; Newman, R. C. *Corros. Sci.* **2004**, *46*, 2117–2133.
20. Deverell, H. E.; Franson, I. A *Mater. Perform* **1989**, *28*, 52–57.
21. Schillmoller, C. M. NiDI Technical Report 10019: Alloy Selection for Caustic Soda Service Nickel Development Institute, 1988.
22. Horn, E.M. *Corrosion and Environmental Degradation*; Schutze, M., Ed.; Wiley-VCH: London, 2000; Vol. 2, pp 69–111.
23. Leander, D. Corrosion Characteristics of Different Stainless Steels, Austenitic and Duplex, in NaOH Environments Stainless Steel World Conference Maastricht, Stainless Steel World, 2003.
24. Cassagne, T.; Combrade, P. *Innovation Stainless Steels* **1993**, *3*, 215–220.
25. Yasuda, M.; Fukumoto, K.; Koizumi, H.; Ogata, Y.; Hine, F. *Corrosion* **1987**, *43*, 492–498.
26. Paul, L.; Alves, H. *Corrosion 2004, Alloy 33: A Versatile Alloy for Mineral Acid and Other Applications*, New Orleans, LA; NACE International: Houston, TX, 2004; Paper 04226.
27. Rondelli, G.; Vincenti, B.; Sivieri, E. *Corros. Sci.* **1997**, *39*, 1037–1049.

28. Rondelli, G.; Vincenti, B. *Mater. Corros.* **2002**, *53*, 813–819.
29. Dean, S.; Abayarathna, D.; Kane, R. D. *Corrosion'98, Stress Corrosion Cracking of 304L Stainless Steel and 1.25Cr 0.5Mo in Caustic Environments in Syn Gas Service*, San Diego, CA; NACE International: Houston, TX, 1998; Paper 590.
30. Turnbull, A.; Psaila-Dombrowski, M. *Corros. Sci.* **1992**, *33*, 1925–1966.
31. Yasuda, M.; Takeya, F.; Hine, F. *Corrosion* **1983**, *39*, 10.
32. Staehle, R. W.; Gorman, J. A. *Corrosion* **2004**, *60*, 5–63.
33. Newman, R. C.; Gendron, T. S.; Scott, P. M. *Proceedings of the Ninth International Symposium on Environmental Degradation of Materials in Nuclear Power Systems – Water Reactors*, Warrendale; TMS-AIME: Warrendale, PA, 2000; pp 79–93.
34. Coull, Z. L.; Newman, R. C. *Corrosion 2007, Selective Dissolution and Oxidation Zones in Ni-Cr-Fe Space and Their Relationship to Stress Corrosion Cracking*, Nashville, TN; NACE International: Houston, TX, 2007; Paper 07487.
35. Heubner, U.; Kohler, M. *High Alloy Materials for Aggressive Environments*; ThyssenKrupp VDM GmbH 1998; VDM Report No 26.
36. Rebak, R. B. In *Corrosion 2005, Environmentally Assisted Cracking of Commercial Ni-Cr-Mo Alloys: A Review*, Houston, TX; NACE International: Houston, TX, 2005; Paper 05457.
37. Popa, M. V.; Vasilescu, E.; Mirza-Rosca, I.; Gonzalez, S.; Llorente, M. L.; Drob, P.; Anghel, M. *Eurocorr'97,* Trondheim: Norway, 1997; Vol. 2, pp 687–692.

2.22 Corrosion in Hydrogen Halides and Hydrohalic Acids

J. A. Richardson
Anticorrosion Consulting, 5 Redhills Lane, Durham DH1 4AL, UK

Abbreviations

AHF Anhydrous hydrogen fluoride
AES Auger electron spectroscopy
ASTM American Society for Testing Materials
CPVC Chlorinated polyvinyl chloride
ECTFE Ethylene chlorotrifluoroethylene
EIS Electrochemical impedance spectroscopy
EPDM Ethylene propylene diene terpolymer
ETFE Ethylene tetrafluoroethylene
FEP Fluorinated ethylene propylene
GRP Glass reinforced plastic
HAZ Heat affected zone
HIC Hydrogen-induced cracking
HSC Hydrogen stress cracking
PE Polyethylene
PFA Perfluoroalkoxy
PP Polypropylene
ppm Parts per million
PTFE Polytetrafluoroethylene
PVC Polyvinyl chloride
PVDF Polyvinylidene fluoride
PWHT Postweld heat treatment
SCC Stress corrosion cracking
SCE Saturated calomel electrode
SHE Standard hydrogen electrode
SIMS Secondary ion mass spectrometry
SOHIC Stress-oriented hydrogen-induced cracking
TOF Time of flight
XPS X-ray photoelectron spectroscopy

Symbols

E_{corr} Corrosion potential (V)
i_{corr} Corrosion current density (mA cm^{-2})

2.22.1 Introduction

The hydrohalic acids comprise the group of monobasic acids that are formed when the hydrogen halide gases are dissolved in water, with strengths in descending order:

hydroiodic (HI) > hydrobromic (HBr) > hydrochloric (HCl) > hydrofluoric (HF)

The differences in strength arise from the differing sizes of the relevant halide ion. Thus, the largest iodide anion interacts relatively weakly with its complementary proton, leaving it relatively free to dissociate. In practice, the differences amongst HI, HBr, and HCl acids are not great, and all are strong acids that are highly dissociated in water at all concentrations to

release protons and the corresponding halide anion as follows:

$$HX = H^+ + X^-$$

The result, in relation to the pHs of hydrohalic acids, is that at the higher end of the pH range the solutions are effectively acidified waters. Thus, the pHs 2, 3, and 4 are equivalent to concentrations of 360, 36, and 3.6 ppm (w/w) hydrochloric acid, respectively at 25 °C. However, HF acid is an exception to this behavior. There is very strong hydrogen bonding between the nonionized HF and water molecules, and the small fluoride anion imposes a high degree of order on the surrounding protons and water molecules as a result of which HF acid is a very weak acid, the pH of which varies in a complicated manner with concentration as shown in Figure 1.[1]

Hydrogen chloride, bromide, and iodide are all gases at ambient temperatures and available commercially in nominally anhydrous specifications, containing trace quantities of water. Anhydrous hydrogen fluoride (AHF) is a liquid that boils at ~20 °C, and is normally specified as containing < 400 ppm water. The significant commercial grades of the acids are as follows:

1. HI acid, the least significant acid commercially, is supplied at various concentrations, typically 47% or 55% (w/w), but concentrated grades in the range 90–98% (w/w) are also available.
2. Hydrobromic acid is manufactured in significant quantities and supplied at various concentrations, most commonly at ~47–49% (w/w).
3. Hydrochloric acid is manufactured in the greatest quantities at concentrations up to ~38% (w/w), above which the fuming nature of the acid introduces problems in relation to storage and transport. Most bulk commercial grades have concentrations in the range 30–35% (w/w).
4. HF acid is also significant commercially, and is available in bulk at concentrations of 49% and 70% (w/w).

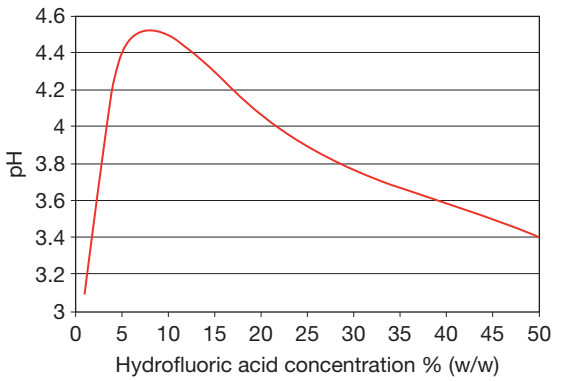


Figure 1 pH of hydrofluoric acid solutions. Adapted from *Hydrofluoric Acid Properties*; Honeywell Specialty Materials Products, 2002.

From the corrosion standpoint, the hydrohalic acids are commonly classified as inherently 'reducing' acids because the only cathodic process that the pure acids can deliver is the reduction of protons to evolve hydrogen:

$$2H^+ + 2e^- = H_2$$

This is in contrast to 'oxidizing' acids such as sulfuric and nitric acid in which reductions of the acid itself, or its constituent species, occur at potentials more noble than hydrogen evolution, and to an extent that can determine the behaviors of materials exposed to the acid, as described in the separate chapters devoted to the two acids in this book. In the case of the hydrohalic acids, their 'reducing' characteristics are compounded by the aggressive properties of the halide anions that inhibit the formation of, and attack preexisting, protective passive layers on metals and alloys. Indeed, both HCl and HF acids find significant commercial applications in the cleaning and pickling of metals and alloys because of their ability to dissolve metal oxides.

In practice, alternative cathodic processes to hydrogen evolution may be available due to the presence of specific contaminants in the acid, in particular oxidants such as dissolved oxygen and metal cations in a higher oxidation state such as ferric, Fe^{3+} or cupric, Cu^{2+} ions. The presence of the corresponding halogen gas, for example chlorine in the case of hydrochloric acid, can also have a significant effect. The effects of oxidizing agents that raise the potentials of metals and alloys are almost invariably detrimental to corrosion performance in the hydrohalic acids, in which metals and alloys are mostly unable or struggle to form protective, passive films, as will emerge below.

The corrosion performances of the various classes of material in hydrohalic acids have been reviewed extensively elsewhere.[2–8] Much of the data relate to hydrochloric and HF acids, reflecting their commercial significance. There is relatively little data for hydrobromic, and virtually none for HI acid in the public domain. Readers whose main interest is to identify 'what works where' are referred to these sources and the relevant chapters on specific materials in this book. In this chapter, the corrosion

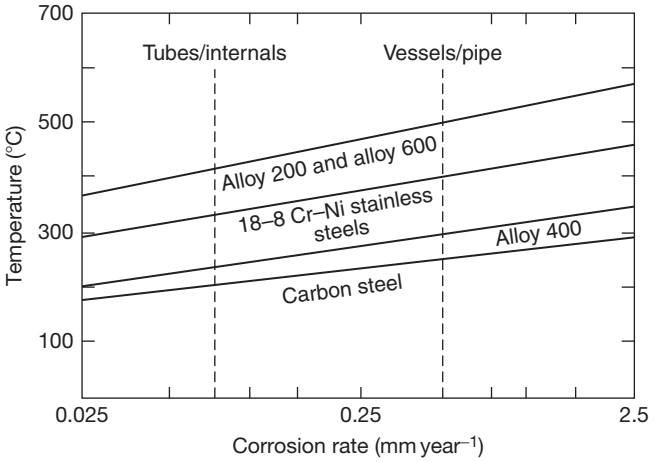


Figure 2 Proposed design limits for various alloys in anhydrous hydrogen chloride. Adapted from Schillmoller, C. M. *Chem. Eng*. **1980**, *87*, 161–163.

performances of materials are reviewed with an emphasis on the principles and mechanisms that underpin their corrosion performances in the acids.

2.22.2 Steels and Cast Irons

Steels react with anhydrous hydrogen halide gases to form the corresponding iron halide and hydrogen. In AHF, the fluoride films are protective and steels have acceptable corrosion rates at temperatures up to ~200 °C and velocities up to 10 m s^{-1}, and arguably to higher temperatures at lower velocities.[3,6] The velocity limitations arise because the fluoride films are vulnerable to detachment which not only increases the corrosion rates, but also adversely affects the operability of equipment due to accumulations of solid iron fluoride in seals and joints in valves, pumps, and elsewhere. In anhydrous hydrogen chloride gas, the growth of a film of $FeCl_2$ obeys parabolic kinetics up to temperatures of 500–600 °C, above which the vaporization of $FeCl_2$ results in a switch to linear, nonprotective kinetics. As in the case of the analogous fluoride films, the chloride films on steels are protective within specific temperature and velocity constraints,[5,9,10] and some widely used design temperature limits for carbon steel, austenitic stainless steels, and specific nickel alloys in anhydrous hydrogen chloride are shown in **Figure 2**.[10]

Aqueous hydrohalic acids are very aggressive toward steels and grey or ductile cast irons. A typical polarization curve for carbon steel in stagnant, aqueous 3M (~10%, w/w) hydrochloric acid is shown in **Figure 3**.

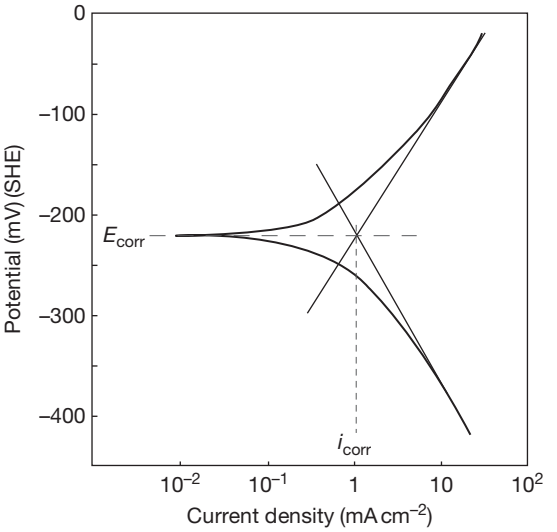


Figure 3 Potentiodynamic polarization curve for carbon steel in stagnant, 3 M (~10%, w/w) hydrochloric acid open to air at ambient temperature. Adapted from Poorqasemi, E.; Abootalemi, O.; Peikari, M.; Haqdar, F.*Corros. Sci* **2009**, *51*, 1043–1054.

The anodic curve shows classic Tafel behavior, characteristic of active dissolution of iron:

$$Fe = Fe^{2+} + 2e^{-}$$

The cathodic curve also shows classic Tafel behavior, characteristic of hydrogen evolution. In more dilute acids, curves are obtained that are time dependent to a small extent, and electrochemical impedance spectroscopic (EIS) measurements suggest the presence of highly nonprotective films that present limited

barriers to anodic dissolution.[11] Such films that do exist are likely to consist of iron chloride, which is very soluble in hydrochloric acid and the degree of protection that they can provide is therefore very marginal. The general experience of steels in aqueous HCl, HBr, and HI acids is of active corrosion across the full concentration range at rates that preclude practical application.

In aqueous HF acid, corrosion rates are generally lower, but acceptable only at concentrations above ~70% (w/w), as shown in **Figure 4**. The corrosion resistance of steels in concentrated aqueous and liquid AHF arises from the spontaneous formation of protective iron fluoride films on steel surfaces that are much less soluble than the corresponding chloride films. Studies of the growth of fluoride films in vapors over dilute HF acids at ambient temperature have shown that the kinetics is linear, and controlled by reaction at the metal–film interface, resulting in the formation of porous, nonprotective films. However, when the HF acid concentration rises to 40% (w/w), the initial film that forms in the vapor phase has been identified as $FeF_2 \cdot 4H_2O$, and it grows with a parabolic dependence on time. With increasing film thickness, a product identified as nonstoichiometric $Fe_2F_5 \cdot 7H_2O$ occurs. The change to parabolic kinetics indicates a switch to diffusion control of reaction rate through a thickening, potentially protective film.[12]

The significance of such films is evident in the polarization curves for two steels in 90% (w/w) acid at 90 °C, shown in **Figure 5**.

In both cases, the anodic kinetics is clearly influenced strongly by diffusion due to the presence of the fluoride film, while the cathodic reactions obey Tafel-type kinetics, characteristic of hydrogen evolution beneath the film.[13] The adherence and protection afforded by the fluoride film has been shown to depend upon several factors. Although the laboratory test in **Figure 5** demonstrates the existence of a film at 90 °C, temperatures and limiting velocities are typically restricted to ~65 °C and ~1.6 m s^{-1} in liquid AHF, reducing to ~30 °C and 0.6 m s^{-1} in 70% (w/w) acid.[6] The carbon and residual element content of the steel, in particular the copper and nickel contents, also appear to influence the persistence of the fluoride films that form and the corresponding corrosion rates that are experienced,[13] and recommendations have been formulated relating to the control of such elements.[14] Corrosion rates have also been shown to be higher in the presence of oxygen in both the laboratory[10] and in the field.[6]

Steels are vulnerable to damage due to hydrogen uptake in HF acid service that is similar to the types of damage that can be experienced in wet H_2S such as hydrogen stress cracking (HSC) of stronger materials and hard welds and heat affected zones (HAZ), and hydrogen-induced cracking (HIC), stress-oriented hydrogen-induced cracking (SOHIC), and blistering of plate materials. General approaches to mitigating these risks are available[7,15] based on the

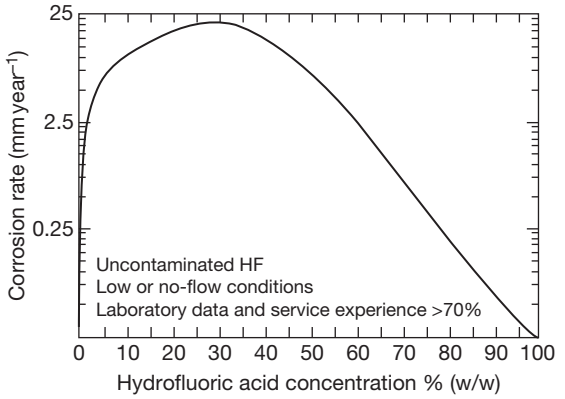


Figure 4 Corrosion rate of carbon steel in hydrofluoric acid at 21–38 °C as a function of concentration. Adapted from NACE International Publication 5A–171 *Materials for Storing and Handling Commercial Grades of Aqueous Hydrofluoric Acid and Anhydrous Hydrogen Fluoride*; NACE International: Houston, TX, 2007.

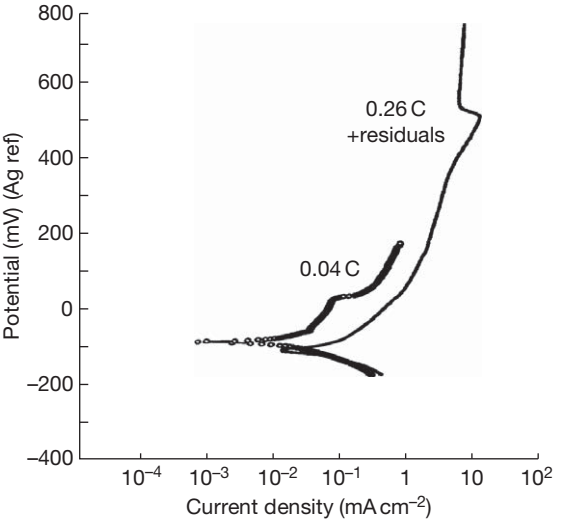


Figure 5 Potentiodynamic polarization curves for steels in 90% (w/w) hydrofluoric acid at 90 °C. Adapted from Chirinos, G.; Turgoose, S.; Newman, R. C. *Effects of Residual Elements on the Corrosion of Steels in HF*, Paper 97513, Corrosion '97; NACE International: Houston, TX, 1997.

practices that have been developed for the control of the similar problems that are experienced in wet H_2S, including the control of base metal chemistry and the hardness of welds and HAZs within threshold levels and, if appropriate, postweld heat treatment (PWHT).[16] The risks of HIC and SOHIC in the base materials are mitigated by controlling steel chemistry and microstructure (including limits on sulfur content and inclusion shape control) and manufacture (in particular rolling conditions), and confirming resistance by appropriate testing.[17–19]

Gray and ductile cast irons have corrosion resistances broadly similar to steels in hydrohalic acids, but find no application in HF acid service, not least in the case of gray irons because of safety concerns relating to their poor ductility. High-silicon cast irons, containing typically 14–16% Si, are resistant to all concentrations of hydrochloric acid up to 40% (w/w) at ambient temperatures and at higher temperatures in more dilute acids, as shown in the isocorrosion curves for some metals and alloys with exceptional corrosion resistance to aqueous hydrochloric acid, described later in **Figure 14**. Their exceptional corrosion resistance is due to the formation of robust, siliceous films that have considerable resistance to erosion and abrasion. However, they find relatively limited application due to their poor ductility that renders them difficult to fabricate and requires them to be protected from thermal and/or mechanical shock.

2.22.3 Stainless Steels

The compositions of the relevant commercial grades of stainless steels are summarized in **Table 1**.

In anhydrous hydrogen chloride gas, stainless steels perform in a similar manner to steels in that they form chloride films that obey parabolic kinetics up to temperatures of 500–600 °C, above which the vaporization of the chlorides results in a switch to linear kinetics. However, the films contain $CrCl_3$ and $NiCl_2$ in addition to, or instead of $FeCl_2$, depending on the alloy and the temperature.[10] They form more slowly and are much more protective than the corresponding films on steels at temperatures below ~500 °C. As a result, the corrosion rates for stainless steels are lower than for steels, and the design temperature limits are correspondingly higher, exemplified by those reported in **Figure 2**.[10] At temperatures above ~500 °C, stainless steels offer no significant advantage over carbon or low alloy steels in anhydrous hydrogen chloride gas.

The performances of austenitic stainless steels in AHF liquid and vapor are good up to a temperature of ~100 °C. All of the basic grades 304 (S30400), 304L (S30403), 316 (S31600) and 316L (S31603), and their cast equivalents show good resistance and find significant application, although due caution to prevailing velocities is necessary as in the case of steels. However, austenitic stainless steels that can develop α-martensite as a result of cold working are vulnerable

Table 1 Compositions of some wrought stainless steels that are relevant to hydrohalic acid applications

UNS no.	*Common name*	*Fe*	*Ni*	*Cr*	*Mo*	*Cu*	*Si*	*Other*
S31803	2205	Balance	4.5–6.5	21–23	2.5–3.5			N
S32205								
S32750	2507	Balance	6–8	24–26	3–4.5	<0.5		N
S32760					3–4	0.5–1.0		W, N
S30400	304	Balance	8–10.5	18–20				<0.08C
S30403	304L	Balance	8–12	18–20				
S34700	347	Balance	9–13	17–19				Nb/Ta < 0.08C
S31600	316	Balance	10–14	16–18	2–3			<0.08C
S31603	316L	Balance	10–14	16–18	2–3			
S31700	317	Balance	11–15	18–20	3–4			<0.08C
S31703	317L	Balance	11–15	18–20	3–4			
S31254	254SMO	Balance	17.5–18.5	19.5–20.5	6–6.5	0.5–1.0		N
S30900	309	Balance	12–15	22–24				Mn
S31000	310	Balance	19–22	24–26				Mn
N08904	904L	Balance	23–28	19–23	4–5	1–2		
N08926	25–6MO	Balance	24–26	19–21	6–7	0.5–1.5		N
S31277	27–7MO	Balance	26–28	20.5–23.0	6.5–8.0	0.5–1.5		Mn, N
N08020	20	Balance	32–35	19–21	2–3	3–4		
N08031	31	Balance	30–32	26–28	6–7	1.4		N
R20033	33	Balance	30–33	31–35	0.5–2.0	0.3–1.2		N

to HSC, and some users favor the higher nickel 316/316L grades for this reason.[8] Above ~100 °C, the rates of corrosion of austenitic stainless steels in AHF increase to unacceptable levels and they are outperformed by carbon steel, as shown in **Figure 6**.

The performances of stainless steels in aqueous hydrohalic acids are generally poor because they are unable to form and sustain the passive, chromium oxide films upon which they depend for corrosion resistance and, as a result, corrode actively. Ferritic and martensitic stainless steels have poor resistance under all conditions. Austenitic and duplex stainless steels offer some limited resistance, as shown in the isocorrosion diagrams for various grades in HCl and HF acids in **Figures 7** and **8** respectively.

The basic 316L (S31600) grade is clearly limited to very dilute acids at relatively low temperatures. The grades with higher levels of nickel, molybdenum, and copper offer progressive extension of the boundaries for stable passivity, but more than 30% nickel and significant contents of molybdenum and copper are required to provide resistance to hydrochloric acid at all concentrations up to 40% (w/w) at close to ambient temperatures, exemplified by the isocorrosion diagram for alloy 31 (N08031) which is shown in **Figure 9**.[20]

Overall, taking into account the additional risks of localized corrosion and stress corrosion cracking (SCC) presented by the halide anions, stainless steels are weak options for aqueous, hydrohalic acid service.

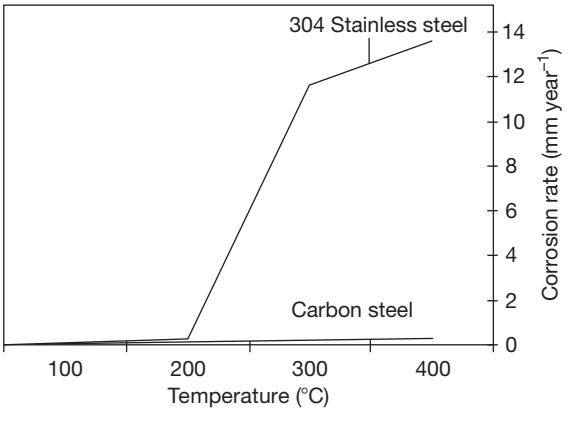


Figure 6 Corrosion rates of carbon steel and type 304 stainless steel as a function of temperature in static, anhydrous hydrofluoric acid. Adapted from Jennings, H. S. Corrosion by Hydrogen Fluoride and Hydrofluoric Acid. In *ASM Handbook Corrosion: Environments and Industries*; ASM International, 2006; pp 690–703, Vol. 13C.

2.22.4 Nickel and Nickel Alloys

The compositions of the relevant commercial grades of nickel and its alloys are summarized in **Table 2**.

The reaction of nickel with anhydrous hydrogen chloride gas as a function of temperature is shown in **Figure 10**.[21]

Evidently, reaction kinetics is parabolic up to a temperature of ~450 °C, associated with the formation of protective $NiCl_2$. In the temperature range of 450–550 °C, there is a transition to linear kinetics associated with the relatively high vapor pressure and progressive evaporation of $NiCl_2$. Alloying, in general, reduces the temperature at which there is a transition to unacceptable evaporation rates. The effects are relatively small for the commercial nickel–chromium

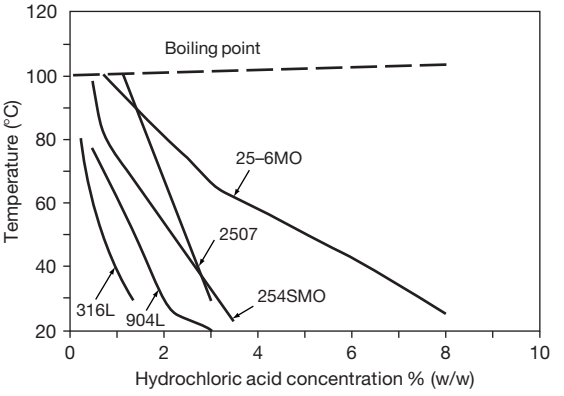


Figure 7 Isocorrosion charts (0.1 mm year^{-1}) for austenitic and duplex stainless steels in hydrochloric acid. Adapted from *Corrosion Handbook, Stainless Steels*; AB Sandvik Steel, 1994.

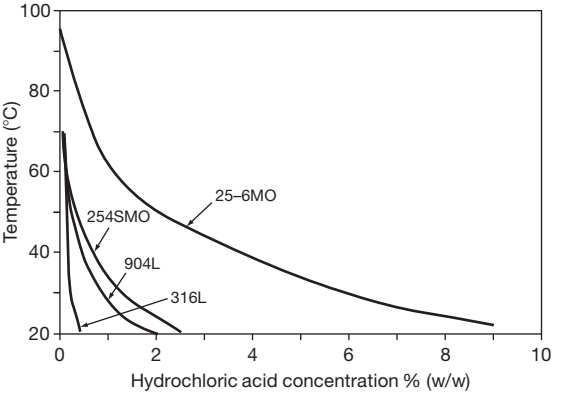


Figure 8 Isocorrosion charts (0.1 mm year^{-1}) for austenitic stainless steels in hydrofluoric acid. Adapted from *Corrosion Handbook, Stainless Steels*; AB Sandvik Steel, 1994.

and nickel–chromium–molybdenum alloys, but more significant for nickel–molybdenum and particularly nickel–copper alloys because of the relatively low melting point and high volatility of copper chloride. This is reflected in the relatively high, proposed, design temperature limits proposed in Figure 2 for nickel 200 (N02200) and alloy 600 (N06600) compared with alloy 400 (N04400) for which the proposed limits are little better than those for carbon steel.[10]

The resistance of nickel and its alloys to AHF gas at elevated temperatures is dependent on the formation of protective metal fluoride films. Nickel and molybdenum fluorides are stable and protective relative to chromium and iron fluorides that have lower melting points and higher volatilities, and this is reflected in the corrosion rates for various metals and alloys in AHF at 600 °C that are charted in Figure 11[22] which links to the data reported in Figure 6.

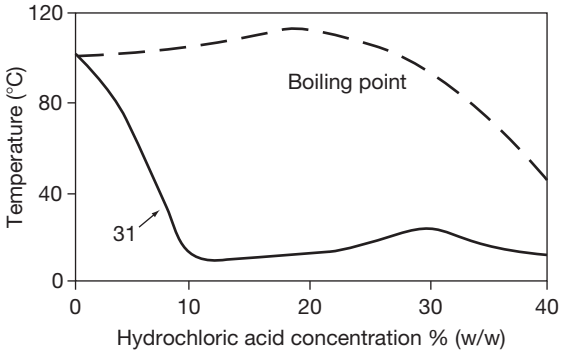


Figure 9 Isocorrosion chart (0.13 mm year^{-1}) for austenitic alloy 31 in hydrochloric acid. Adapted from Agarwal, D. C., Alves, H. *Applications of Alloys 59 (UNS N06059) and 31 (UNS N08031) in Mitigating Corrosion Risks in the CPI and Petrochemical Industries*, Paper 07186, Corrosion 2007; NACE International: Houston, TX, 2007.

Alloys can suffer internal as well as external attack in high temperature hydrogen fluoride related to fluoride diffusivity and solubility in the metal and, in addition to chromium and iron, the elements niobium, tantalum, and titanium also exacerbate the vulnerability of an alloy to internal attack due to the lower melting points and higher volatilities of their fluorides. Figure 11 confirms that the chromium-rich types 309 (S30900) and 310 (S31000) and the niobium/tantalum-containing type 347 (S34700) stainless steel suffer extremely high rates of corrosion. Type 304 (S30400) stainless steel is confirmed as having unacceptable rates relative to steel, and the most resistant materials are nickel 200 (N02200) and nickel alloys 400 (N04400) and 600 (N06600).

In AHF liquid, the leaner nickel alloys such as alloys 20 (N08020), 825 (N08825), and the G family are highly resistant to corrosion at temperatures up to ~125 °C, even under flowing conditions, and find significant application, subject to appropriate mitigation of risks relating to preferential weld corrosion and HSC of heavily cold-worked material.[6] The more

Table 2 Compositions of some wrought nickel alloys that are relevant to hydrohalic acid applications

UNS no.	*Common name*	*Fe*	*Ni*	*Cr*	*Mo*	*Cu*	*Si*	*Other*
N02200	200		>99					<0.15C
N02201	201		>99					<0.02C
N04400	400	<2.5	Balance			28–34		Mn
N05500	500	<2.0	Balance			27–33		Al, Mn
N08825	825	>22	Balance	19.5–23.5	2.5–3.5	1.5–3.0		Ti
N06007	G	18–21	Balance	21–23.5	5.5–7.5	1.5–2.5		Co, Nb
N06985	G-3	18–21	Balance	21–23.5	6–8	1.5–2.5		Co W
N06030	G30	13–17	Balance	28.0–31.5	4–6	1.0–2.4		W
N06600	600	6–10	Balance	14–17				
N06690	690	7–11	Balance	27–31				
N10276	C-276	4–7	Balance	14.5–16.5	15–17			W, Co
N06455	C-4	3	Balance	14–18	14–17			Co
N06022	C-22	2–3	Balance	20.5–22	13–14			W, Co
N06200	C-2000		Balance	23	16	2		
N06059	59		Balance	23	16			
N06686	686	<5	Balance	19–23	15–17			W
N06625	625	<5	Balance	20–23	8–10			Nb
N10001	B	5.5	Balance		28			Co
N10665	B-2	<2	Balance		28			
N10675	B-3	1.5	Balance	1.5	28.5			W, Co
	Hybrid BC-1	2	Balance	15	22			

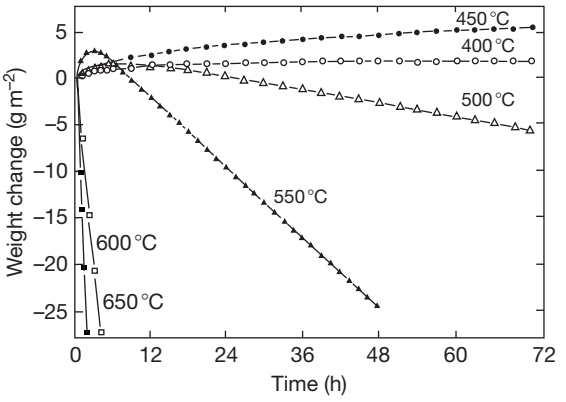


Figure 10 Weight change of nickel exposed to anhydrous hydrogen chloride gas as a function of temperature. Adapted from Ihara, Y.; Ohgame, S.; Sakiyama, K.; Hashimoto, K. *Corros. Sci*. **1982**, *22*, 901.

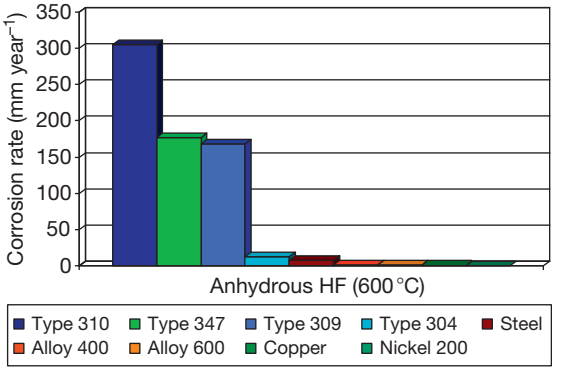


Figure 11 Corrosion rates for various metals and alloys tested for 4–15 h in anhydrous hydrogen fluoride (AHF) at 600 °C. Adapted from Myers, W. R.; Delong, W. B. *Chem. Eng. Prog.* **1948**, *44*, 359.

nickel-rich alloys 600 (N06600), C-22 (N06022), and C-276 (N10276) are even more resistant, as illustrated in **Figure 12** that shows corrosion rates for various metals and alloys tested for 6–40 days in AHF at 15–40 °C.[6]

The nickel–chromium and nickel–chromium–molybdenum alloys are used at temperatures up to at least ~150 °C, but appropriate precautions are also necessary with the nickel-rich alloys to avoid environmental cracking. Thus, the nickel–copper alloys 400 (N04400) and 500 (N05500) have excellent general corrosion resistance but are vulnerable to environmental cracking in the presence of air. Alloy 600 (N06600) has shown a vulnerability to cracking in the presence of chlorine and copper fluoride after cold forming.

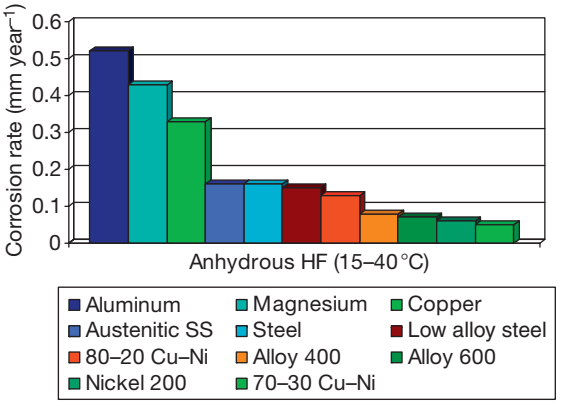


Figure 12 Corrosion rates for various metals and alloys tested for 6–40 days in AHF at 15–40 °C. Adapted from Jennings, H. S. Corrosion by Hydrogen Fluoride and Hydrofluoric Acid. In *ASM Handbook. Corrosion: Environments and Industries*; ASM International, 2006; pp 690–703, Vol. 13C.

Such risks can be mitigated by appropriate heat treatment procedures.[6]

The basic electrochemistry of the classes of alloys of commercial significance for aqueous hydrohalic acid duties is illustrated by the polarization curves in **Figure 13** for nickel and three nickel alloys in aqueous hydrochloric acid.

Evidently, the corrosion potentials and polarization characteristics of nickel 200 (N02200) and the nickel–copper alloy 400 (N04400) in deaerated 5N (~18%, w/w) hydrochloric acid at room temperature are very similar.[23] Both display very active corrosion potentials and anodic polarization behavior characteristic of active dissolution of the base elements with no tendency to passive film formation:

$$Ni = Ni^{2+} + 2e^-$$
$$Cu = Cu^{2+} + 2e^-$$

The polarization curves in **Figure 13** and **Figure 3** confirm the much greater resistance to active, anodic dissolution that nickel displays relative to iron under similar conditions. That is the basis of the generally better performances of nickel-based compared with iron-based alloys in hydrochloric acid. Regarding the specific effects of alloying with copper, the curves in **Figure 13** suggest slightly more polarization of both anodic and cathodic processes in the case of alloy 400 (N04400) than for nickel 200 (N02200), but this depends very much on acid concentration and temperature. Alloy 400 (N04400) is more sensitive to increases in concentration and less sensitive to increases in temperature than nickel 200. In practice, both alloys

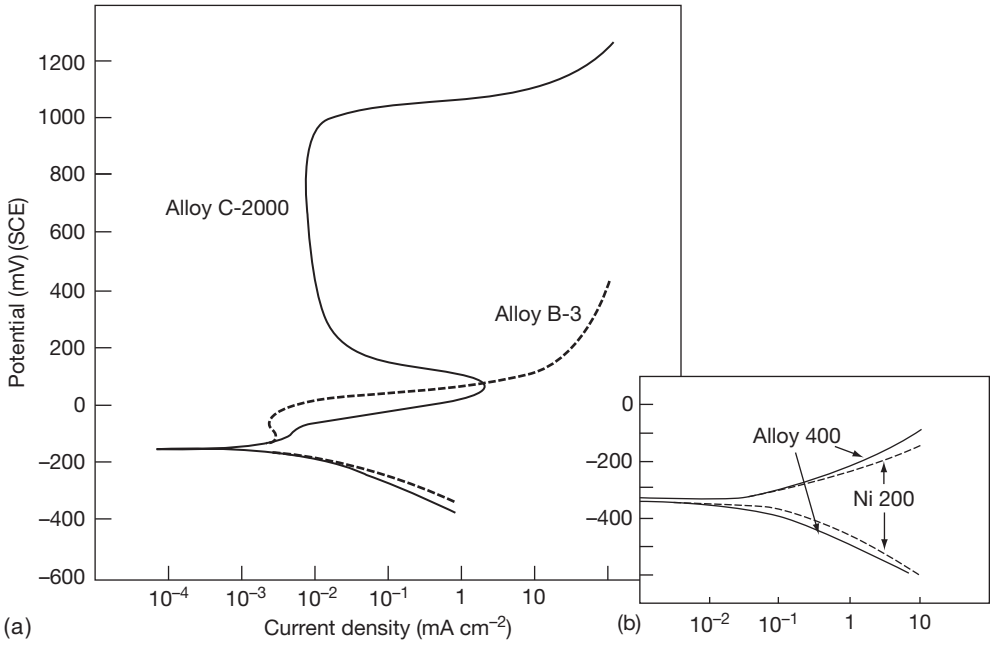


Figure 13 Polarization behavior of (a) alloys B-3 and C-2000 in deaerated 20% (w/w) hydrochloric acid at 25 °C. Adapted from Nacera, S. M.; Crook, P.; Klarstrom, D. L.; Rebak, R. B. *Effect of Ferric Ions on the Corrosion Performance of Nickel Alloys in Hydrochloric Acid Solutions*, Paper 04430, Corrosion 2004: NACE.

have acceptable resistance to corrosion in relatively dilute (<∼10%, w/w) acids at ambient temperatures but at higher concentrations and temperatures, and particularly if oxygen or other oxidizing agents are present, their application is limited.

The more important alloying elements for resistance to aqueous hydrochloric acid are molybdenum and chromium, exemplified by the polarization curves in **Figure 13** for the nickel–molybdenum alloy B-3 (N10675) and the nickel–chromium–molybdenum alloy C-2000 (N06200) in deaerated 20% (w/w) hydrochloric acid at 25 °C.[24] Both alloys exhibit relatively active corrosion potentials, albeit more noble than for nickel 200 (N02200) or the nickel–copper alloy 400 (N04400) under similar conditions. Active dissolution is clearly inhibited in the nickel–molybdenum alloy B-3 (N10675) relative to nickel 200 (N02200), arising from the significant contribution of molybdenum to overall dissolution:

$$Mo = Mo^{3+} + 3e^{-}$$

The mechanism via which molybdenum inhibits dissolution has been variously attributed to the relatively high Mo–Mo bond energy that blocks active sites[25,26] and the promotion of salt film formation due to the relative insolubility of molybdenum chlorides.[27] It has been suggested that the pseudo-passive behavior, exemplified by the anodic polarization curve for alloy B-3 (N10675) in **Figure 13**, may be due to the formation of a molybdenum dioxide (MoO_2) and/or trioxide (MoO_3) film[24] that breaks down under further anodic polarization. Whatever the mechanism, the inhibiting effect of molybdenum on the anodic dissolution of nickel in hydrochloric acid is of significant practical benefit, exemplified by the isocorrosion diagrams in **Figure 14** that show the performance of alloy B-2 (N10665) relative to other metals and alloys with exceptional resistance to aqueous hydrochloric acid.[28] However, the absence of the capacity of nickel–molybdenum alloys to passivate, which is apparent in the anodic polarization curve for alloy B-3 (N10675) in **Figure 13** signals a sensitivity to the presence of oxidizing agents which limits their practical application, as described later.

In contrast to alloy B-3 (N10675), the anodic polarization curve for the chromium-containing alloy C-2000 (N06200) in **Figure 13** shows a pronounced capacity to passivate, which offers the prospect of useful corrosion performance in environments that have sufficient oxidizing capacity to exceed the critical current density for passivation and raise the potential into the passivation range. The active portion of the curve describes active dissolution of the alloy to produce, in addition to Ni^{2+} and Mo^{3+}, soluble

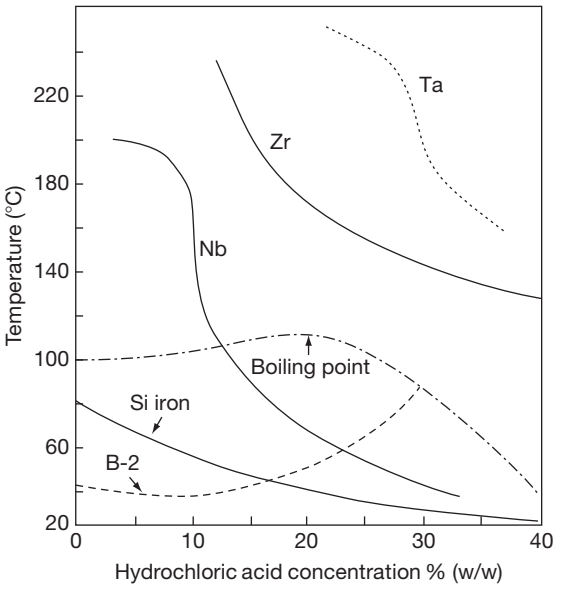


Figure 14 Isocorrosion charts (0.13 mm year^{-1}) for various metals and alloys in hydrochloric acid. Adapted from Hunkeler, F. J. Tantalum and Niobium. In *Process Industries Corrosion*; Moniz, B. J., Pollock, W. J., Eds.; NACE International: Houston, TX, 1986.

chromous ions as follows:

$$Cr = Cr^{3+} + 3e^{-}$$

The active dissolution curve is less polarized than alloy B-3 (N10675) due to its lower molybdenum content. At a potential in the 50–100 mV range, the current density drops by more than two orders of magnitude, signaling the formation of a protective passive film. The passive films that form on nickel–chromium and nickel–chromium–molybdenum alloys have been shown to be thin (<10 nm), and to consist mostly of chromium oxide, Cr_2O_3. An Auger electron spectroscopic (AES) study showed the presence of outer nickel and iron oxides over an inner chromium oxide layer on alloy 600 (N06600), but nickel oxide only in the case of alloy C-4 (N06455).[29] An X-ray photoelectron spectroscopic (XPS) study differentiated inner chromium oxide and outer hydroxide layers.[30] A study using both XPS and time of flight secondary ion mass spectrometry (TOF SIMS) suggested a threshold concentration of chromium in an alloy of ~20% to promote strong passivity, above which thicker oxides developed with layered structures consisting of inner Cr–Ni and outer Cr–Mo oxides. Molybdenum and tungsten exert little discernible effect on passivation at potentials below ~200 mV, when Cr^{3+} is the only possible chromium dissolution product. However, they appear to suppress passive film dissolution at potentials > ~500 mV, when transpassive oxidation to Cr^{6+} is possible.[31]

The relative performances of nickel–chromium and nickel–chromium–molybdenum alloys in aqueous hydrochloric acid will depend on their compositions, as illustrated in the isocorrosion diagrams in **Figure 15** for a range of alloys relative to the stainless steel 25–6MO (N08926) (note that these are isocorrosion diagrams for a corrosion rate of 0.5 mm year^{-1} which is not an acceptable corrosion rate for most practical applications).[32]

Clearly, all of the alloys have significant limitations and none perform as well as the nickel–molybdenum, B family of alloys at concentrations above ~25% (w/w) pure acid. Alloy 825 (N08825) is the leanest in relation to both nickel and molybdenum contents, and is, therefore, the poorest performer. Alloy 625s (N06625) poor performance relates mainly to its relatively low molybdenum content. The remainder has similar resistances, the marginal differences amongst them reflecting mainly their relatively high, but differing molybdenum and tungsten contents.

There is relatively little published data on the behavior of nickel alloys in aqueous hydrobromic acid, but such data suggest that their relative performances are broadly similar to those in hydrochloric acid. However, hydrobromic acid appears to be less aggressive than hydrochloric acid, despite being a stronger acid, exemplified by the isocorrosion diagrams (0.1 mm year^{-1} for alloy C-22 (N06022) in hydrochloric and hydrobromic acids at concentrations up to ~5M that are shown in **Figure 16**.[33]

In practice, the performances of nickel alloys in aqueous hydrohalic acids are very often determined by the presence of oxygen or other oxidizing agents that can accelerate markedly the corrosion rates of nickel–molybdenum alloys, or reduce the corrosion rates of chromium-containing alloys by promoting passivity. For example, corrosion rates for alloy B (N10001) tested in aerated and deaerated aqueous hydrochloric acid at 65–70 °C are shown in **Figure 17**, and the pronounced accelerating effects of aeration, particularly at concentrations below ~25% (w/w), are evident.[34]

The most pronounced effects are produced by stronger oxidizing agents such as ferric ions, the effects of which on the corrosion of alloys B-3 (N10675) and C-2000 (N06200) in deaerated 20% (w/w) hydrochloric acid at 25 °C are summarized in **Figure 18**.

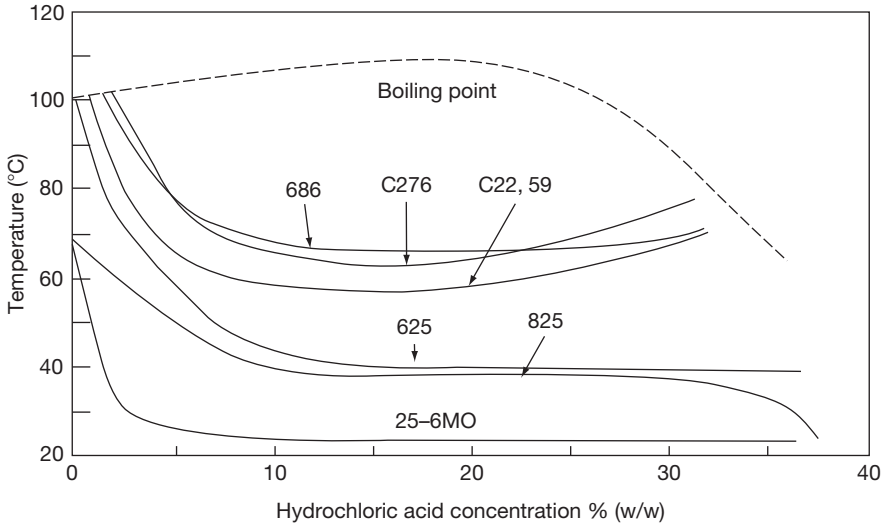


Figure 15 Isocorrosion charts (0.5 mm year^{-1}) for nickel alloys in hydrochloric acid. Adapted from *High Performance Alloys for Resistance to Aqueous Corrosion*; Publication SMC-026, Special Metals Corporation, 2000.

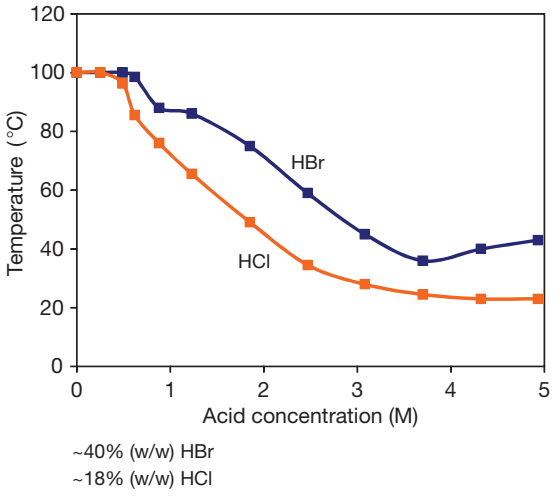


Figure 16 Isocorrosion charts (0.1 mm year^{-1}) for alloy C-22 in hydrochloric and hydrobromic acids at concentrations up to ~5 M. Adapted from Meck, N. S.; Pike, L.; Crook, P. Corrosion Performance of a New Age-Hardenable Ni–Cr–Mo Alloy, Paper 08181, Corrosion 2008; NACE International: Houston, TX, 2008.

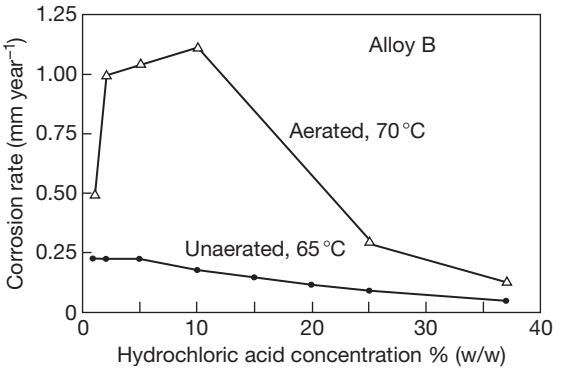


Figure 17 Corrosion rates for alloy B tested in aqueous hydrochloric acid at 65–70 °C. Adapted from Friend, W. Z. *Corrosion of Nickel and Nickel-Base Alloys*; John Wiley & Sons: London, 1980.

Evidently, the addition of up to ~200 ppm Fe^{3+} increases the corrosion rates of both alloys by increasing their corrosion potentials in the active potential range. Further additions produce progressive increases in the corrosion potential and corrosion rate of alloy B-3 (N10675) which remains active throughout the full potential range, as revealed in the polarization curves for the alloy with and without 1000 ppm Fe^{3+} in Figure 18. In the case of alloy C-2000 (N06200), further additions also produce progressive and significant increases in the corrosion potential, but reductions in the corrosion rate by promoting passivation. This is apparent in the polarization curve in Figure 18 for alloy C-2000 (N06200) in the presence of 1000 ppm Fe^{3+} which shows no evidence of an active/passive transition because the material is spontaneously passivated at the outset of the test, as confirmed by its relatively high corrosion potential and the correspondingly low applied current densities required to raise the potential above the corrosion potential.[24]

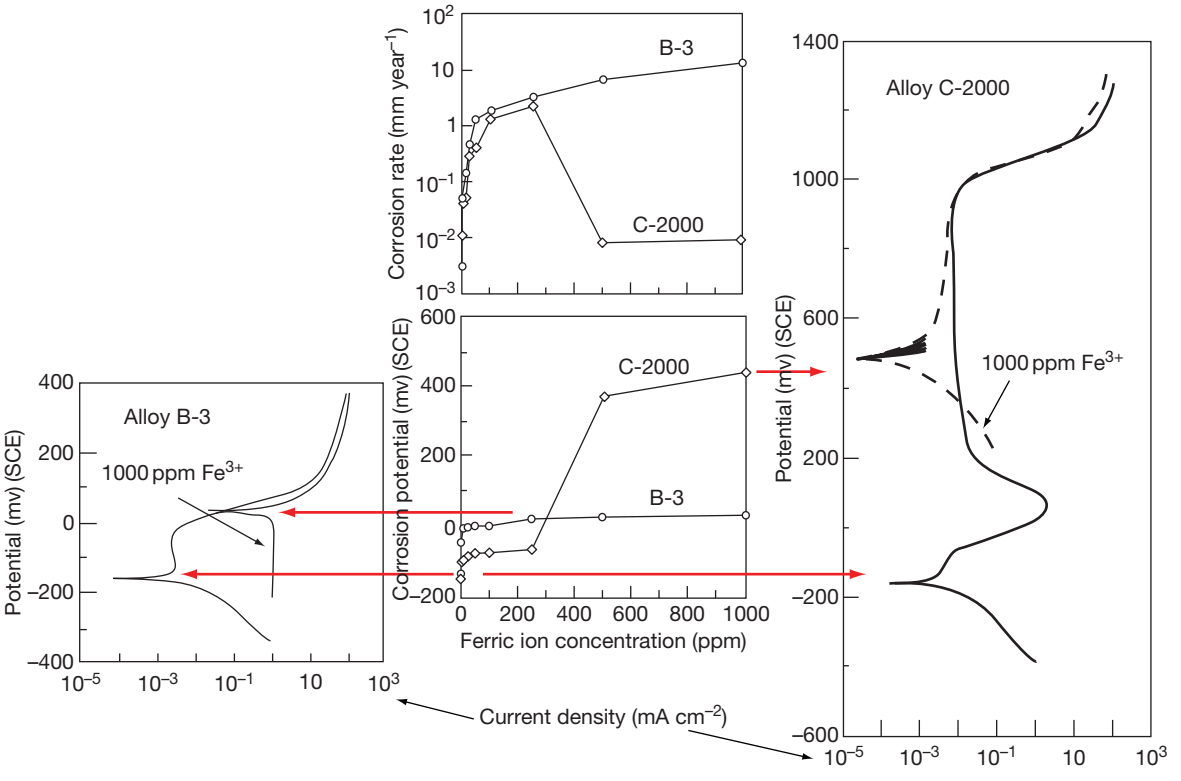


Figure 18 Corrosion rates, corrosion potentials, and polarization behaviors of alloys B-3 and C-2000 in deaerated 20% (w/w) hydrochloric acid with various additions of ferric ions at 25 °C. Adapted from Nacera, S. M.; Crook, P.; Klarstrom, D. L.; Rebak, R. B. *Effect of Ferric Ions on the Corrosion Performance of Nickel Alloys in Hydrochloric Acid Solutions*, Paper 04430, Corrosion 2004; NACE International: Houston, TX, 2004.

Nickel and its alloys are among the more corrosion resistant materials in aqueous HF acid, and their relative performances up to the relevant atmospheric boiling point are summarized relative to the stainless steel type 316L (S31603) in the isocorrosion diagrams in **Figure 19** (note that these are isocorrosion diagrams for a corrosion rate of 0.5 mm year^{-1} which is not an acceptable corrosion rate for most practical applications).[35]

Corrosion rates increase generally with concentration and temperature, and the nickel–copper and nickel–molybdenum alloys outperform significantly the nickel–chromium–molybdenum alloys in the pure acid. Behavior is influenced strongly by the formation and persistence or otherwise of pseudo-passive, fluoride films. However, as with the other acids, the presence of oxygen and other oxidizing agents play a key role and can dominate materials performance, particularly in the cases of nickel–copper and nickel–molybdenum alloys. As a result, the performances of alloys can vary significantly between liquid and vapor phases as shown in the corrosion rates for various nickel alloys tested in and above 10% (w/w) HF acid at 24 and 76 °C that are charted in **Figure 20**.[36]

Thus, all the alloys perform relatively well in the liquid phase but nickel 200 (N02200) and alloys 400 (N04400) and 600 (N06600) perform very poorly in the vapor phase at both temperatures. The corrosion rates of alloy C-22 (N06022) also increase in the vapor phase, but not to the same extent as the other alloys that were tested.

These effects are all due to the increased availability of oxygen in the thin, liquid films that condense in the vapor phase. In the case of the chromium-free alloys 200 (N02200) and 400 (N04400) or chromium-deficient alloy 600 (N06600), corrosion product accumulation in the liquid films promotes progressively increased rates of corrosion. Thus, in the case of alloy 400 (N04400), cuprous ions (Cu^+) that are the initial products of dissolution are oxidized by oxygen to the cupric state (Cu^{2+}) in which they induce further

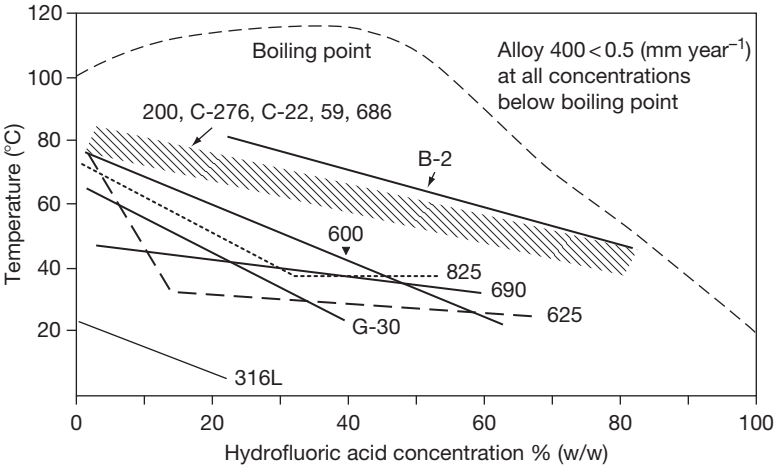


Figure 19 Isocorrosion charts (0.5 mm year^{-1}) for nickel alloys in aqueous hydrofluoric acid. Adapted from Crum, J. R.; Smith, G. D.; McNallan, M. J.; Hirnyj, S. *Characterisation of Corrosion Resistant Materials in Low and High Temperature HF Environments*, Paper 382, Corrosion '99; NACE International: Houston, TX, 1999.

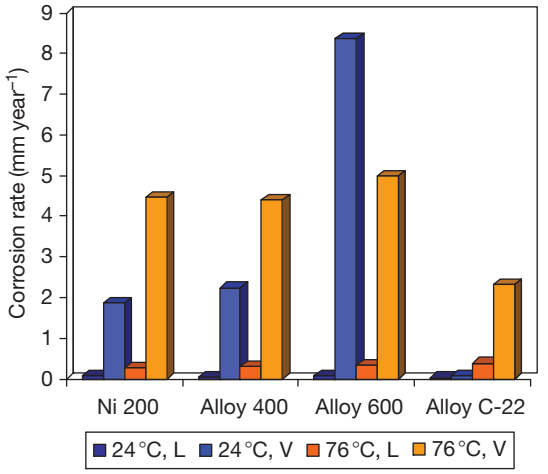


Figure 20 Corrosion rates for various nickel alloys tested in (liquid, *L*) and above (vapor, *V*) 10% (w/w) hydrofluoric acid at 24 °C and 76 °C. Adapted from Pawel, S. J. *Corrosion* **1994**, *50*, 963–971.

increases in corrosion potential and corrosion rate. In contrast, the chromium-containing alloy C-22 (N06022) performs better in the vapor phase because passive film formation, albeit relatively nonprotective, is possible. Clearly, this is a significant consideration in selecting nickel alloys for aqueous HF acid duties. Other significant factors include:

1. Niobium has an adverse effect on the corrosion resistance of nickel alloys, exemplified by the relatively poor performance of alloy 625 (N06625) that contains a nominal 3.7% niobium, as shown in Figure 19.
2. Segregation of specific elements, not least niobium, can result in preferential corrosion of welds in nickel alloys.[6]
3. SCC is a significant threat to nickel alloys in aqueous HF acid. Nickel 200 (N02200) and alloys 400 (N04400), 500 (N05500), 600 (N06600), B-2 (N10665), and C-276 (N10276) have all been reported as vulnerable under specific environmental conditions, in particular in vapors over 48% (w/w) acid. The risks of SCC are exacerbated by cold work, and the presence of oxygen and Cu^{2+} and can be mitigated by appropriate heat treatment.[6]

The very different behaviors of the nickel–molybdenum B and nickel–chromium–molybdenum C families of alloys have stimulated interest in alloy compositions that might provide both the resistance to anodic dissolution of the B family and the capacity to passivate of the C family. For example, a proprietary alloy Hybrid BC-1 has recently been developed[37] containing 22% molybdenum and 15% chromium that outperforms significantly the other members of the C family of alloys in aqueous hydrochloric and hydrobromic acid solutions, and maintains its performance in the presence of oxygen. For example, corrosion rates for the new alloy and various members of the C family of alloys in tests in hydrochloric acid at concentrations in the range 5–20% (w/w) at ~80 °C are shown in Figure 21. Clearly, the

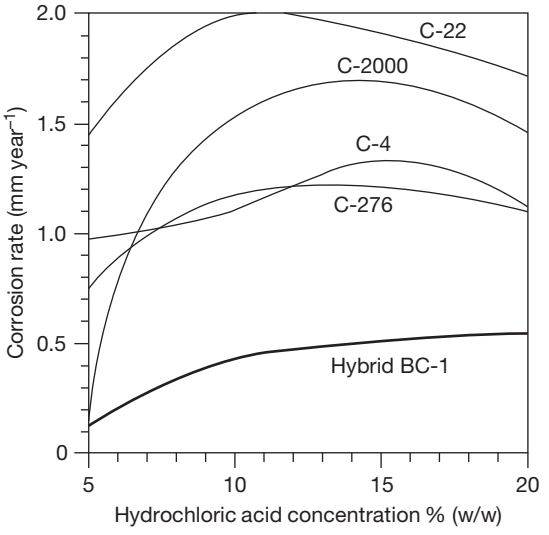


Figure 21 Corrosion rates for various nickel–chromium–molybdenum alloys tested in aqueous hydrochloric acid at 79 °C. Adapted from Crook, P.; Meck, N. S.; Koon, N. E. *The Corrosion Characteristics of a Uniquely Versatile Nickel Alloy*, Paper 08190, Corrosion 2008; NACE International: Houston, TX, 2008.

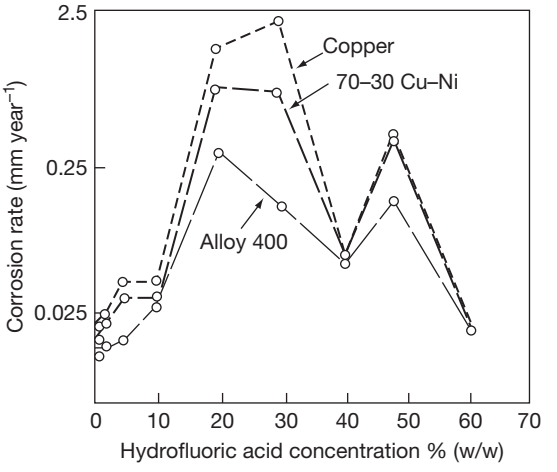


Figure 22 Corrosion rates for copper alloys in static, aqueous hydrofluoric acid under air at 60 °C. Adapted from Materials Selector for Hazardous Chemicals. In *Hydrogen Fluoride and Hydrofluoric Acid*; MTI Publication MS-4, Materials Technology Institute of the Process Industries Inc., 2000; Vol. 4.

new alloy performs significantly better than any of the other alloys tested, but the prevailing rates of ~0.5 mm year^{-1} that apply across much of the concentration range are of limited practical benefit.

2.22.5 Other Metals and Alloys

2.22.5.1 Copper and Copper Alloys

Copper has limited resistance to anhydrous hydrogen chloride gas, with a suggested upper temperature limit for continuous service below 100 °C. Copper and some of its alloys are, however, much more resistant to AHF gas and liquid, as shown in **Figures 11** and **12**. Copper has been used at ambient temperatures to handle AHF liquid and vapor and the cupro–nickels (C71000, C71500), in particular, have useful resistance, and have been used for heat exchanger tubing because of their relative resistance to flowing AHF.[6]

Copper might be expected to resist aqueous hydrochloric acid because it is relatively noble and does not normally displace hydrogen from acid solutions. However, its resistance is reduced significantly in the presence of air or other oxidizing agents, not least its own corrosion product, Cu^{2+}, which can stimulate corrosion autocatalytically. Silicon bronzes (for example, C65500) have found some use at concentrations < 20% (w/w) at moderate temperatures[5] but, in general, applications for copper and its alloys in aqueous hydrochloric acid are limited.

Some copper alloys have useful resistance to aqueous HF acid but only at relatively low velocities in the absence of oxygen or other oxidizing agents. Corrosion rates for some copper alloys in static, aqueous HF acid under air at 60 °C are shown in **Figure 22**.[3]

Other tests have shown increased rates of corrosion in vapors over aqueous acids due to the presence of oxygen. Copper, cupronickels (C71000, C71500), and aluminum/aluminum–silicon bronzes have all found some application in aqueous acids, but there have been problems with dealloying and SCC.[6]

2.22.5.2 Titanium and Titanium Alloys

Titanium and its alloys are dependent on the formation of protective oxide films consisting mostly of titanium dioxide, TiO_2, for their resistances to corrosion. Such films require the presence of oxygen or water to form, and as a general rule anhydrous conditions are best avoided. Titanium and its alloys thus have useful resistance to hydrogen chloride gas containing water but anhydrous gas or condensates from water-containing gas can present significant corrosion risks, depending on the temperature.[2]

One of the conditions for protective film formation in aqueous hydrohalic acids, namely the availability of a

cathodic process to elevate the corrosion potential above the potential for passive oxide formation, is satisfied under most conditions because this potential is below the hydrogen evolution potential. However, the solubility of the passive film increases with concentration and temperature. It has also been shown that the protection afforded by the film depends on the conditions under which it forms, in particular the formation rate.[38] Titanium dioxide is attacked strongly by HF acid to a degree that precludes the application of titanium and its alloys in acidic environments containing HF acid or fluorides. The performance of titanium and its alloys in other hydrohalic acids is dominated by the stability of the titanium dioxide film, and applications are limited to relatively weak acids.

Figure 23 shows isocorrosion curves (0.13 mm $year^{-1}$) for commercially pure titanium (ASTM grade 2, R50400), a titanium–nickel–molybdenum alloy (ASTM grade 12, R53400), and a titanium–palladium alloy (ASTM grade 7, R52400) in naturally aerated hydrochloric acid at concentrations up to 35% (w/w).[39] Evidently, pure titanium is restricted to concentrations below ~5% (w/w) even at ambient temperatures. Alloying with palladium significantly improves the corrosion resistance arising from the well established cathodic modification effects of palladium on promoting and sustaining the passivity of titanium, which are described elsewhere in this book.

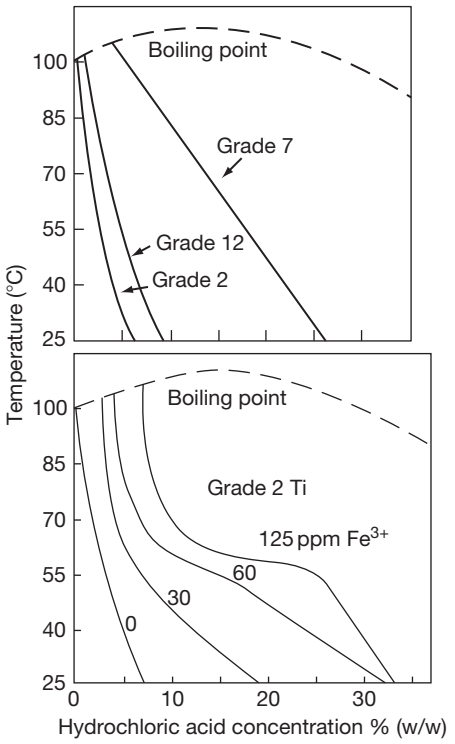


Figure 23 Isocorrosion charts (0.13 mm $year^{-1}$) for titanium and various titanium alloys in hydrochloric acid and in hydrochloric acid containing various levels of ferric ions. Adapted from *Corrosion Resistance of Titanium*; Timet Metals Corporation, 1997.[39]

There is little published data relating to the performance of titanium and its alloys in hydrobromic acid. Apart from the order of hydrohalic acid strengths, it is known that the breakdown potentials of titanium in aqueous halide-containing solutions are in the order: chlorides > bromides > iodides,[40] so hydrobromic acid might be expected to be more aggressive than hydrochloric acid. Notwithstanding that, titanium and its alloys do find commercial application in processes that involve exposure to hydrobromic acid, such as in the manufacture of terephthalic acid from *p*-xylene using metal bromide catalysts.

Oxidizing agents in the acid such as dissolved oxygen, chlorine, nitric and chromic acids, and metal cations in higher oxidation states such as ferric, Fe^{3+} or cupric, Cu^{2+} ions also promote significant increases in corrosion resistance by promoting and sustaining passivity. For example, the effects of progressive additions of up to 125 ppm ferric ions on the performance of commercially pure titanium, (ASTM grade 2, R50400), in aqueous hydrochloric acid is also shown in **Figure 23**. Evidently, the addition of 125 ppm ferric ions extends the acceptable performance boundary to match the performance of the titanium–palladium alloy, (ASTM grade 7, R52400), at lower temperatures.[39] However, even the temporary absence of oxidizing agents can result in a loss of passivity and the addition of palladium to titanium provides a much more robust mechanism for the mitigation of corrosion risk than additions of oxidizing agents to the acid.

2.22.5.3 Zirconium and Zirconium Alloys

Zirconium and its alloys are similar to titanium in their dependence on oxide films, composed mainly of ZrO_2, to resist corrosion. They are less resistant to anhydrous hydrogen halide gases than titanium and they find no practical application. Zirconium dioxide is also attacked strongly by HF acid to a degree that also precludes the application of zirconium and its alloys in acidic environments containing HF acid or fluorides.

The performance of zirconium and its alloys in other aqueous hydrohalic acids is superior to titanium because the zirconia film is much less soluble than the corresponding titania film on titanium.

Indeed, zirconium is second only to tantalum in its resistance to aqueous hydrochloric acid, exemplified by the isocorrosion diagram for commercial purity zirconium (Zr702, R60702) in **Figure 14**. The performances of the commercial alloys containing tin (Zr704, R60704) and niobium (Zr705, R60705) are broadly similar, but they may display inferior resistance toward the performance boundary for commercial purity zirconium (Zr702, R60702) in **Figure 14**. The main threats to the successful performance of zirconium and its alloys in hydrochloric acid are:

1. Strongly oxidizing conditions can result in the potential exceeding the breakdown potential, resulting in the initiation of localized corrosion. Aeration does not present a significant threat in this respect but stronger oxidizing agents such as cupric and ferric ions in sufficient quantities can stimulate significant pitting attack. Galvanic coupling with nobler materials such as graphite can present a similar threat.
2. Galvanic coupling to more active materials can result in the uptake of cathodically produced hydrogen, resulting in the embrittlement and hydriding of zirconium and its alloys, with a consequential loss of ductility.

These risks and their mitigation are described in more detail in the chapter on zirconium and its alloys in this book.

As with titanium, there is little published data relating to the performance of zirconium and its alloys in hydrobromic acid. However, in contrast to titanium, the breakdown potentials of zirconium in aqueous halide-containing solutions are in the order: iodides > bromides > chlorides[41] and it might thus be expected to be less vulnerable to oxidizing agents than in hydrochloric acid.

2.22.5.4 Tantalum and Tantalum Alloys

Tantalum and its alloys are similar to zirconium and titanium in their dependence on the formation of a protective oxide, specifically tantalum pentoxide (Ta_2O_5), to resist corrosion. Tantalum is resistant to hydrogen chloride and bromide gases up to at least 150 °C, and even at higher temperatures if the gases contain some water. However, in keeping with the other reactive metals, tantalum's protective oxide is attacked strongly by anhydrous and aqueous HF acid to a degree that precludes its application in acidic environments containing HF acid or fluorides.

Tantalum pentoxide is very stable in aqueous hydrochloric acid, and tantalum (R05200) has exceptional corrosion resistance to the acid, as illustrated in **Figure 14**. Niobium is similar to tantalum, being protected from corrosion by a film of niobium pentoxide, Nb_2O_5. The film is less protective than tantalum pentoxide, as a result of which niobium is considerably less resistant to corrosion in hydrochloric acid than tantalum or zirconium, as shown in **Figure 14**. Binary alloys of tantalum and niobium deliver performances intermediate between those of the parent metals. The performance of the commercial 2.5% tungsten alloy (R05252) is broadly similar to tantalum metal.

The main threat to the successful performance of tantalum and its alloys in hydrochloric acid is hydrogen embrittlement due to uptake of cathodic hydrogen from the corrosion of tantalum itself or a more active metal to which tantalum is coupled electrically. Tantalum has a high solubility for hydrogen and forms hydrides. It can absorb hydrogen at ambient temperatures but embrittlement arising from the self corrosion of tantalum and its alloys is usually associated with higher temperatures at which significant corrosion rates are experienced. The rate of embrittlement depends on the corrosion rate but the effects are cumulative, and it has been reported[5] in 25% (w/w) hydrochloric acid at 190 °C in which the corrosion rate is no more than 0.025 mm $year^{-1}$.

Hydrogen embrittlement of tantalum can be prevented by electrical isolation from more active equipment items using insulated flanges or by coupling with a more noble metal with a low hydrogen overvoltage such as platinum, either mechanically, or by resistance or spot welding, or brush plating.[2] An area ratio of ~1000:1 Ta:Pt has been shown to be effective, not only in eliminating embrittlement, but also in reducing the corrosion rates of both tantalum and the noble metal, as shown in **Figure 24**.[42] The presence of oxidizing agents also prevents hydrogen embrittlement by raising the potential of the tantalum. Thus, embrittlement does not occur in chlorine-saturated hydrochloric acid.[2]

Tantalum is also highly resistant to hydrobromic acid but there is much less data in the public arena than for hydrochloric acid. Like titanium, the breakdown potentials of tantalum in aqueous halide-containing solutions are in the order: chlorides > bromides > iodides,[40] so tantalum might be expected to be more vulnerable to the presence of oxidizing agents in hydrobromic than in hydrochloric acid.

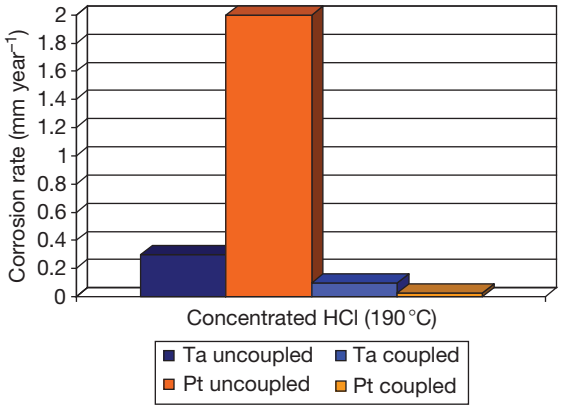


Figure 24 Corrosion rates for uncoupled and coupled tantalum and platinum in concentrated hydrochloric acid at 15–40 °C. Adapted from Bishop, C. R.; Stern, M. *Hydrogen Embrittlement of Tantalum in Aqueous Media*; 17th Annual NACE Conference; National Association of Corrosion Engineers, 1966.

2.22.5.5 Noble Metals

Silver, gold, and platinum are highly resistant to the anhydrous hydrogen halides. In anhydrous hydrogen chloride, upper temperature limits of 230, 870, and 1200 °C have been suggested for silver, gold, and platinum, respectively.[5] Silver has found significant application in equipment for AHF in the absence of oxidizing agents and sulfides.[6]

Silver has good but not unlimited resistance to the aqueous hydrohalic acids below the atmospheric boiling point, depending on the temperature, velocity, and the presence or otherwise of oxidizing agents. Gold and platinum are highly resistant below the atmospheric boiling point, although gold is vulnerable to the presence of very strong oxidants such as ferric ions or nitric acid, and platinum is slightly attacked in 36% (w/w) acid at the boiling point[5,6] and more significantly so at higher temperatures, exemplified by the data in Figure 24.

2.22.6 Nonmetallic Materials

Given the limitations of most classes of alloys that might be considered for handling hydrohalic acid, and the high cost of the more resistant materials such as nickel, zirconium, and tantalum alloys, it is inevitable that nonmetallic materials play major roles in hydrohalic acid applications. The application of these materials is covered in more detail in the relevant chapters in this book but the more significant materials are as follows.

2.22.6.1 Thermoplastic and Reinforced Thermosetting Materials

In practice, chemical compatibility is but one factor that determines the suitability or otherwise of a thermoplastic material for a specific application. Other factors such as mechanical and fabrication properties, thermal expansion characteristics, permeation properties, etc. are significant design considerations. In practice, other than for small scale equipment, thermoplastics are used more commonly as linings on stronger substrates rather than in solid form, because of reliability and integrity concerns. Permeation is a particular issue in the case of HF acid at higher concentrations and temperatures, and needs to be managed by, for example, the use of manifolded vent holes through the metal[7] in the case of lined metal systems.

Most thermoplastics of commercial significance find applications in the handling of hydrohalic acids. Readers interested in specific applications should consult materials suppliers and the more detailed sources are referenced,[2–8] but the more significant materials are as follows:

1. Polyethylene (PE), depending upon its molecular weight, and polypropopylene (PP) have useful resistance. PE tends to be restricted to temperatures close to ambient, but PP has been used at much higher temperatures as a lining material in the handling, for example, of hydrochloric acid.
2. Polyvinyl chloride (PVC) and chlorinated polyvinyl chloride (CPVC) also find uses at lower temperatures, for example, as linings in glass reinforced plastic tanks for the storage of hydrochloric acid. Unplasticized grades are normally specified because plasticized grades have lower chemical resistance.
3. Fluoroplastics are highly resistant to hydrohalic acids, depending upon their degree of fluorination. The highly fluorinated materials polytetrafluoroethylene (PTFE), fluorinated ethylene propylene (FEP), and perfluoroalkoxy (PFA) are resistant to all concentrations of the acids up to and well-beyond their atmospheric boiling points. The less highly fluorinated materials ethylene chlorotrifluoroethylene (ECTFE), ethylene tetrafluoroethylene (ETFE), and polyvinylidene fluoride (PVDF) are also highly resistant but tend to have lower temperature limits for service than the fully fluorinated grades.

Reinforced polyester, epoxy, phenolic, and furane resins all find application in the handling of hydrohalic

acids. Glass reinforced materials (GRP) based on polyester or vinyl ester resins, possibly lined with PVC (so called dual laminate construction), are widely used for the storage and handling of hydrochloric acid at temperatures up to ~90 °C, depending on the acid concentration and the specific resin. Glass is attacked by HF acid but constructions using epoxy vinyl esters and corrosion resistant barriers such as carbon veils can be used for handling weaker acids at lower temperatures.

2.22.6.2 Elastomers

Natural rubber is used widely as a lining for carbon steel in the storage, transport, and handling of hydrochloric acid. Softer rubbers tend to be restricted to lower temperatures, but hard rubber can tolerate temperatures as high as ~90 °C. Natural rubbers are limited to ambient temperatures in HF acid. Chloro- or bromobutyl rubbers are preferred as linings on carbon steel for the storage and transport of up to 70% (w/w) HF acid, up to maximum service temperatures that depend on concentration.

Other elastomers such as butyl, ethylene propylene diene terpolymer (EPDM), and chlorosulphonated PE (such as Hypalon) find niche applications in hydrohalic acids where, for example, the resistance of natural rubber is limited by factors such as the presence of organic contaminants, higher temperatures, or flexibility requirements. Additives such as silica or magnesia compounds are unsuitable for HF acid service, and have to be substituted by materials such as carbon in special grades.

Fluoro- and perfluoroelastomers provide the best resistance to hot and/or concentrated hydrohalic acids, but their high cost tends to limit their use to smaller components such as gaskets and other sealing components.

2.22.6.3 Inorganic Materials

Borosilicate glass is highly resistant to hydrochloric and hydrobromic acids up to temperatures well beyond their atmospheric boiling points, and solid or glass lined equipment finds significant use for handling the acids. However, glass is heavily attacked by HF acid for which it is unsuitable.

Impervious graphite, filled with impregnants such as phenolic resins, PTFE, or carbon, is highly resistant to hydrohalic acids and finds significant use in heat exchange duties for temperatures well-beyond atmospheric boiling. Carbon–carbon composites have excellent resistance to hydrohalic acids and found application for internal components, particularly in HF acid duties. Carbon bricks with appropriate mortars and fluoroplastic membranes have found applications as linings for steel in high temperature duties. Silicon carbide also has excellent hydrohalic acid resistance and has been used in heat exchange duties.

References

1. *Hydrofluoric Acid Properties*; Honeywell Specialty Materials Products, 2002: http://www51.honeywell.com/sm/hfacid/common/documents/HF_Properties_Graph_technical_doc.pdf
2. Materials Selector for Hazardous Chemicals. *Hydrochloric Acid, Hydrogen Chloride and Chlorine;* MTI Publication MS-3: Materials Technology Institute of the Process Industries Inc.: St Louis, MO, 1999; Vol. 3.
3. Materials Selector for Hazardous Chemicals. *Hydrogen Fluoride and Hydrofluoric Acid*; MTI Publication MS-4: Materials Technology Institute of the Process Industries Inc.: St Louis, MO, 2000; Vol. 4.
4. DECHEMA Corrosion Handbook. *Hydrochloric Acid, Nitric Acid*; Wiley-VCH: London, 2004; Vol. 2.
5. Crum, J. R. Corrosion by Hydrogen Chloride and Hydrochloric Acid. In *ASM Handbook. Corrosion: Environments and Industries*; ASM International: Materials Park, OH, 2006; pp 682–689, Vol. 13C.
6. Jennings, H. S. Corrosion by Hydrogen Fluoride and Hydrofluoric Acid. In *ASM Handbook. Corrosion: Environments and Industries*; ASM International: Materials Park, OH, 2006; pp 690–703, Vol. 13C.
7. NACE International Publication 5A-171. *Materials for Storing and Handling Commercial Grades of Aqueous Hydrofluoric Acid and Anhydrous Hydrogen Fluoride*; NACE International: Houston, TX, 2007.
8. Jennings, H. S. *Materials for HF Service in the New Millenium,* Paper 01345, Corrosion 2001; NACE International: Houston, TX, 2001.
9. Foroulis, Z. A. *Anticorros. Methods Mater.* **1988**, *35*(11), 4–12.
10. Schillmoller, C. M. *Chem. Eng.* **1980**, *87*, 161–163.
11. Poorqasemi, E.; Abootalemi, O.; Peikari, M.; Haqdar, F. *Corros.Sci.* **2009**, *51*, 1043–1054.
12. Crouse, P. L. *Corros. Sci.* **1993**, *34*, 295–305.
13. Chirinos, G.; Turgoose, S.; Newman, R. C. *Effects of Residual Elements on the Corrosion of Steels in HF,* Paper 97513, Corrosion '97; NACE International: Houston, TX, 1997.
14. Gysbers, A.; Hasim, H.; Clarida, D.; Chirinos, G.; Marsh, J.; Palmer, J. *Specification for Carbon Steel Materials for Hydrofluoric Acid Alkylation Units,* Paper 03651, Corrosion 2003; NACE International: Houston, TX, 2003.
15. API Recommended Practice RP751. *Safe Operation of Hydrofluoric Acid Alkylation Units*; American Petroleum Institute: Washington, DC, 1999.
16. NACE Publication 8X194. *Materials and Fabrication Practices for New Pressure Vessels Used in Wet H_2S Refining Service*; NACE International: Houston, TX, 2006.
17. NACE Standard Test Method TM0177–2005. *Laboratory Testing of Metals for Resistance to Sulphide Stress Cracking and Stress Corrosion Cracking in H_2S Environments*; NACE International: Houston, TX, 2005.
18. NACE Standard Test Method TM0284–2003. *Evaluation of Pipeline and Pressure Vessel Steels to Hydrogen Induced Cracking*; NACE International: Houston, TX, 2003.

19. NACE Standard Test Method TM0103–2003. *Laboratory Test Procedures for Evaluation of SOHIC Resistance of Plate Steels Used in Wet H_2S Service*; NACE International: Houston, TX, 2003.
20. Agarwal, D. C.; Alves, H. *Applications of Alloys 59 (UNS N06059) and 31 (UNS N08031) in Mitigating Corrosion Risks in the CPI and Petrochemical Industries,* Paper 07186, Corrosion 2007; NACE International: Houston, TX, 2007.
21. Ihara, Y.; Ohgame, S.; Sakiyama, K.; Hashimoto, K. *Corros. Sci.* **1982**, *22*, 901.
22. Myers, W. R.; Delong, W. B. *Chem. Eng. Prog.* **1948**, *44*, 359.
23. Trabanelli, G.; Zucchi, F.; Felloni, L. *Corros. Sci.* **1965**, *5*, 211.
24. Nacera, S. M.; Crook, P.; Klarstrom, D. L.; Rebak, R. B. *Effect of Ferric Ions on the Corrosion Performance of Nickel Alloys in Hydrochloric Acid Solutions,* Paper 04430, Corrosion 2004; NACE International: Houston, TX, 2004.
25. Marcus, P. *Corros. Sci.* **1994**, *36*, 2155.
26. Newman, R. C. *Corros. Sci.* **1985**, *25*, 331–341.
27. Schneider, A.; Kuron, D.; Hofmann, S.; Kirchhein, R. *Corros. Sci.* **1990**, *31*, 191.
28. Hunkeler, F. J. Tantalum and Niobium. In *Process Industry Corrosion*; Moniz, B. J., Pollock, W. J., Eds.; NACE International: Houston, TX, 1986; pp 545–549.
29. Lorang, G.; Jallerat, N.; Vu Kwang, K.; Langeron, J.-P. *Surf. Interface Anal.* **1990**, *16*, 325.
30. Jabs, T.; Borthen, P.; Strehblow, H.-H. *J.Electrochem. Soc.* **1997**, *144*, 1231–1243.
31. Lloyd, A. C.; Noel, J. J.; McIntyre, N. S.; Shoesmith, D. W. *Electrochim. Acta.* **2004**, *49*, 3015–3027.
32. *High Performance Alloys for Resistance to Aqueous Corrosion*; Publication SMC-026, Special Metals Corporation, 2000.
33. Meck, N. S.; Pike, L.; Crook, P. *Corrosion Performance of a New Age-Hardenable Ni-Cr-Mo Alloy,* Paper 08181, Corrosion 2008; NACE International: Houston, TX, 2008.
34. Friend, W. Z. *Corrosion of Nickel and Nickel-Base Alloys*; John Wiley & Sons, 1980.
35. Crum, J. R.; Smith, G. D.; McNallan, M. J.; Hirnyj, S. *Characterisation of Corrosion Resistant Materials in Low and High Temperature HF Environments,* Paper 382, Corrosion '99; NACE International: Houston, TX, 1999.
36. Pawel, S. J. *Corrosion* **1994**, *50*, 963–971.
37. Crook, P.; Meck, N. S.; Koon, N. E. *The Corrosion Characteristics of a Uniquely Versatile Nickel Alloy,* Paper 08190, Corrosion 2008; NACE International: Houston, TX, 2008.
38. Blackwood, D. J.; Peter, L. M. *Electrochim. Acta.* **1989**, *34*, 1505–1511.
39. *Corrosion Resistance of Titanium*; Timet Metals Corporation, 1997.
40. Dugdale, I.; Cotton, J. B. *Corros. Sci.* **1964**, *4*, 397–411.
41. Yau; Te-Lin *Understanding Corrosion Behavior From Electrochemical Measurements,* Paper 04227, Corrosion 2004; NACE International: Houston, TX, 2004.
42. Bishop, C. R.; Stern, M. *Hydrogen Embrittlement of Tantalum in Aqueous Media*, 17th Annual NACE Conference; National Association of Corrosion Engineers: Houston, TX, 1966.

2.23 Corrosion in Sulfuric Acid

J. A. Richardson
Anticorrosion Consulting, 5 Redhills Lane, Durham DH1 4AL, UK

Abbreviations

ASTM American Society for Testing Materials
CPVC Chlorinated polyvinyl chloride
CR Corrosion rate (mm year^{-1})
ECTFE Ethylene chlorotrifluoroethylene
ETFE Ethylene tetrafluoroethylene
FEP Fluorinated ethylene propylene
FRP Fiber reinforced plastic
MSE Mercurous sulfate electrode
ppm Parts per million
PE Polyethylene
PFA Perfluoroalkoxy
PP Polypropylene
PTFE Polytetrafluoroethylene
PVC Polyvinyl chloride
PVDF Polyvinylidene fluoride
SCE Saturated calomel electrode
SHE Standard hydrogen electrode
XPS X-ray photoelectron spectroscopy

Symbols

$C_{surface}$ Solubility of ferrous sulfate in the acid (mass %)
C_{bulk} Initial concentration of ferrous sulfate in the acid (mass %)
d Diameter (m)
E Electrode potential (V)
i Current density (m A cm^{-2})
M_{H2O} Molarity of free water
$M_{species}$ Molarity of species
ρ Density (kg m^{-3})
T Temperature (K)
u Velocity (m s^{-1})
μ Viscosity (cp)

2.23.1 Introduction

Sulfuric acid (H_2SO_4) is a major feedstock for many chemical processes and is manufactured in greater quantities than any other chemical, most commonly

by the contact process, in which sulfur or sulfur-containing waste is burned to form sulfur dioxide (SO_2), which is converted to sulfur trioxide (SO_3) by contact with a vanadium catalyst and absorbed in acid to increase its strength. The structure of sulfuric acid solutions as a function of concentration is summarized in **Figure 1**.[1] From the corrosion standpoint, acids are conventionally classified into three concentration ranges:

1. Weak acids with concentrations $<\sim$25% by weight (w/w). The dominant species are water molecules in which the acid is dissociated into protons (H^+) and bisulfate (HSO_4^-) and sulfate (SO_4^{2-}) anions. Below a concentration of ~5% (w/w), acid strengths are commonly defined by pH. At the higher end of the pH range, the solutions are effectively acidified waters. Thus, pHs of 2, 3, and 4 are equivalent to concentrations of 490, 49, and 4.9 ppm (w/w) sulfuric acid respectively at 25 °C.
2. Medium-strength acids with concentrations in the range 25–80% (w/w), toward the upper end of which the thermodynamic activity of water decreases to low values while acid molecules are dissociated into protons and bisulfate (HSO_4^-) anions.
3. Concentrated acids, with strengths $>\sim$80% (w/w), which contain little free water. At the lower end of the concentration range, acid molecules are dissociated into protons and bisulfate (HSO_4^-) anions but undissociated H_2SO_4 molecules dominate as the concentration increases. This includes all of the significant commercial grades of acid, the commonest of which have strengths within the range 93–99% (w/w). Oleums are fuming acids that contain an excess of uncombined SO_3, defined in terms of either the free SO_3 content or the concentration of sulfuric acid equivalent to the contained H_2SO_4 plus the acid content that would form if the SO_3 were reacted with water to form additional acid. Thus, 20% oleum is equivalent to 104.5% (w/w) sulfuric acid.

Sulfuric acid can be either 'reducing' or 'oxidizing' depending on its concentration. These are rather loose terms, not least because their common usage invariably fails to define the material to which the adjective applies. Conventionally, a reducing acid is one in which cathodic activity is confined to the reduction of protons to evolve hydrogen:

$$2H^+ + 2e^- = H_2$$

An oxidizing acid is one in which reductions of the acid itself, or its constituent species, occur at potentials more noble than hydrogen evolution and to an extent that can determine the behaviors of materials exposed to the acid. The terms have some validity as descriptors of the hydro-halic acids as inherently reducing and nitric acid as inherently oxidizing, but otherwise they are rather imprecise terms that are potentially misleading. Notwithstanding that, sulfuric acid exhibits inherently reducing or oxidizing properties, depending

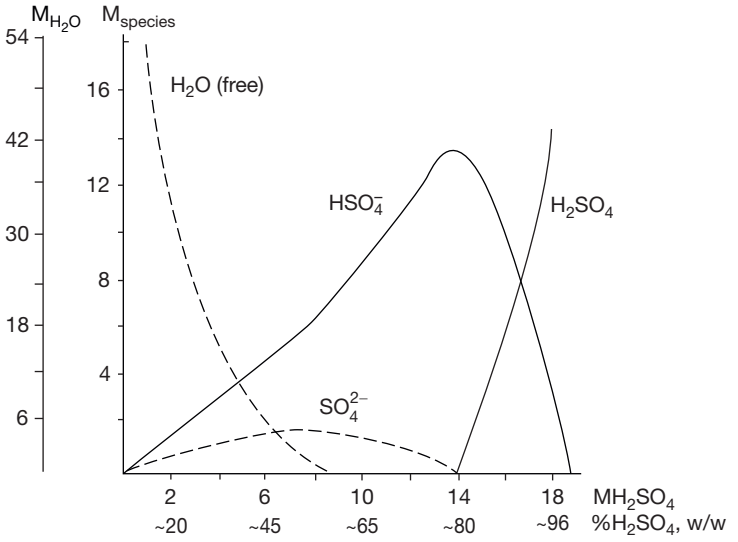


Figure 1 Structure of sulfuric acid solutions. Adapted from Hammer, W. J. *The Structure of Electrolyte Solutions*; J. Wiley, 1959.

on the concentration and temperature. Thus, weak sulfuric acid is inherently reducing but concentrated sulfuric acid is inherently oxidizing, because both bisulfate anions and undissociated acid molecules contain sulfur in the S^{6+} oxidation state that can be reduced to sulfur dioxide, sulfur, or hydrogen sulfide, in which sulfur is in the S^{4+}, S^{0}, and S^{2-} oxidation states, respectively. These reductions can be described stoichiometrically as follows:

$$HSO_4^- + 3H^+ + 2e^- = SO_2 + 2H_2O$$
$$H_2SO_4 + 2H^+ + 2e^- = SO_2 + 2H_2O$$
$$HSO_4^- + 7H^+ + 2e^- = S + 4H_2O$$
$$H_2SO_4 + 6H^+ + 6e^- = S + 4H_2O$$
$$HSO_4^- + 9H^+ + 8e^- = H_2S + 4H_2O$$
$$H_2SO_4 + 8H^+ + 8e^- = H_2S + 4H_2O$$

Medium-strength acids can be inherently reducing or oxidizing. Dillon[1] has pointed out that 25% (w/w) acid is reduced to sulfur by finely divided nickel at ambient temperatures and by alloy 400 (Ni–30%Cu) at the boiling point, and on that basis, would be classified as oxidizing. However, 25% (w/w) acid corrodes steels and stainless steels rapidly at and above ambient temperatures supported by hydrogen evolution, which classifies it as reducing. As a general rule, hotter, more concentrated acids in the medium concentration range are more likely to exhibit oxidizing properties.

In practice, the extent to which a particular strength of acid is reducing or oxidizing at a particular temperature is also influenced and possibly dominated by the presence or otherwise of specific contaminants, in particular, oxidants, such as dissolved oxygen, nitric acid, and metal cations in a higher oxidation state, such as Fe^{3+} or Cu^{2+}, that can be beneficial or detrimental to corrosion performance, depending upon the circumstances. However, the presence of reducing agents, such as sulfites and sulfides, is usually detrimental, and halides, notably the ubiquitous chlorides, can have significant adverse effects on the corrosion performances of specific materials, particularly in weak- and medium-strength acids.

The corrosion performances of the various classes of material in sulfuric acid have been reviewed extensively elsewhere[2–7], and readers whose main interest is to identify 'what works where' are referred to these sources and the relevant chapters on specific materials in this book. An exhaustive literature survey of corrosion in sulfuric acid has also been published.[8] In this chapter, the corrosion performances of materials are reviewed with an emphasis on the principles and mechanisms that underpin their corrosion performances in sulfuric acid.

2.23.2 Steels and Cast Irons

Steels and gray or ductile cast irons are active and corroded rapidly by weak- and medium-strength acids. However, they have useful corrosion resistance in concentrated acids under static or low flow conditions as illustrated by the corrosion rate (CR) data in **Figure 2** relating to the performances of steel and gray or ductile cast irons in acid concentrations above 80% (w/w) at a temperature of 27 °C under stagnant conditions.[2]

The corrosion resistance arises from the spontaneous formation of protective films on steel and cast iron surfaces in concentrated acids. The potential to form such films is illustrated in **Figure 3** that shows anodic polarization curves for gray cast iron in stagnant 20% (w/w) sulfuric acid at 30 and 60 °C.[9] Gray cast iron corrodes rapidly at its natural corrosion potential in 20% (w/w) sulfuric acid at both temperatures, but the CR is reduced significantly in stages by anodic polarization, initially as the potential increases through a less protective sulfation region into a much more protective passivation region. The sulfation process involves the formation of a supersaturated

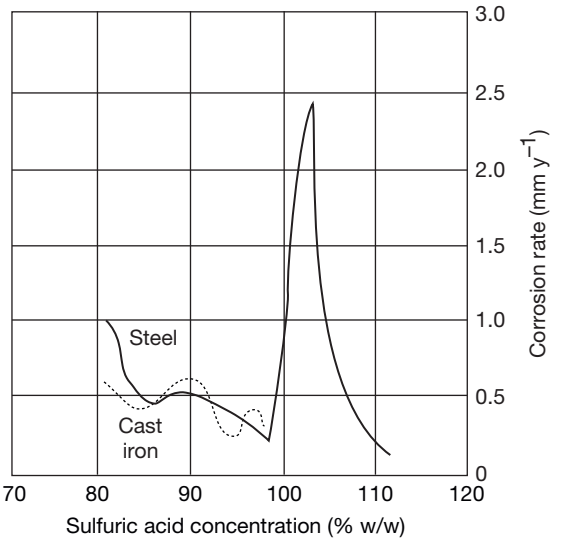


Figure 2 Corrosion rates of steel and cast iron in concentrated sulfuric acid under stagnant conditions. Adapted from Materials Selector for Hazardous Chemicals. *Concentrated Sulphuric Acid and Oleum*, 2nd Ed.; MTI Publication MS-1, Materials Technology Institute of the Process Industries Inc., 2005; Vol. 1.

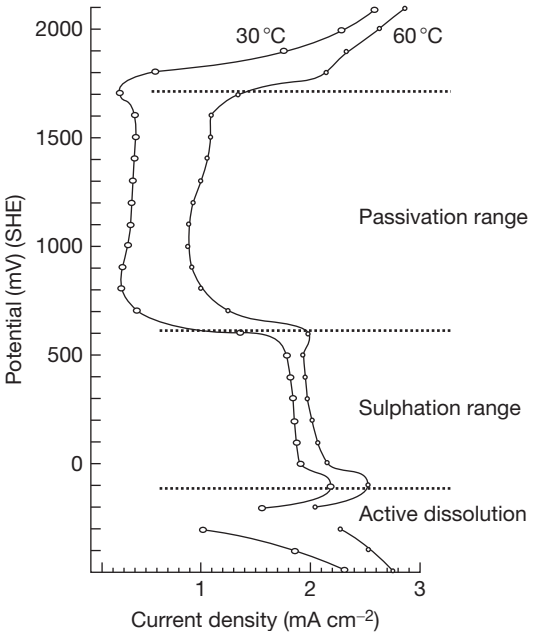


Figure 3 Anodic polarization curves for grey cast iron in 20% (w/w) sulfuric acid. Adapted from Breakell, J. E. *The Corrosive Wear of Cast Irons in Marine Diesel Engine Cylinder Environments;* Ph. D. Thesis, University of Manchester, October, 1980.[9]

solution of ferrous sulfate, from which a crystalline film nucleates and grows. The degree of hydration of this film is thought to vary with acid strength as the concentration of water reduces from the heptahydrate ($FeSO_4{\cdot}7H_2O$) in dilute acids through the monohydrate ($FeSO_4{\cdot}H_2O$) in the 60–70% (w/w) concentration range to $2FeSO_4{\cdot}H_2SO_4{\cdot}H_2O$ in the 75–85% (w/w) concentration range and $FeSO_4{\cdot}3H_2SO_4$ in 96% (w/w) acid.[10] This is the archetypal process underpinning the more general salt-film formation on metals, which may (as in this case) limit the CR by providing a diffusion barrier to reactive species.

As electrochemical potentials increase further into the passivation range, the ferrous salts are thought to oxidize progressively to ferric salts such as $Fe_2(SO_4)_3{\cdot}H_2O$ and $Fe_2(SO_4)_3{\cdot}5H_2SO_4$ in concentrated solutions and ferric oxide centers nucleate and grow in the mixed ferrous–ferric salt film.[10] The passive film has been shown by X-ray photoelectron spectroscopy (XPS) to consist of a hydrated oxy-hydroxide layer across all acid concentrations. In weak acids, the film is thin as a result of the reaction of iron with water. As the acid concentration increases, progressively thicker, less hydrated films with higher contents of oxide phase are formed containing sulfates (S^{6+}) at their surfaces but sulfur species in oxidation state S^{2-} in their bulks, confirming that acid reduction plays an integral role in the passivation process.[11]

In practice, the natural, open circuit potentials of steels and gray or ductile cast irons in concentrated acids are confined to the sulfation potential range, and their corrosion behaviors are determined essentially by the dissolution rates of the iron sulfate films on their surfaces. These are determined in practice by mass transport control of the diffusion of dissolved iron sulfate away from the surface in the prevailing conditions of acid concentration, temperature, velocity, and the presence or otherwise of dissolved iron in the acid. Convective mass transfer rates have been modeled for turbulent flow in a variety of geometries and for the case of a steel pipe protected by a layer of iron sulfate that remains constant with time (ie any loss of the layer is compensated for by further corrosion of the steel), the following expression for calculating the corrosion rate (CR) can be derived.[12]

$$\mathrm{CR} = 1.42 \times 10^{-4}\, T^{0.654} u^{0.913} \rho^{1.567} \mu^{-1.221} d^{-0.087} \left(C_{\mathrm{surface}} - C_{\mathrm{bulk}}\right)$$

Since T, u, ρ, μ and d, and C_{surface} are all known for the concentration of acid at the temperature in question, the CR can be calculated if C_{bulk} is known. C_{bulk} could be assumed to be zero, in which case a maximum CR would be calculated. However, in practice commercial acid always contains some ferrous sulfate, having been produced and handled in carbon steel equipment, and therefore, C_{bulk} can either be measured, or assumed to be typically ~100 ppm (~0.01% w/w), in which case a rate correspondingly lower than the zero case is calculated. In practice, an excellent correlation between calculated and observed CRs in pipes is obtained, confirming that the corrosion of steel is dominated by mass transport control of the diffusion of dissolved iron sulfate away from the corroding surface.[12]

The CRs in **Figure 2** underpin the widespread use of carbon steels and gray/ductile irons for the storage and piping of commercial grades of sulfuric acids with concentrations in the range 93–99% (w/w) at near-ambient temperatures. However, the protective sulfate films are very vulnerable to disturbance. Velocities in steel piping systems have to be restricted[2] to ~0.6 m s^{-1} or less. In steel storage tanks, if roof inlet nozzles are too close to the shell wall, during filling, turbulence adjacent to the wall can promote rapid dissolution of the protective iron sulfate film, resulting in accelerated local corrosion. Bubbles of cathodically produced hydrogen ascend vertically, disrupting the

protective film, causing vertical corrosion grooves at CRs up to several millimeters per year, as shown in Figure 4. Gray and ductile cast irons perform in a manner similar to that of steels in concentrated acids, as confirmed in Figure 2. However, gray irons have slightly better temperature and velocity tolerances that have been attributed to an ennoblement of corrosion potential by the graphite flake network, resulting in a sulfate film that is more resistant to dissolution and disturbance.

High-silicon cast irons, containing typically 14–16% Si (e.g., F47003), are very resistant to all concentrations of sulfuric acid up to 100% (w/w) at temperatures up to their boiling points at atmospheric pressure. Their corrosion resistance is due to the formation of robust, siliceous films that have considerable resistance to erosion and abrasion. Unfortunately, their exceptional resistance finds relatively limited application becauce of their poor ductility that renders them difficult to fabricate and requires them to be protected from thermal and/or mechanical shock.

Electrochemical anodic protection can be used to reinforce passivity steel in concentrated acids, as described in the separate chapter on anodic protection in this book. CRs can be reduced by up to 80%, which can be beneficial, for example, in controlling the iron content of acids in steel storages. However, anodic protection of gray cast irons can result in cracking by a mechanism similar to that described in oleum in the following paragraph.

In oleum, steel experiences relatively high rates of corrosion in the 100–101% (w/w) acid concentration range at near-ambient temperatures that drop to acceptable levels as the concentration increases, arguably beyond ~102.5% and certainly beyond 105% acid (w/w). Ductile irons behave similarly, but gray flake irons cannot be used in oleum, because they are vulnerable to cracking, triggered, it is thought, by the oxidation of graphite flake surfaces by free SO_3 resulting in penetrating volumetric expansion and associated strains. High-silicon cast irons are also rapidly attacked by SO_3 in oleum.

2.23.3 Stainless Steels

2.23.3.1 Mechanisms of Protection

The compositions of some of the more important stainless steels for sulfuric acid applications are summarized in Table 1. The behavior of stainless steels in sulfuric acid is determined by the presence or otherwise of passive surface films that are composed mostly of chromium oxide formed from the chromium in the alloy. The compositions and structures of the films can be modified by other alloying elements, in particular, molybdenum and silicon, resulting in changes in the

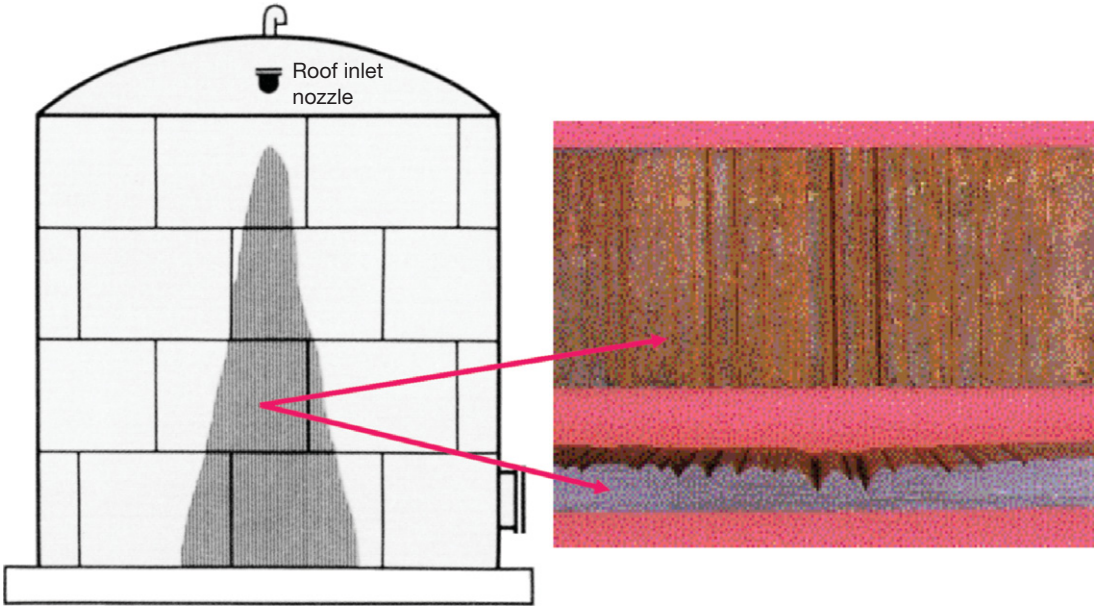


Figure 4 Hydrogen grooving of the wall of a concentrated sulfuric acid storage tank. Adapted from Materials Selector for Hazardous Chemicals *Vol. 1: Concentrated Sulphuric Acid and Oleum (2nd ed.)*; MTI Publication MS-1, Materials Technology Institute of the Process Industries Inc., 2005.

Table 1 Compositions of some wrought stainless steels that are significant for sulfuric acid applications

UNS no	*Common name*	*Fe*	*Ni*	*Cr*	*Mo*	*Cu*	*Si*	*Other*
S43000	430	Balance		16–18				
S31803	2205	Balance	4.5–6.5	21–23	2.5–3.5			N
S32205								
S32750	2507	Balance	6–8	24–26	3–4.5	<0.5		N
S32760					3–4	0.5–1.0		W,N
S30400	304	Balance	8–10.5	18–20				<0.08
S30403	304L	Balance	8–12	18–20				
S31600	316	Balance	10–14	16–18	2–3			<0.08
S31603	316L	Balance	10–14	16–18	2–3			
S31700	317	Balance	11–15	18–20	3–4			<0.08C
S31703	317L	Balance	11–15	18–20	3–4			
S30601		Balance	17–18	17–18			5–5.6	
S31254	254SMO	Balance	17.5–18.5	19.5–20.5	6–6.5	0.5–1.0		N
S31000	310	Balance	19–22	24–26				
N08904	904L	Balance	23–28	19–23	4–5	1–2		
N08926	25–6MO	Balance	24–26	19–21	6–7	0.5–1.5		N
S31277	27–7MO	Balance	26–28	20.5–23.0	6.5–8.0	0.5–1.5		Mn, N
N08020	20	Balance	32–35	19–21	2–3	3–4		
N08031	31	Balance	30–32	26–28	6–7	1.4		N
R20033	33	Balance	30–33	31–35	0.5–2.0	0.3–1.2		N

behaviors of the films under specific conditions of concentration and temperature.

The potential benefits that alloying with chromium provides are illustrated by the polarization curves in Figure 5 for iron, chromium, and type 430 (S43000) ferritic stainless steel that were obtained in agitated, 93.5% (w/w) sulfuric acid at 60 °C.[13] All the three anodic curves show an initial limiting current plateau preceding an active–passive transition as the potential is increased. The limiting current in all the cases results from the convective diffusion of Fe^{2+} and Cr^{2+} into the bulk liquid from a saturated sulfate salt layer. However, the plateau persists over a much wider potential range for iron than for chromium and the type 430 stainless steel, both of which show well developed passivity at much lower potentials and at current densities several orders of magnitude lower than for iron.

XPS studies of the passive layer on chromium in sulfuric acid[14,15] have shown it to consist of chromium oxy-hydroxides across the full range of acid concentrations. As in the case of iron, in weak- and medium-strength acids, passivation results from the reaction of chromium with water molecules from the acid:

$$Cr + 2H_2O = CrOOH + 3H^+ + 3e^-$$

However, in concentrated acids, bisulfate (HSO_4^-) anions and undissociated acid molecules provide the oxygen:

$$Cr + HSO_4^- = CrOOH + SO_2 + e^-$$

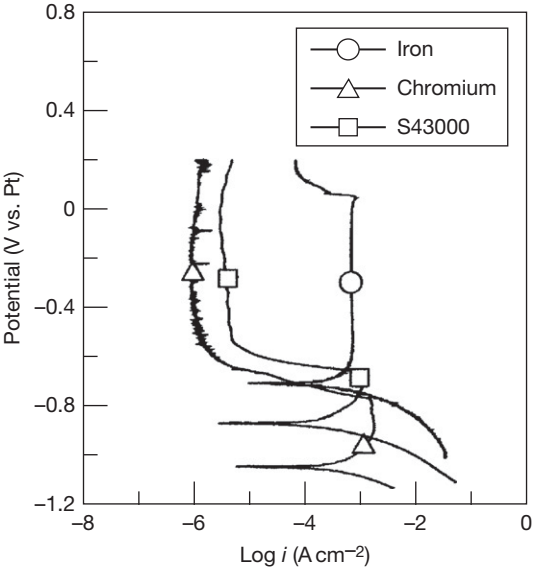


Figure 5 Potentiodynamic anodic polarization behavior of iron, chromium and S43000 (Fe~18%Cr) ferritic stainless steel in stirred 93.5% (w/w) sulfuric acid at 60 °C. Adapted from Kish, J. R.; Ives, M. B.; Rodda, J. R. *Corros. Sci.* **2003**, *45*, 1571–1594.

$$Cr + H_2SO_4 = CrOOH + SO_2 + H^+ + e^-$$

The films formed in concentrated acids contain sulfur species in lower oxidation states (S^{2-}) than in the acid (S^{6+}), confirming that acid reduction plays an integral role in the passivation process.

The passive films that form on stainless steels in sulfuric acid are very similar in structure and thickness to the chromium oxy-hydroxide films that form on chromium metal.[16,17] However, the formation or otherwise of the films depends on the presence and concentration of elements other than chromium in the base material. Thus, it is clear from the polarization curve shown in **Figure 5** for type 430 stainless steel that the steel does not passivate spontaneously under the conditions tested, because the reductions of bisulfate anions and undissociated acid molecules cannot occur at sufficient rates to exceed the limiting current density for passivation. However, under the same conditions, type 304 (S30400) austenitic stainless steel exhibits spontaneous, periodic oscillations of corrosion potential between the active and passive states, as shown in **Figures 6(a)** and **6(b)** on different time bases.[18] These oscillations have been studied by several workers[19–22], and it is generally agreed that they result from the formation of an intermediate corrosion product in the active state that promotes passivation, which is subsequently redissolved resulting in reactivation. Since only nickel-containing steels exhibit oscillations, the intermediate corrosion product is believed to be a nickel salt, probably nickel sulfide, and a sequence of reactions that give rise to the oscillations has been proposed in four stages, as in **Figure 6(b)**:

1. The potential decrease from zero time to point one results from the dissolution of the nickel sulfide film and subsequently, the passive film.
2. The rapid potential drop between points one and two is associated with the active dissolution of the alloying elements as Fe^{2+}, Ni^{2+}, and Cr^{3+} supported by acid reduction to sulfide, S^{2-}.
3. The gradual rise in potential between points two and three is associated with the nucleation and growth of nickel sulfide on the active surface.
4. Nickel sulfide increases the exchange current density for acid reduction to levels at which passivation occurs spontaneously, signaled by a rapid rise in potential between points three and four.

This sequence is the basis for the significant improvements in corrosion resistance of stainless steels in sulfuric acid that result from alloying with nickel at modest levels, although as the level in the alloy increases, nickel's effect in retarding active dissolution of the alloy also becomes significant. Three other alloying elements bring additional improvements:

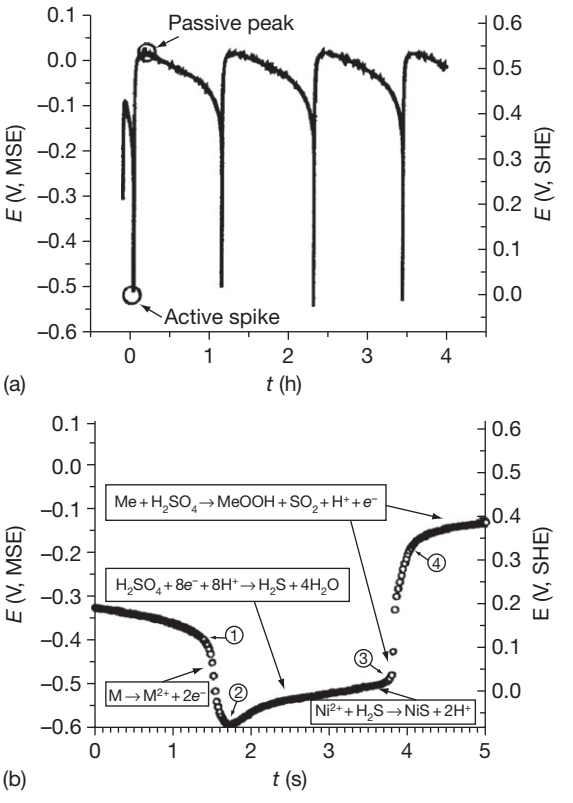


Figure 6 Potential of type 304L austenitic stainless steel in 93.5% (w/w) sulfuric acid at 60 °C as a function of time (a) typical active passive oscillations (b) dominant reactions in a single oscillation. Adapted from Li, Y.; Ives, M. B.; Coley, K. S. *Corros. Sci.* **2006**, *48*, 1560–1570.

1. Copper promotes the formation of passive films, partly by retarding active dissolution and reducing the critical current density required to form a film, but mostly by depolarizing cathodic processes, including hydrogen evolution, in the form of copper films that redeposit rapidly on active surfaces.[23] The beneficial effect is similar to that delivered by nickel sulfide films, and increases up to a level of 2–3% copper, as illustrated in **Figure 7**.
2. Molybdenum also promotes the formation of passive films by depolarizing cathodic processes, in particular hydrogen evolution. It also retards active dissolution in sulfuric acid by the formation of a molybdenum-rich, stable oxide film in the active region that acts as a diffusion barrier to metal dissolution, thereby reducing the critical current density and lowering the potential required to form a passive film.[24] Molybdenum also modifies the composition and structure of the passive film, being incorporated in the Mo^{6+} oxidation state as

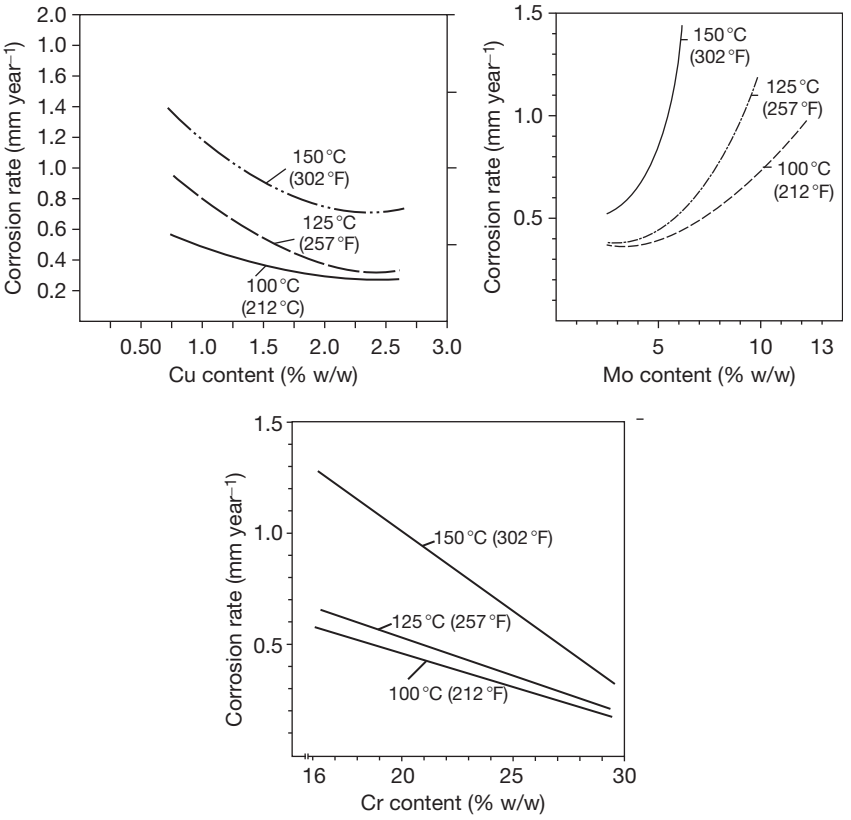


Figure 7 Influence of copper, molybdenum, and chromium contents on corrosion rates of stainless steels and nickel alloys in stagnant 95% sulfuric acid solutions at 100–150 °C. Adapted from Coppolechia, V. D.; Renner, M. H. W.; Rockel, M. B. *Corrosion Resistance of Stainless Steels and Nickel Alloys in Concentrated Sulphuric Acid*; Paper 189, Corrosion '86, NACE International, 1986.

either MoO_3 or MoO_4^{2-}, particularly in the outer regions.[24,25] However, molybdenum increases the vulnerability of the alloy and film to transpassive dissolution,[25] and as a result, molybdenum-containing alloys become vulnerable to corrosion in very concentrated acids and oleum, particularly as temperatures increase, and at levels above ~5% molybdenum, as illustrated in Figure 7.

3. Silicon additions, typically in the range 5–6% (w/w), modify the composition and structure of the passive film, producing silicon-rich films that confer high resistance to corrosion and erosion corrosion in hot acids with concentrations in the 93–99% (w/w) range. However, corrosion resistance drops steeply on either side of this range, and as in the case of high-silicon cast irons, the silicon-rich films are attacked by SO_3 in oleums, rendering the materials unsuitable for service.

The ease and sustainability of passivation of stainless steels is also a function of the chromium content itself, as illustrated in Figure 7. This has an influence on the range of sulfuric acid conditions that specific grades of stainless steel can withstand. Most of the widely used materials, such as alloys 304L (S30403), 316L (S31603), 904L (N08904), and 20 (N08020) contain ~20% chromium but are limited by concentration and/or temperature to varying degrees, depending on their contents of other alloying elements. Increased tolerance of temperature and/or concentration is provided by alloys, such as 31 (N08031) and 33 (R20033), that contain ~27% and ~31% chromium, respectively.

Oxidizing agents and aggressive anions exert strong influences on the corrosion of stainless steels in sulfuric acid. Dissolved oxygen has a pronounced effect on the presence or otherwise of passivity, particularly in relatively weak acids, as confirmed in Figure 8, which demonstrates that rates of corrosion for types 304 (S30400), 316 (S31600), 317 (S31700), and 310 (S31000) stainless steels are all higher at

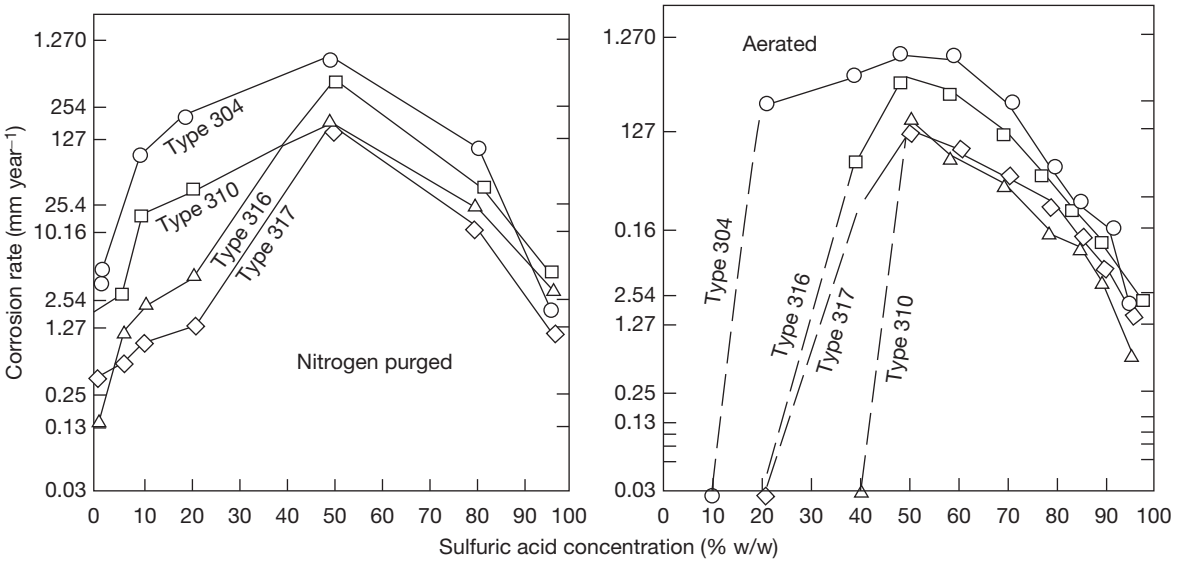


Figure 8 Effects of aeration on corrosion rates of stainless steels in sulfuric acid solutions at 70 °C. Adapted from Phelps, E. H., Vreeland, D. C. *Corrosion* **1957**, *13*, 619–624.

concentrations below 30–40% acid at 70 °C when oxygen is absent.[26] **Figure 9(a)** shows the beneficial effects of additions of Fe^{3+} and Cu^{2+} as sulfates on the corrosion of type 316L (S31603) stainless steel in 75% and 85% acids at 100 °C.[27] **Figure 9(b)** confirms the dramatic effect of sufficient (2%) nitric acid that effectively stifles the corrosion of type 316L (S31603) stainless steel in 30% sulfuric acid at 90 °C, but also illustrates that insufficient oxidizing agent, in this case nitric acid (1%), can accelerate, rather than inhibit, corrosion.[28] Other oxidizing agents have been shown to have similar effects, including Cr^{6+} added as chromic oxide or chromates and Sn^{2+} added as sulfate, which is effective by being reduced to the metallic state.[7] **Figure 9(c)** shows the significant accelerating effects of 5% sodium chloride additions on the CRs of alloys 316L (S31603), 20 (N08020) and 825 (N08825) in 10% acid at 66 °C, and in practice, much lower levels of chloride can exert significant accelerating effects on corrosion.[7]

The passive films that form on stainless steels are in general much less vulnerable to acid velocity than the sulfate films that form on steels. Thus, maximum velocities up to ~3.3 m s^{-1} are recommended for stainless piping systems for commercial grades of sulfuric acids with concentrations in the range 93–99% (w/w) at near-ambient temperatures compared with ~0.6 m s^{-1} for steels.[2] Brubaker[4] has pointed out that the sensitivity of CRs of passive stainless steels to velocity depends on how close they are operating to the transitions to activity or transpassivity, with 'solidly' passive stainless steels tolerating relatively high velocities. A good example of this is shown in **Figure 10** that shows CRs for different stainless steels in 98.7% (w/w) sulfuric acid at 100 °C as a function of velocity. Clearly, the CRs of alloys 316 (S31600) and 825 (N08825) are very dependent on velocity, whereas the CR of the proprietary alloy 1815LCSi, an austenitic stainless steel containing 4% silicon, is independent of velocity, reflecting the tenacious nature of the passive film formed on high Si-containing stainless steels under these conditions.

Some of the main factors that influence the formation and stability of passive films on stainless steels are summarized schematically in **Figure 11**. The anodic polarization curves show the familiar active dissolution, passive, and transpassive regions. As the temperature and/or the concentration of aggressive anions, such as chloride increase, the capacity of materials to passivate is squeezed by extended active dissolution and less protective passive regions. Under nonoxidizing conditions, the prevailing cathodic process, hydrogen evolution, anchors the corrosion potential in the active dissolution region of the anodic polarization curve of the material. Under oxidizing conditions, cathodic processes arising from reduction of the acid itself or from the presence of oxidizing agents, such as dissolved oxygen, nitric acid, or metal cations in higher oxidation states, such as Fe^{3+} or Cu^{2+}, are able to promote passivity.

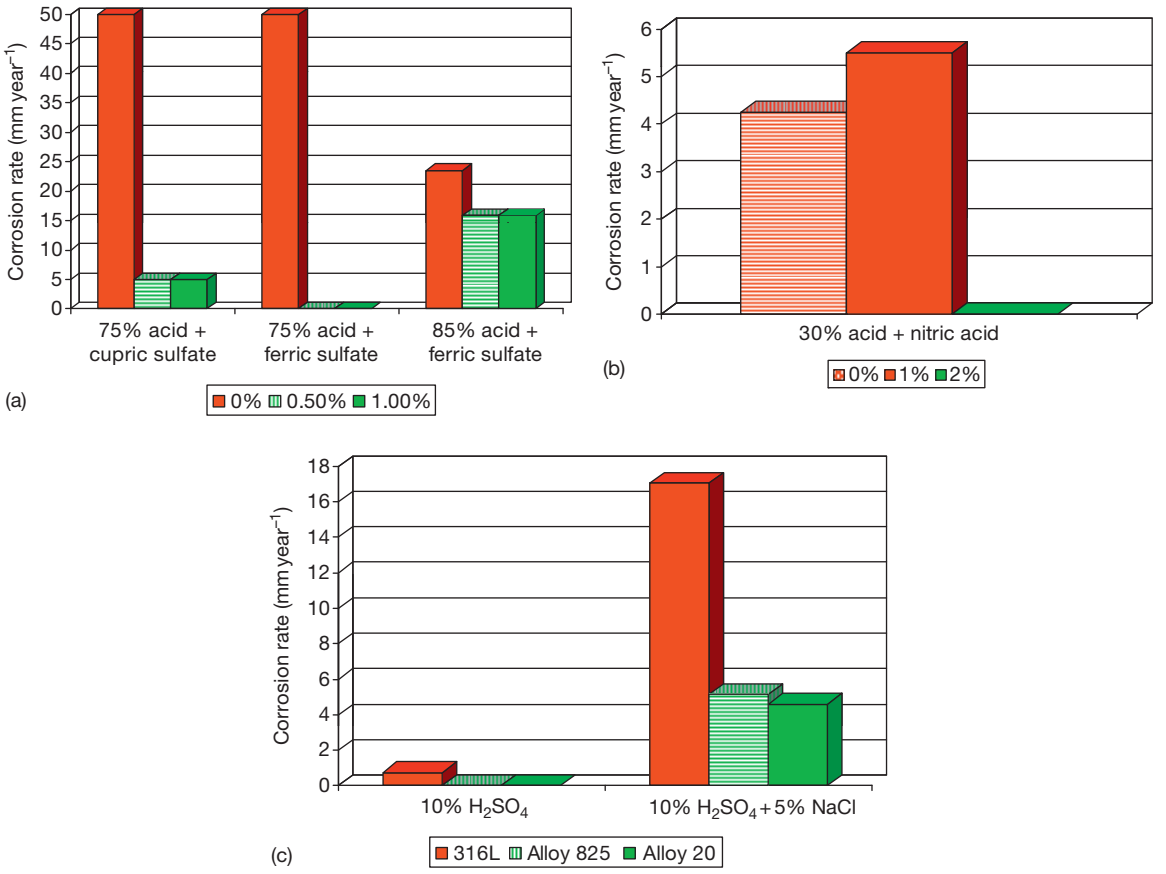


Figure 9 Effects of contaminants on the corrosion of 316L stainless steel and other alloys in sulfuric acid (a) Cu^{2+} and Fe^{3+} additions at 100 °C, adapted from Berglund, G.; Martenson, C. *Application of a High Alloyed Stainless Steel in Sulphuric Acid Environments*; Paper 21, Corrosion '87, NACE International, 1986 (b) HNO_3 additions at 93 °C, adapted from Kiefer, G. C.; Renshaw, W. G. *Corrosion* **1950**, *6*, 235–244 (c) Cl^- additions at 66 °C. Adapted from *The Corrosion Resistance of Nickel-Containing Alloys in Sulphuric Acid and Related Compounds*; Corrosion Engineering Bulletin 1, The International Nickel Company, 1983.

Under strongly oxidizing conditions, the cathodic polarization curve intersects the anodic polarization curves at relatively high potentials well into the passive region, and stable passivity is promoted that may well survive the moderate increases in temperature and/or the concentrations of aggressive anions, such as chloride. However, further increases may move the corrosion potential into the transpassive region in the case of temperature, or above a film breakdown potential in the case of aggressive ions, such as chloride, resulting in accelerated corrosion.

Moderately oxidizing conditions may be sufficient to promote passivity at low temperatures and/or concentrations of aggressive anions, such as chloride. However, moderate increases in either can result in multiple intersections of the cathodic and anodic polarization curves, which are indicative of unstable passivity. Under such borderline conditions, the corrosion potential may oscillate between the passive and active regions, and the CR will lie between the rates that are characteristic of passive and active behavior, depending on the period and frequency of the oscillations.

2.23.3.2 Performances of Stainless Steels

The performances of some of the more important stainless steels for sulfuric acid applications are summarized in **Figures 12–14**. The basic austenitic grades, such as 304L (S30403) and 316L (S31603), provide useful resistance to relatively weak, and in particular, strong acids where they find frequent use for storage and transport equipment at concentrations >~93% (w/w). The molybdenum-containing 316L (S31603) grade expands the region of stable

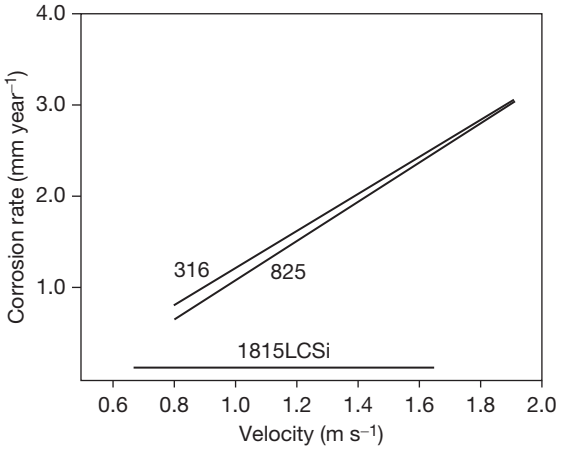


Figure 10 Corrosion rates of different stainless steels in 98.7% (w/w) sulfuric acid at 100 °C as a function of velocity. Adapted from Coppolechia, V. D., Renner, M. H. W., Rockel, M. B. *Corrosion Resistance of Stainless Steels and Nickel Alloys in Concentrated Sulphuric Acid*; Paper 189, Corrosion '86, NACE International, 1986.

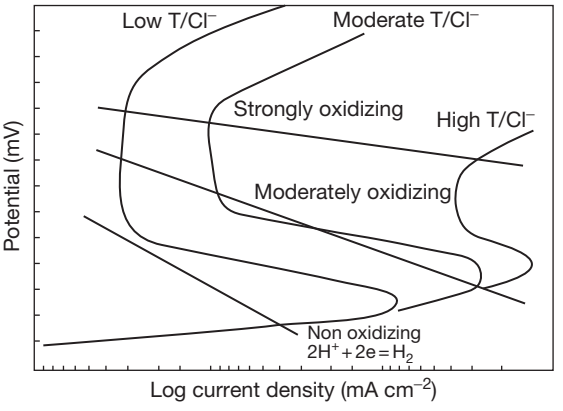


Figure 11 Schematic polarization characteristics of stainless steels in sulfuric acid as a function of temperature (T), chloride content (Cl^-) and the inherent oxidizing potential of the acid, and/or the presence of oxidizing contaminants, such as dissolved oxygen, nitric acid or metal cations in higher oxidation states, such as Fe^{3+} or Cu^{2+}.

passivity relative to type 304L (S30403) and is a candidate, for example, for service in the 90–93% (w/w) concentration range at ambient temperatures, where 304L (S30403) is at greater risk of unstable passivity. In concentrated acids, the upper temperature limit for stable passivity increases with concentration from ~40 °C at 93% to above 100 °C above 99% (w/w), in line with the much greater oxidizing potential of the more concentrated acid. Anodic protection, which is the subject of a separate chapter in this book, can

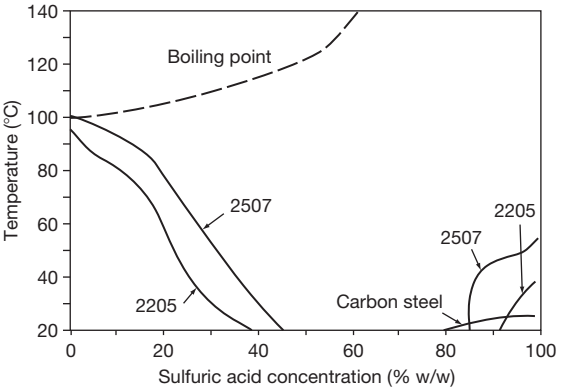


Figure 12 Isocorrosion charts (0.1 mm year^{-1}) for carbon steel and duplex stainless steels in naturally aerated sulfuric acid. Adapted from *Corrosion Handbook, Stainless Steels;* AB Sandvik Steel, 1994.

be of significant benefit in reducing CRs in concentrated acids, as illustrated for type 316 (S31600) stainless steel in **Figure 15**, and is used routinely for the protection of concentrated acid coolers in sulfuric acid plants.

In weak acids, the molybdenum-containing 316L (S31603) grade shows significantly expanded resistance relative to type 304L (S30403) that has a limited resistance to solutions other than very dilute solutions at temperatures close to ambient, and both grades are vulnerable to loss of passivity in the absence of oxygen. The duplex grades, 2205 (S32205) and 2507 (S32750) grades offer significantly enhanced resistance to weaker acids than the basic austenitic grades, in line with their higher contents of chromium, molybdenum, and copper. However, none of the basic austenitic or duplex grades contains sufficient nickel to provide serviceable corrosion performances in medium-strength acids.

Stainless steel grades that offer useful corrosion resistance to medium strength as well as weak and concentrated acids, contain a minimum of ~18% nickel, supplemented by significant molybdenum and/or copper contents. The grades that are leaner in nickel, relying more on molybdenum and copper to promote passivation, such as alloys 254SMO (S31254), 25–6MO (N08926), and 27–7MO (S31277), perform particularly well in weaker acids but have lower temperature limits for stable passivity as the concentration rises and relatively poor performances in concentrated acids. The stainless steel grades that offer the most stable passivity across the whole concentration range are alloys 20 (N08020) and 31 (N08031), which contain a minimum of ~30% nickel

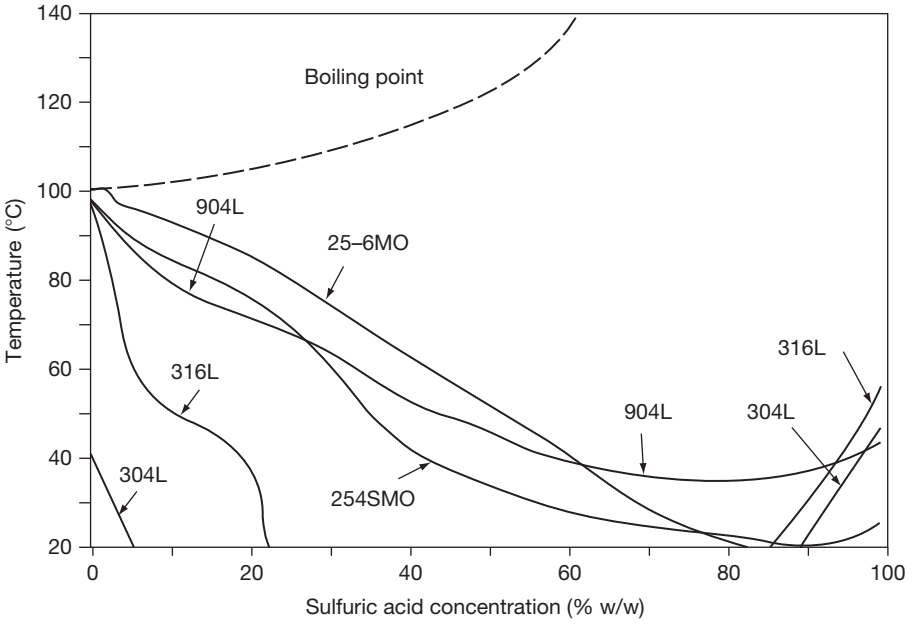


Figure 13 Isocorrosion charts (0.1 mm year^{-1}) for austenitic stainless steels in naturally aerated sulfuric acid. Adapted from *Corrosion Handbook, Stainless Steels;* AB Sandvik Steel, 1994.

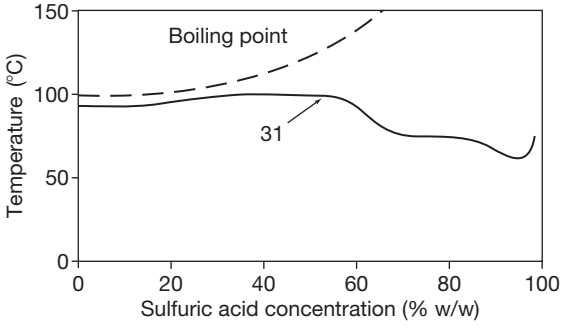


Figure 14 Isocorrosion chart (0.1 mm year^{-1}) for austenitic alloy 31 in lightly aerated sulfuric acid. Adapted from *Material Data Sheet No 4031*, Thyssenkrupp VDM, 2002.

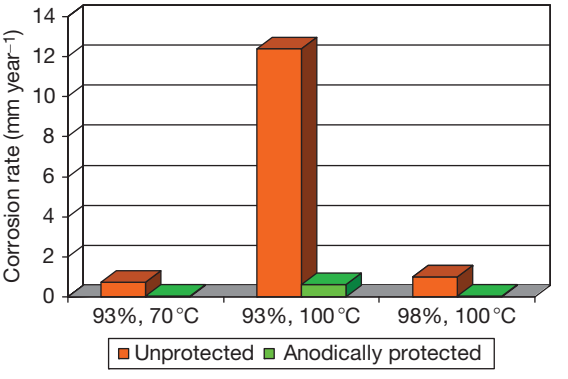


Figure 15 Effects of anodic protection on corrosion rates of type 316 stainless steel in 93–98% sulfuric acid at 70 and 100 °C. Adapted from Walker, R. Anodic protection. In Corrosion, 3rd ed.; Shreir, L. L., Jarman, R. A., Burstein, G. T., Eds.; Butterworth-Heinemann, Oxford, 10.155-10.170, 1994.

supplemented by molybdenum and copper additions, exemplified by the temperature limits for stable passivity of alloy 31 (N08031) shown in **Figure 14**.

Acid strengths above 98% (w/w) and oleum are very oxidizing, and the CRs of most stainless steels are correspondingly lower. However, alloys with particularly high molybdenum contents are vulnerable to transpassive corrosion and can underperform the more basic grades of stainless steel, as illustrated in the ranking of various materials in tests under static conditions in 99% (w/w) acid at 100–120 °C shown in **Figure 16**. Evidently, the most resistant material was type 304L (S30403) stainless steel that outperformed not only the molybdenum-containing alloys 904L (N08904), 20 (N08020), C-276 (N10276), and B-2 (N10665), but also the silicon-containing material S30601. Notwithstanding that, high-silicon stainless steels, such as S30601, have found significant applications in equipment handling concentrated acids at elevated temperatures, as long as there is no risk of oleum formation, to which their resistance is poor. Brubaker[4] has drawn attention to the care that is required in selecting stainless steels for acid concentrations above 98% (w/w) where variations in velocity, oxidant

content and transient acid dilution can adversely affect performance. For oleum duties in particular, stainless steels with high chromium and low molybdenum contents are least vulnerable to such risks. Thus, alloy 33 (R20033) provides very high resistance to hot, concentrated, oxidizing acids, as illustrated in Figure 17, and this extends into the oleum concentration range.

2.23.4 Nickel Alloys

2.23.4.1 Mechanisms of Protection

The CR of nickel in aerated sulfuric acid as a function of concentration at room temperature is shown in Figure 18.[29] Evidently, CRs peak at ~1.5 mm year^{-1} in ~80% (w/w) acid in the concentration range 0–96% (w/w) acid, which, although unacceptable from the engineering standpoint, is indicative of the much slower rates of active dissolution of nickel in weak and medium-strength acids compared with iron. Nickel can be used in deaerated sulfuric acid solutions at low temperatures, but in practice, its significant applications in sulfuric acid are alloy form, and the compositions of some of the more important alloys for sulfuric acid service are summarized in Table 1.

The mechanisms of protection of nickel alloys in sulfuric acid are illustrated by the anodic polarization curves for nickel and binary alloys of nickel with copper, molybdenum and chromium in deaerated sulfuric acid in Figure 19.[30–33] The active region for pure nickel is generally much smaller and the passivation potential lower than for iron, but the current density in the passive range remains relatively high, indicative of relatively unprotective passivity. XPS studies of the passive film indicate a duplex structure with an inner NiO and outer $Ni(OH)_2$ layers, the solubility of the latter increasing as the pH drops in weak acids.[34] At higher potentials, transpassive dissolution of nickel occurs, particularly at grain boundaries. The effects of alloying elements are as follows:

1. Copper additions move the corrosion and the active/passive transition potentials to more noble values and increase the current densities required to form and sustain passivity, so that in the case of the curve for alloy 400 (N04400), the commercial ~30% copper alloy, shown in Figure 19(a), the remaining passivity is even less protective than the relatively unprotective films that are characteristic of pure nickel.[31] These curves not only suggest the significant benefits of alloying with copper at potentials in the active region, but also that in the passive region, alloying with copper alone is deleterious to corrosion resistance.

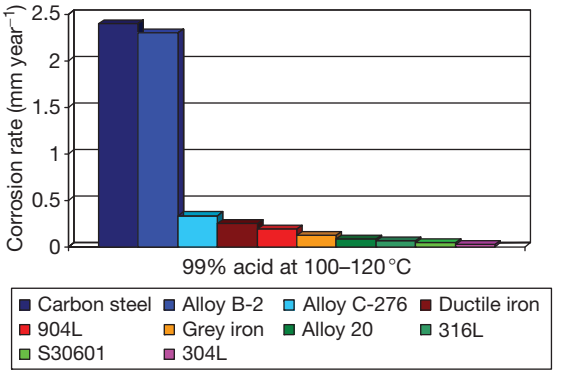


Figure 16 Corrosion rates for various materials under static conditions in 99% (w/w) acid at 100–120 °C. Adapted from Brubaker, S. K., H. Corrosion by Sulfuric Acid. In *ASM Handbook Vol. 13C Corrosion: Environments and Industries*; ASM International, 2006; pp 659–667.

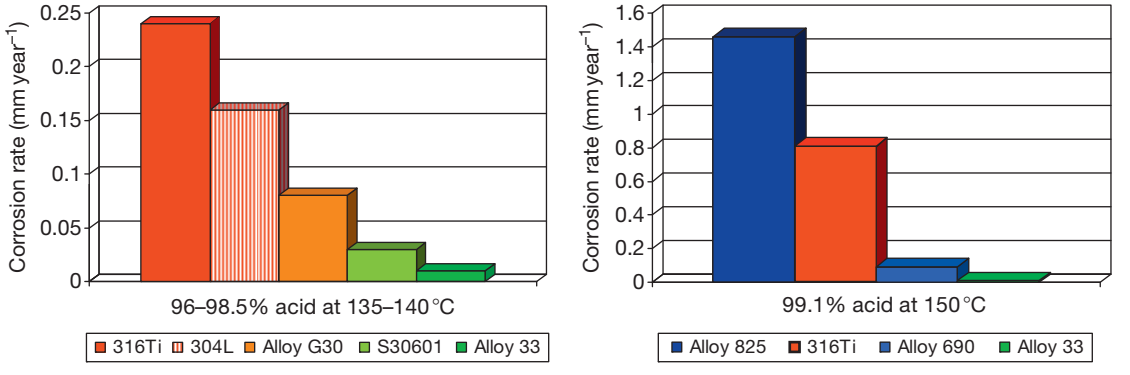


Figure 17 Corrosion rates for various materials in hot, concentrated sulfuric acids. Adapted from Renner, M. H.; Michalski-Vollmer, D. *Stainless Steel World* **2000**, *12*, 53–61.

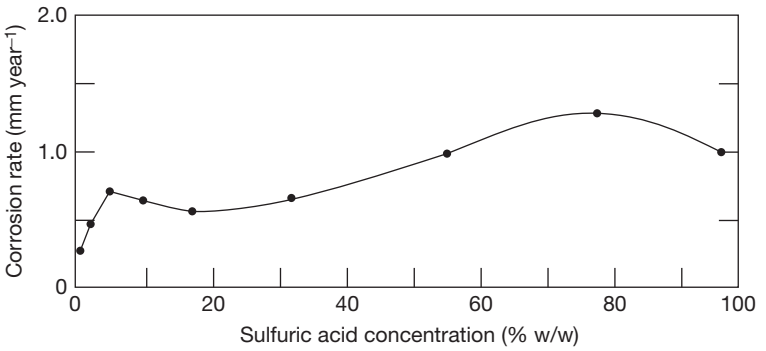


Figure 18 Corrosion rate of nickel in aerated sulfuric acid at room temperature. Adapted from Fink, C. G.; DeCroly, C. M. *Trans. Am. Electrochem. Soc.* **1929**, *56*, 239.

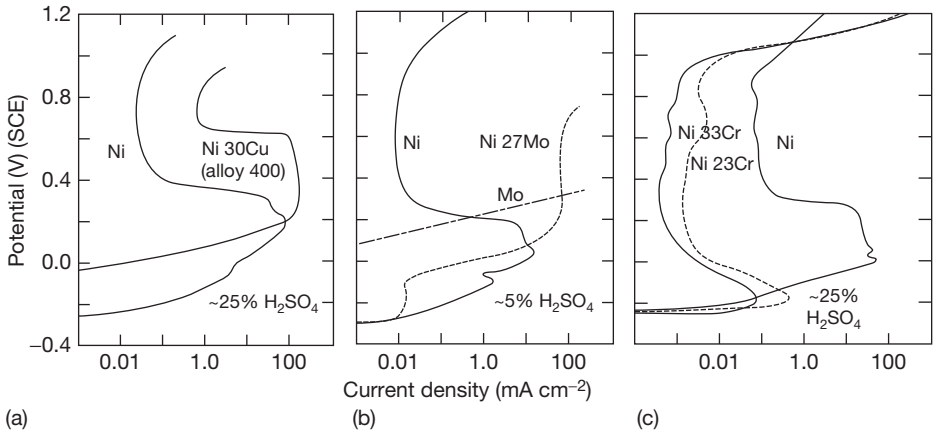


Figure 19 Anodic polarization curves in deaerated sulfuric acid at 25 °C for nickel and (a) nickel–copper. Adapted from Trabanelli, G.; Zucchi, F.; Felloni, L. *Corros. Sci.* **1965**, *5*, 211. (b) nickel–molybdenum. Adapted from Greene, N. D. *The Passivity of Nickel and Nickel-Base Alloys*; Proceedings of the 1st International Congress on Metallic Corrosion; Butterworth: London, 1962. (c) Nickel–chromium alloys, adapted from Myers, J. R.; Beck, F. H.; Fontana, M. G. *Corrosion* **1965**, *21*, 277.

2. Molybdenum additions have effects similar to those of copper additions. The polarization curve for pure molybdenum shown in **Figure 19(b)**, shows no evidence of passive behavior in ~5% (w/w) acid. Alloys with molybdenum additions of up to ~20% (w/w) continue to display passive characteristics but with progressively increasing passive current densities with respect to pure nickel.[32] However, at the 27% (w/w) level that is representative of the commercial B family of alloys, there is no evidence of passivation in the polarization curve shown in **Figure 19(b)** and the current density increases progressively with potential in a manner similar to that of pure molybdenum. As with copper, the main benefit of molybdenum alloying is in the active region, and in the passive region, molybdenum is also deleterious to corrosion resistance.
3. Chromium additions confer very pronounced benefits in both active and passive regions of the polarization curve, as shown in **Figure 19(c)**. In practice, levels of chromium of ~9% or more produce significant reductions in the current density in the active region, the potential at which passivation occurs, and the passive current density.[33] At the 23% and 33% (w/w) levels depicted in **Figure 19(c)**, chromium reduces the potential required to form passivity by ~400 m V and the current densities required to form and sustain passivity by up to two orders of magnitude relative to pure nickel.
4. Silicon additions in the range 5–6% (w/w) modify the composition and structure of the passive

film on nickel–chromium alloys by a similar mechanism as in the case of stainless steels, producing silicon-rich films that confer high resistance to corrosion and erosion corrosion in hot acids with concentrations in the 93–99% (w/w) range.

Oxidizing agents and aggressive anions exert influences on the corrosion of nickel alloys that are similar to those observed with stainless steels in sulfuric acid. Thus, the effects of dissolved oxygen and oxidizing agents on nickel–chromium alloys are broadly beneficial in that they help to promote and sustain the formation of protective passive films. However, dissolved oxygen has a significant accelerating effect on the corrosion of nickel metal and nickel–copper and nickel–molybdenum alloys that do not passivate in the acid. For example, CRs for alloy 400 (N04400) across the concentration range in gently moving acid at 30 °C and alloy B at various acid concentrations and temperatures[7] are shown in **Figure 20(a)** and **(b)**. The accelerating effects of air in weak and medium-strength acids and the dramatic accelerations in CR in strongly oxidizing, concentrated/hot acids are evident in both cases. Other oxidizing agents, such as metal cations in higher oxidation states, in particular Fe^{3+} or Cu^{2+}, chromates, nitrates/nitrites, and peroxides exert similar effects, depending on the alloy. Thus, the CR of alloy B-2 (N10665) in boiling 30% (w/w) acid is increased dramatically by Fe^{3+} additions at levels in the range of 10s–100s ppm, whereas the CR of alloy C-276 (N10276) is, if anything, lowered under similar conditions as shown in **Figure 21(a)** and **21(b)**.[7]

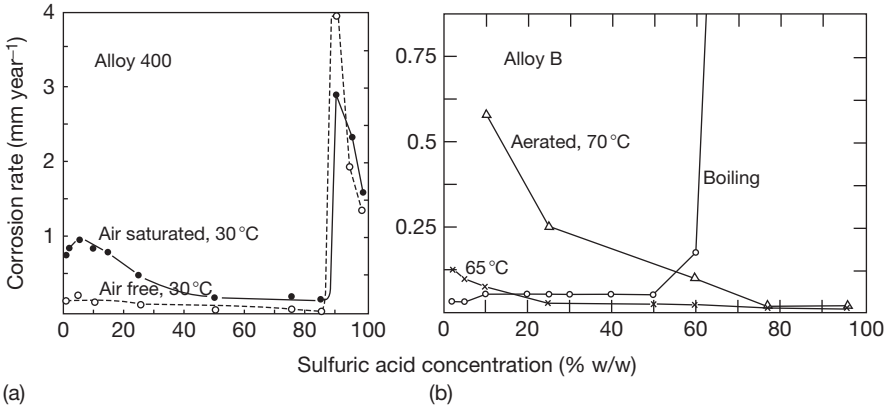


Figure 20 Effects of aeration on corrosion rates in sulfuric acid of (a) alloy 400 and (b) alloy B. Adapted from *The Corrosion Resistance of Nickel-Containing Alloys in Sulphuric Acid and Related Compounds*; Corrosion Engineering Bulletin 1, The International Nickel Company, 1983.

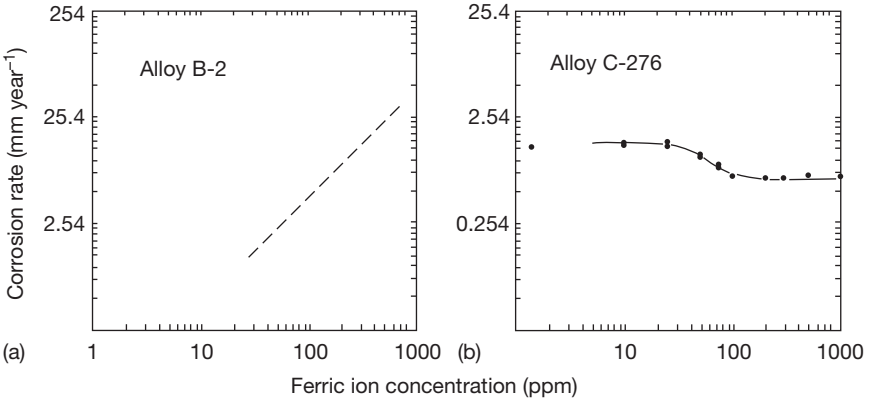


Figure 21 Effects of additions of Fe^{3+} on the corrosion rate of alloys B-2 and C-276 in boiling 30% (w/w) sulfuric acid. Adapted from *The Corrosion Resistance of Nickel-Containing Alloys in Sulphuric Acid and Related Compounds*; Corrosion Engineering Bulletin 1, The International Nickel Company, 1983.

Chloride contaminants have a general accelerating effect on the CRs but the magnitude of the effect varies, depending on the alloy, as illustrated in Figure 22 that shows CRs for alloys B-3 (N10675) and C-2000 (N06200) in boiling weak and medium-strength acids in the presence and absence of 200 ppm chloride.[35] Although the presence of chloride increases the CRs of both alloys at all concentrations, the effects are significantly greater for alloy C-2000 (N06200), which depends on passivity for its corrosion resistance. In contrast, for alloy B-3 (N10675) depends on its intrinsic resistance to anodic dissolution in the nonoxidizing environment that is only marginally affected by the presence of a relatively small amount of chloride.

The sensitivity of CRs to acid velocity also varies with the alloy. The CRs of nickel–copper and nickel–molybdenum alloys that do not form passive films are sensitive to the relatively low velocities that determine the rates of mass transfer of oxidants and soluble corrosion products to/from the alloy surface, as illustrated for alloy 400 (N04400) in 5% (w/w) acid in Figure 23. Also evident in Figure 23 is a limiting CR beyond which velocity has little effect, because the relevant anodic and cathodic processes are activation controlled. In contrast, the passive films that form on chromium-containing alloys are less vulnerable to relatively low acid velocities away from conditions close to transitions to activity or transpassivity and the films that form on silicon-containing alloys, as in the case of stainless steels, are relatively resistant to acid velocity.

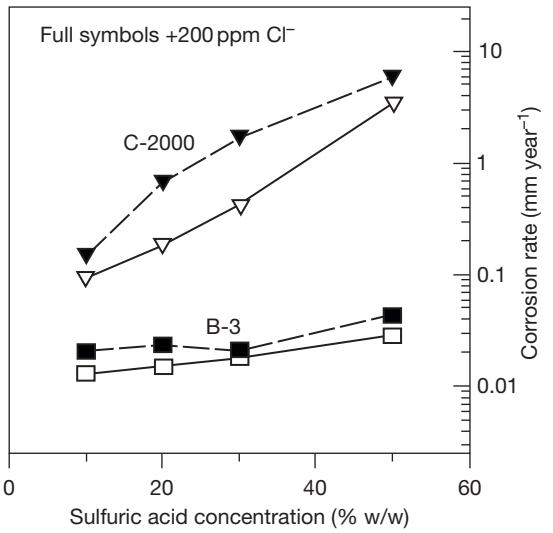


Figure 22 Effects of chloride contamination on the corrosion rates of alloy B-3 and C2000 in boiling sulfuric acid. Adapted from Rebak, R. B.; Crook, P. *Influence of Alloying Elements, Temperature and Electrolyte Composition on the Corrosion Behaviour of Nickel-based Alloys*, Paper 00499, Corrosion 2000; NACE International, 2000.

2.23.4.2 Performances of Nickel Alloys

The compositions of some wrought nickel alloys that are significant for sulfuric acid applications are summarized in Table 2. The substantial nickel contents of all the alloys provide significant improvements in corrosion resistance relative to the iron-based alloys across most of the concentration range. The nickel–copper alloy 400 (N04400) can be used to handle weak- and medium-strength acids up to ~60% concentration (w/w) and ~95 °C in the absence of air and other oxidizing agents, but other materials are generally preferred and the alloy has no application in the strong acid concentration range. The B family of alloys show the greatest resistances across a wide range of temperatures and concentrations in the pure acid, exemplified by alloy B-2 (N10665) in Figure 24, but in practice are very vulnerable to the presence of even small quantities of oxidizing agents that can promote corrosion at unacceptable rates, and therefore, great care is needed in their application.

The most versatile alloys for handling sulfuric acid are the passivable nickel–chromium and nickel–chromium–molybdenum alloys that range from alloy 825 (N08825), which bridges the gap to stainless steels, to the more nickel-rich 690 (N06690) and the

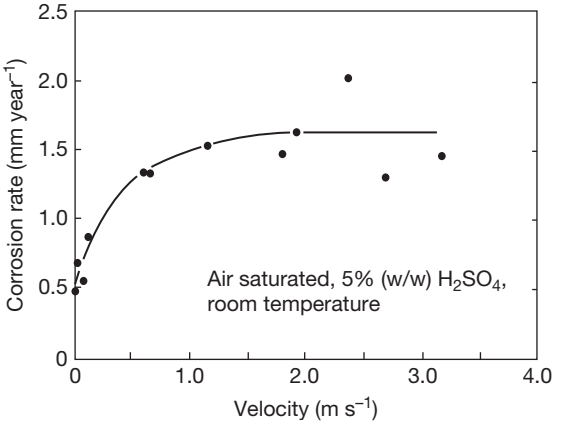


Figure 23 Effect of velocity on the corrosion rate of alloy 400 in sulfuric acid. Adapted from *The Corrosion Resistance of Nickel-Containing Alloys in Sulphuric Acid and Related Compounds*; Corrosion Engineering Bulletin 1, The International Nickel Company, 1983.

Table 2 Compositions of some wrought nickel alloys that are significant for sulfuric acid applications

UNS no	*Common name*	*Fe*	*Ni*	*Cr*	*Mo*	*Cu*	*Si*	*Other*
N08825	825	>22	Balance	19.5–23.5	2.5–3.5	1.5–3.0		Ti
N06007	G	18–21	Balance	21–23.5	5.5–7.5	1.5–2.5		Co, Nb
N06985	G-3	18–21	Balance	21–23.5	6–8	1.5–2.5		Co W
N06030	G30	13–17	Balance	28.0–31.5	4–6	1.0–2.4		W
N06690	690	7–11	Balance	27–31				
N10276	C-276	4–7	Balance	14.5–16.5	15–17			W,Co
N06455	C-4	3	Balance	14–18	14–17			Co
N06022	C-22	2–3	Balance	20.5–22	13–14			W, Co
N06200	C-2000		Balance	23	16	2		
N06059	59		Balance	23	16			
N06686	686	<5	Balance	19–23	15–17			W
N06625	625	<5	Balance	20–23	8–10			Nb
N10001	B	5.5	Balance		28			Co
N10665	B-2	<2	Balance		28			
N10675	B-3	1.5	Balance	1.5	28.5			W, Co
N04400	400	<2.5	Balance			28–34		
	Lewmet 55	16	Balance	31		3		Co
	Lewmet 66	16	Balance	32	4	3	0.5	Co, Mn
	Ilium B		Balance	28	8	5	3.5	
	D205	6	Balance	20	2.5	2	5	

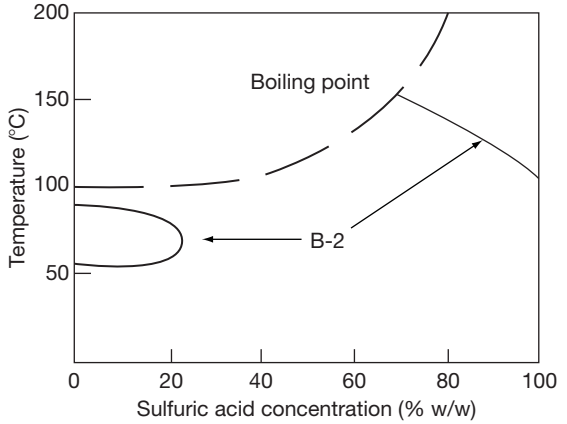


Figure 24 Isocorrosion chart (0.13 mm year^{-1}) for Ni–Mo alloy B-2 in sulfuric acid. Adapted from *The Corrosion Resistance of Nickel-Containing Alloys in Sulfuric Acid and Related Compounds*; Corrosion Engineering Bulletin 1, The International Nickel Company, 1983.

C family of alloys. The performances of alloys 825 (N08825), 625 (N06625), G (N06007), and C-276 (N10276), shown in **Figure 25**, bear comparison with the equivalent performances of the more highly alloyed stainless steels that are exemplified in **Figures 13 and 14**. The leanest alloy 825 (N08825) offers little advantage over the 904L/26–5MO/27–7MO grades of stainless steel. The C family of alloys, exemplified by alloy C-276 (N10276) in **Figure 25**, provides the most consistent resistance to temperature in weak and medium-strength acids into the strong acid concentration range, and their performances are generally enhanced by the presence of oxidizing agents. However, they are vulnerable to the presence of chlorides, particularly in more concentrated acids at higher temperatures.[2,4]

A number of proprietary alloys have found significant usage for niche applications, particularly in relatively hot, concentrated acids. Lewmet alloys 55 and 66, and Ilium B are cast alloys containing relatively high levels of chromium and copper, and in the case of Ilium B, a significant level of silicon, and they are used for pumps and valves. Alloy D-205 is a silicon-containing nickel alloy that was developed to withstand acids in the concentration range of 93–99% (w/w) at temperatures of 90 °C and above, depending on acid strength, and which is more easily formed in thin sections than the equivalent stainless steels for components such as heat exchanger plates.

2.23.5 Other Metals and Alloys

2.23.5.1 Aluminum and Aluminum Alloys

Aluminum and its alloys have limited resistance to sulfuric acid other than in weak acids at relatively low temperatures and find no significant applications as constructional materials. However, thick, porous coatings can be formed on aluminum and some of its

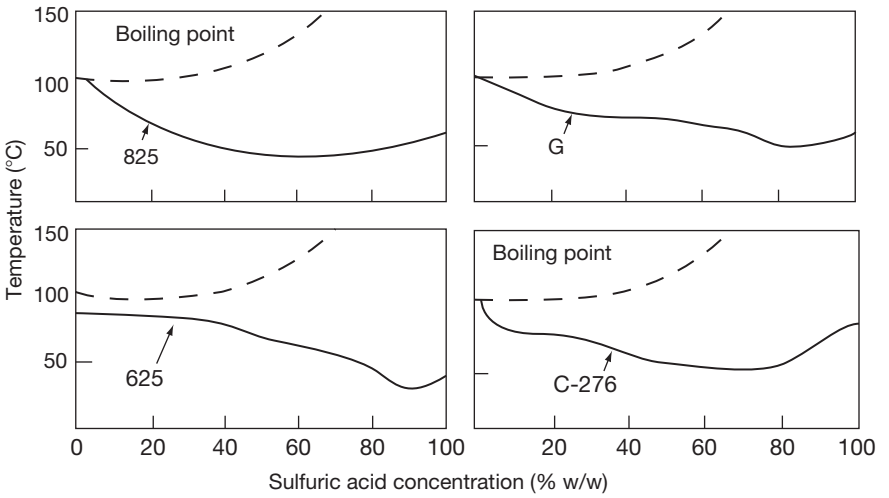


Figure 25 Isocorrosion charts (0.13 mm year^{-1}) for Ni–Cr–Mo alloys in sulfuric acid. Adapted from *The Corrosion Resistance of Nickel–Containing Alloys in Sulphuric Acid and Related Compounds*; Corrosion Engineering Bulletin 1, The International Nickel Company, 1983.

alloys by anodic polarization to voltages up to ~25 V in 5–10% sulfuric acid. These are of commercial significance for corrosion protection and decoration in architectural and other applications and are the subject of a separate chapter in this book.

2.23.5.2 Copper and Copper Alloys

Copper is relatively noble and does not normally displace hydrogen from acid solutions and thus has a relatively good resistance to sulfuric acid in nonoxidizing conditions. However, its resistance is reduced significantly in the presence of air or other oxidizing agents, not least its own corrosion product, Cu^{2+}, which can stimulate corrosion autocatalytically. The brasses are vulnerable to dezincification in sulfuric acid solutions. The tin, silicon, and aluminum bronzes have useful resistances in weak and medium-strength acids at moderate temperatures,[2,4] but, in general, applications for copper and its alloys in sulfuric acid are limited.

2.23.5.3 Titanium and Titanium Alloys

Titanium and its alloys are dependent on passive film formation to resist corrosion in sulfuric acid. One of the conditions for film formation, viz. the availability of a cathodic process to elevate the corrosion potential above the potential for passive oxide formation, is satisfied across the concentration range, even in weak acids, because this potential is below the hydrogen evolution potential. However, film formation also depends on the availability of free water, the concentration of which is progressively limited in concentrated acids and in weak and medium-strength acids, the solubility of the passive film increases with the concentration and temperature and dominates performance. It has been shown that the protection afforded by the film depends on the conditions under which it forms, in particular the formation rate.[36] As a result, the resistances of titanium and its alloys are limited to relatively weak acids.

Figure 26(a) shows isocorrosion curves (0.13 mm year^{-1}) for commercially pure titanium (ASTM grade 2, R50400), a titanium–nickel–molybdenum alloy (ASTM grade 12, R53400), and a titanium–palladium alloy (ASTM grade 7, R52400) in naturally aerated sulfuric acid. Evidently, pure titanium is restricted to concentrations below ~5% (w/w) even at ambient temperatures. Alloying with palladium significantly improves the corrosion resistance at lower temperatures but, as shown in Figure 26(b), not much beyond ~5% (w/w) at the boiling temperature. The improvements in corrosion resistance of ASTM grade 7 (R52400) arise from the well-established cathodic modification effects of palladium on promoting and sustaining the passivity of titanium, described elsewhere in this book.

Oxidizing agents in the acid, such as dissolved oxygen, chlorine, nitric and chromic acids, and metal cations in higher oxidation states, such as Fe^{3+} or Cu^{2+}, also promote significant increases in corrosion resistance by promoting and sustaining passivity. Some of the largest, titanium-clad process vessels ever

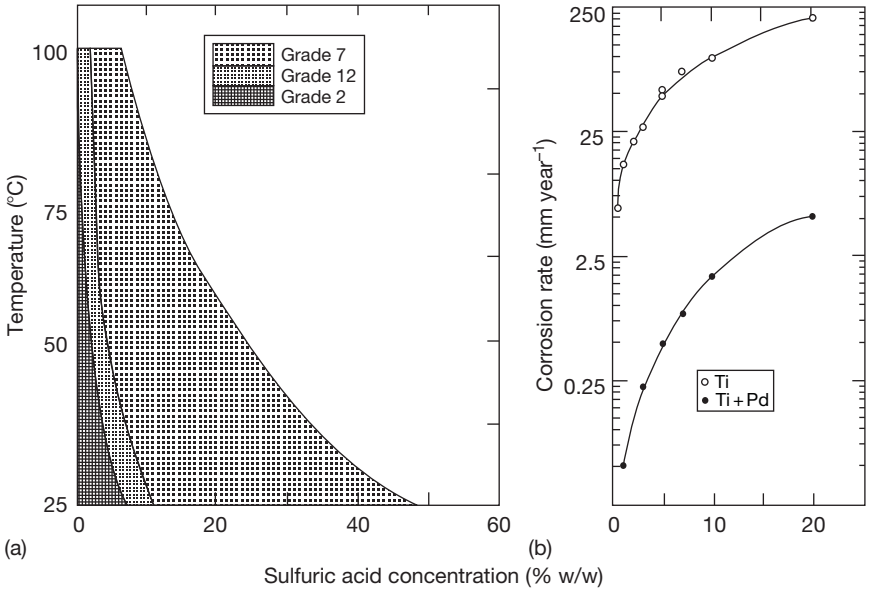


Figure 26 (a) Isocorrosion chart (0.13 mm year^{-1}) for ASTM grades 2 (commercially pure Ti), 7 (Ti 0.12–0.25%Pd) and 12 (Ti 0.6–0.9%Ni 0.2–0.4%Mo) alloys in naturally aerated sulfuric acid (b) corrosion rates of Ti and Ti 0.2%Pd alloy in boiling acid. Adapted from *Handbook of Corrosion Data*; ASM International, 1995.

constructed are used for the acid leaching of ores to recover gold, nickel, and other metals, involving exposure to several % (w/w) sulfuric acid in the presence of significant partial pressures of oxygen at temperatures as high as 200–250 °C.[37] However, even the temporary absence of oxidizing agents can result in a loss of passivity and a risk of very high CRs, and this inevitably limits the wider application of titanium and its alloys in sulfuric acid applications.

2.23.5.4 Lead and Lead Alloys

The lead–sulfuric acid system is the basis of the ubiquitous lead-acid accumulator battery in which lead electrodes are exposed to acid in the concentration range of 27–39% (w/w) when fully charged. As a result, the prevailing anodic and cathodic processes have been studied intensively. The anodic reaction during discharge is the dissolution of lead in the divalent state followed by the precipitation of lead sulfate on the anode, and this is also the mechanism via which lead and its alloys resist corrosion in sulfuric acid. The effects of temperature and concentration on the CR of lead in stagnant sulfuric acid are shown in **Figure 27**.[38] Evidently, lead is very resistant to corrosion at temperatures up to boiling in 50% (w/w) acid and ~80 °C in 85% (w/w) acid, beyond which the film dissolves at rates that increase progressively. However, as in the case of the equivalent films on steels, sulfate films on lead are vulnerable to erosion and abrasion resulting in excessive rates of attack, particularly at conditions close to the boundary for acceptable corrosion. For this reason, lead has often been used as a corrosion resistant membrane behind acid-resistant brick linings that separate the lead from any potentially damaging velocity and abrasion effects.

Historically, lead and its alloys found significant application in equipment for the manufacture and storage of sulfuric acid, but its use has declined because of its relatively poor mechanical properties, and health and toxicity concerns, resulting in a shortage of skilled craftsmen.

2.23.5.5 Zirconium and Zirconium Alloys

Zirconium and its alloys are similar to titanium and its alloys in their dependence on passive film formation to resist corrosion in sulfuric acid. The key difference is that the passive films on zirconium and its alloys are much less soluble in sulfuric acid across the weak and medium-strength concentration range. The effects of concentration on the anodic polarization behavior of zirconium near the boiling point are shown in **Figure 28(a)**.[39] Evidently, at concentrations below 20% (w/w), zirconium shows no active region and no tendency to transpassive breakdown at potentials significantly above 1 V relative to the saturated

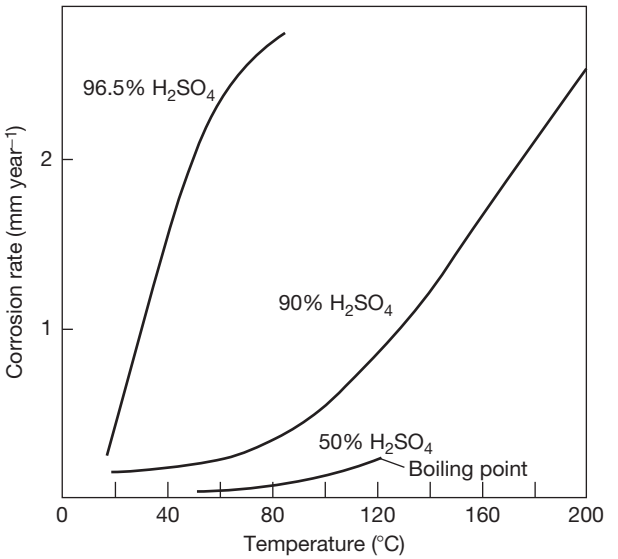


Figure 27 Effects of temperature and concentration on corrosion rates of lead in sulfuric acid. Adapted from Hofmann, W. *Lead and Lead Alloys: Properties and Technology*; Vibranos, G.,Ed.; Springer-Verlag, 1970.

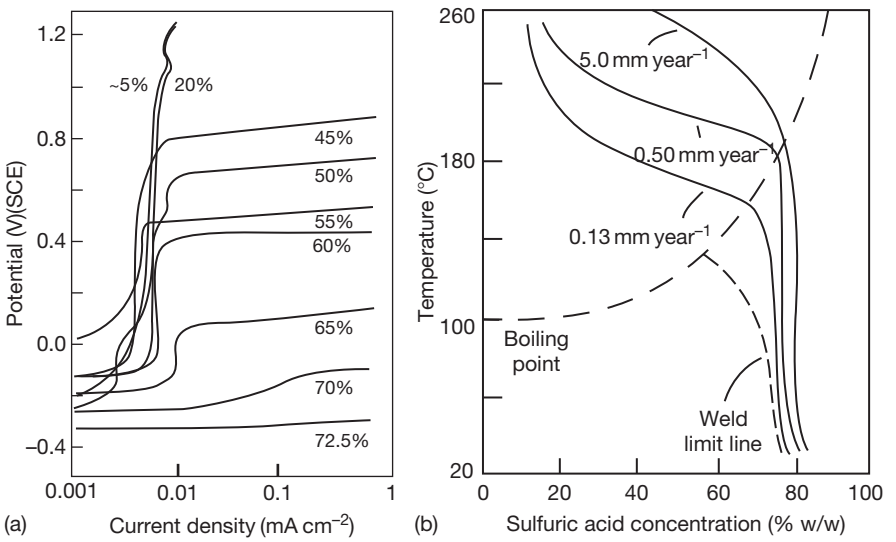


Figure 28 Effects for zirconium in sulfuric acid of (a) concentration (% w/w) on anodic polarization behavior near the boiling point. Adapted fromYau, Te-Lin *Understanding Corrosion Behaviour From Electrochemical Measurements*; Paper 04227, Corrosion 2004; NACE International, 2004 (b) temperature and concentration on corrosion rate. Adapted from *Handbook of Corrosion Data*; ASM International, 1995.

calomel electrode (SCE). At higher concentrations, transpassive breakdown behavior is exhibited at potentials that reduce with increasing concentration and at 70% (w/w) concentration and beyond, zirconium reverts to active behavior with no evidence of stable passivity.

This behavior is reflected in the performance of commercial purity zirconium (Zr702, R60702) in sulfuric acid as a function of concentration and temperature, which is summarized in **Figure 28(b)**.[40] Evidently, at concentrations up to 20% (w/w), zirconium resists attack at temperatures well above the boiling point, beyond 200 °C. The limiting temperature for passivity decreases as the concentration increases beyond 20% (w/w) but remains above the boiling point in pure acid up to >60% (w/w).

In 70–80% (w/w) acid, the limiting temperature for passivity falls off rapidly with increasing concentration due to the formation of nonprotective zirconium sulfate films. Figure 28(b) indicates that the resistance of zirconium welds to >55% (w/w) acid is marginally lower than for parent material due to the presence and morphology of second-phase particles, related to tin content, but appropriate heat treatment can restore the corrosion resistance of weld metal. The performances of the commercial grades containing tin (Zr704, R60704) and niobium (Zr705, R60705) are broadly similar, except in the near boiling region above ~60% (w/w) acid concentration, where Zr702 outperforms Zr704 and Zr705.

The polarization curves in Figure 28(a) suggest that zirconium and its alloys may be vulnerable to the presence of oxidizing agents in higher concentration acids that might elevate the potential above the corresponding transpassive breakdown potential. In practice, zirconium is relatively tolerant to oxidizing agents over most of its useful concentration range but in >60% (w/w) acid, becomes sensitive to the presence of oxidizing agents.[41] Chlorides can also accelerate CRs if oxidizing agents are present, and zirconium and its alloys are acutely sensitive to the presence of even small quantities of fluorides in acids of all strengths.

Zirconium is vulnerable to stress corrosion cracking in 64–69% (w/w) sulfuric acid, the mechanism and remedies for which have been described elsewhere.[42]

2.23.5.6 Tantalum and Tantalum Alloys

Tantalum forms very stable, oxide films in sulfuric acid across the concentration range. For example, films formed galvanostatically in 95% (w/w) acid at 25 °C continue to thicken to support an applied potential of ~50 V before breaking down.[43] It is thus not surprising that tantalum (R05200) has an exceptional corrosion resistance to sulfuric acid, as illustrated by the CR data in Figure 29, which shows that tantalum has exceptionally low CRs at concentrations up to 98% (w/w) acid at temperatures up to ~150 °C and corrodes very slowly in 98% (w/w) acid at temperatures up to ~200 °C. However, tantalum is much less resistant to fuming acids, as illustrated by the data for oleum in Figure 29.[42] Long-term tests have suggested temperature limits for unalloyed tantalum of 200 °C and 150 °C for acid concentrations of 96 and 97% (w/w), respectively, and temperature limits for the 2.5% tungsten alloy (R05252) of 210 °C and 175 °C for acid concentrations of 96 and 97.5% (w/w), respectively. The same tests have suggested that the tungsten alloy has superior resistance to hot, concentrated acids that contain oxidizing agents such as nitrates.[44]

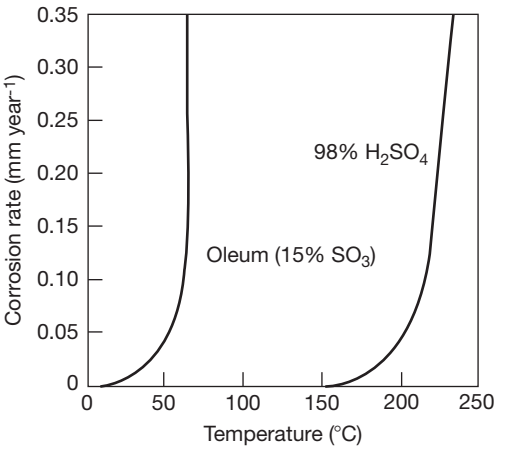


Figure 29 Effects of temperature on corrosion rates of tantalum in 98% (w/w) sulfuric acid and oleum. Adapted from *Handbook of Corrosion Data*; ASM International, 1995.

A guide to the uniquely wide ranging corrosion resistance of tantalum and its alloys to sulfuric acid is illustrated by the test data in Figure 30 for a range of metals and alloys in boiling acids across the concentration range up to 80% (w/w). It is clear that only tantalum and its alloys provide acceptable resistance across the full concentration range tested.

2.23.6 Nonmetallic Materials

Given the limitations of most classes of alloy that might be considered for handling sulfuric acid, and the high cost of the more resistant materials, such as nickel, zirconium, and tantalum alloys, it is inevitable that nonmetallic materials play significant roles in sulfuric acid applications. The application of these materials is covered in more detail in the relevant chapters in this book, but the more significant materials are as follows:

2.23.6.1 Organic Materials

Concentrated sulfuric acid is a strong oxidizing and dehydrating agent, and hence, organic materials have inevitable performance limitations in concentrated and hot acids.

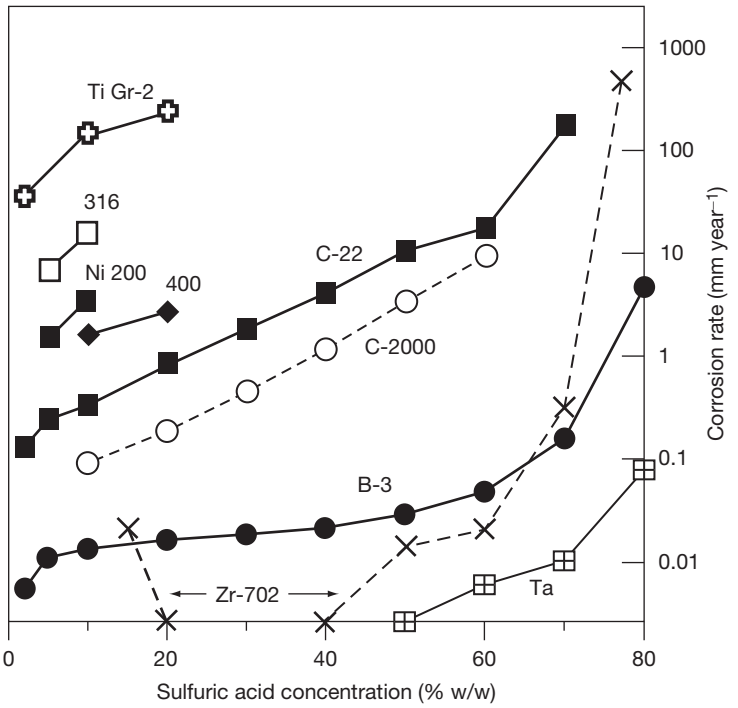


Figure 30 Effects of concentration on corrosion rates in boiling sulfuric acid. Adapted from Rebak, R. B.; Crook, P. *Influence of Alloying Elements, Temperature and Electrolyte Composition on the Corrosion Behaviour of Nickel-based Alloys*, Paper 00499, Corrosion 2000; NACE International, 2000.

In relation to chemical compatibility, the major thermoplastics perform broadly as follows:

1. Polyethylene (PE), depending upon its molecular weight, can resist weak and medium-strength sulfuric acid at temperatures up to ~60 °C, depending on concentration and concentrated acids at ambient temperature.
2. Polypropylene (PP) is more resistant than PE and can tolerate weak acids up to ~90 °C, and medium-strength and concentrated acids at temperatures that reduce with increasing concentration.
3. Polyvinyl chloride (PVC) and chlorinated polyvinyl chloride (CPVC) resist concentrated acids at ambient temperatures and medium and weak acids at temperatures that increase as the concentration decreases.
4. Fluoroplastics are very resistant to sulfuric acid, depending upon their degree of fluorination. The highly fluorinated materials polytetrafluoroethylene (PTFE), fluorinated ethylene propylene (FEP), and perfluoroalkoxy (PFA) are resistant to all concentrations of sulfuric acid and oleum at temperatures up to at least 200 °C. The less highly fluorinated materials ethylene chlorotrifluoroethylene (ECTFE) and ethylene tetrafluoroethylene (ETFE) resist up to 98% (w/w) acid at temperatures up to ~150 °C. Polyvinylidene fluoride (PVDF) also resists up to 98% (w/w) acid but at temperatures that vary from ~120 °C at the lower end of the concentration range to ~50 °C at 98% (w/w) acid.[2,4]

In practice, chemical compatibility is but one factor that determines the suitability or otherwise of a thermoplastic material for a specific application. Other factors, such as mechanical and fabrication properties, thermal expansion characteristics, permeation properties, etc., are significant design considerations. In practice, except for small scale equipment, thermoplastics are used more commonly as linings on stronger substrates rather than in solid form, because of reliability and integrity concerns.

The more common natural and synthetic elastomers are unsuitable for concentrated acids. Thus, butyl rubber and neoprene are generally limited to weak and medium-strength acids, but chlorosulfonated PE elastomers (e.g., Hypalon) can be used in concentrated acids. Some grades of fully fluorinated elastomers (e.g., some Kalrez grades) can be used in

concentrated acids and even oleum, depending on the filler materials.

The common glass reinforced polyester and vinyl ester resins find significant use in the storage and handling of weak acids up to temperatures of ~90 °C and medium-strength acids up to lower temperatures, depending on the acid strength. However, these materials are not suitable for concentrated acids.

Thin paint coatings are unsuitable for the protection of substrates that would corrode at significant rates if exposed. However, some high-performance coatings, for example, heat-cured phenolic coatings, are used to prevent contamination of concentrated acids that are stored in steel containers.

2.23.6.2 Inorganic Materials

Borosilicate glass is highly resistant to acids across the concentration range up to temperatures well beyond their boiling points and solid- or glass-lined equipment finds significant use for handling sulfuric acid.

Impervious graphite, filled with impregnants such as phenolic resins or PTFE, finds significant use in heat exchange duties and is resistant to all but the most oxidizing conditions at temperatures well beyond boiling. Silicon carbide also has an excellent sulfuric acid resistance and has been used in heat exchange duties.

Acid-resistant brick linings on steel are widely used to handle sulfuric acid. Below ~70% (w/w) concentration, corrosion-resistant membranes are required behind the brick, which can be metallic, such as lead, or nonmetallic, such as fluoroplastic linings or asphaltic or other coatings. In concentrated acids, the bricks can be used directly onto steel, although in practice, fluoroplastic or other membranes are often installed. Carbon bricks find occasional use in very hot, concentrated acid duties.

References

1. Hammer, W. J. *The Structure of Electrolyte Solutions*; J. Wiley: London, 1959.
2. Materials Selector for Hazardous Chemicals In *Concentrated Sulphuric Acid and Oleum*, 2nd ed.; MTI Publication MS-1, Materials Technology Institute of the Process Industries Inc: St. Louis, MO, 2005; Vol. 1.
3. DECHEMA Corrosion Handbook. *Sulphuric Acid*; Wiley-VCH: London, 2008; Vol. 11.
4. Brubaker, S. K. H. Corrosion by Sulphuric Acid. In *ASM Handbook Corrosion: Environments and Industries;* ASM International: Materials Park, OH, 2006; Vol. 13C, pp 659–667.
5. NACE Recommended Practice RP0391–2001. *Materials for Handling and Storage of Commercial Concentrated (90–100%) Sulphuric Acid at Ambient Temperatures*; NACE International: Houston, TX, 2001.
6. NACE Standard Practice SP0254–2006. *Design, Fabrication and Inspection of Storage Tank Systems for Concentrated Fresh and Process Sulphuric Acid and Oleum at Ambient Temperatures*; NACE International: Houston, TX, 2006.
7. *The Corrosion Resistance of Nickel-Containing Alloys in Sulphuric Acid and Related Compounds*; Corrosion Engineering Bulletin 1, The International Nickel Company, 1983.
8. Renner, M. H. W. *Metallic Materials for Concentrated Sulphuric Acid Service*, Paper 01342, Corrosion 2001; NACE International: Houston, TX, 2001.
9. Breakell, J. E. *The Corrosive Wear of Cast Irons in Marine Diesel Engine Cylinder Environments*; Ph. D. Thesis, University of Manchester, October, 1980.
10. Williamson, R. C.; Hines, J. G. *Corros. Sci.* **1964**, *4*, 221–235.
11. Mazurkiewicz, B. *Electrochim. Acta* **1993**, *38*, 495–502.
12. Dean, S. W.; Grab, G. D. *Mater. Perform.* **1985**, 21–25.
13. Kish, J. R.; Ives, M. B.; Rodda, J. R. *Corros. Sci.* **2003**, *45*, 1571–1594.
14. Stypula, B.; Banas, J. *Electrochim. Acta* **1993**, *38*, 2309–2314.
15. Stypula, B.; Stoch, J. *Corros. Sci.* **1994**, *36*, 2159–2167.
16. Bera, S.; Rangarajan, S.; Narasimhan, S. V. *Corros. Sci.* **2000**, *42*, 1709–1724.
17. Olsson, C. O. A.; Landolt, D. *Electrochim. Acta* **2003**, *48*, 1093–1104.
18. Li, Y.; Ives, M. B.; Coley, K. S. *Corros. Sci.* **2006**, *48*, 1560–1570.
19. Chang, Y. S. *Periodic Active-Passive Corrosion Behaviour*; Ph. D. Thesis, Fitzwilliam College, Cambridge University, 1984.
20. Mitsuhashi, R. *Corros. Eng.* **1990**, *39*, 89.
21. Renner, M. H. W. *Corrosion Behaviour of Stainless Steels and Nickel Alloys in Hot, Concentrated Static and Flowing Sulphuric Acid*; Ph. D. Thesis, University of Teesside, 1991.
22. Li, Y.; Ives, M. B.; Coley, K. S.; Rodda, J. R. *Corros. Sci.* **2004**, *46*, 1969–1979.
23. Hermas, A. A.; Ogura, K.; Adachi, T. *Electrochim. Acta* **1995**, *40*, 837–844.
24. Pardo, A.; Merino, M. C.; Coy, A. E.; Viejo, F.; Arrabal, R.; Matykina, E. *Corros. Sci.* **2008**, *50*, 780–794.
25. Bojinov, M.; Fabricius, G.; Laitinen, T.; Makela, K.; Saario, T.; Sundholm, G. *Electrochim. Acta* **2001**, *46*, 1339–1358.
26. Phelps, E. H.; Vreeland, D. C. *Corrosion* **1957**, *13*, 619–624.
27. Berglund, G.; Martenson, C. Application of a High Alloyed Stainless Steel in Sulphuric Acid Environments; Paper 21, Corrosion '87; NACE International: Houston, TX, 1986.
28. Kiefer, G. C.; Renshaw, W. G. *Corrosion* **1950**, *6*, 235–244.
29. Fink, C. G.; DeCroly, C. M. *Trans. Am. Electrochem. Soc.* **1929**, *56*, 239.
30. Friend, W. Z. *Corrosion of Nickel and Nickel-Base Alloys*; John Wiley & Sons: New York, 1980.
31. Trabanelli, G.; Zucchi, F.; Felloni, L. *Corros. Sci.* **1965**, *5*, 211.
32. Greene, N. D. *The Passivity of Nickel and Nickel-Base Alloys*; Proceedings of the 1st International Congress on Metallic Corrosion; Butterworth: London, 1962.
33. Myers, J. R.; Beck, F. H.; Fontana, M. G. *Corrosion* **1965**, *21*, 277.
34. Hoppe, H. W.; Strehblow, H. H. *Surf. Interface Anal.* **2004**, *14*, 121–131.

36. Blackwood, D. J.; Peter, L. M. *Electrochim. Acta* **1989**, *34*, 1505–1511.
37. Banker, J. G. *Hydrometallurgical Applications of Titanium Clad Steel*; Conference Reactive Metals in Corrosive Applications, Sun River, OR, Sept. 12–16 1999.
38. Hofmann, W. *Lead and Lead Alloys: Properties and Technology*; Vibranos, G., Ed.; Springer: Verlag, 1970; pp 274.
39. Yau, Te-Lin *Understanding Corrosion Behavior From Electrochemical Measurements*, Paper 04227, Corrosion 2004; NACE International: Houston, TX, 2004.
40. *Handbook of Corrosion Data*; ASM International: Materials Park, OH, 1995.
41. Yau, Te-Lin *Effects of Impurities in H_2SO_4 on the Corrosion Resistance of Zirconium*; ASTM STP 830, ASTM, 1984, pp 203.
42. Fitzgerald, B. J.; Yau, Te-Lin *The Mechanism and Control of Stress Corrosion Cracking of Zirconium in Sulfuric Acid*, Paper 092, 12th International Corrosion Congress, Houston, TX, September 19–24 1993.
43. Shimizu, K.; Brown, G. M.; Habazaki, H.; Kobayashi, K.; Skeldon, P.; Thompson, G. E.; Wood, G. C. *Corros. Sci.* **1998**, *40*, 963–973.
44. Coscia, M.; Renner, M. H. W. *Mater. Perform.* **1998**, *37*, 52–57.

2.24 Corrosion in Nitric Acid

G. O. H. Whillock and S. E. Worthington
National Nuclear Laboratory Ltd., B170, Sellafield, Seascale, Cumbria CA20 1PG, UK

Glossary

18/13/1 A niobium-stabilized stainless steel used in UK nuclear reprocessing plant, conforming approximately to the top end of the AISI 347 specification range.

End grain The cross-section perpendicular to the hot-working direction of a fabricated component.

Grain dropping Dislodgement of metal grains as they are undermined by intergranular

corrosion. Dropped stainless steel grains have the appearance of fine soot.

Magnox A magnesium alloy containing 0.8% aluminum, used to clad uranium metal fuel for CO_2-cooled Magnox reactors.

NAG 18/10L A proprietary stainless steel used in UK nuclear reprocessing plant, similar to AISI 304L but with tighter control of residuals.

Radiolysis Chemical change resulting from the absorption of ionizing radiation.

R-SUS 304ULC A proprietary stainless steel used in Japanese nuclear reprocessing plant. Similar to NAG 18/10L.

20Cr25Ni1Nb A proprietary stainless steel used in UK Advanced Gas-Cooled Reactor fuel.

Uranus 65 A proprietary stainless steel, conforming approximately to AISI 310L, but with added niobium. Used in UK and French nuclear reprocessing plant.

Abbreviations

AISI American Iron and Steel Institute
AR As-received
Gy Grays (unit of absorbed radiation dose; $1\ Gy = 1\ J\ kg^{-1}$)
mpy Thousandths of an inch per year
M Molar or metal
NAG Nitric acid grade
SCC Stress corrosion cracking
ss Stainless steel

Symbols

E Electrode potential
i Current density
V_{SCE} Electrode potential, volts with respect to the Standard calomel electrode
X(Y) Element X present in valency state Y; for example, Cr(VI) is $Cr_2O_7^{2-}$ or CrO_4^{2-}

2.24.1 Uses of Nitric Acid

Nitric acid is used in salt formation, preparation of organic nitrates and nitro compounds, and oxidation of organic compounds.[1] Major applications include the production of agricultural fertilizers and reprocessing of nuclear fuels.

2.24.2 Chemical Attack by Nitric Acid

2.24.2.1 Pure Nitric Acid

At greater than about 96–97% concentration, HNO_3 chemistry is dominated by the NO_2^+ ion. The acid undergoes no proton transfer to water molecules and it cannot be considered as an acid in the protionic sense at this concentration.[2]

The initial electron transfer reaction with a metal, M:

$$M + NO_2^+ \rightarrow M^+ + NO_2 \qquad [1]$$

leads to the overall reaction

$$M + 2HNO_3 \rightarrow MNO_3 + NO_2 + H_2O \qquad [2]$$

with nitrates produced at the metal surface.[2]

The reaction of some metals with pure HNO_3 is given in Table 1.

There is no correlation between apparent reactivity and position of the metal in the periodic table. In some instances, there is evidence of cohesive film formation (e.g., Ca) while some metals are either inert (e.g., Pt) or form protective oxide films (e.g., Al).

2.24.2.2 Aqueous Nitric Acid

In practice, most nitric acid corrosion occurs in aqueous mixtures of <96% concentration.

2.24.2.2.1 Base metals in aqueous nitric acid

These generally dissolve extremely rapidly to form nitrates, cathodically discharging hydrogen, which reacts with the nitric acid to give compounds such as ammonia, ammonium nitrate, hydroxylamine nitrate, nitrogen, and nitrous oxide.[3]

2.24.2.2.2 Noble metals in aqueous nitric acid

The dominant cathodic reaction in this case is the reduction of the nitrate ion. This is a complex process, proceeding via an autocatalytic mechanism involving a charge transfer reaction and various chemical reactions leading to the regeneration of the active species (nitrous acid). The autocatalytic nature of the reaction can readily be demonstrated by simple observation of the corrosion rate of copper, for example, in nitric acid solutions: the rate, as evinced by NO_x evolution, is initially low, but accelerates rapidly if the copper is placed in contact with a surface, thus forming a crevice, which traps the nitrous acid generated, or if a little nitrite is added

Table 1 Reaction of pure HNO_3 with metals

Solution complete within			*Evidence of inhibition by film*	*No evidence of reaction*
1 day	*1 week*	*1 month*		
Mg	Cr powder	Fe	Ca	Be
Sc		Mn	Cu	Al
V	Ni		Ag	Ti
Co	Cd		Bi	Pt
Zn	U			Sn
Hg				

Reprinted with permission from Addison, C. C. *Chem. Rev.* **1980**, *80*, 21–39. Copyright 1980 American Chemical Society.

to the solution; conversely, the rate is retarded by agitation or stirring of the acid, which displaces the nitrous acid generated, or by the addition of a chemical such as urea, which impedes the autocatalytic reaction by destroying nitrous acid.[3]

Two basic mechanisms have been proposed: one heterogeneous[4] leading to NO_2 formation and one homogeneous[5] leading to NO. Both mechanisms have been shown to be consistent with experimental data collected over different potential ranges.[6] More recently, a modified mechanism has been suggested, which results in two limiting cases, one applicable for low nitric acid concentrations (<8 M) leading to NO formation, the other for high nitric acid concentrations (>8 M) leading to NO_2.[7] According to this scheme, the major species involved in the reduction of nitric acid at low overvoltages are nitrous acid (HNO_2), NO, and NO_2, and the charge transfer reaction is the reduction of HNO_2:

$$HNO_2 + H^+ + e^- \rightarrow NO + H_2O \qquad [3]$$

The product of the overall cathodic reaction is HNO_2, thus giving rise to autocatalysis:

$$3H^+ + NO_3^- + 2e^- \rightarrow HNO_2 + H_2O \qquad [4]$$

Whatever the precise mechanism, the important point is that nitrate reduction is autocatalytic, being strongly affected by HNO_2, NO, and NO_2. Reactions [3] and [4] indicate that H^+ is required, but in all but the most dilute solutions it is the nitrate concentration that is important; thus, the addition of neutral nitrate salts to nitric acid generally increases the liquor corrosivity.

Iron can show both base and noble metal characteristics in its corrosion reactions. It is attacked at an increased rate in crevices, but its corrosion products include nitrous oxide, ammonia, and nitrogen as well as nitrogen dioxide and nitric oxide. In concentrated acids of >~40%, after an initial reaction, iron becomes passive (although this can be disrupted relatively easily, by agitation, for example), whereas in dilute acids, reactions are rapid.[8]

2.24.2.2.3 Passive metals

Highly reactive metals such as chromium, titanium, hafnium, niobium, and zirconium rapidly form thin insoluble surface oxides in air. These films may grow slightly on immersion in oxidizing acids, but then prevent further corrosion or reduce it to negligible rates in a wide range of nitric acid conditions. These metals or their alloys, as in the case of stainless steels, where alloying of iron with chromium confers much greater passivity, are the most important class of metals for the industrial handling of nitric acid.

This behavior is illustrated for stainless steels in **Figure 1**.[9] The polarization curve shows three distinct regions of active, passive, and transpassive corrosion. The effect of temperature on the anodic reaction is important, as demonstrated in **Figure 2**. Although increasing the temperature of stainless steel in nitric acid produces only a small increase in corrosion potential (due to augmentation of the cathodic reaction), the anodic polarization curve changes significantly; the passive region is effectively lost, and therefore, a small increase in potential leads to significantly increased corrosion. This is pertinent when considering the effect of oxidizing species or galvanic effects, which will therefore be more pronounced at high temperatures (see **Sections 2.24.4.7** and **2.24.5.4.1**).

2.24.3 Materials for Nitric Acid Containment

The primary materials used for nitric acid environments are stainless steels, titanium, and zirconium. Aluminum can also be used in some circumstances. The austenitic stainless steel, AISI 304L, and its

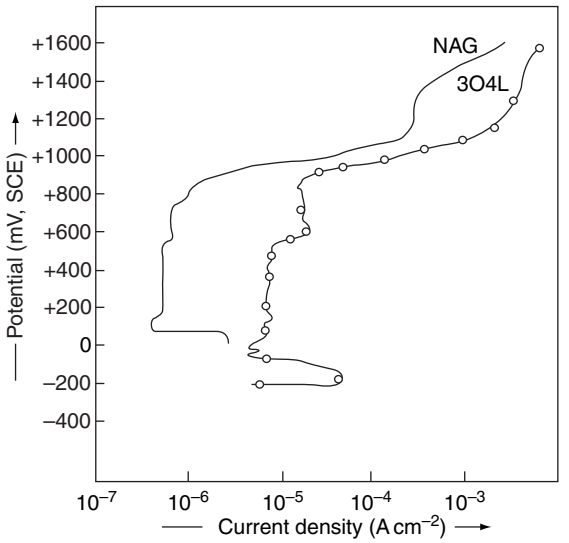


Figure 1 Anodic polarization of stainless steel in 1 M nitric acid at room temperature. Reproduced from Kain, V.; Shinde, S.; Gadiyar, H. J. *Mater. Eng. Perform.* **1994**, *3*, 699–705, with permission from ASM International.

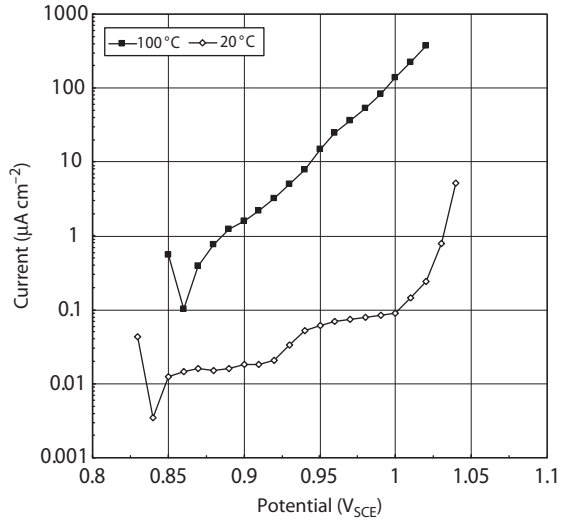


Figure 2 Effect of temperature on polarization behavior of NAG 18/10L stainless steel in 8M nitric acid (measured by step-wise potentiostatic polarization, 10 mV steps each of 1 h duration). Takeuchi, M.; Whillock, G. O. H. BNFL unpublished work.

variants are now the most commonly used, although much early plant has been made from stabilized austenitic steels such as AISI 321, AISI 347, and 18/13/1. In more oxidizing conditions, higher alloy steels such as AISI 310L or possibly high Si stainless may be preferred. In the most aggressively oxidizing conditions, titanium, zirconium, niobium, hafnium, or tantalum may be used. In the past, titanium has been the most widely used, but the use of zirconium in preference has increased in recent years. The others find only occasional use, despite their excellent corrosion resistance, because of high cost and fabrication difficulties. Table 2 provides indicative corrosion rates for some of these metals.

Table 2 Indicative corrosion rates of various metals subject to testing in boiling 65% nitric acid

Metal	*Corrosion rate (mm y^{-1})*
18/13/1 stainless steel[a]	0.2–0.5
NAG 18/10L stainless steel[b]	0.1–0.2
NAG 18/10L stainless steel[c]	0.1–0.3
310L stainless steel	0.05–0.1
Ti[d]	0.7
Zr	<0.001
Ta	<0.001

[a]Similar to AISI 347.
[b]As-received.
[c]Sensitized.
[d]The high corrosion rate is due to the test conditions employed (see Section 2.24.3.2).
Source: Quayle, B. E. BNFL unpublished work.

2.24.3.1 Stainless Steels

The industrial use of stainless steel for nitric acid containment seeks to ensure that it operates in the passive regime where corrosion rates are typically <0.01 mm year^{-1}. However, it is also possible to operate stainless steel plant successfully for extended periods in the potential regime between passivity and full transpassivity where intergranular corrosion occurs and the corrosion rate is significant (see Section 2.24.4.1). In practice, all stainless steel plant handling hot nitric acid operates in the intergranular corrosion regime.

Corrosion rates show Arrhenius behavior over a large range of acid concentrations and temperatures, as illustrated in Figure 3, indicating that electron transfer is the rate-determining step in the corrosion mechanism.

Austenitic stainless steels are the most widely used for nitric acid containment. Reasonable or good nitric acid resistance can be achieved with ferritic, super ferritic, and ferritic–austenitic (duplex) steels, but they are generally less resistant than the austenitics, and factors such as compositional variation, heat treatment, and welding can lead to reduction in corrosion resistance.[11] Ferritic stainless steels with >14% Cr resist corrosion in nitric acid at room temperature, and high Cr/Mo ferritics such as 26Cr1Mo, and 29Cr4Mo resist nitric acid up to 65% boiling. Duplex

stainless steels have similar resistance to ferritics, but the ferritic phase can be preferentially attacked.[11]

AISI 300 series stainless steels (i.e., austenitic) are adequate for a wide range of applications up to about 70% acid concentration. AISI 304L is now commonly used and improvements in its performance can be achieved by improved manufacturing processes to reduce elements such as Si, P, S, and C, as illustrated in Figure 1. Here, a nitric acid grade version of 304L shows improved corrosion resistance in the passive and lower transpassive region. However, it is doubtful whether significant corrosion differences persist under more aggressive oxidizing conditions where higher chromium steels such as 310L may be required.[12] Alternatively, in extremely oxidizing conditions stainless steels containing Si at >∼4% (e.g., Uranus S1N) may find application, although such steels are more problematic to join because some ferrite needs to be maintained in the welds to avoid cracking.[13]

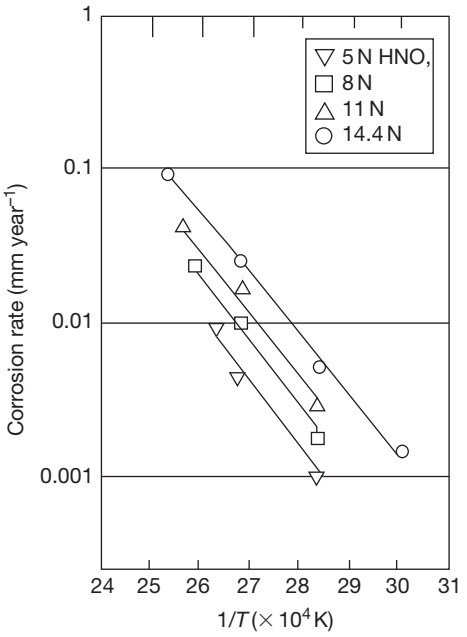


Figure 3 Corrosion rate of stainless steel (R-SUS 304ULC) as functions of temperature and nitric acid molarity. Reproduced from Onoyama, M.; Nakata, M.; Hirose, Y.; Nakagawa, Y. In *Proceedings of the RECOD'91*, Japan Atomic Industrial Forum: Japan, 1991; Vol II, pp 1066–1071.

2.24.3.2 Titanium

Titanium has a high corrosion resistance to a range of nitric acid concentrations, provided there is sufficient concentration of dissolved titanium or other oxidizing ions present.[14,15] In the absence of such dissolved ions, corrosion rates can reach or exceed 1 mm year^{-1}. Hence, corrosion rates in the vapor regions of operating plant handling condensing nitric acid, where continuous washing prevents the accumulation of dissolved titanium, can be significant even though the metal below liquor is unaffected. Titanium is accordingly no longer recommended for use in condensing nitric acid. Owing to the sensitivity of titanium corrosion rates to the presence of dissolved titanium, account must be taken of the suppressing effect of corrosion product accumulation when assessing corrosion rate data. The susceptibility of titanium to corrosion in pure acid can be mitigated by alloying with tantalum; at >5% Ta, corrosion is significantly reduced.[14] These effects are illustrated by the corrosion data in Table 3.

Table 3 Effect of test conditions on the corrosion rate of titanium and titanium–tantalum alloys

Metal	*Nitric acid concentration plus additions*	*Liquor renewal (h)*	*Corrosion rate (mm year^{-1})*	*Source*
Ti	6 M	14	0.6	(a)
	6 M	65	0.43	(b)
	6 M	280	0.15	(a)
	6 M + 0.1% Fe^{3+}	65	0.14	(b)
	6 M + 0.1% Ce^{4+}	65	<0.01	(b)
	40%	24	0.75	(c)
	40% + 40 mg l^{-1} Ti	24	0.05	(c)
	Condensing acid	24	0.33	(c)
	40%	4	0.74	(c)
Ti–6.1%Ta	40%	4	0.02	(c)

[a]Quayle, B. E. BNFL unpublished work.
[b]Shimogori, K.; Satoh, H.; Kamikubo, F. In *Titanium Science and Technology*; Proceedings of the 5th International Conference on Titanium, Munich, Germany, 10–14 Sept; 1984; Luthering, G., Ed., Vol. 4, pp 1111–1114.
[c]Takamura, A.; Arakawa, K.; Moriguchi, Y. In *The Science Technology and Application of Titanium*; Jafee, R. I., Promisel, N. E., Eds.; Pergamon Press: Oxford, 1970; pp 209–216.

Titanium can also be susceptible to enhanced corrosion in both weld metal and heat-affected zones. This has generally been attributed to the redistribution of iron impurities during the welding operation so that the iron originally present in discrete stringers or well-distributed particles is deposited on grain boundaries either as iron–titanium intermetallic compounds or as β-phase. The formation of a continuous grain boundary network can lead to intergranular corrosion and grain dropping in certain conditions. This is likely to become more significant in marginally passive environments where these iron-rich phases may become soluble. To avoid this problem, it is now recommended that iron levels in titanium are specified to be <0.05%.[16]

2.24.3.3 Zirconium, Niobium, Hafnium, and Tantalum

These materials are generally highly resistant to nitric acid corrosion (see **Table 2**), although operating experience of these materials is less extensive than that of titanium. There are no reported materials issues affecting corrosion performance, although stress corrosion cracking (SCC) of zirconium cannot be entirely discounted in extremely oxidizing conditions (see **Section 2.24.4.4**). In addition, zirconium is very susceptible to trace levels of fluoride ions (see **Section 2.24.5.5**). There have been no reported instances of enhanced vapor phase corrosion and they do not require the presence of dissolved ions to achieve passivity.

2.24.3.4 Aluminum

The corrosion rate of aluminum and its alloys is appreciable in all but very concentrated nitric acid, reaching 3–5 mm year^{-1} at room temperature in 30% acid, whereas rates of <0.1 mm year^{-1} can be achieved in >85% acid. This is illustrated in **Figure 4**[17]. The corrosion rate is strongly temperature dependent and increases with the level of copper impurity. However, a number of aluminum alloys containing magnesium silicon or manganese have corrosion rates comparable with those of pure aluminum in 98.5% acid.[18] As would be expected from its decreased corrosion resistance in dilute acids, aluminum corrodes more quickly where dilute acid condensates are formed above concentrated acids.[19] Hence, although aluminum can be suitable for room temperature storage of concentrated nitric acid, the vapor regions of storage vessels may undergo increased corrosion.

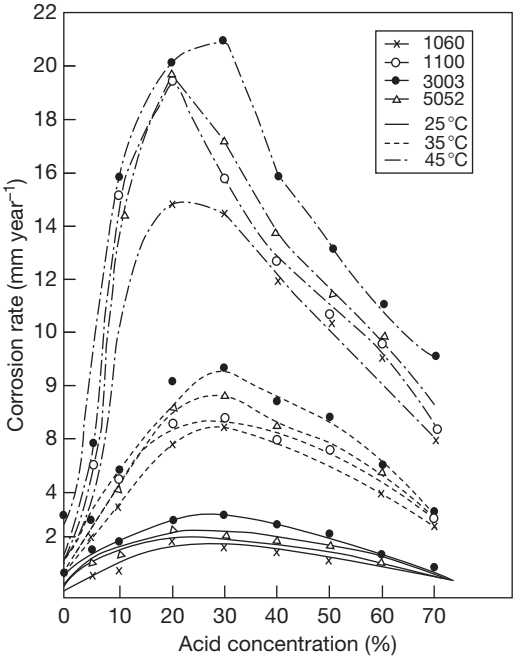


Figure 4 Corrosion rate of various aluminum alloys as a function of nitric acid concentration and temperature. Reproduced from Singh, D. D. N.; Chaudhary, R. S.; Agarwal, C. V. *J. Electrochem. Soc.* **1982**, *129*, 1869–1874, with permission from The Electrochemical Society.

2.24.4 Corrosion Mechanisms of Passive Metals in Nitric Acid

2.24.4.1 General Corrosion and Intergranular Corrosion

For stainless steels, full transpassivity occurs when the Cr_2O_3-based passive film becomes unstable in extremely oxidizing conditions, allowing corrosion to proceed by the formation of soluble CrO_3; this occurs at ~1.0 V_{SCE}.[20] However, grain boundaries are more vulnerable to corrosion. This is largely due to impurities in the metal, with the segregation of silicon and phosphorus being particularly implicated.[21,22] The transpassive corrosion process therefore proceeds primarily by grain boundary dissolution, but is manifest as general corrosion owing to grain dropping. Owing to the high corrosion rates encountered, stainless steels are unsuitable for use in extremely oxidizing conditions where transpassive corrosion occurs.

In the potential band between full passivity and transpassivity (i.e., between ~0.9 and ~1.0 V_{SCE}), intergranular corrosion occurs. This is the principal mechanism affecting stainless steels handling nitric

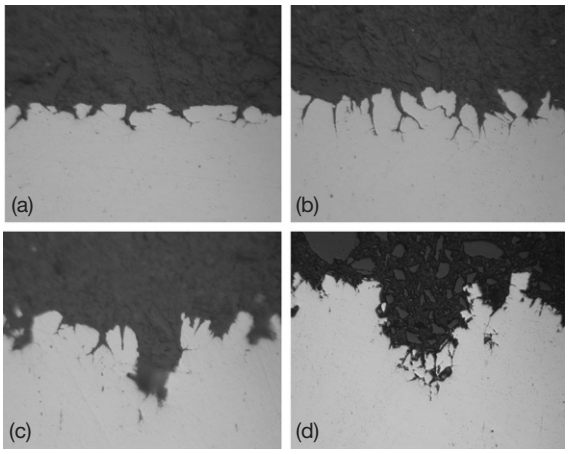


Figure 5 Morphology of intergranular corrosion of 18/13/1 stainless steel exposed to 8 M nitric acid containing 25 g l^{-1} Fe(III) at 100 °C. The sections show penetration into (a) plate, (b) side and (c) end surfaces after 340 h exposure. (d) Shows end grain pitting after 1000 h exposure. Magnification: (a)–(c) ×360; (d) ×150. Dunnett, B. F. BNFL unpublished work.

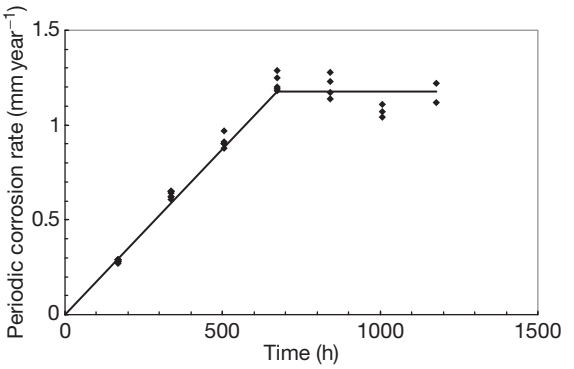


Figure 6 Effect of immersion time on the corrosion rate of 18/13/1 stainless steel in 8 M nitric acid containing 25 g l^{-1} Fe(III) at a solution temperature of 100 °C. Reproduced from Dunnett, B. F.; Whillock, G. O. H. *Corrosion* **2003**, *59*, 274–283, with permission from NACE International.

acid solutions which are too oxidizing for the maintenance of passivity. Intergranular attack increases significantly if the steel has become sensitized by heat treatment,[23] but sensitization is not a prerequisite, intergranular corrosion occurring in the fully solution-annealed condition even in very clean stainless steels providing potentials in excess of ~0.9 V_{SCE} are achieved in hot solutions. The intergranular corrosion rate can be strongly anisotropic, depending on the orientation of the exposed face to the underlying microstructure (see also **Section 2.24.4.2**). **Figure 5** illustrates this. The effects tend to be less pronounced in modern ultraclean stainless steels, but are still evident.

General corrosion curves derived for stainless steels from weight loss are broadly characterized by two regions, as illustrated in **Figure 6**.[24] In the early stages, grain boundaries are etched, but no grains are lost, and weight loss is by chemical dissolution only. The apparent corrosion rate rises as the grain boundaries are penetrated more deeply because a greater area of metal is opened to corrosion. As time progresses, grains begin to drop out of the metal surface as their boundaries are completely corroded. Eventually, a relatively stable corrosion rate is reached, which is a combination of chemical dissolution and grain dropping. For AISI 304L stainless steel and related steels, grain dropping accounts for only ~25% of the total metal loss, chemical dissolution of the grain boundary regions accounting for the rest.[25] It is therefore important in deriving corrosion rates from analysis of dissolved species to allow for grain dropping.

Corrosion rates derived from weight loss averaged over the specimen area will tend to be underestimated in the early stages of corrosion. In practice, the corrosion curve is an artifact of the measurement process; if measurements of penetration are made, then it is found that the rate of penetration is constant and corresponds to the corrosion rate derived from weight loss once steady grain dropping is established.[24] The time to reach this state depends on the corrosion resistance of the steel, its grain size, and the corrosivity of the medium.

This problem can be addressed either by numerical models to predict the long-term rate from short-term data[26,27] or by pre-corroding test samples in aggressive media to establish grain dropping before reverting to the required test media.[24,28] The validity of numerical models is still to be fully established. Pre-corrosion appears to be a valid method, although it appears that care needs to be taken in selecting the pre-corrosion conditions, otherwise reasonably prolonged testing in the liquor of interest may be required in order to establish the true corrosion rate. In practice, it is found that the overall metal loss by intergranular corrosion is fairly uniform even in plants operating with relatively aggressive environments. Hence, extrapolation of the steady-state corrosion rate to predict plant life is valid, provided there are no localized mechanisms acting, such as crevice/underdeposit corrosion (see **Section 2.24.4.3**).

Intergranular corrosion in stainless steels can be suppressed by alloying with silicon.[13] At >~4% Si, intergranular corrosion is suppressed, but general

corrosion rates at most nitric acid concentrations are higher at least in the early stages of corrosion compared with the equivalent AISI 300 series stainless steel. In particular, high Si stainless steels show a significant improvement in corrosion resistance in super-azeotropic solutions. However, welding of these alloys can be problematical; ferrite needs to be maintained in the welds to avoid cracking, and welds can be susceptible to enhanced corrosion.[13]

Intergranular corrosion of other passive metals such as zirconium and tantalum has not been reported. Of these metals, only titanium shows a transpassive breakdown in normal operating conditions, although a form of transpassivity (breakaway oxidation) can be induced in zirconium and its alloys under extreme conditions (in excess of $\sim 2\ V_{SCE}$).[29] Titanium can also exhibit intergranular corrosion at welds; this propensity tending to be associated with the presence of β-phase (see Section 2.24.3.2).

2.24.4.2 End Grain Corrosion

The rate of intergranular penetration depends on the orientation of the exposed face to the underlying microstructure. In general, the corrosion rate increases in the order: plate < side < end (see Figure 5). The effects of end grain corrosion are clearly seen in boiling nitric acid testing (the Huey test, see Section 2.24.6) if specimens are cut so as to vary the percentage of exposed end grain in a susceptible material. Figure 7 shows an example, clearly demonstrating a significant effect for end grain fractions exceeding ~10%.

In addition to this pronounced effect of end grain on the general corrosion rate, localized pitting-like attack can develop parallel to the hot-working direction. This is known as 'end grain' or 'tunneling' corrosion. Figure 8 shows an example. Older stabilized steels appear to be most vulnerable, and the phenomenon is extremely variable even within the same grade of steel. However, extremely clean stainless steels such as R-SUS 304ULC and Uranus 65 also exhibit end grain pitting.[30,31] In such steels, the mechanism is considered to be a consequence of the exposure of segregated material in the original ingot, which has been distributed, producing bands with more susceptible grain boundaries.[31] In stabilized steels, stringers of magnesium sulfide and precipitates of niobium or titanium carbide mark out the regions of maximum segregation as they solidify late from the melt. The outer regions of the ingots solidify first, and hence contain little impurity segregation conferring superior corrosion resistance to the as-formed plate surface.

In practice, end grain corrosion of stainless steels is usually more of a problem for corrosion testing (see Section 2.24.6) than for operating plant where most end grain is welded over in pipe or plate joints, although issues can arise, for instance, with forgings where outcropping end grain can be present.

End grain corrosion of other passive metals has not generally been reported, although at least one instance on titanium is known.

2.24.4.3 Crevice Corrosion

Although crevice corrosion is not widely reported in nitric acid, austenitic stainless steels are known to be susceptible to this mechanism in certain circumstances. One such example is corrosion testing of heat transfer surfaces, where corrosion focuses at

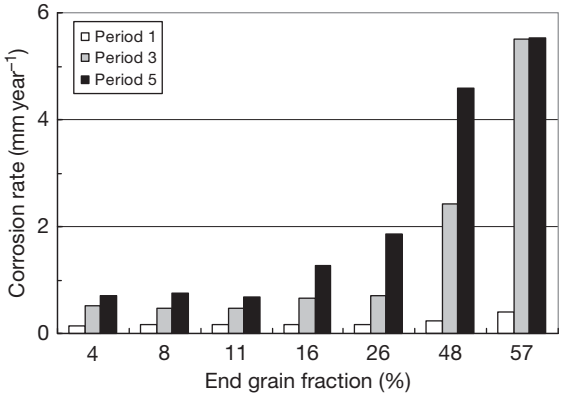


Figure 7 Effect of end grain fraction on the corrosion rate obtained from successive 48 h test periods of 18/13/1 stainless steel in boiling 65% nitric acid. Quayle, B. E. BNFL unpublished work.

Figure 8 End grain corrosion pits: 18/13/1 stainless steel stirrer paddle in nitric acid media. Green, D. S. BNFL unpublished work.

the mouth of the crevice on the external surface; for instance, under nucleate boiling at a heat flux of 50 kW m^{-2}, the depth of attack was twice that on the free surface.[32] Crevice corrosion appears to be possible only in more aggressive conditions, although indications of it have been found in isothermal tests at temperatures as low as 90 °C,[33] and in a nitric acid reboiler where corrosion rates increased from ~0.02 to 0.1 mm $year^{-1}$ on a heat transfer surface at ~90 °C in 12 M nitric acid. The mechanism may have several origins depending on the circumstances. For heat-transfer surfaces, the inhomogeneity provided by the crevice may simply lead to increased local surface temperatures, and hence increased local corrosion rate. The entrapment of stainless steel corrosion products is also possible (see **Section 2.24.5.4.1**), as is entrapment of NO_x species (see **Section 2.24.5.3**). The mechanism is oxidative rather than reductive in contrast to crevice corrosion brought about by chloride ions in water systems. The operation of a crevice in nitric acid systems is capable of adversely affecting the potential outside the crevice and increases the external corrosion rate. Care is therefore required in interpreting corrosion test data where the operation of a crevice is suspected.

Crevice corrosion has not been found on zirconium, titanium, or niobium under similar heat-transfer conditions in nitric acid. However, it has been recorded on zirconium in laboratory trials in acid contaminated by fluoride ions (see also **Section 2.24.5.5**) and also under cathodic polarization. It is probable that here the phenomenon is one of active corrosion in localized reducing conditions in the crevice. Crevice corrosion has not been reported for these metals in operating plant.

2.24.4.4 Stress Corrosion Cracking

Despite the intergranular nature of nitric acid attack of stainless steels, SCC of stainless steels in nitric acid is almost unknown. This is not due to a simple absence of stress, as postfabrication stresses at welds in austenitic stainless steels invariably approach yield. SCC is known to occur in weak nitric acid solutions and is facilitated by severe sensitization, the only known example in practice being intergranular SCC of irradiated stainless steel nuclear fuel cladding exposed to moist air, the corrosive agent being generated by nitrogen radiolysis (see **Section 2.24.5.6**). Slow strain rate testing reveals a link between the oxidizing power of the acid (i.e., the repassivation rate) and the metal rupturing rate (i.e., the strain rate). SCC can occur in a critical regime where repassivation kinetics are such that the crack tip remains active; otherwise, crack blunting occurs at one extreme, with the absence of sufficient crack advancement at the other. The absence of SCC of stainless steel in stronger nitric acid solutions is accordingly attributed to too rapid repassivation rates relative to practically achievable strain rates.

SCC has also has been induced in zirconium by slow strain rate testing.[34] However, this only occurs above the azeotrope and is therefore unlikely to be relevant other than in extremely oxidizing conditions. SSC of zirconium has not been reported in plant operating in nitric acid media.

2.24.4.5 Corrosion Fatigue

Normal air fatigue processes can be exacerbated in the presence of corrosive media, leading to corrosion fatigue. As in SCC, the process is complex, requiring conditions where a crack tip is continuously formed by fatigue and its propagation is accelerated by corrosion of the exposed metal. However, in passivating media, crack blunting is also possible.[35] Test rig data indicate that corrosion fatigue of stainless steel in nitric acid solutions is feasible in some circumstances, indicating that careful assessment may be required. However, corrosion fatigue in nitric acid plant appears not to have been encountered in practice.

There is no indication that zirconium or similar passive metals are susceptible to corrosion fatigue in nitric acid.

2.24.4.6 Erosion Corrosion and Fretting Corrosion

These mechanisms proceed in passive metals by removing the protective oxide film, leading to oxidation of the underlying reactive metal. Erosion corrosion requires the surface oxide to be scoured by a slurry of abrasive particles. The erosion corrosion rate depends on a variety of factors, including particle weight and sharpness, but is especially sensitive to particle velocity, with attrition rates typically increasing as $(\text{velocity})^n$ where n is ~4.[36] A dynamic process ensues in which oxide is continuously removed and regrown. In metal–environment combinations where transient metal dissolution rather than simple repassivation of the scoured surface is possible, considerable enhancement of metal loss can occur.[37] For stainless steel abraded in water and in dilute nitric acid, the attrition rates have been found to be similar.

Although repassivation kinetics will differ for the two media, they will be rapid in both cases, and therefore, unless the erosion rate approached the repassivation rate, it is probable that full repassivation occurs at any given point before it is re-eroded, and therefore, no difference between the attrition rates in the two media is apparent.

It is, however, anticipated that a corrosive element will be more significant in erosion corrosion at transpassive potentials or where nitric acid is contaminated with halides (see Section 2.24.5.5) or other species, which lead to active dissolution of bared metal.

Fretting corrosion is possible when loaded surfaces move against each other in a corrodant. The process depends on loading, displacement, and cycling frequency and tends to be exacerbated if debris remains between the moving faces. Some studies in nitric acid show high attrition rates, but the relative contribution of wear and corrosion was not established and it is likely that similar processes and considerations to erosion corrosion apply.

2.24.4.7 Galvanic Corrosion

Galvanic corrosion effects can occur in nitric acid in much the same way as other media. The consequence depends on the strength and direction of the resulting potential shift and may thus be beneficial, neutral, or deleterious to the half-couple of interest. Coupling with an actively corroding base metal, such as Magnox or uranium, tends to reduce the corrosion rate of stainless steels in hot nitric acid solutions. The behavior of aluminum is more complex, ranging from slight reduction in the corrosion rate of stainless steel in concentrated nitric acid solutions,[38] to marked enhancement in the stainless steel corrosion rate if the potential shift is sufficient to move stainless steel into its active region.[39] Caution is therefore needed in predicting the effect of coupling with base metals. Deleterious effects on stainless steel arise from coupling with lead[40] and noble metals, such as, for example, platinum catalyst fragments (fertilizer plant) or fission product solids rich in platinum metals (nuclear reprocessing plant). Less obvious is the effect of coupling stainless steels with markedly different chromium contents; coupling 18% and 25% Cr stainless steels (i.e., 304L and 310L) results in significant increase in corrosion rate of the former and decrease in corrosion rate of the latter in hot nitric acid solution. This arises because the corrosion potential of 25% Cr stainless steel is ~20 mV higher than that of 18% Cr stainless steel in hot nitric acid, illustrating the extreme sensitivity of corrosion rate to potential close to the onset of full transpassivity. In less oxidizing liquors, the effects of coupling noble metals to stainless steel are much less pronounced.[41]

Connections between other passive metals tend not to be problematic. For example, connections between stainless steel and zirconium plant have minimal effects on the corrosion of the junction below transpassivity; similarly, measurements on titanium/stainless steel couples show no enhanced corrosion.[42]

2.24.5 Factors Affecting Corrosion Processes in Nitric Acid

2.24.5.1 Welding

Sensitization can occur during welding of stainless steels even if they are stabilized or low carbon grades, although the effects are much less pronounced than is the case for nonstabilized or high carbon steels. The formation of chromium-rich carbides at grain boundaries leaves chromium-depleted steel in their vicinity, creating a band of corrosion susceptible material running parallel to the weld but displaced a few millimetres from it. Such zones are more susceptible to intergranular corrosion in hot nitric acid, leading to heat-affected zone corrosion (weld decay). In niobium-stabilized stainless steels, the enhancement appears to be modest, amounting to at most ~10% of the parent metal corrosion rate. Figure 9 shows an example.

Knife-line corrosion is also possible in some circumstances. This is manifest as a sharp band of corrosion penetrating along the weld–parent metal interface where temperatures exceed 1150 °C during

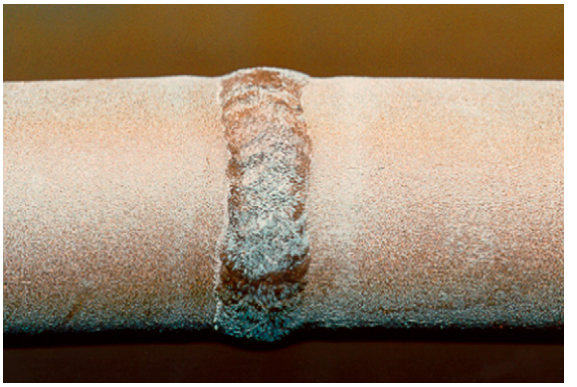


Figure 9 Weld metal and heat affected zone corrosion: 18/13/1 stainless steel tube (steam-heated) in nitric acid media. Green, D. S. BNFL unpublished work.

the welding process. In most nitric acid solutions, this is an issue primarily for stabilized stainless steels such as AISI 321, since titanium carbides are susceptible to dissolution releasing carbon into solid solution, which then reacts with chromium, resulting in a chromium-deficient narrow band at the weld interface. Niobium-stabilized stainless steels such as AISI 347 and 18/13/1 do not exhibit knife-line corrosion in most nitric acid solutions, but are susceptible in mixed acid systems such as nitric acid/fluoride.

In general, welds in austenitic stainless steels are at least as corrosion resistant in nitric acid as the parent metal (see **Figure 9**). Their coarse-grained structure may delay the onset of grain dropping relative to the parent material, but overall penetration rates are similar. Occasionally preferential attack of weld roots may be seen.

Titanium welds can be susceptible to enhanced corrosion, both the weld metal and the heat-affected zone, in hot nitric acid solutions (see **Section 2.24.3.2**).

2.24.5.2 Cold Work

Most information on the effects of cold work on corrosion in nitric acid relates to austenitic stainless steels. The effect of cold work is intimately linked to the thermal and mechanical history of the steel and can lead to an increase or reduction in corrosion rate depending on the circumstances. Data can appear to be contradictory, and the mechanisms by which cold work affects corrosion are not fully explained, although processes such as carbon diffusion, precipitation, and martensite formation appear to be involved.[43,44] The relevance of the process by which cold work is achieved appears not to have been assessed and both compression by cold rolling, and tension by tensometer have been used in studying the phenomenon.

Cold work on the as-received material has generally been found to produce little effect or a marginal decrease in corrosion resistance; steels with an intrinsically lower corrosion resistance (including high Si stainless steels) and greater susceptibility to sensitization tending to fall into the latter category.[43,45–47] **Figure 10** illustrates this.

The order of cold working and thermal treatments seems to be important. Austenitic stainless steels, either sensitised[46] or partly sensitized as a result of cooling after prolonged solution treatment,[44,48] have been found to exhibit improved corrosion resistance when subsequently cold worked, although at very high levels of cold work the corrosion resistance can begin to decrease again due to martensite formation.[44]

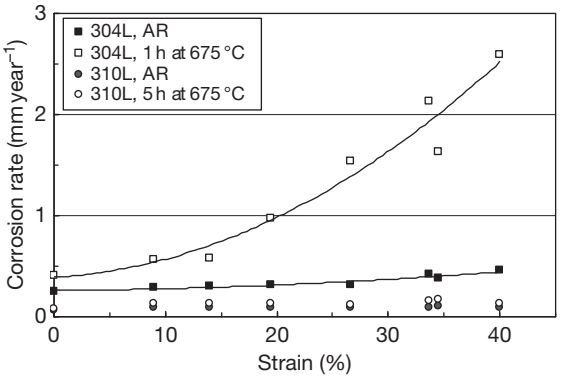


Figure 10 Effect of cold work and subsequent thermal treatment on the corrosion resistance of austenitic stainless steels in boiling 65% nitric acid. Quayle, B. E. BNFL unpublished work.

Conversely, if austenitic stainless steels are first cold worked and then sensitized, those steels having a susceptibility to sensitization show decreased corrosion resistance with increasing levels of cold work (see **Figure 10**), although the effect can reach a peak at relatively low deformation for some stainless steels.[43]

In practice, cold work is only likely to be problematical, at least for austenitic stainless steels, if they are in fact subsequently sensitized, for instance, by welding or by operation for prolonged periods at temperatures exceeding ~300 °C.[43]

2.24.5.3 NO_x Gases

NO_x gases dissolve in aqueous nitric acid producing nitrous acid. The resulting concentration of nitrous acid depends on temperature and it has been shown that steady state is established relatively quickly owing to solution saturation.[49] In the absence of other dissolved species, the redox potential of a given aqueous nitric acid solution is determined by its nitrous acid content and is amenable to calculation over a wide temperature range for a wide range of nitric acid concentrations.[50] Increasing nitrous acid concentration leads to an essentially perfect Nernstian decrease in solution redox potential. Nitrous acid is thus more reducing in nature than nitric acid. Metals such as zirconium exhibit the same behavior as platinum, the corrosion potential decreasing upon NO_x addition. However, because nitrous acid catalyses nitrate reduction (see **Section 2.24.2.2.2**), the effect on stainless steels is a significant rise in the metal's corrosion potential (see **Figure 11**). These opposing effects on solution redox potential and stainless steel corrosion potential are explained in **Figure 12**.

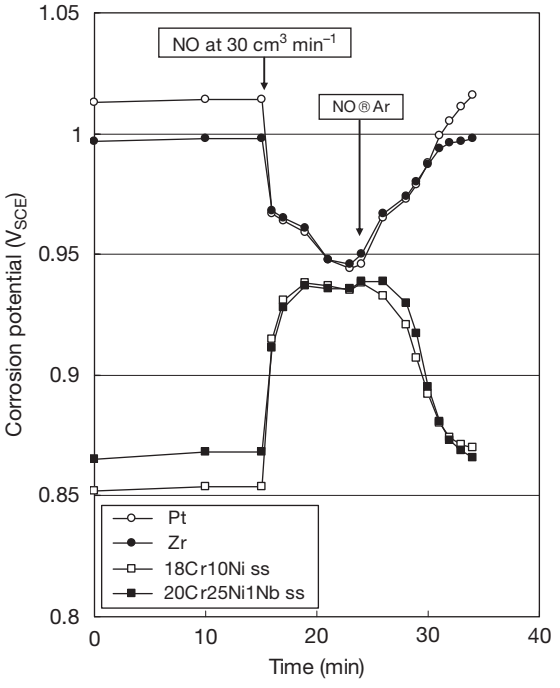


Figure 11 Effect of NO_x purging on the corrosion potential of the indicated metals in 8 M nitric acid at 100 °C Takeuchi, M.; Whillock, G. O. H. BNFL unpublished work.

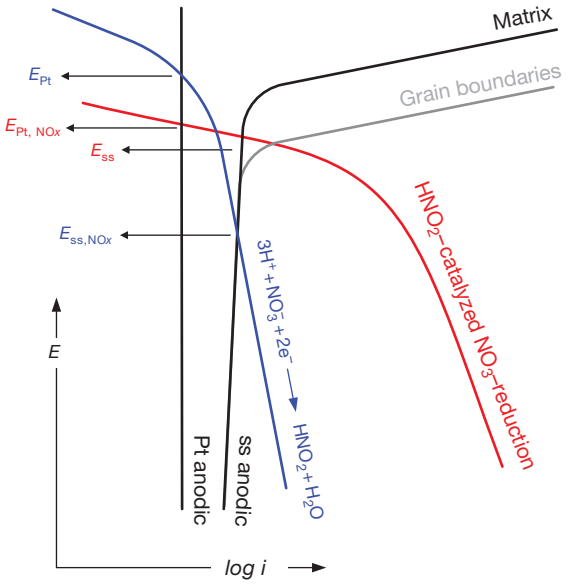


Figure 12 Evans diagram accounting schematically for the observed effect of NO_x purging on corrosion processes occurring in nitric acid. Adapted from Takeuchi, M.; Whillock, G. O. H. *Br. Corros. J.* **2002**, *37*, 199–205, with permission from Maney Publishing.

Because NO_x addition increases the corrosion potential of stainless steels in pure aqueous nitric acid solutions, the corrosion rate rises.[49] However, in the presence of more powerful oxidants, such as Cr(VI) or Ce(IV), NO_x purging significantly lowers the corrosion rate of stainless steel by reducing these strong accelerators to innocuous valency states. A similar effect was observed by Truman,[51] indicating the importance of considering the effect of NO_x gases in experimental design (see **Section 2.24.6**). In addition, the operating pressure and state of solution boiling could be of significance in some circumstances (see **Section 2.24.5.7**); lowering the partial pressure of NO_x gases would favor solution oxidation reactions, which would have the potential to exert a significant effect in solutions containing dissolved species such as chromium and cerium, particularly when the conditions were otherwise just insufficiently oxidizing to give significant equilibrium concentrations of Cr(VI) and Ce(IV) (see **Sections 2.24.5.4.1** and **2.24.5.6**).

2.24.5.4 Dissolved Oxidizing Species and Corrosion Products

2.24.5.4.1 Stainless steel

Many dissolved species capable of existing in more than one valency state can increase the corrosion rate of stainless steel in nitric acid; for instance in boiling 12 M nitric acid, Cr(VI), Ce(IV), Au(III), V(V), Ru(VIII), and I(VII) are strong corrosion accelerators.[52] The effect of oxidizing species like these is to provide a facile cathodic reaction, essentially replacing nitrate reduction, which is capable of polarizing stainless steel to a potential where breakdown of the passivating film occurs (specifically, Cr_2O_3 to CrO_3 conversion). It was proposed that any species that forms a redox couple lying in the range 1.05–1.30 V_{SCE} will produce the same effect, and hence would act as a strong accelerator.[52] This appears to provide a means of assessing the likely effect of species that cannot be tested easily (e.g., Pu, Np, Tc, etc).

However, this theory is incomplete in that it fails to predict the accelerating effect of a number of dissolved species, notably Fe(III) (see **Figure 13**[53]), Np(VI), Ru(III), and Pu(VI), the redox potentials of which are all too low to induce transpassivity, ranging between 0.53 V_{SCE} (Fe(II)/(III)) and 0.89 V_{SCE} (Np(V)/Np(VI)). The failure of Evans's approach[52] is due in part to oversimplification of the anodic metal dissolution reaction: commercial stainless steels do not remain passive until Cr_2O_3 to CrO_3 conversion

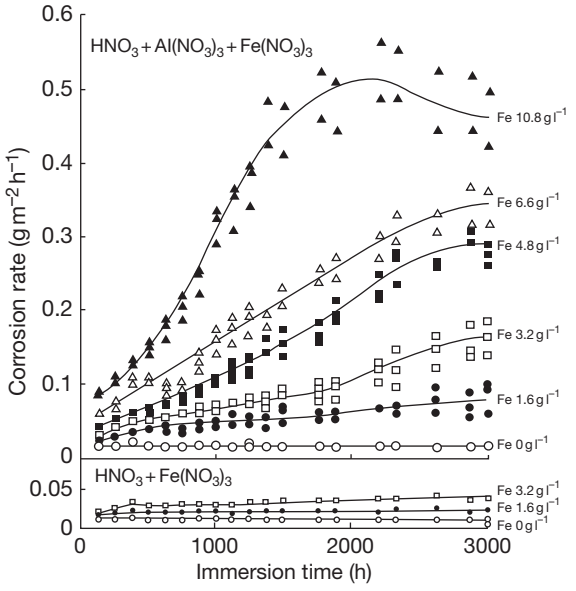


Figure 13 Effect of dissolved iron on the corrosion rate of R-SUS 304ULC stainless steel in (i) 8 M nitric acid containing 1.53M (Al+Fe) and (ii) 9 M nitric acid, both at 100 °C. Reproduced from Hirose, Y.; Suzuoki, A.; Sukekawa, M. In *1994 Fall Meeting of Atomic Energy Society of Japan*; Japan Atomic Energy Society: Japan, 1994; 70, with permission from Atomic Energy Society of Japan.

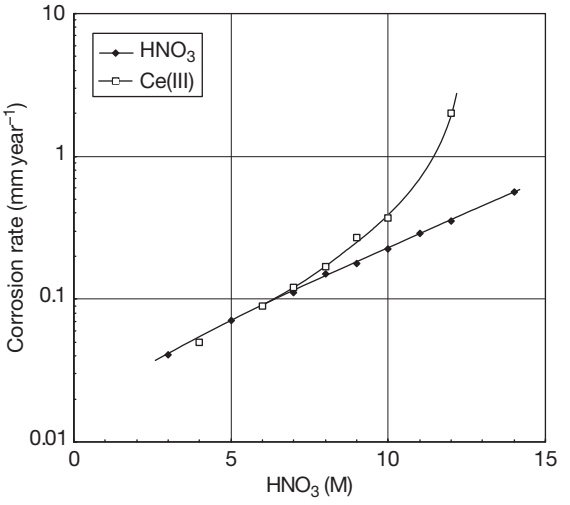


Figure 14 Effect of 0.01 M Ce (added as Ce(III)) on the corrosion rate of 18/13/1 stainless steel in boiling nitric acid solutions. Jones, F. BNFL unpublished work.

occurs; instead, grain boundary attack sets in at significantly lower potential (see **Figure 12**). Thus, corrosion accelerators can be considered to fall into two classes:

- species capable of polarizing stainless steel to very high potential, sufficient to promote Cr_2O_3 to CrO_3 conversion (i.e., Cr(VI), Ce(IV), Au(III), Ru (VIII) and I(VII));
- species capable of polarizing stainless steel to slightly lower potential where intergranular attack occurs (i.e., Fe(III), V(V), Ru(III), Np(VI), and Pu (VI)), but Cr_2O_3 to CrO_3 conversion does not (except perhaps at very high concentrations).

The mechanism whereby species belonging to the latter type act to elevate the corrosion potential of stainless steel has not been rigorously established. However, it is probably due to catalysis of nitrate reduction. This has previously been suggested to account for the action of dissolved iron[54] and neptunium,[55] but kinetic studies demonstrating this unequivocally have not been reported. The catalysis mechanism, if this is correct, is obscure, as Fe(II) should be unobtainable even transiently in hot nitric acid; it may be the case that the thermodynamic properties of the transitional complex have not been properly considered.

An important feature of species such as dissolved iron is that they will produce a persistent effect, because, in nitric acid solutions of sufficient strength and temperatures to pose a corrosion threat, they will be ever-present in their highest oxidation state. Conversely, deliberate additions of species such as Cr(VI) and Ce(IV) would not be capable of creating a persistent effect (provided the solution volume to metal surface area ratio was sufficient to preclude a complicating effect from released iron corrosion product) except in extremely oxidizing conditions where the oxidation of the reduced species proceeds reasonably quickly, because the reduced forms (i.e., Cr(III) and Ce(III)) are not corrosion accelerators *per se*, but can be converted to their oxidized forms by reaction with nitric acid if the conditions are sufficiently oxidizing. **Figure 14** illustrates this effect for dissolved cerium (added as Ce(III)), indicating that cerium accelerates corrosion in boiling nitric acid stronger than ~7 M. However, at lower temperature (100 °C), 8 M nitric acid can be tolerated without significant effect of dissolved cerium, since these conditions are just insufficiently oxidizing for the production of Ce(IV) in solution. But at still lower temperature (70 °C), significant production of Ce(IV) occurs in 10 M nitric acid, but Cr(III) does not act as an accelerator under these conditions because the oxidation of Cr(III) is more difficult than the oxidation of Ce(III) owing to

the more complicated chemical rearrangements required. These examples demonstrate the complex effect of temperature, nitric acid (i.e., nitrate) concentration, and dissolved species on solution oxidizing power, and hence the corrosive effect on stainless steels.

Although in hot, concentrated solutions, nitric acid is usually the most powerful oxidant present, it is not capable of polarizing stainless steel to as high a potential as can be achieved in the presence of species such as Fe(III), Cr(VI), etc.; this must be due to the kinetics of the various reduction reactions, as illustrated in **Figure 15**. Redox potentials alone therefore provide no certain guide to the corrosive effect of dissolved species in nitric acid solutions. The valency state of potential corrosion accelerators is of critical importance (i.e., the equilibrium concentration of the most oxidized form) and this is governed by the balance between the solution oxidation rate and the reduction rate by reaction with stainless steel. It is therefore a prerequisite of successful corrosion testing that the valency states that would be present in an operating plant liquor are maintained during the course of testing.

2.24.5.4.2 Other passive metals

Similar considerations can be applied to other metals as to stainless steels in terms of the effect of dissolved species on polarization behavior. The need for oxidants to be present for titanium to achieve passivity has already been noted (**Section 2.24.3.2**). Although the very passive metals such as zirconium and tantalum will be polarized by the presence of dissolved species, the absence of a low breakdown potential in nitric acid ensures that this is of little consequence for corrosion in most circumstances. Hence, these materials are generally suitable for use in oxidizing plant conditions where stainless steel would suffer an unacceptably high corrosion rate.

2.24.5.5 Dissolved Reducing Species

Dissolved ions such as chloride and fluoride in sufficient quantity can lead to weakening of passivity in nitric acid. The process is dynamic, depending on the relative concentration of halide to nitrate and appears to occur more readily at lower temperatures as the passivating tendency of the nitric acid decreases.

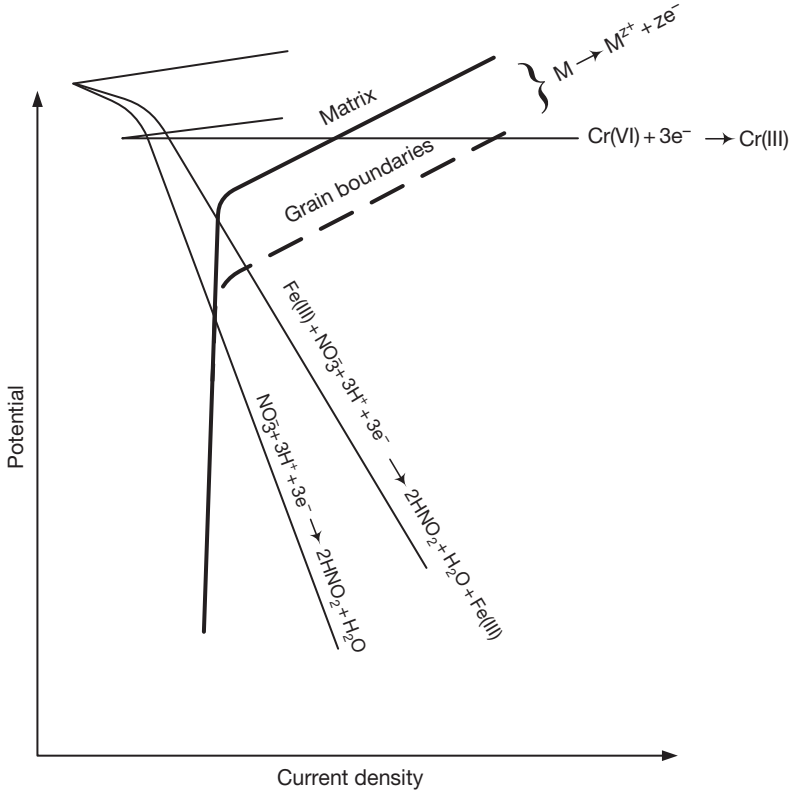


Figure 15 Schematic Evans diagram, showing increased corrosion rate due to Fe(III) and Cr(VI).

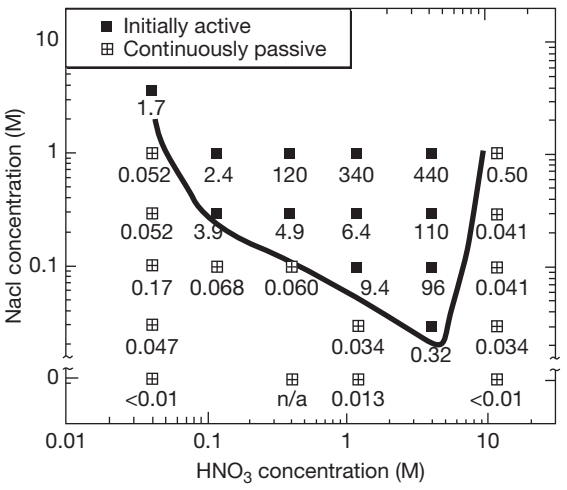


Figure 16 Plot of corrosion state and average corrosion rate (mpy) determined from immersion tests for 1 week of 304 stainless steel in aerated HNO_3/NaCl liquors at 22 °C. Reproduced from Kolman, D. G.; Ford, D. K.; Butt, D. P.; Nelson, T. O. *Corros. Sci.* **1997**, *39*, 2067–2093, with permission from Elsevier.

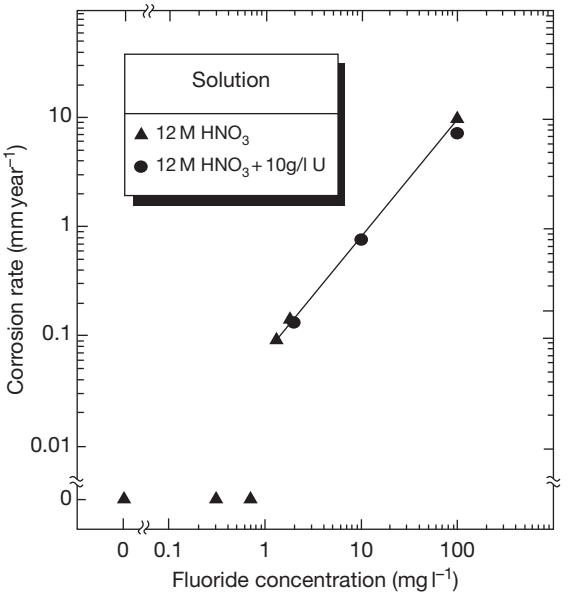


Figure 17 Effect of fluoride ion additions on the corrosion rate of zirconium 702 in boiling nitric acid solutions. Herbert, D. BNFL unpublished work.

The breakdown in passivity is readily indicated by a fall in the corrosion potential, which is intermittent at low concentrations (i.e., downward spikes are observed, increasing in frequency and duration as the halide concentration increases), but sustained at high concentrations leading to rapid active corrosion. Figure 16 maps out regions of active corrosion and passivity for AISI 304 stainless steel in nitric acid/chloride mixtures.[56] On the borderline between activity and passivity, sustained active corrosion can be induced by the application of ultrasound.[57] Although SCC, corrosion fatigue, and erosion/fretting corrosion appear not to have been reported for nitric acid/halide systems, susceptibility to all such chemicomechanical processes is expected.

An effect of chloride ions on corrosion in nitric acid is not reported for other passive metals such as zirconium and tantalum. However, zirconium, and, to much lesser extent, tantalum are sensitive to attack by fluoride ions in nitric acid. For zirconium, a significant effect is observed as low as 1 $\mu g\,ml^{-1}$ F^-, as illustrated in Figure 17. For tantalum and Ta–40Nb alloy, an increase in corrosion rate is not seen for fluoride concentrations $<10\,\mu g\,ml^{-1}$.[58] The effect of fluoride ions can be reduced by the addition of ions such as Zr^{4+} or Al^{3+}, which form strong fluoride complexes, although in high fluoride concentrations, the effect appears to be difficult to suppress completely. The corrosion of zirconium is thus amenable to self-inhibition to some extent, although this would not be expected in a frequently-washed situation such as condensing vapor regions. Table 4 illustrates the major effects observed.

2.24.5.6 Radiation

An overview of the basic effects of ionizing radiation of all types on corrosion processes is given by Byalobzhevskii.[59] For nuclear reprocessing plant, the most important result of radiation is the formation of a large number of short-lived radical species, which react to produce stable molecular products. This is known as radiolysis. Radiolysis of the nitrate ion and water produces the following molecular products, although the reaction chains are long and complex involving a great number of intermediate species[60,61]:

$$HNO_3(aq) \rightarrow HNO_2, H_2O_2, NO_x, H_2 \quad [5]$$

$$H_2O \rightarrow H_2O_2, H_2, H_3O^+ \quad [6]$$

Radiolytic reactions will be affected by the nitrate concentration, dissolved species, temperature and the radiation type. At high dose rates, appreciable levels of nitrous acid (up to $\sim 0.5\,g\,l^{-1}$) are formed, though the nitrous acid would be expected to become increasingly unstable at higher temperatures. Although the action of hydrogen peroxide has not been exhaustively assessed, it probably has little direct impact on

corrosion in hot nitric acid as it decomposes by reaction with nitric acid, producing nitrous acid.

Although data are sparse, the overall effect of the irradiation of nitric acid on corrosion appears to be similar to that produced by nitrous acid. Analogous to the effect of NO_x purging (Figure 11), in pure nitric acid solutions the corrosion potential of stainless steel is increased and hence the corrosion rate is increased slightly, as illustrated by Figure 18.[62] The effects of irradiation are more pronounced when nitric acid contains dissolved oxidizing species where reduction of oxidizing species such as Cr(VI) and Ce(IV) can profoundly reduce corrosion rates by maintaining the corrosion potential below transpassivity.[52] Whether the oxidation by the acid or reduction by irradiation predominates depends on the nitric acid concentration and temperature, increase of either favoring oxidation, and radiation dose rate, increase of which favors reduction. The effect of radiation on the corrosion of stainless steels in nitric acid containing the oxidizing species Cr(VI), Ce(IV) and Ru(III) is demonstrated in Figure 19.[63] It is notable that in the case of Ru(III) the effect is less than for Cr(VI) and Ce(IV).

By analogy with the behavior of Ce(IV) and Cr(VI), it is expected that other species of the same type (i.e., Ru(VIII), I(VII) and Au(III)[52]) would also be reduced by high radiation fields and thus exert negligible corrosive effect towards stainless steels. Redox potential considerations indicate that nitrous acid is incapable of reducing either Fe(III) or V(V); hence, no effect of radiation on their corrosive effect results.

Because the corrosion resistance of metals such as zirconium and tantalum is unaffected by oxidizing species in solution (Section 2.24.5.4.2), no effect of radiation is apparent. No effect of radiation on the corrosion resistance of titanium has been reported, although a deleterious effect might be expected if reliance were placed on oxidizing ions, rather than dissolved titanium, for the maintenance of passivity.

Table 4 Effect of fluoride ion concentration and various additions on the corrosion rate of zirconium 702 in boiling nitric acid solutions

Solution			*Corrosion rate (mm y^{-1})*	
HNO_3 (M)	F^-(mg l^{-1})	Additions	Period 1	Period 2
12	0	–	0.001	0.001
	0.1	–	18.9	10.5
	0.1	10 g l^{-1} U	13.9	8.0
	0.1	3.5 g l^{-1} Bi	12.9	7.3
	0.1	1.1 g l^{-1} Al	1.6	1.4
3	1.8	–	29.6	29.6
	1.8	1.8 g l^{-1} Zr	19.1	20.1
	0.5	2.4 g l^{-1} Zr	1.7	1.4
	0.5	4.8 g l^{-1} Zr	0.002	0.001

Two 24 h test periods without solution renewal.
Source: Herbert, D. BNFL unpublished work.

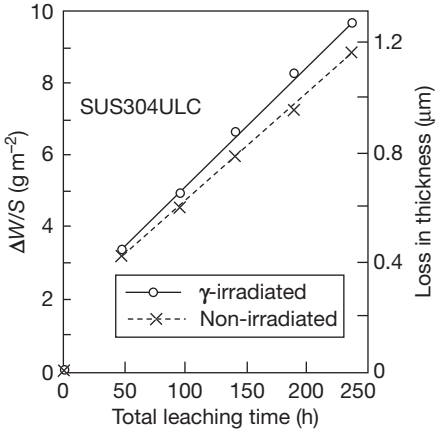


Figure 18 Effect of Co-60 radiation (40 kGy h^{-1}) on the corrosion rate of SUS 304ULC stainless steel in boiling (at atmospheric pressure) 9 M nitric acid. Reproduced from Yamamoto, T.; Tsukui, S.; Okamoto, S.; Nagai, T.; Takeuchi, M.; Takeda, S.; Tanaka, Y. *J. Nucl. Mater.* **1996**, *228*, 162–167, with permission from Elsevier.

2.24.5.7 Solution Boiling

The effects, if any, of boiling on corrosion processes on passive metals appear to be minimal in simple nitric acid systems. This is demonstrated by the failure of Arrhenius data to show a marked discontinuity at the boiling point (see, e.g., Figure 3). Other studies found no difference between corrosion rates measured in boiling (under reduced pressure) and nonboiling nitric acid solutions containing dissolved iron.

However, in a more complex nitric acid solution containing many dissolved species a reduction in corrosion rate is apparently found in the boiling solution. Such systems are very corrosive and there are several possible mechanisms which could occur once boiling commences including enhanced mass transport, reduction in dissolved NO_x by mechanical stripping from solution, removal of reactants and corrosion products from the reacting metal surface by agitation, local concentration effects in nucleating bubbles, limitation of corrodant access to the metal surface by gas blanketing and facilitation of interaction between dissolved species. It may be the case

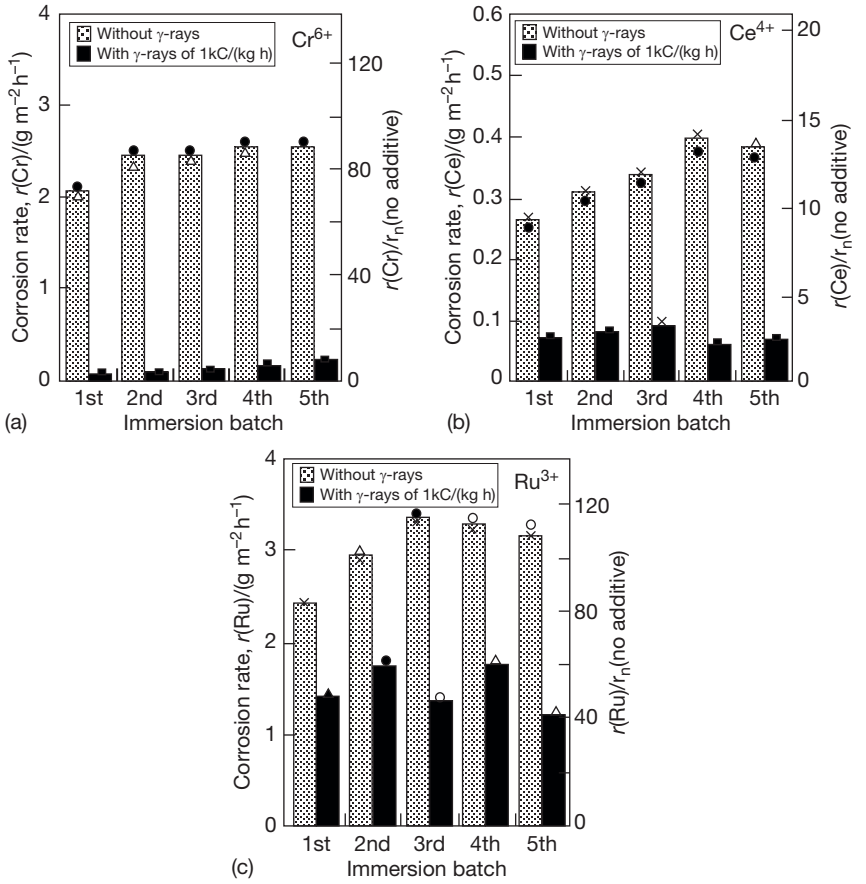


Figure 19 Effect of Co-60 radiation (40 kGy h^{-1}) on the corrosion rate of SUS 304ULC stainless steel in boiling 9 M nitric acid containing: (a) 0.02M Cr(VI); (b) 0.01 M Ce(IV); (c) 0.004 M Ru(III). Reproduced from Yamamoto, T.; Tsukui, S.; Okamoto, S.; Nagai, T.; Takeuchi, M.; Takeda, S.; Tanaka, Y. *J. Nucl. Sci. Technol.* **1998**, *35*, 353–356, with permission from Atomic Energy Society of Japan.

that removal of iron and chromium corrosion products from the corrosion surface by agitation reduces corrosion.

In practice, individual systems may need to be tested if the possibility of a boiling effect is of interest.

2.24.5.8 Heat Transfer

It would not be expected that the transfer of heat *per se* would affect corrosion processes and to date heat flux has not been demonstrated to have an effect on the corrosion process which cannot be assigned to the metal/liquor interface temperature. However, consequent processes such as changes of interface temperatures, modification of the Helmholtz layer, effects on diffusion/mass transport, solute concentration and gas blanketing under boiling could affect corrosion processes.

Data, which are limited, are principally derived from either heated disc[32,64,65] or heated tube rigs[25,66] and can be difficult to interpret. The former suffer from crevice corrosion effects on stainless steel, and the latter from difficulties in maintaining controlled and even heating.

In most simple nitric acid systems, the predominant factor determining the corrosion rate is the effective interface temperature; this will lie between that of the metal heat transfer surface and that of the liquid, but can be difficult to determine with precision. So at a constant metal surface temperature the presence of cooler acid at the metal surface under heat flux reduces the corrosion rate compared to a fully isothermal system.[32,64]

However, it is plausible that oxidation processes such as Cr(III) to Cr(VI) could proceed on heat transfer surfaces with a consequent increase in corrosion providing the surface is of sufficient temperature

even if the liquors are significantly cooler. When boiling occurs on the heat transfer surface, the effects may be further complicated. Although no effect of nucleate boiling on austenitic stainless steel corrosion was found at a heat flux of 9 kW m^{-2},[32] other workers indicate that boiling improves heat transfer and leads to a reduction in metal surface temperature with a consequent reduction in corrosion rate.[64] Conversely, other workers suggest an increased oxidizing effect on heated surfaces, generally consistent with solute rejection from nucleating voids.[65] It is evident that further, carefully controlled, studies are required to resolve the debate.

Industrial experience tends to support the view that corrosion processes on heated surfaces can be affected relative to isothermal surfaces, but that the effects can vary from increasing the corrosion rate where temperatures are increased at the heat transfer face to decreasing them where heating and boiling are so rapid as to create gas-blanketing effects.

2.24.5.9 Liquor-Line and Vapor Regions

No significant liquor-line corrosion effects have been found for stainless steel exposed to nitric acid liquors, either in operating plant or in test conditions.[25] Significant effects would not be expected on other passive metals.

Enhanced corrosion effects can occur in vapor–condensate phases, however. This can occur for instance where condensing films run onto heated tubes in boiler systems and are reevaporated allowing high solute concentrations to build up. In some circumstances transpassivity can be reached with consequent rapid corrosion. Instances have occurred in situations as diverse as stainless steel nitric acid transport tankers and industrial acid boilers. For stainless steels, such thin liquor film effects are considered largely to be due to corrosion product build-up, since both Fe(III) and Cr(VI) are corrosion accelerators (**Section 2.24.5.4.1**). Conversely, titanium is prone to corrosion in condensing systems owing to continuous removal of its corrosion products (-**Section 2.24.3.2**), as is aluminum (**Section 2.24.3.4**).

In mixed acid systems, enhanced corrosion can occur in vapor regions if preferential partitioning of the active species occurs. For stainless steels in nitric acid/chloride solutions, the formation of volatile nitrosyl chlorides may be implicated.[67] Preferential partitioning of fluoride to vapor occurs, hence enhanced corrosion of zirconium in vapor regions above nitric acid/fluoride liquors might be expected; however, the absence of reports of any such effect in operating plant points to the presence of fluoride complexants that prevent or limit fluoride partitioning.

2.24.6 Corrosion Testing Methods

2.24.6.1 Ranking Tests

These are intended to indicate general trends of corrosion resistance for different materials or to show the effect of, for example, different heat treatments. They are also used to screen different batches of material procured for plant construction to ensure consistency of supply.

For stainless steels in nitric acid, the Huey test,[68] which is covered in both ASTM[69] and European standards,[70] is the most frequently used procedure for the assessment of intergranular or general corrosion of stainless steel. It can also form the basis of assessing general corrosion rates of other metals intended for nitric acid service. In this test, a prepared coupon is subjected to five successive 48 h test periods in boiling 65% nitric acid with the solution renewed at each period, the periodic corrosion rate being determined by weight loss. It is noted that the acid specification is more stringently controlled in the ASTM specifications with F^- limited to <0.1 ppm as opposed to <3 ppm in the European standard. This could be significant for some passive materials such as zirconium.

Even with such a simple test, care must be taken to ensure consistency of apparatus (e.g., the type of condenser used affects the results obtained[51]) and the selection of a consistent ratio of liquor volume to metal surface area, since this affects the rate of corrosion product build-up. Maintenance of a constant end, to side, to face grain ratio is essential in materials that have any susceptibility to end grain corrosion; in particular, care is need in the selection of a consistent coupon design which avoids the exposure of excessive or varying end grain fraction (see **Figure 7**). Although not recommended in the test procedures, some assessment and recording of the attack on the end grain is usually worthwhile. Additional periods of testing may be required to discriminate between similar materials.

2.24.6.2 Prediction of In-Service Corrosion Rates

Owing to the dependence of the corrosion rate of stainless steels in nitric acid on microstructural

orientation (see **Section 2.24.4.1**), corrosion rates derived from periodic weighing of immersed coupons can significantly overestimate in-service corrosion rates and can also be subject to considerable variability if different coupon designs are used; the latter in particular prevents meaningful comparison of test data in a great many instances. Conversely, owing to the progressive nature of intergranular attack, short-term tests can lead to significant underestimation of in-service corrosion rates. These problems can be overcome in a number of ways:

- direct measurement of the corrosion rate by metallurgical sectioning transverse to the surface of interest (this works only for relatively short test times otherwise grain dropping removes the initial surface)[30];
- short-term weight loss measurements made on specially-prepared coupons where the surface of interest has been corroded to the grain dropping stage, but all other surfaces are freshly-ground[24];
- weight loss measurements made on a series of coupons specially prepared to have very different ratios of plate, side and end surfaces, all these previously having been corroded to the grain-dropping stage.[30]

In addition, it is possible to seal exposed end grain by overlaying it with weld metal,[71] but this can produce undesirable metallurgical changes and usually requires excessively large coupons unless laser techniques are employed.

All these methods are time consuming, but are necessary where a high confidence of plant life prediction is required. However, none are effective at predicting pit depths where tunneling occurs on the end grain,[30] so prediction of forging life, for instance, is difficult.

Great care must of course be paid to reproducing anticipated plant conditions as closely as possible owing to the many factors which affect the corrosion rate (see **Section 2.24.5**).

References

1. Miles, F. D. *Nitric Acid Manufacture and Uses*; Oxford University Press: London, 1961; Chapter VI, pp 61–73.
2. Addison, C. C. *Chem. Rev.* **1980**, *80*, 21–39.
3. Evans, U. R. *An Introduction to Metallic Corrosion*, 3rd ed.; Edward Arnold: London, 1981; Chapter III, pp 70–72.
4. Vetter, K. J. *Electrochemical Kinetics – Theoretical and Experimental Aspects*; Academic Press: New York, 1967; pp 490–493.
5. Schmid, G.; Lobeck, M. A. *Z. Elektrochem. Ber. Busenges physik. Chem.* **1969**, *73*, 89–199.
6. Razygraev, V. P.; Lebedeva, M. V.; Kabakchi, S. A.; Ponomareva, E. Yu.; Balovneva, R. S.; Lobanova, L. P. *J. Appl. Chem. USSR* **1988**, *61*, 67–73.
7. Balbaud, F.; Sanchez, G.; Santarini, G.; Picard, G. *Eur. J. Inorg. Chem.* **2000**, 665–674.
8. Pakhomov, V. S. *Trans. Mosk. Inst. Khim. Mashinostr.* **1975**, *67*, 102–109.
9. Kain, V.; Shinde, S.; Gadiyar, H. *J. Mater. Eng. Perform.* **1994**, *3*, 699–705.
10. Onoyama, M.; Nakata, M.; Hirose, Y.; Nakagawa, Y. In *Proceedings of the RECOD'91,* Japan Atomic Industrial Forum: Japan, 1991; Vol II, pp 1066–1071.
11. Kreysa, G.; Eckermann, R. Eds. *Dechema Corrosion Handook: Corrosive Agents and their Interaction with Materials*; VCH Publishers: New York, 1991; pp 142–147.
12. Blom, U.; Kvarnback, B. *Mater. Perform.* **1975**, *14*, 43–46.
13. Donat, H.; Schäfer, K. *Schweissen Schneiden* **1975**, *27*, 343–351.
14. Takamura, A.; Arakawa, K.; Moriguchi, Y. In *The Science Technology and Application of Titanium*; Jafee, R. I., Prom, N. E., Eds.; Pergamon Press: Oxford, 1970; pp 209–216.
15. Shimogori, K.; Satoh, H.; Kamikubo, F. In *Titanium Science and Technology*; Proceedings of the 5th International Conference on Titanium, Munich, Germany, Sept 10–14; 1984; Luthering, G., Ed., Vol. 4, pp 1111–1114.
16. Bomberger, H. B. In *Titanium and Zirconium,* Proceedings of the 3rd International Conference, STP 830, 1984; 143.
17. Singh, D. D. N.; Chaudhary, R. S.; Agarwal, C. V. *J. Electrochem. Soc.* **1982**, *129*, 1869–1874.
18. Horn, E. M.; Schoeller, K.; Dölling, H. *Werkst. Korros.* **1990**, *41*, 308–329.
19. Zhuravleva, L.; Rebrunov, V. P. *Zasch. Met.* **1970**, *6*, 224–227.
20. Pourbaix, M. *Atlas of Electrochemical Equilibria in Aqueous Solutions*, 2nd ed.; NACE: Houston, 1974; Chapter 10, pp 256–271.
21. Hosoi, K.; Yokosuka, T.; Honda, T.; Yoshida, T.; Masaoka, I. In *Proceedings of the RECOD 91,* Japan Atomic Industrial Forum: Japan, 1991; Vol. 2, 1060.
22. Kolotyrkin, Y. M.; Kasparova, O. V.; Bogolyubski, S. D. In *Proceedings International Congress on Metallic Corrosion;* National Research Council of Canada: Canada, 1984; Vol. 3, 130.
23. Cihal, V. *Intergranular Corrosion of Steels and Alloys*; Elsevier: Amsterdam, 1984; Chapter 3, pp 67–129.
24. Dunnett, B. F.; Whillock, G. O. H. *Corrosion* **2003**, *59*, 274–283.
25. Rowland, S. M.Sc. thesis, University of Manchester: UK, 1998.
26. Kiuchi, K.; Yanagihara, T.; Kato, T.; Takizawa, M. *JAERI Rev.* **2001**, *10*, 211–213.
27. Ohno, A.; Isoo, H.; Akashi, M. In *Plant Aging and Life Predictions of Corrodible Structures*; Proceedings of the International Symposium, Sapporo, Japan, May 15–18, 1995; Shoji, T., Shibata, T., Eds.; NACE: Houston, TX, 1997; pp 869–874.
28. Balbaud, F.; Sanchez, G.; Fauvet, P.; Santarini, G.; Picard, G. *Corros. Sci.* **2000**, *42*, 1685–1707.
29. Clark, A. T.; Meyer, L. H.; Owen, J. H.; Rust, F. G. Idaho Chemical Processing Plant Report DP-647, 1961.
30. Whillock, G. O. H.; Dunnett, B. F.; Takeuchi, M. *Corrosion* **2005**, *61*, 58–67.
31. Nagano, H.; Kajimura, H. In *Proceedings of the 12th International Corrosion Congress,* September 19–24, 1993 NACE International: Houston, TX, 1993; pp 1341–1352.

32. Andon, R. J. L.; Pemberton, R. C.; Thomas, J. G. N.; Harrison, J. M.; Worthington, S. E.; Shaw, R. D. *Br. Corros. J.* **1986**, *21*, 119–128.
33. Maslov, V. A.; Semenova, A. A. *Svar. Proiz.* **1980**, *11*, 29–30.
34. Nagano, H.; Kajimura, H. *Corros. Sci.* **1996**, *38*(5), 781–792.
35. Turnbull, A. In *Comprehensive Structural Integrity (Fracture of Materials from Nano to Macro), Environment-Assisted Fracture*; Petit, J., Scott, P., Eds.; Elsevier, 2003; Vol. 6, pp 163–210.
36. Hutchings, I. M. The Erosion of Materials by Liquid Flow; Materials Technology Institute of the Chemical Industries Inc.: Columbus, USA, 1986; MTI publication No. 25.
37. Burstein, G. T.; Ashley, G. W.; Marshall, P. I.; Misra, R. D. K. In *Proceedings of the 6th International Conference on Erosion by Liquid and Solid Impact* Field, R., Corney, J., Eds.; Cambridge, UK; 1983.
38. Miki, M. *Corros. Eng. (Boshoku Gijutsu)* **1983**, *32*, 701–706.
39. Brunet, S.; Coriou, H.; Grall, L.; Mahieu, C.; Pelras, M.; Sanatine, O. *Mém. Sci. Rev. Métallurg.* **1970**, *LXVII, 12*, 781–786.
40. Gorelik, G. N.; Anikonova, E. A. *Zasch. Met.* **1970**, *6*, 416–419.
41. Casarini, G.; Colonna, C.; Songa, T. *Metall. Ital.* **1970**, *62*, 183–186.
42. Shvarts, G. L.; Glazkova, S. A.; Moroz, V. A. *Chem. Petroleum Eng.* **1986**, *22*, 25–29.
43. Cihal, V. *Intergranular Corrosion of Steels and Alloys*; Elsevier: Amsterdam, 1984; Chapter 4, pp 144–156.
44. Hahin, C.; Stoss, R. M.; Nelson, B. H.; Reucroft, P. J. *Corrosion* **1976**, *32*, 229–238.
45. Uhlig, H. H. Ed. *Corrosion Handbook*; John Wiley & Sons: New York, 1948; pp 159–160.
46. Mahla, E.; Nielsen, N. *Trans. ASM* **1951**, *43*, 290–297.
47. Herbsleb, G.; Jäkel, U.; Schwaab, P. *Werkst. Korros.* **1990**, *41*, 170–176.
48. Hahin, C. *Corrosion* **1982**, *38*, 116–119.
49. Takeuchi, M.; Whillock, G. O. H. *Br. Corros. J.* **2002**, *37*, 199–205.
50. Takeuchi, M.; Whillock, G. O. H. *Zairyo-to-Kankyo* **2002**, *51*, 549–554.
51. Truman, J. E. *J. Appl. Chem.* **1954**, *4*, 273–283.
52. McIntosh, A. B.; Evans, T. E., *Peaceful Uses of Atomic Energy*; In Proceedings of the 2nd International Conference, Geneva, Switzerland, Sept 1958; UN: Geneva, Switzerland, 1958; pp 206–225.
53. Hirose, Y.; Suzuoki, A.; Sukekawa, M. In *1994 Fall Meeting of Atomic Energy Society of Japan;* Japan Atomic Energy Society: Japan, 1994; 70.
54. Cleland, G. E. Ph.D. Thesis, University of Newcastle upon Tyne, 1998.
55. Motooka, T.; Kiuchi, K. *J. Nucl. Sci. Technol.* **2002**, *3*, 367–370.
56. Kolman, D. G.; Ford, D. K.; Butt, D. P.; Nelson, T. O. *Corros. Sci.* **1997**, *39*, 2067–2093.
57. Whillock, G. O. H.; Harvey, B. F. *Ultrason. Sonochem.* **1996**, *3*, 111–118.
58. Klas, W.; Herpers, U.; Reich, M.; Michel, R.; Droste, R.; Holm, R.; Horn, E.-M.; Müller, G. *Werkst. Korros.* **1991**, *42*, 570–575.
59. Byalobzheskii, A. V. *Radiation Corrosion*; Israel Programme for Scientific Translations Ltd.: Jerusalem, Israel, 1970.
60. Kazanjian, A. R.; Miner, F. J.; Brown, A. K.; Hagan, P. G.; Berry, J. W. *Trans. Faraday Soc.* **1970**, *66*, 2192–2198.
61. Spinks, J. W. T.; Woods, R. J. *An Introduction to Radiation Chemistry*, John Wiley: New York, 1990.
62. Yamamoto, T.; Tsukui, S.; Okamoto, S.; Nagai, T.; Takeuchi, M.; Takeda, S.; Tanaka, Y. *J. Nucl. Mater.* **1996**, *228*, 162–167.
63. Yamamoto, T.; Tsukui, S.; Okamoto, S.; Nagai, T.; Takeuchi, M.; Takeda, S.; Tanaka, Y. *J. Nucl. Sci. Technol.* **1998**, *35*, 353–356.
64. Honda, T.; Hirose, Y.; Sasada, Y.; Furutani, Y.; Suzuki, K.; Fujita, H. In *Proceedings of the RECOD 91,* Japan Atomic Industrial Forum: Japan, 1991; Vol. 2, pp 558–563.
65. Kiuchi, K.; Kikuchi, M. In *Plant Aging and Life Predictions of Corrodible Structures*; Proceedings of the International Symposium, Sapporo, Japan, May 15–18, 1995; Shoji, T., Shibata, T., Eds.; NACE: Houston, TX, 1997; pp 861–867.
66. Yanagihara, T.; Kiuchi, K. *JAERI Rev.* **2001**, *010*, 208–210.
67. Shashi Gupta, M. V. D.; Pandey, G. N. *Corros. Prev. Control* **1986**, *33*, 47–50.
68. Huey, W. R. *Trans. Am. Soc. Steel Treating* **1930**, *18*, 1126–1143.
69. *Annual Book of ASTM Standards*; ASTM International: West Conshohocken, PA, 1985; pp 7–10.
70. ISO 3651–1; Determination of Resistance to Intergranular Corrosion of Stainless Steels – Part 1: Austenitic and Ferritic-Austenitic (Duplex) Stainless Steels – Corrosion Test in Nitric Acid Medium by Measurement of Loss in Mass (Huey Test), ISO 1998.
71. Jeng, J. Y.; Quayle, B. E.; Modern, P. J.; Steen, W. M.; Bastow, B. D. *Corros. Sci.* **1993**, *35*, 1289–1296.

2.25 Corrosion in Acid Gas Solutions

S. Nesic
Institute for Corrosion and Multiphase Technology, Ohio University, Athens, OH, USA
W. Sun
ExxonMobil Upstream Research Company, Houston, TX, USA

Abbreviations

CR Corrosion rate (mm year^{-1})

Symbols

A Surface area of the steel (m^2)

A*/*V Surface to volume ratio (m^{-1})

$A_{(FeCO_3)}$ Constant in the Arrhenius-type equation for $k_{r(FeCO_3)}$

A_{H_2S} Solid state diffusion kinetic constant for H_2S through mackinawite film, $A_{H_2S} = 2.0 \times 10^{-5}$ mol m^{-2} s^{-1}

A_{H^+} Solid state diffusion kinetic constant for H^+ through mackinawite film, $A_{H^+} = 4.0 \times 10^{-4}\ \text{mol m}^{-2}\,\text{s}^{-1}$

A_{CO_2} Solid state diffusion kinetic constant for CO_2 through mackinawite film, $A_{CO_2} = 2.0 \times 10^{-6}\ \text{mol m}^{-2}\,\text{s}^{-1}$

$b_{a(Fe)}$ Anodic Tafel slope for Fe oxidation (V)

$b_{c(H^+)}$ Cathodic Tafel slope for H^+ ion reduction (V)

$b_{c(H_2CO_3)}$ Cathodic Tafel slope for H_2CO_3 reduction (V)

$b_{c(H_2O)}$ Cathodic Tafel slope for H_2O reduction (V)

$B_{(FeCO_3)}$ Constant in the Arrhenius-type equation for $k_{r(FeCO_3)}$ (kJ mol^{-1})

c_{CO_2} Bulk aqueous concentration of CO_2 (kmol m^{-3})

$c_{CO_3^{2-}}$ Bulk aqueous concentration of CO_3^{2-} ions (kmol m^{-3})

$c_{Fe^{2+}}$ Bulk aqueous concentration of Fe^{2+} ions (kmol m^{-3})

c_{H^+} Bulk aqueous concentration of H^+ ions (kmol m^{-3})

$c_{s(H^+)}$ 'Near-zero' concentration of H^+ underneath the mackinawite film at the steel surface, set to 1.0×10^{-7} (kmol m^{-3})

$c_{HCO_3^-}$ Bulk aqueous concentration of HCO_3^- ions (kmol m^{-3})

$c_{H_2CO_3}$ Bulk aqueous concentration of H_2CO_3 (kmol m^{-3})

c_{H_2S} Bulk aqueous concentration of H_2S (kmol m^{-3})

c_{HS^-} Bulk aqueous concentration of HS^- ions (kmol m^{-3})

c_i Bulk aqueous concentration of a given aqueous species (kmol m^{-3})

$c_{i(H_2S)}$ Aqueous concentration of H_2S at the inner sulfide film/outer sulfide layer interface (kmol m^{-3})

$c_{S^{2-}}$ Bulk aqueous concentration of S^{2-} ions (kmol m^{-3})

$c_{s(H_2S)}$ 'Near-zero' aqueous concentration of H_2S underneath the mackinawite film at the steel surface, set to 1.0×10^{-7} (kmol m^{-3})

$c_{o(H_2S)}$ Aqueous concentration of H_2S at the outer sulfide layer/solution interface (kmol m^{-3})

$c_{s(CO_2)}$ Aqueous concentration of CO_2 underneath the mackinawite film at the steel surface

d Characteristic dimension for a given flow geometry (m)

d_p Diameter of a pipe (m)

d_c Diameter of a rotating cylinder (m)

D Diffusion coefficient of a given species ($\text{m}^2\,\text{s}^{-1}$)

$D_{H_2CO_3}$ Aqueous diffusion coefficient of H_2CO_3 ($\text{m}^2\,\text{s}^{-1}$)

$D_{ref(H_2CO_3)}$ Reference aqueous diffusion coefficient of H_2CO_3, $D_{ref,H_2CO_3} = 1.3 \times 10^{-9}\ \text{m}^2\,\text{s}^{-1}$ at 25°C

D_{H^+} Aqueous diffusion coefficient for H^+

$D_{ref(H^+)}$ Reference aqueous diffusion coefficient for H^+, $D_{ref(H^+)} = 2.80 \times 10^{-8}\ \text{m}^2\,\text{s}^{-1}$ at 25 °C

D_{H_2S} Aqueous diffusion coefficient for dissolved H_2S

D_{CO_2} Aqueous diffusion coefficient for dissolved CO_2, $D_{CO_2} = 1.96 \times 10^{-9}$, $\text{m}^2\,\text{s}^{-1}$

E Potential (V)

E_{corr} Corrosion (open circuit) potential (V)

$E_{rev(Fe)}$ Reversible potential of Fe oxidation, $E_{rev(Fe)} = -0.488$ V

$E_{rev(H^+)}$ Reversible potential for H^+ ion reduction (V)

$E_{rev(H_2CO_3)}$ Reversible potential for H_2CO_3 reduction (V)

$E_{rev(H_2O)}$ Reversible potential for H_2O reduction (A m^{-2})

$f_{H_2CO_3}$ Flow factor for the chemical reaction boundary layer

F Faraday's constant, $F = 96\,485\ \text{C mol}_e^{-1}$

$Flux_{H_2S}$ Flux of H_2S ($\text{kmol m}^{-2}\,\text{s}^{-1}$)

$Flux_{H^+}$ Flux of H^+ ions ($\text{kmol m}^{-2}\,\text{s}^{-1}$)

$Flux_{CO_2}$ Flux of CO_2 ($\text{mol m}^{-2}\,\text{s}^{-1}$)

$H_{sol(CO_2)}$ Henry's constant for dissolution of CO_2 (bar kmol m^{-3})

ΔH_{Fe} Activation enthalpy for Fe oxidation, $\Delta H_{Fe} = 50\ \text{kJ mol}^{-1}$

$\Delta H_{(H^+)}$ Activation enthalpy for H^+ ion reduction, $\Delta H_{(H^+)} = 30\ \text{kJ mol}^{-1}$

$\Delta H_{(H_2CO_3)}$ Activation enthalpy for H_2CO_3 reduction, $\Delta H_{(H_2CO_3)} = 57.5\ \text{kJ mol}^{-1}$

$\Delta H_{(H_2O)}$ Activation enthalpy for H_2O reduction, $\Delta H_{(H_2O)} = 30\ \text{kJ mol}^{-1}$

i Current density (A m^{-2})

i_{corr} Corrosion current density (A m^{-2})

$i_{a(Fe)}$ Anodic current density of iron oxidation (A m^{-2})

$i_{c(H^+)}$ Cathodic current density for H^+ ion reduction (A m^{-2})

$i_{c(H_2CO_3)}$ Cathodic current density for H_2CO_3 reduction (A m^{-2})

$i_{c(H_2O)}$ Cathodic current density for H_2O reduction (A m^{-2})

$i^d_{lim(H^+)}$ Mass transfer (diffusion) limiting current density for H^+ ion reduction (A m^{-2})

$i^r_{lim(H_2CO_3)}$ Chemical reaction limiting current density for H_2CO_3 reduction (A m^{-2})

$i_{o(Fe)}$ Exchange current density of iron oxidation ($A\,m^{-2}$)
$i_{o(H^+)}$ Exchange current density for H^+ ion reduction ($A\,m^{-2}$)
$i_{o(H_2CO_3)}$ Exchange current density for H_2CO_3 reduction ($A\,m^{-2}$)
$i_{o(H_2O)}$ Exchange current density for water reduction ($A\,m^{-2}$)
$i^{ref}_{o(Fe)}$ Reference exchange current density of Fe oxidation, $i^{ref}_{o(Fe)} = 1\ A\,m^{-2}$
$i^{ref}_{o(H^+)}$ Reference exchange current density of H^+ oxidation, $i^{ref}_{o(H^+)} = 0.03\ A\,m^{-2}$ at $T_{c,ref} = 25\,^{\circ}C$ and pH 4
$i^{ref}_{o(H_2CO_3)}$ Reference exchange current density for H_2CO_3 reduction, $i^{ref}_{o(H_2CO_3)} = 0.06\ A\,m^{-2}$ at $T_{c,ref} = 25\,^{\circ}C$, pH 5, and $c_{H_2CO_3,ref} = 10^{-4}$ $kmol\,m^{-3}$
$i^{ref}_{o(H_2O)}$ Reference exchange current density for H_2O reduction, $i^{ref}_{o(H_2O)} = 3 \times 10^{-5} A\,m^{-2}$ at $T_{c,ref} = 20\,^{\circ}C$
$i_{\alpha(H^+)}$ Charge transfer current density for H^+ ion reduction ($A\,m^{-2}$)
$i_{\alpha(H_2CO_3)}$ Charge transfer current density for H_2CO_3 reduction ($A\,m^{-2}$)
I Ionic strength $kmol\,m^{-3}$
k^{b}_{hyd} Backward reaction rate of H_2CO_3 dehydration reaction ($1\,s^{-1}$), k^{b}_{hyd}=k^{f}_{hyd}/K_{hyd}
k^{f}_{hyd} Forward reaction rate for the CO_2 hydration reaction ($1\,s^{-1}$)
$k_{m(H^+)}$ Aqueous mass transfer coefficient for H^+ ($A\,m^{-2}$)
$k_{m(H_2CO_3)}$ Aqueous mass transfer coefficient for H_2CO_3 ($A\,m^{-2}$)
$k_{m(H_2S)}$ Aqueous mass transfer coefficient for H_2S ($A\,m^{-2}$)
$k_{m(CO_2)}$ Aqueous mass transfer coefficient for CO_2 ($A\,m^{-2}$)
$k_{r(FeCO_3)}$ Kinetic constant in the ferrous carbonate precipitation rate equation ($1\,mol^{-1}\,s^{-1}$)
K_{hyd} Equilibrium hydration constant for CO_2, $K_{hyd} = k^{f}_{hyd}/k^{b}_{hyd} = 2.58 \times 10^{-3}$
K_{bi} Equilibrium constant for dissociation of HCO_3^- ($kmol\,m^{-3}$)
K_{bs} Equilibrium constant for dissociation HS^- ($kmol\,m^{-3}$)
K_{ca} Equilibrium constant for dissociation of H_2CO_3 ($kmol\,m^{-3}$)
K_{hs} Equilibrium constant for dissociation H_2S ($kmol\,m^{-3}$)
$K_{sol(H_2S)}$ Solubility constant for dissolution of H_2S ($kmol\,m^{-3}\,bar^{-1}$)
$K_{sol(CO_2)}$ Solubility constant for dissolution of CO_2 ($kmol\,m^{-3}\,bar^{-1}$)
$K_{sp(FeCO_3)}$ Solubility product constant for ferrous carbonate ($kmol\,m^{-3}\,bar^{-1}$)
$K^{mackin}_{sp(FeS)}$ Solubility product constant for mackinawite ($kmol\,m^{-3}\,bar^{-1}$)
m_{os} Mass of the outer sulfide layer (kg)
M_{Fe} Molecular mass of iron ($kg\,kmol^{-1}_{Fe}$)
M_{FeS} Molecular mass of ferrous sulfide ($kg\,mol^{-1}_{FeS}$)
n Number of electrons used in reducing or oxidizing a given species ($kmol_e\,kmol^{-1}$)
p_{CO_2} Partial pressure of CO_2 (bar)
p_{H_2S} Partial pressure of H_2S (bar)
R Electrochemical reaction rate ($kmol\,m^{-2}\,s^{-1}$)
R_{FeCO_3} Precipitation rate for iron carbonate ($kmol\,m^{-3}\,s^{-1}$)
R Universal gas constant, $R = 8.314\ J\,mol^{-1}\,K^{-1}$
Re Reynolds number, $Re = v\rho_{H_2O}d/\mu_{H_2O}$
Sc Schmidt number of a given species, $Sc = \mu_{H_2O}/(\rho_{H_2O}D)$
Sh_p Sherwood number of a given species for a straight pipe flow geometry, $Sh_p = k_m d_p/D$
Sh_r Sherwood number of a given species for a rotating cylinder flow geometry, $Sh_r = k_m d_c/D$
$SS_{(FeCO_3)}$ Supersaturation of iron carbonate
ST Scaling tendency
T_c Temperature (°C)
$T_{c,ref}$ Reference temperature, $T_{c,ref} = 25\,^{\circ}C$
T_f Temperature (°F)
T_k Temperature (K)
v Water characteristic velocity ($m\,s^{-1}$)
z_i Species charge of various aqueous species
$\delta_{m(H_2CO_3)}$ Thickness of the mass transfer layer for H_2CO_3 (m)
$\delta_{r(H_2CO_3)}$ Thickness of the chemical reaction layer for H_2CO_3 (m)
δ_{os} Thickness of the outer sulfide layer (m), $\delta_{os} = m_{os}/(\rho_{FeS}A)$
Δt Time interval (s)
μ_{H_2O} Water dynamic viscosity (Pa s)
$\mu_{H_2O,ref}$ Reference water dynamic viscosity (Pa s) at a reference temperature, $\mu_{H_2O,ref} = 1.002 \times 10^{-4}$ Pa s at $20\,^{\circ}C$
$\zeta_{H_2CO_3}$ Ratio of the mass transfer layer and chemical reaction thicknesses for H_2CO_3
ϵ Outer sulfide layer porosity

ψ Outer sulfide layer tortuosity factor
ρ_{H_2O} Density of water (kg m^{-3})
ρ_{Fe} Density of iron (kg m^{-3})
ρ_{FeS} Density of ferrous sulfide (kg m^{-3})

2.25.1 Introduction

As oil and gas emerge from the geological formation, they are always accompanied by some water and varying amounts of 'acid gases': carbon dioxide, CO_2, and hydrogen sulfide, H_2S. This is a corrosive combination, which affects the integrity of mild steel. This has been known for over 100 years; aqueous CO_2 and H_2S corrosion of mild steel still represents a significant problem for the oil and gas industry.[1] Although corrosion resistant alloys that are able to withstand this type of corrosion exist, mild steel is often the most cost effective construction material used in this industry for these applications. All the pipelines, many wells, and much of the processing equipment in the oil and gas industry are built out of mild steel. The cost of equipment failure due to internal CO_2/H_2S corrosion is enormous, both in terms of direct costs such as repair costs and lost production, as well as in indirect costs such as environmental cost, impact on the downstream industries, etc.

The following section summarizes the degree of understanding of the so-called 'sweet' CO_2 corrosion and the so-called 'sour' or H_2S corrosion of mild steel exposed to aqueous environments. It also casts the knowledge in the form of mathematical equations whenever possible. This should enable corrosion engineers and scientists to build entry level corrosion simulation and prediction models.

2.25.2 Aqueous CO_2 Corrosion of Mild Steel

Aqueous CO_2 corrosion of carbon steel is an electrochemical process involving the anodic dissolution of iron and the cathodic evolution of hydrogen. The overall reaction is

$$Fe + CO_2 + H_2O \rightarrow FeCO_3 + H_2 \qquad [1]$$

CO_2 corrosion of mild steel is reasonably well understood. A number of chemical, electrochemical, and transport processes occur simultaneously. They are briefly described below.

2.25.2.1 Chemistry of CO_2 Saturated Aqueous Solutions – Equilibrium Considerations

CO_2 gas is soluble in water:

$$CO_{2(g)} \overset{K_{sol}}{\Leftrightarrow} CO_2 \qquad [2]$$

For ideal gases and ideal solutions in equilibrium, Henry's law can be used to calculate the aqueous concentration of dissolved CO_2, c_{CO_2}, given that the respective concentration in the gas phase (often expressed in terms of partial pressure, p_{CO_2}) is known:

$$H_{sol(CO_2)} = \frac{1}{K_{sol(CO_2)}} = \frac{p_{CO_2}}{c_{CO_2}} \qquad [3]$$

The CO_2 solubility constant, $K_{sol(CO_2)}$, is a function of temperature, T_f, and ionic strength, I[2]:

$$K_{sol(CO_2)} = \frac{14.5}{1.00258} \times 10^{-(2.27+5.65\times10^{-3}T_f - 8.06\times10^{-6}T_f^2 + 0.075I)} \qquad [4]$$

Ionic strength, I, can be calculated as

$$I = \frac{1}{2}\sum_i c_i z_i^2 = \frac{1}{2}(c_1 z_1^2 + c_2 z_2^2 + \cdots) \qquad [5]$$

The concentration of CO_2 in the aqueous phase is of the same order of magnitude as the one in the gas phase. For example, at $p_{CO_2} = 1$ bar, at 25 °C, the gaseous CO_2 concentration is ~4 mol l^{-1} (kmol^{-1} m^{-3}) while in the water it is about 3 mol l^{-1}. Since the solubility of CO_2 decreases with temperature, at 100 °C, the respective concentrations are 3.3 mol l^{-1} in the gas and 1.1 mol l^{-1} in water.

A rather small fraction (about 1 in 500) of the dissolved CO_2 molecules hydrates to make a 'weak' carbonic acid, H_2CO_3:

$$CO_2 + H_2O \overset{K_{hyd}}{\Leftrightarrow} H_2CO_3 \qquad [6]$$

due to a relatively slow forward (hydration) rate. Assuming that the concentration of water remains unchanged, the equilibrium concentration $c_{H_2CO_3}$ is determined by:

$$K_{hyd} = \frac{c_{H_2CO_3}}{c_{CO_2}} \qquad [7]$$

The equilibrium hydration/dehydration constant, $K_{hyd} = 2.58 \times 10^{-3}$, does not change much across the typical temperature range of interest (20–100 °C).[3]

Carbonic acid is considered to be 'weak' because it only partially dissociates in water to produce

hydronium, H^+ ions and bicarbonate ions, HCO_3^-:

$$H_2CO_3 \overset{K_{ca}}{\Leftrightarrow} H^+ + HCO_3^- \quad [8]$$

The HCO_3^- dissociates further to give some more H^+ and carbonate ion, CO_3^{2-}:

$$HCO_3^- \overset{K_{bi}}{\Leftrightarrow} H^+ + CO_3^{2-} \quad [9]$$

The respective equilibrium relations can be written as

$$K_{ca} = \frac{c_{H^+}\, c_{HCO_3^-}}{c_{H_2CO_3}} \quad [10]$$

$$K_{bi} = \frac{c_{H^+}\, c_{CO_3^{2-}}}{c_{HCO_3^-}} \quad [11]$$

The equilibrium constants can be calculated as functions of temperature T_f, and ionic strength, I as[2]

$$K_{ca} = 387.6 \times 10^{\left(-(6.41-1.594\times10^{-3}\,T_f+8.52\times10^{-6}\,T_f^2-3.07\times10^{-5}p\times14.7-0.4772I^{0.5}+0.118I\right)\right)} \quad [12]$$

$$K_{bi} = 10^{\left(-(10.61-4.97\times10^{-3}\,T_f+1.331\times10^{-5}\,T_f^2-2.624\times10^{-5}p\times14.7-1.166I^{0.5}+0.3466I\right)\right)} \quad [13]$$

One can use the equations above to calculate the pH for a pure aqueous CO_2 saturated system. Assuming that the concentration of CO_2 (or partial pressure, p_{CO_2}) in the gas phase is known, one can calculate the concentration of aqueous/dissolved CO_2 c_{CO_2}, via eqn [3]. Then the concentration $c_{H_2CO_3}$ can be determined via eqn [7]. However, in the remaining two eqns [12] and [13], there are three unknowns: c_{H^+}, $c_{HCO_3^-}$, and $c_{CO_3^{2-}}$, and therefore one more equation is needed to close the system: a constraint that describes charge conservation, that is, electroneutrality of the solution. Clearly, chemical reactions [8] and [9], which involve ions, always remain balanced with respect to charge and therefore one can write

$$c_{H^+} = c_{HCO_3^-} + 2c_{CO_3^{2-}} \quad [14]$$

Now, the system of equations is closed and concentrations of all the aqueous species can be determined, including the c_{H^+} and the corresponding pH. The pH of pure water as a function of p_{CO_2} at room temperature is shown in **Figure 1**.

If there are other ions in the aqueous solution, such as for example Fe^{2+} produced by corrosion of steel, then eqn [14] is extended to read

$$2c_{Fe^{2+}} + c_{H^+} = c_{HCO_3^-} + 2c_{CO_3^{2-}} \quad [15]$$

By inspecting the equations above, one can see that, as iron dissolution causes an increase in $c_{Fe^{2+}}$, it is accompanied by a decrease of c_{H^+} due to the cathodic reaction and a corresponding increase in pH. Other cations and anions as well as other chemical reactions can be introduced into the mix in a similar way. An example of a CO_2 aqueous species distribution as a function of pH for an open system is given in **Figure 2**.

It is worth noting that this simple water chemistry calculation procedure is valid only for the case when the concentration of gaseous CO_2, i.e., the partial

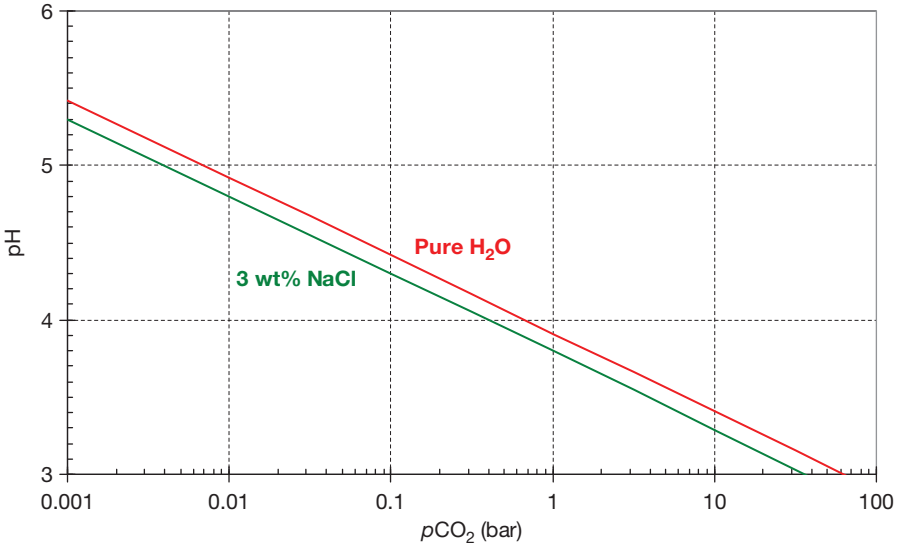


Figure 1 Calculated pH of a pure aqueous solution saturated with CO_2 as a function of partial pressure of CO_2; $T = 25\,^{\circ}C$, 1 wt% NaCl.

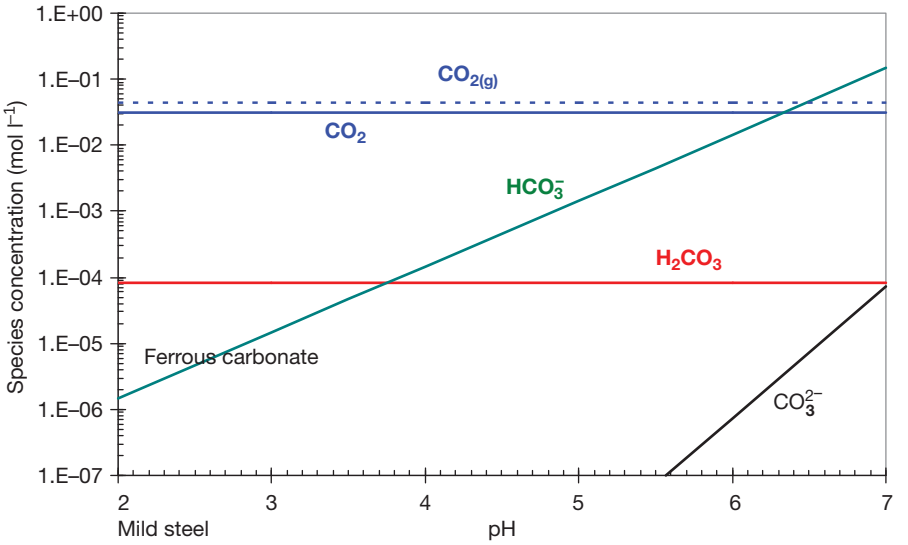


Figure 2 Calculated carbonic species concentrations as a function of pH for a CO_2 saturated aqueous solution; p_{CO_2}=1 bar, 25 °C, 1 wt% NaCl.

pressure, p_{CO_2} is known, constant, and independent from what is happening in the aqueous phase. This is often referred to as an *open* system. It is relevant to field situations where there is an overwhelming amount of CO_2 in the gas phase (such as seen in wet gas lines, multiphase pipelines, gas/liquid separators, etc.). In the lab setting, this condition is easily achieved by continuous purge of a vessel with gaseous CO_2.

In contrast, there are many systems where there is a limited amount of CO_2 in the gas phase compared to the amount in the liquid phase, such as in oil well tubing, oil transportation lines, liquid/liquid separators, etc. In the lab, aqueous systems with a limited gas phase are frequently found in high-pressure autoclaves and flow loops. Consequently they are often referred to as *closed* systems, and in principle can have varying gas/liquid volume ratios. An *open* system can be seen as a *closed* system with an infinitely large gas/liquid volume ratio. In *closed* systems, the concentration of gaseous CO_2, that is, the partial pressure, p_{CO_2}, is not known explicitly and typically depends on the aqueous chemistry. In mathematical terms, this means that there is one more unknown: p_{CO_2}, and therefore one needs one more equation to be able to solve for species concentrations. The extra equation comes from the additional constraint: in a *closed* system, the total amount of carbonic species remains constant; they are just redistributed between the gas and aqueous phases as conditions change. When one accounts for this, an extra equation is obtained:

$$n_{CO_2(g)} + n_{CO_2(aq)} + n_{H_2CO_3(aq)} + n_{HCO_3^-(aq)} + n_{CO_3^{2-}(aq)} = \text{Const.} \quad [16]$$

where n denotes the number of moles of a particular species in a gaseous or aqueous phase of a *closed* system.

The dissociation steps [8] and [9] are very fast compared to all other processes occurring simultaneously in corrosion of mild steel, thus preserving chemical equilibrium. However, the CO_2 dissolution reaction [2] and the hydration reaction [6] are much slower. When such chemical reactions proceed slowly, other faster processes (such as electrochemical reactions or diffusion) can lead to local nonequilibrium in the solution.

Either way, the occurrence of chemical reactions can significantly alter the rate of electrochemical processes at the surface and the rate of corrosion. This is particularly true when, due to high local concentrations of species, the solubility limit of salts is exceeded and precipitation of a surface layer occurs. In a precipitation process, heterogeneous nucleation occurs first on the surface of the metal or within the pores of an existing layer since homogenous nucleation in the bulk requires a much higher concentration of species. Nucleation is followed by crystalline layer growth.

Under certain conditions, surface layers can become very protective and reduce the rate of corrosion.

In CO_2 corrosion, when the concentrations of Fe^{2+} and CO_3^{2-} ions exceed the solubility limit, they form solid ferrous carbonate according to

$$Fe^{2+} + CO_3^{2-} \overset{K_{sp(FeCO_3)}}{\Leftrightarrow} FeCO_{3(s)} \quad [17]$$

where the solubility product constant for ferrous carbonate $K_{sp(FeCO_3)}$ is[4]

$$K_{sp(FeCO_3)} = 10^{-(59.3498 - 0.041377T_k - (2.1963/T_k) + 24.5724 \log T_k + 2.518I^{0.5} - 0.657I)} \quad [18]$$

Actually ferrous and carbonate ions are frequently found in the aqueous solution at concentrations much higher than predicted by the equilibrium $K_{sp(FeCO_3)}$. This is termed *supersaturation* and is a necessary condition before any substantial precipitation can occur. The ferrous carbonate supersaturation, $SS_{(FeCO_3)}$, is defined as:

$$SS_{(FeCO_3)} = \frac{c_{Fe^{2+}}\, c_{CO_3^{2-}}}{K_{sp(FeCO_3)}} \quad [19]$$

The precipitation process can be seen as the process of the solution returning to equilibrium and is driven by the magnitude of supersaturation. The rate of the precipitation ($\Re_{FeCO_3}$) is therefore often expressed as

$$\Re_{FeCO_3} = k_{r(FeCO_3)} \frac{A}{V} K_{sp(FeCO_3)} \left(SS_{(FeCO_3)} - 1\right) \quad [20]$$

where $k_{r(FeCO_3)}$ is a kinetic constant, which can be derived from the experimental results as a function of temperature, using an Arrhenius' type equation[5]

$$k_{r(FeCO_3)} = \exp\left(A_{(FeCO_3)} - \frac{B_{(FeCO_3)}}{RT_k}\right) \quad [21]$$

where $A_{(FeCO_3)} = 28.2$ and $B_{(FeCO_3)} = 64\,851\,J\,mol^{-1}$.

SEM images of a crystalline ferrous carbonate layer formed on a mild steel substrate are shown in Figure 3. The ferrous carbonate layer can slow down the corrosion process by presenting a diffusion barrier for the species involved in the corrosion process, thereby changing the conditions at the steel surface. The effective protectiveness of a solid ferrous carbonate layer depends on its porosity, which hangs in the balance of the precipitation rate and the underlying corrosion rate. For high precipitation rates, and low corrosion rates, a dense and protective ferrous carbonate layer is obtained, and conversely, low precipitation rates and high corrosion rates lead to the formation of porous unprotective ferrous carbonate layers. A non-dimensional parameter termed 'scaling tendency' (ST) can be used to quantify the relative rates of precipitation ($\Re_{FeCO_3}$) and corrosion (CR) expressed in the same volumetric units:

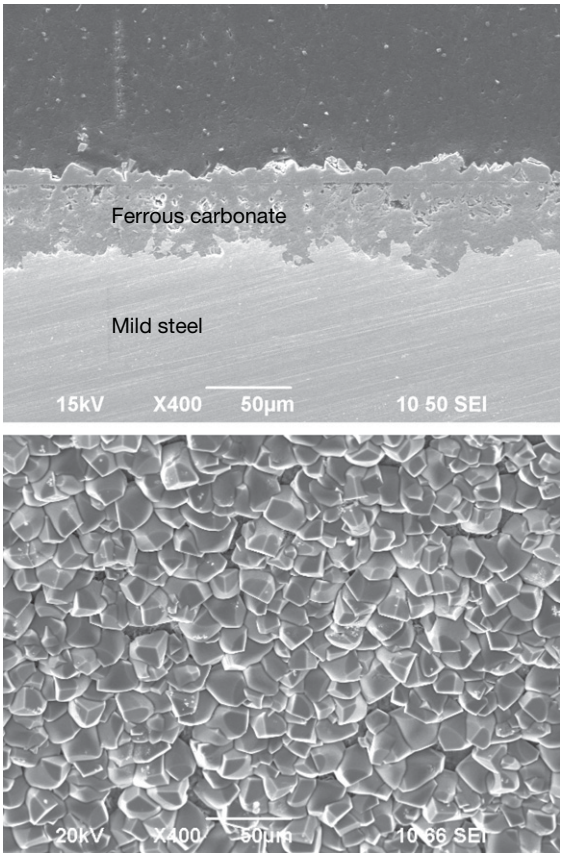


Figure 3 SEM images showing a cross-section and a top view of a ferrous carbonate layer formed on mild steel; 80 °C, pH 6.6, $p_{CO_2} = 0.5$ bar, stagnant conditions.

$$ST = \frac{\Re_{FeCO_3}}{CR} \quad [22]$$

For $ST \ll 1$, porous and unprotective films are likely to form. Conversely, for $ST \geq 1$, conditions become favorable for the formation of dense, protective ferrous carbonate films. However, the use of scaling tendency is not as straightforward as it appears, as it requires simultaneous calculation of $\Re_{FeCO_3}$ and corrosion CR.

In some cases, other salts can be detected in the surface layers that form on mild steel. In high temperature CO_2 corrosion magnetite (Fe_3O_4) has been detected. In the presence of oxygen, a ferric oxide hematite (Fe_2O_3), which is very insoluble but offers little protection from corrosion forms. Oilfield scales also include calcium carbonate, calcium sulfate,

barium sulfate, strontium sulfate, etc. The presence of calcium carbonate, in particular, can have a beneficial effect upon corrosion and upon the stability of the $FeCO_3$ scale. Finally, in the presence of H_2S, various types of sulfides form as discussed in a separate section below.

2.25.2.2 Electrochemistry of Mild Steel Corrosion in CO_2 Saturated Aqueous Solutions

The electrochemical dissolution of iron in a water solution:

$$Fe \rightarrow Fe^{2+} + 2e^- \qquad [23]$$

is the dominant anodic reaction in CO_2 corrosion. The reaction is pH dependent in acidic solutions with a reaction order with respect to OH^- between 1 and 2, decreasing toward 1 and 0 at $pH > 4$, which is the typical range for CO_2 corrosion. Measured Tafel slopes are typically 30–80 mV. This subject, which is still somewhat controversial with respect to the mechanism, has been reviewed for acidic corrosion[6,7] and CO_2 solutions.[8]

The presence of CO_2 increases the rate of corrosion of mild steel in aqueous solutions primarily by increasing the rate of the hydrogen evolution reaction. It is well known that in *strong* acids, which are fully dissociated, the rate of hydrogen evolution occurs according to

$$2H^+ + 2e^- \rightarrow H_2 \qquad [24]$$

and is, for the case of mild steel corrosion, limited by the rate at which H^+ ions are transported from the bulk solution to the steel surface (mass transfer limitation). In CO_2 solutions, where typically $pH > 4$, this limiting flux would be small, and therefore it is the presence of H_2CO_3 which enables hydrogen evolution at a much higher rate. Thus, for $pH > 4$, the presence of CO_2 leads to a much higher corrosion rate than would be found in a solution of a *strong* acid at the same pH.

This can be readily explained by considering that the homogenous dissociation of H_2CO_3, as given by reaction [8], serves as an additional source of H^+ ions, which are subsequently adsorbed at the steel surface and reduced according to reaction [24].[1] A different pathway is also possible, where the H_2CO_3 first adsorbs at the steel surface followed by heterogeneous dissociation and reduction of the H^+ ion. This is often referred to as 'direct' reduction of carbonic acid[9–11] and is written as

$$2H_2CO_3 + 2e^- \rightarrow H_2 + 2HCO_3^- \qquad [25]$$

Clearly, the addition of the reactions [8] and [24] gives the reaction [25] proving that the overall reaction is the same and the distinction is only in the pathway, that is, in the sequence of reactions. The rate of reaction [25] is limited primarily by the slow hydration step [6][11,12] and in some cases by the slow CO_2 dissolution reaction [2].

It can be conceived that in CO_2 solutions at $pH > 5$ the direct reduction of the bicarbonate ion becomes important[13]:

$$2HCO_3^- + 2e^- \rightarrow H_2 + 2CO_3^{2-} \qquad [26]$$

which seems plausible, as the concentration of HCO_3^- increases with pH and can exceed that of H_2CO_3 as seen in **Figure 2**. However, it is difficult to distinguish experimentally the effect of this particular reaction pathway for hydrogen evolution from the two previously discussed (eqns [8] and [25]). In addition, evidence exists that suggests that the rate of this reaction is comparatively low and can be neglected. For example, as the pH increases, the amount of HCO_3^- increases as well (see **Figure 2**), suggesting that the corrosion rate should follow the same trend, if one is to believe that the direct reduction of the bicarbonate ion [26] is a significant cathodic reaction. Experimental evidence does not support this scenario and shows the opposite trend: the corrosion rate actually decreases with an increasing pH, even if no protective ferrous carbonate layer forms.

Hydrogen evolution by direct reduction of water:

$$2H_2O + 2e^- \rightarrow H_2 + 2OH^- \qquad [27]$$

is always possible, but is comparatively very slow and is important only at $p_{CO_2} \ll 0.1$ bar and $pH > 6$.[14,15] Therefore, this reaction is rarely a factor in practical CO_2 corrosion situations.

The various electrochemical processes described above can be quantified using the well established electrochemical theory. The rate of the electrochemical reactions, $\Re$ in kmol m^{-2} s^{-1}, can be readily expressed in terms of current density, i in A m^{-2}, since the two are directly related: for example, during hydrogen evolution [24] for every kmol of H^+ 1 kmol of electrons is used ($n = 1\ kmol_e\ kmol^{-1}$), while for every kmol of iron dissolved [23] two

kmoles of electrons are used ($n = 2\,\mathrm{kmol_e\,kmol^{-1}}$). Therefore, one can write

$$i = nF\Re \qquad [28]$$

2.25.2.2.1 Oxidation of iron

In the corrosion of mild steel, the oxidation (dissolution) of iron [23] is the dominant anodic reaction. The anodic dissolution of iron at the corrosion potential (and up to 200 mV above) is under charge transfer control. Thus, pure Tafel behavior can be assumed close to the corrosion potential:

$$i_{a(Fe)} = i_{o(Fe)} 10^{(E_{corr} - E_{rev(Fe)})/b_{a(Fe)}} \qquad [29]$$

The exchange current density of iron oxidation is a function of temperature:

$$i_{o(Fe)} = i^{ref}_{o(Fe)} \exp\left(\frac{-\Delta H_{Fe}}{R}\left(\frac{1}{T_c + 273.15} - \frac{1}{T_{c,ref} + 273.15}\right)\right) \qquad [30]$$

The Tafel slope of this reaction is given by

$$b_{a(Fe)} = \frac{2.303R(T_c + 273.15)}{1.5F} \qquad [31]$$

2.25.2.2.2 Reduction of hydronium ion

In general, the H^+ ion reduction reaction [24] can be either under charge transfer or mass transfer (diffusion) control, therefore, one can write:

$$\frac{1}{i_{c(H^+)}} = \frac{1}{i_{\alpha(H^+)}} + \frac{1}{i^d_{lim(H^+)}} \qquad [32]$$

The charge transfer current density can be calculated by

$$i_{\alpha(H^+)} = i_{o(H^+)} \times 10^{(-E_{corr} - E_{rev(H^+)})/b_{c(H^+)}} \qquad [33]$$

The exchange current density $i_{o(H^+)}$ is a function of pH and temperature. The pH dependence is

$$\frac{\partial \log i_{o(H^+)}}{\partial \mathrm{pH}} = -0.5 \qquad [34]$$

The temperature dependence of the exchange current density can be calculated via an Arrhenius-type relation:

$$\frac{i_{o(H^+)}}{i^{ref}_{o(H^+)}} = \exp\left(-\frac{\Delta H_{(H^+)}}{R}\left(\frac{1}{T_c + 273.15} - \frac{1}{T_{c,ref} + 273.15}\right)\right) \qquad [35]$$

The reversible potential for H^+ reduction $E_{rev(H^+)}$ is a function of temperature and pH:

$$E_{rev(H^+)} = -\frac{2.303R(T_c + 273.15)}{F}\mathrm{pH} \qquad [36]$$

The cathodic Tafel slope $b_{c(H^+)}$ is calculated as

$$b_{c(H^+)} = \frac{2.303R(T_c + 273.15)}{0.5F} \qquad [37]$$

The limiting mass transfer current density $i^d_{lim(H^+)}$ is related to the rate of transport of H^+ ions from the bulk of the solution through the boundary layer to the steel surface:

$$i^d_{lim(H^+)} = k_{m(H^+)} F c_{H^+} \qquad [38]$$

where the mass transfer coefficient, $k_{m(H^+)}$ can be calculated from a correlation of the Sherwood, Reynolds, and Schmidt numbers as explained in the following section.

2.25.2.2.3 Reduction of carbonic acid

The carbonic acid reduction reaction [25] can be under charge transfer control or limited by the slow chemical reaction–hydration step [6], preceding it.[11,12] The rate of this reaction in terms of current density is

$$\frac{1}{i_{c(H_2CO_3)}} = \frac{1}{i_{\alpha(H_2CO_3)}} + \frac{1}{i^r_{lim(H_2CO_3)}} \qquad [39]$$

The charge transfer current density $i_{\alpha(H_2CO_3)}$ is calculated as

$$i_{\alpha(H_2CO_3)} = i_{o(H_2CO_3)} \times 10^{(-E_{corr} - E_{rev(H_2CO_3)})/b_{c(H_2CO_3)}} \qquad [40]$$

The exchange current density $i_{o(H_2CO_3)}$ depends on pH, H_2CO_3 concentration, and temperature:

$$\frac{\partial \log i_{o(H_2CO_3)}}{\partial \mathrm{pH}} = 0.5 \qquad [41]$$

$$\frac{\partial \log i_{o(H_2CO_3)}}{\partial c_{H_2CO_3}} = 1 \qquad [42]$$

$$\frac{i_{o(H_2CO_3)}}{i^{ref}_{o(H_2CO_3)}} = \exp\left(-\frac{\Delta H_{(H_2CO_3)}}{R}\left(\frac{1}{T_c + 273.15} - \frac{1}{T_{c,ref} + 273.15}\right)\right) \qquad [43]$$

The cathodic Tafel slope $b_{c(H_2CO_3)}$ is

$$b_{c(H_2CO_3)} = \frac{2.303R(T_c + 273.15)}{0.5F} \qquad [44]$$

Since the reductions of H_2CO_3 and H^+ are equivalent

thermodynamically, the reversible potential for H_2CO_3 reduction $E_{rev(H_2CO_3)}$ is calculated as

$$E_{rev(H_2CO_3)} = -\frac{2.303R(T_c + 273.15)}{F}pH \quad [45]$$

The chemical reaction limiting current density $i^r_{lim(H_2CO_3)}$ can be calculated from[16]:

$$i^r_{lim(H_2CO_3)} = Fc_{CO_2}f_{H_2CO_3}\sqrt{D_{H_2CO_3}K_{hyd}k^f_{hyd}} \quad [46]$$

The diffusion coefficient for carbonic acid $D_{H_2CO_3}$ as a function of temperature can be calculated using Einstein's relation:

$$D = D_{ref}\left(\frac{T_c + 273.15}{T_{c,ref} + 273.15}\right)\left(\frac{\mu_{H_2O,ref}}{\mu_{H_2O}}\right) \quad [47]$$

where T is temperature and μ is dynamic viscosity. The forward reaction rate for the CO_2 hydration reaction k^f_{hyd} is calculated as

$$k^f_{hyd} = 10^{169.2-53.0\log(T_c+273.15)-(11715/(T_c+273.15))} \quad [48]$$

The flow factor $f_{H_2CO_3}$ is

$$f_{H_2CO_3} = \coth\zeta_{H_2CO_3} \quad [49]$$

where

$$\zeta_{H_2CO_3} = \frac{\delta_{m(H_2CO_3)}}{\delta_{r(H_2CO_3)}} \quad [50]$$

and

$$\delta_{m(H_2CO_3)} = \frac{D_{H_2CO_3}}{k_{m(H_2CO_3)}} \quad [51]$$

$$\delta_{r(H_2CO_3)} = \sqrt{\frac{D_{H_2CO_3}}{k^b_{hyd}}} \quad [52]$$

The carbonic acid mass transfer coefficient $k_{m(H_2CO_3)}$ is discussed in Section 2.25.2.3.

2.25.2.2.4 Reduction of water

Unless water is mixed with methanol or glycol to prevent hydrate formation or somehow diluted otherwise, it can be assumed that water molecules are present in virtually unlimited quantities at the steel surface, and the reduction rate of H_2O is controlled by the charge-transfer process and, hence, pure Tafel behavior:

$$i_{c(H_2O)} = i_{o(H_2O)}10^{-(E_{corr}-E_{rev(H_2O)})/b_{c(H_2O)}} \quad [53]$$

Since the reduction of H_2O and H^+ are equivalent thermodynamically, they have the same reversible potential at a given pH:

$$E_{rev(H_2O)} = -\frac{2.303R(T_c + 273.15)}{F}pH \quad [54]$$

The exchange current density for water reduction $i_{o(H_2O)}$ depends on temperature:

$$\frac{i_{o(H_2O)}}{i^{ref}_{o(H_2O)}} = \exp\left(-\frac{\Delta H_{(H_2O)}}{R}\left(\frac{1}{T_c + 273.15} - \frac{1}{T_{c,ref} + 273.15}\right)\right) \quad [55]$$

The Tafel slope for H_2O reduction was found to be the same as that for H^+ reduction:

$$b_{c(H_2O)} = \frac{2.303R(T_c + 273.15)}{0.5F} \quad [56]$$

2.25.2.3 Transport Processes in CO_2 Corrosion of Mild Steel

From the description of the electrochemical processes above, it is clear that certain species in the solution are 'produced' at the metal surface (e.g., Fe^{2+}) while others are depleted (e.g., H^+). The established concentration gradients lead to molecular diffusion of the species toward and away from the surface. In cases when the diffusion processes are much faster than the electrochemical processes, the concentration change at the metal surface is small. In contrast, when the diffusion is unable to 'keep up' with the rate of the electrochemical reactions, the concentration of species at the metal surface can become very different from that in the bulk solution. The rate of the electrochemical processes depends on the concentration of the reactants at the surface. Therefore, there exists a two-way coupling between the electrochemical processes at the metal surface (corrosion) and processes in the adjacent solution layer (i.e., diffusion in the boundary layer). The same is true for chemical reactions, which interact with both the transport and electrochemical processes in a complex way.

In most practical systems, the water solution moves with respect to the metal surface. Therefore, the effect of convection on transport processes cannot

be ignored. Turbulent eddies can penetrate deep into the hydrodynamic boundary layer and significantly alter the rate of species transport to and from the surface. Very close to the surface no turbulence can exist and the species are transported solely by diffusion. The effect of turbulent flow is captured most easily by using the concept of mass transfer coefficient, described below.

In turbulent flow of dilute ideal solutions, a mass transfer coefficient k_m for a given species (H^+ ions, H_2CO_3 etc.) can be calculated from a correlation, such as the straight pipe correlation of Berger and Hau[25]:

$$Sh_p = 0.0165 Re^{0.86} Sc^{0.33} \quad [57]$$

or the rotating cylinder correlation of Eisenberg *et al.*[26]:

$$Sh_r = 0.0791 Re^{0.7} Sc^{0.356} \quad [58]$$

or any other similar correlation for the flow geometry at hand. It should be noted that most of the mass transfer correlations found in the literature (including the two listed above) are suited only for single-phase flow. Therefore, extension of this approach to multiphase flow situations needs to be done with careful consideration.

Overall, CO_2 corrosion of mild steel is not very sensitive to flow, at least not so when compared to mild steel corrosion in strong acids. This is due to the fact that the main corrosive species in CO_2 corrosion is H_2CO_3 which can easily be depleted due to a slow chemical step which precedes it: the hydration reaction [6]. Therefore, the limiting rate of CO_2 corrosion is primarily affected by the rate of this chemical reaction [46], which is a function of temperature and CO_2 partial pressure and not very sensitive to flow.

2.25.2.4 Calculation of Mild Steel CO_2 Corrosion Rate

Leading to this point, the main processes underpinning CO_2 corrosion were defined: the speciation of the aqueous CO_2 solution using the thermodynamic approach outlined in **Section 2.25.2.1**, the electrochemical theory described in **Section 2.25.2.2**, and the transport processes as covered in **Section 2.25.2.3**. Using this information, the corrosion rate of mild steel can now be calculated. The unknown corrosion potential E_{corr} in [33], [40], [53], and [29] can be found from the current (charge) balance equation at the steel surface:

$$i_{c(H^+)} + i_{c(H_2CO_3)} + i_{c(H_2O)} = i_{a(Fe)} \quad [59]$$

which expresses the simple fact that at steady state all the electrons generated by the oxidation processes are consumed by the sum of the reduction processes. By substituting the expressions for the various currents given by eqns [33], [40], [53], and [29] into eqn [59] a single nonlinear equation is now obtained with E_{corr} as the only unknown, which can be easily solved. When the calculated value of E_{corr} is now returned to eqns [33], [40], [53], and [29], the rate of each individual reaction can be explicitly computed. This also includes the corrosion current density obtained from eqn [29]:

$$i_{corr} = i_{a(Fe)} \quad [60]$$

Finally, the CO_2 corrosion rate is recovered by using Faraday's law:

$$CR = \frac{i_{corr} M_{Fe}}{\rho_{Fe} nF} \quad [61]$$

where M is the molecular mass and ρ is the density. If the unit amperes per square meters is used for the corrosion current density i_{corr}, then conveniently the corrosion rate for iron and steel expressed in millimeter per year takes almost the same numerical value, precisely, $CR = 1.155 i_{corr}$.

2.25.2.5 Successes and Limitations of Modeling of Aqueous CO_2 Corrosion of Mild Steel

Evidence that our basic understanding of the processes underlying CO_2 corrosion of mild steel is reasonably sound can be found by comparing the predictions made by the mechanistic model outlined above with experimental values. In **Figure 4**, below, one can see the comparison of a potentiodynamic sweep obtained in the experiments and the one predicted by the model. Many other comparisons of the predicted and measured corrosion rates are given in the following section, where the effect of key factors in CO_2 corrosion of mild steel is discussed.

Despite the relative progress we have made in understanding and modeling of aqueous CO_2 corrosion of mild steel, many questions persist. One is the issue of localized CO_2 corrosion, which is still a topic of intense ongoing research. Effect of other factors such as steel metallurgy, organic acids, oxygen, multiphase flow, and inhibitors are challenges that need further effort. Some of those are discussed in the following sections.

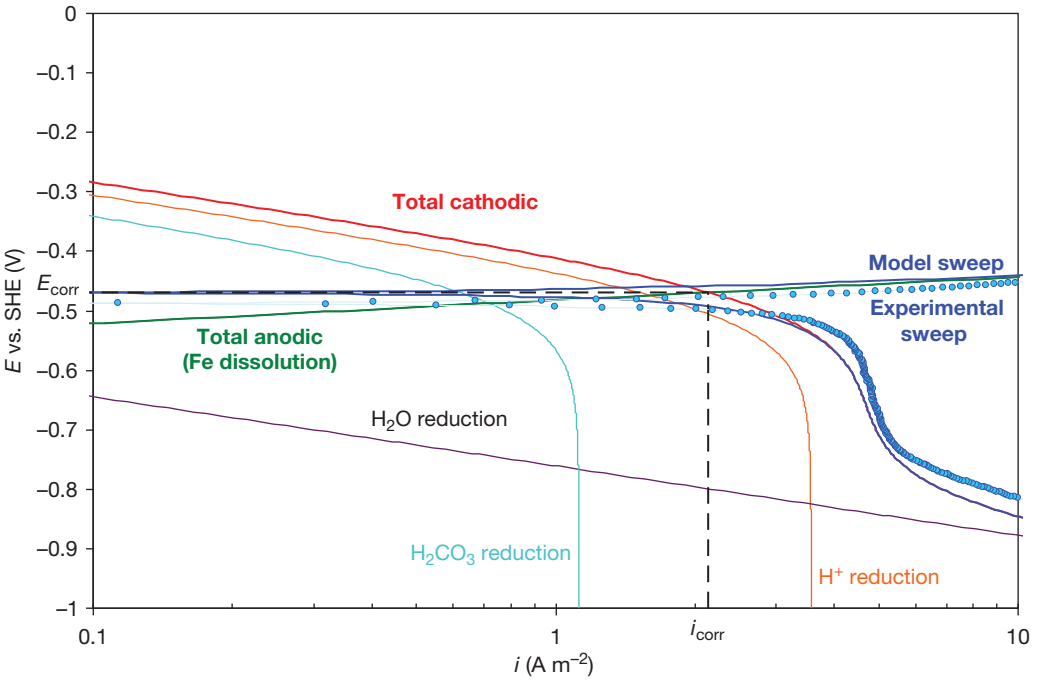


Figure 4 Potentiodynamic sweep, experimental (points) vs. model (lines); 20 °C, p_{CO_2} = 1 bar, pH 4, 2 m s^{-1}.

2.25.2.6 Key Factors Affecting Aqueous CO_2 Corrosion of Mild Steel

Armed with the understanding and the ability to calculate CO_2 corrosion rates, as described in the sections above, in this section, the effect of key factors which affect the rate of CO_2 corrosion are discussed, and the predictions made by the model are compared to empirical results.

2.25.2.6.1 Effect of pH

The pH has a significant influence on the CO_2 corrosion rate. Lower pH leads to higher corrosion rates and vice versa, just like in many other acidic solutions. Typical pH in CO_2 saturated condensed water is about pH 4 while in buffered brines, one frequently encounters $5 < pH < 7$. At pH 4 or below, direct reduction of H^+ ions, reaction [24], is important, particularly at lower partial pressures of CO_2, when direct reduction of carbonic acid, reaction [25], can be ignored. In that case, the pH has a direct effect on the corrosion rate. Another important effect of pH is indirect and relates to how pH changes conditions for the formation of ferrous carbonate layers. Higher pH ($5 < pH < 7$) results in a decreased solubility of ferrous carbonate and leads to an increased precipitation rate and a higher scaling tendency. The effect of various pH and supersaturations are shown in Figure 5. At lower supersaturations obtained at the lower pH of 6, shown in Figure 5, the corrosion rate does not change much with time, even if some ferrous carbonate precipitation occurs, reflecting the fact that a relatively porous, detached and unprotective layer is formed (low scaling tendency ST). The higher pH of 6.6 results in higher supersaturation, faster precipitation, and formation of more protective ferrous carbonate, reflected by a rapid decrease of the corrosion rate with time. There are other indirect effects of pH, and by almost all accounts, higher pH leads to a reduction of the corrosion rate, making the 'pH stabilization' (meaning: pH increase) technique an attractive way of managing CO_2 corrosion. The drawback of this technique is that it can lead to excessive scaling and can rarely be used with formation water systems.

2.25.2.6.2 Effect of CO_2 partial pressure

In the case of scale-free CO_2 corrosion, an increase of p_{CO_2} typically leads to an increase in the corrosion rate. The commonly accepted explanation is that with increasing p_{CO_2} the concentration of H_2CO_3 increases and accelerates the cathodic reaction, eqn [25], and ultimately the corrosion rate. The detrimental effect of p_{CO_2} at a constant pH is illustrated in Figure 6. The model described above reasonably

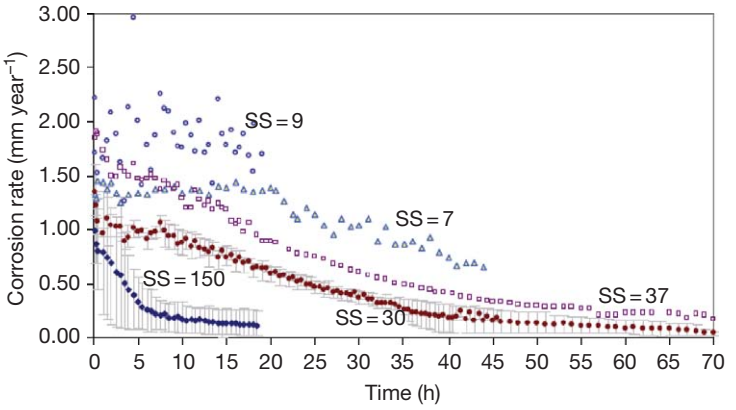


Figure 5 Effect of ferrous carbonate supersaturation $SS_{(FeCO_3)}$ on corrosion rate obtained at a range of pH 6.0–6.6, for 5 ppm $< c_{Fe^{2+}} <$ 50 ppm at $T = 80\,^{\circ}C$, under stagnant conditions. Error bars represent minimum and maximum values obtained in repeated experiments. Data taken from Chokshi *et al.*[17]

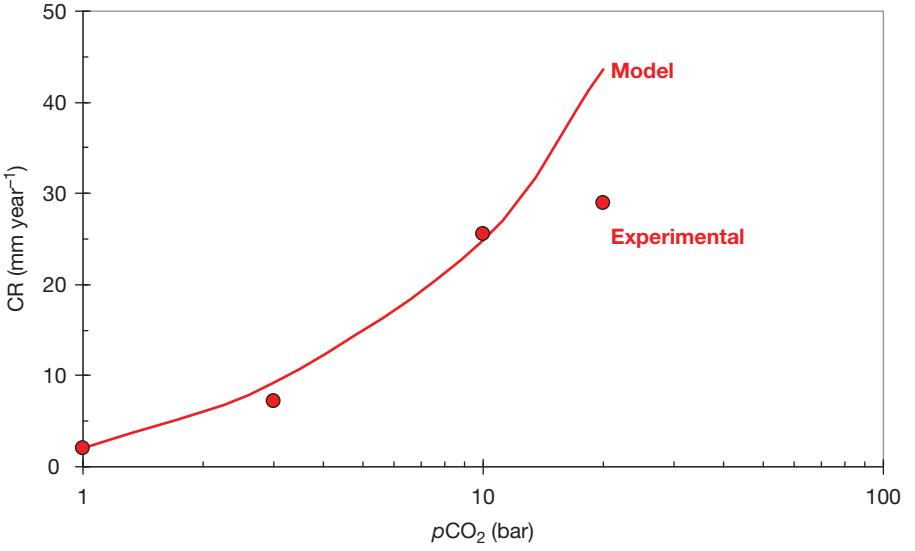


Figure 6 The effect of CO_2 partial pressure, p_{CO_2} on bare steel corrosion rate, comparison of experimental results (points) and model (line); 60 °C, pH 5, 1 m s^{-1}, 100 mm ID single-phase pipe flow.

captures well this trend up to approximately $p_{CO_2} = 10$ bar. However, when other conditions are favorable for the formation of ferrous carbonate layers, increased p_{CO_2} can have a beneficial effect. At a high pH, higher p_{CO_2} leads to an increase in bicarbonate and carbonate ion concentration and a higher supersaturation which accelerates precipitation and protective layer formation. The effect of p_{CO_2} on the corrosion rate in the presence of ferrous carbonate precipitation is illustrated in Figure 7 where in stratified wet gas flow, corrosion rate is reduced both at top and bottom of the pipe with the increase partial pressure of CO_2.

2.25.2.6.3 *Effect of temperature*

Temperature accelerates all the processes involved in corrosion: electrochemical, chemical, transport, etc. One would expect then that the corrosion rate steadily increases with temperature, and this is the case at low pH when precipitation of ferrous carbonate or other protective layers does not occur. An example is shown Figure 8. The situation changes markedly when solubility of ferrous carbonate is exceeded, typically at a higher pH. In that case, increased temperature rapidly accelerates the kinetics of precipitation and protective layer formation, decreasing the corrosion rate. The peak in the

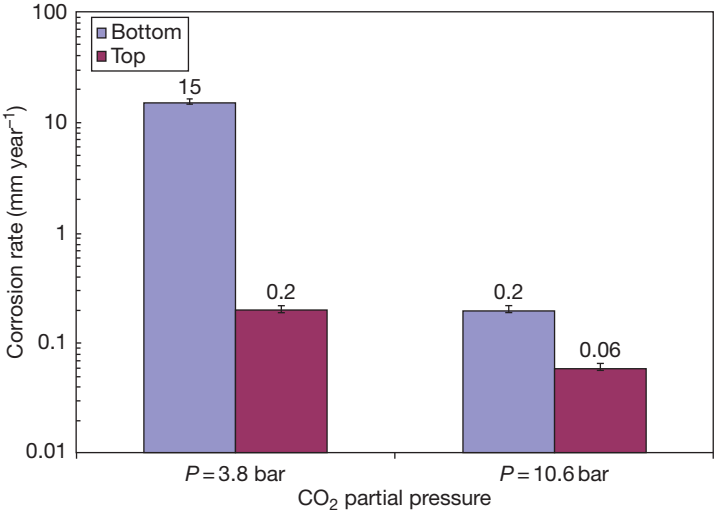


Figure 7 Experimental measurements of the corrosion rate at the top and bottom of the pipe in stratified gas–liquid flow showing the effect of CO_2 partial pressure, p_{CO_2} on formation of ferrous carbonate layer. Test conditions: 90 °C, pH 6, 100 mm ID, $V_{sg} = 10\,m\,s^{-1}$, $V_{sl} = 0.1\,m\,s^{-1}$. Data taken from Sun and Nešić.[18]

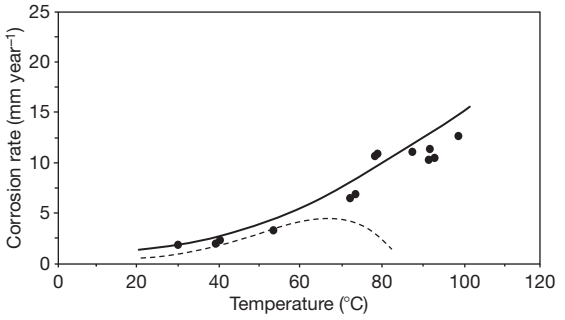


Figure 8 The effect of temperature on CO_2 corrosion rate of mild steel; pH 4, $p_{CO_2} = 1$ bar, 100 mm ID single phase pipe flow. Points are experimental values and the solid line is the model. The dotted line is a model simulation of the same conditions at pH 6.6 accounting for protective ferrous carbonate film formation.

corrosion rate is usually seen between 60 and 80 °C depending on water chemistry and flow conditions as shown in Figure 8 (dotted line).

2.25.2.6.4 *Effect of flow*

There are two main ways in which flow may affect CO_2 corrosion, which can be distinguished based on whether or not other conditions are conducive to protective layer formation or not.

In the case of corrosion where protective layers do not form (typically at low pH as found in condensed water and in the absence of inhibitors), the main role of turbulent flow is to enhance transport of species toward and away from the metal surface. This may lead to an increase in the corrosion rate as illustrated in Figure 9. At lower pH 4, the effect is much more pronounced as the dominant cathodic reaction is direct H^+ ion reduction [24], which is under mass transfer control (see eqn [38]).

When protective ferrous carbonate layers form (typically at higher pH in produced water) or when inhibitor films are present on the steel surface, the above-mentioned effect of flow becomes insignificant as the main resistance to corrosion is now in the surface layer or inhibitor film. In this case, the effect of flow is to interfere with the formation of protective surface layers or to remove them once they are in place, often leading to an increased risk of localized attack.

The two flow accelerated corrosion effects discussed above are frequently aggravated by flow disturbances such as valves, constrictions, expansions, bends, etc. where local increases of near-wall turbulence and wall-shear stress are seen. However, flow can lead to onset of localized attack only when given the 'right' set of circumstances as discussed in a separate heading below.

The effect of multiphase flow on CO_2 corrosion is complicated by the different flow patterns that exist, the most common being *stratified*, *slug*, and *annular-mist* flow. In the liquid phase, water and oil can flow separated or mixed with either phase being continuous with the other flowing as a dispersed phase. Different flow patterns lead to a variety of steel surface wetting mechanisms: stable water wetting, stable oil wetting, intermittent wetting, etc., which

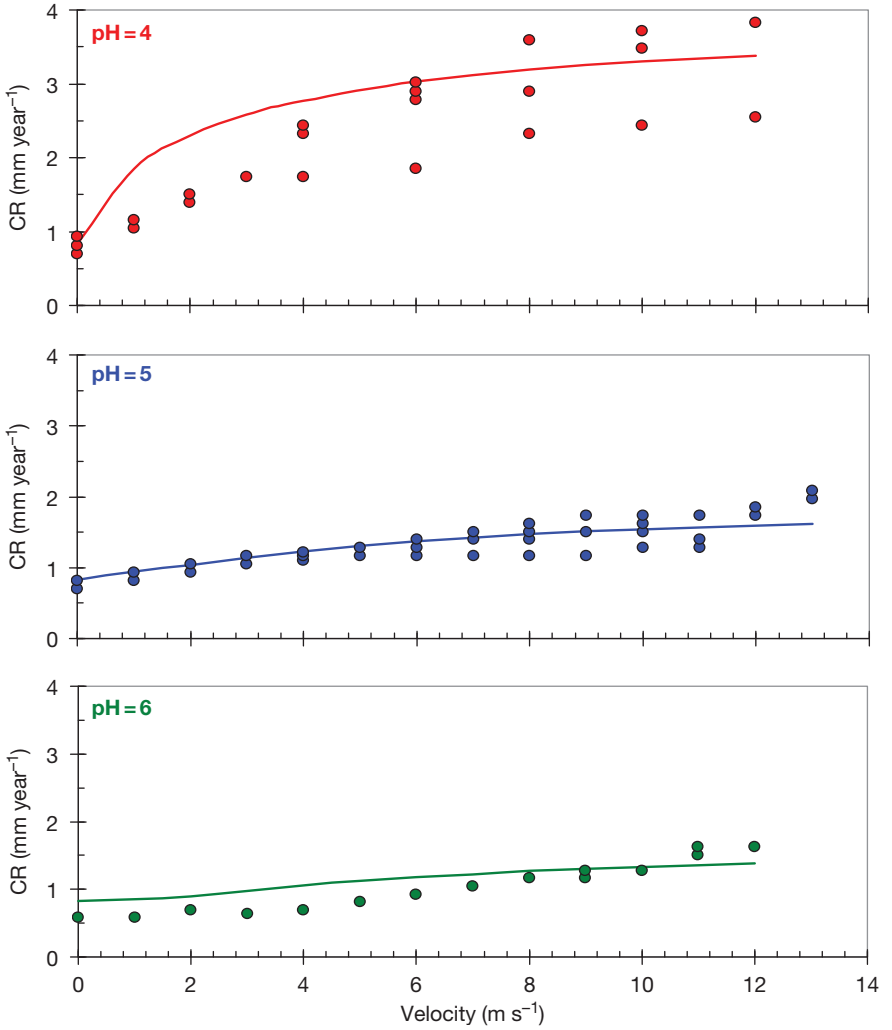


Figure 9 Predicted (line) and experimentally measured corrosion rates (points) showing the effect of velocity in the absence of ferrous carbonate layers. Test conditions: 20 °C, p_{CO_2} = 1 bar, 15 mm ID single-phase pipe flow. Experimental data taken from Nešić *et al.*[19]

greatly affect corrosion. In annular mist flow, the liquid droplets move at high velocity and can lead to protective layer damage at points of impact such as bends, valves, tees, constrictions/expansions, and other pipe fittings. Slug flow can lead to significant short-lived fluctuations in the wall-shear stress, which can help remove a protective surface layer of ferrous carbonate or possibly affect an inhibitor film.

2.25.2.6.5 Effect of corrosion inhibition

The two most common sources of corrosion inhibition need to be considered:

(a) inhibition by addition of corrosion inhibitors and
(b) inhibition by components present in the crude oil.

Corrosion inhibitors

Describing the effect of corrosion inhibitors is not a straightforward task due to the enormous complexity of the subject. Quantifying them and predicting their behavior are even harder. There is a plethora of approaches in the open literature, varying from the use of simple *inhibitor factors* and *inhibition efficiencies* to the application of complicated *molecular modeling* techniques to describe inhibitor interactions with the steel surface and ferrous carbonate layer. A middle-of-the-road approach is based on the assumption that corrosion protection is achieved by *surface coverage*, that is, that the inhibitor adsorbs onto the steel surface and slows down one or more electrochemical reactions by 'blocking.' The degree of protection is

assumed to be directly proportional to the fraction of the steel surface blocked by the inhibitor. In this type of model, one needs to establish a relationship between the surface coverage θ and the inhibitor concentration in the solution c_{inh}. This is most commonly done by the use of adsorption isotherms.

Corrosion inhibition by crude oil

It has been known for a while that CO_2 corrosion rates seen in the field in the presence of crude oil are much lower than those obtained in laboratory conditions where crude oil was not used or synthetic crude oil was used. One can identify two main effects of crude oil on the CO_2 corrosion rate.

The first is a *wettability* effect and relates to a hydrodynamic condition where crude oil entrains the water and prevents it from wetting the steel surface (continuously or intermittently).

The second effect is *corrosion inhibition* by components of the crude oil that reach the steel surface either by direct contact or by first partitioning into the water phase. Various surface active organic compounds found in crude oil (typically oxygen, sulfur and nitrogen containing molecules) have been identified to directly inhibit corrosion of mild steel in CO_2 solutions.

2.25.2.6.6 Effect of organic acids

The low molecular weight organic acids are primarily soluble in water and can lead to corrosion of mild steel. Higher molecular weight organic acids are not water soluble, but are typically soluble in the oil phase and pose a corrosion threat at higher temperatures in the refineries. Acetic acid CH_3COOH (denoted as HAc in the text below) is the most prevalent low molecular weight organic acid found in brines. Other acids typically found in the brine are propionic, formic, etc.; however, their behavior and corrosiveness is very similar to that of HAc and therefore HAc can be used as a 'surrogate' for all the organic acids found in the brine. HAc is a weak acid; however, it is stronger than H_2CO_3 (pK_a 4.76 vs. 6.35 at 25 °C), and it is the main source of H^+ ions when the two acid concentrations are similar. The effect of HAc is particularly pronounced at higher temperatures and low pH when the abundance of undissociated HAc can increase the CO_2 corrosion rate dramatically as seen in Figure 10. Solid iron acetate does not precipitate in the pH range of interest since its solubility is much higher than that of ferrous carbonate. There are some indications that the presence of organic acids impairs the protectiveness of ferrous carbonate layers; however, the mechanism is still not clear.

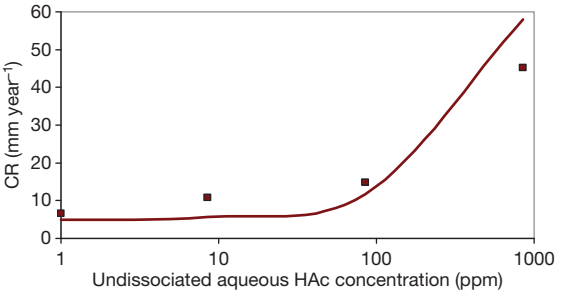


Figure 10 Predicted (line) and experimentally measured data (points) showing the effect of the concentration of undissociated acetic acid (HAc) on the CO_2 corrosion rate, 60 °C, $p_{CO_2} = 0.8$ bar, pH 4, 12 mm OD rotating cylinder flow at 1000 rpm. Experimental data taken from Sun *et al.*[20]

2.25.2.6.7 Effect of glycol/methanol

Glycol and methanol are often added to flowing systems in order to prevent hydrates from forming. The quantities are often significant (50% of total liquid phase is not unusual). In the very few studies available, it has been assumed that the main 'inhibitive' effect of glycol/methanol on corrosion comes from dilution of the water phase, which leads to a decreased activity of water. However, there are many unanswered questions such as the changes in mechanisms of CO_2 corrosion in water/glycol mixtures which are yet to be discovered.

2.25.2.6.8 Effect of condensation in wet gas flow

When transporting humid natural gas, due to the cooling of the stream, condensation of water vapor occurs on the internal pipe wall. The condensed water is pure and, due to dissolved CO_2, typically has a pH < 4. This leads to the so-called *top-of-the-line corrosion* (TLC) scenario. If the rate of condensation is high, plenty of acidic water flows down the internal pipe walls leading to a very corrosive situation. If the condensation rate is low, the water film is not renewed and flows down very slowly and the corrosion process can release enough Fe^{2+} to raise the local pH and saturate the solution, leading to the formation of protective ferrous carbonate layer. The layer is often protective; however, incidents of localized attack in TLC were reported.[21] Either way, the stratified or stratified-wavy flow regime, typical for TLC, does not lead to a good opportunity for inhibitors to reach the upper portion of the internal pipe wall and protect it. A very limited

range of corrosion management options for TLC exists. To qualitatively and quantitatively describe the phenomenon of corrosion occurring at the top of the line, a deep insight into the combined effect of the chemistry, hydrodynamics, thermodynamics, and heat and mass transfer in the condensed water is needed. A full description exceeds the scope of this review, and the interested reader is directed to some recent articles published on this topic.[21,22]

2.25.2.6.9 *Nonideal solutions and gases*

In many cases produced, water has very high dissolved solids content (>10 wt%). At such high concentrations, the infinite dilution theory used above does not hold, and corrections need to be made to account for solution nonideality. A simple way to account for the effect on nonideal homogenous water chemistry is to correct the equilibrium constants by using the concept of ionic strength as indicated above. This approach seems to work well only for moderately concentrated solution (up to a few weight percentage of dissolved solids). For more concentrated solutions, a more accurate way is to use activity coefficients as described by Anderko *et al.*[23] The effect of concentrated solutions on heterogeneous reactions such as precipitation of ferrous carbonate and other layers is still largely unknown. Furthermore, it is unclear how the highly concentrated solutions affect surface electrochemistry. Some experience suggests that corrosion rates can be dramatically reduced in very concentrated brines; nevertheless a more systematic study is needed.

At very high total pressure, the gas–liquid equilibria cannot be accounted for by Henry's law. A simple correction can be made by using a fugacity coefficient, which accounts for nonideality of the CO_2/natural gas mixture[24] and can be obtained by solving the equation of state for the gas mixture. Those cases, in which critical point for CO_2 is approached or exceeded, warrant a separate analysis and are not covered by the considerations discussed above.

2.25.2.7 Localized CO_2 Corrosion of Mild Steel in Aqueous Solutions

As illustrated above, significant progress has been achieved in understanding uniform CO_2 corrosion, without or with protective layers, and hence successful uniform corrosion models can be built. However, much less is known about localized CO_2 corrosion. It is thought that one of the main factors that 'triggers' localized attack is flow, tempered by other environmental variables such as pH, temperature, partial pressure of CO_2, etc. It seems that localized attack occurs when the conditions are such that partially protective ferrous carbonate layers form. It is well known that when fully protective ferrous carbonate forms, low general corrosion rates are obtained and vice versa: when no protective layers form, a high rate of general corrosion is seen. It is when the corrosive environment is 'in between,' in the so-called 'gray zone,' that localized attack can be initiated most often by some extreme flow conditions. There are many combinations of environmental and metallurgical parameters that define the grey zone, making this sound like a difficult proposal. However, there is a single parameter which is easy to calculate: ferrous carbonate supersaturation, $SS_{(FeCO_3)}$ (see eqn [19] above), which can be successfully used as a good delineator for the gray zone and as such as a predictor for the probability for localized attack. When bulk ferrous carbonate supersaturation is in the range $0.5 < SS_{(FeCO_3)} < 2$ there is a risk of localized attack. The further away the solution is from these boundaries, the lower the risk. The scaling tendency ST (see equation [22] above) is conceptually even better suited as a predictor of localized corrosion risk, however, its calculation is much more difficult and uncertain as it involves calculation of both the uniform corrosion rate and the precipitation rate.

Based on mostly anecdotal evidence (field experience), the presence of H_2S and HAc is related to the onset of localized attack, however, little is understood about how and when this may happen.

2.25.3 Aqueous H_2S Corrosion of Mild Steel

Corrosion of mild steel in the presence of hydrogen sulfide (H_2S) also represents a significant problem for the oil and gas industry.[27–33] Many fields have been developed that in addition to CO_2 have high concentrations of H_2S. In CO_2/H_2S corrosion of mild steel, thermodynamic considerations suggest that both ferrous carbonate and ferrous sulfide layers could theoretically form on the steel surface. However, studies have demonstrated that the formation of the ferrous sulfide layers is dominant and presents one of the most important factors governing the H_2S corrosion rate. The sulfide layer growth depends primarily on the kinetics of the corrosion process as it is described below.

Despite the relative abundance of experimental data on H_2S corrosion of steel, most of the literature is still confusing and somewhat contradictory.

Therefore, the mechanism of H_2S corrosion remains much less understood when compared to that of CO_2 corrosion. This uncertainty makes it more difficult to develop a model to predict the corrosion rate of mild steel in H_2S saturated aqueous solution.

2.25.3.1 Chemistry of H_2S Saturated Aqueous Solutions – Equilibrium Considerations

Similar to CO_2 discussed above, H_2S gas is also soluble in water:

$$H_2S_{(g)} \overset{K_{H_2S}}{\Leftrightarrow} H_2S \quad [62]$$

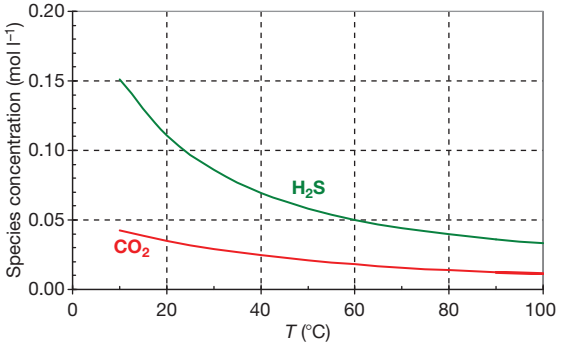


Figure 11 Calculated solubility of H_2S and CO_2 as a function of temperature; 25 °C, $p_{H_2S} = 1$ bar, $p_{CO_2} = 1$ bar.

where K_{H_2S} is the solubility constant of H_2S in mol l^{-1} bar^{-1}:

$$K_{sol(H_2S)} = \frac{c_{H_2S}}{p_{H_2S}} \quad [63]$$

and can be found from[34]

$$K_{sol(H_2S)} = 10^{(634.27+0.2709T_K-0.11132\times10^{-3}T_K^2-(16719/T_K)-261.9\log T_K)} \quad [64]$$

As shown in Figure 11, the solubility of H_2S decreases with temperature, as it is observed for CO_2. However, for the same partial pressure and temperature, the concentration of dissolved H_2S actually exceeds that in the gas phase as shown in Figure 12.

Aqueous H_2S is another weak acid which partly dissociates in two steps:

$$H_2S \overset{K_{hs}}{\Leftrightarrow} H^+ + HS^- \quad [65]$$

$$HS^- \overset{K_{bs}}{\Leftrightarrow} H^+ + S^{2-} \quad [66]$$

where K_{hs} is the dissociation constant of H_2S:

$$K_{hs} = \frac{c_{H^+}c_{HS^-}}{c_{H_2S}} \quad [67]$$

and can be calculated as[35]

$$K_{hs} = 10^{782.43945+0.361261T_K-1.6722\times10^{-4}T_K^2-(20565.7315/T_K)-142.741722\ln T_K} \quad [68]$$

and K_{bs} is the dissociation constant of HS^-:

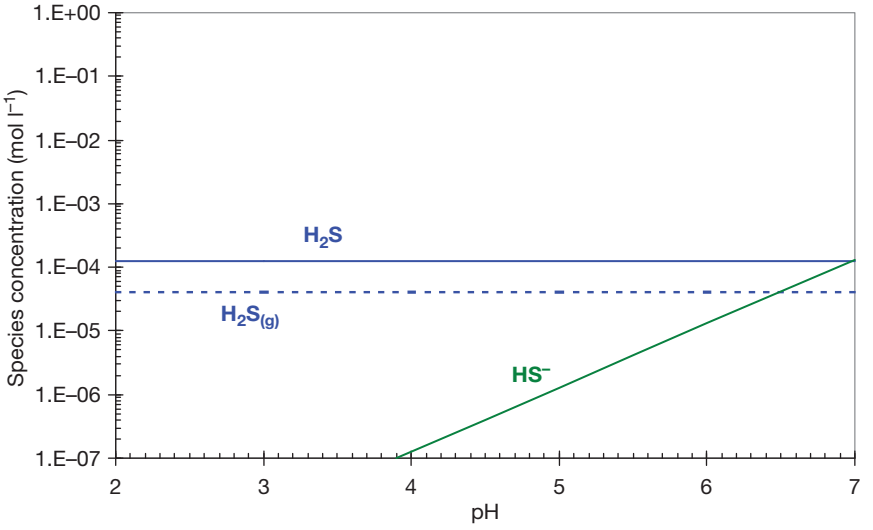


Figure 12 Calculated sulfide species concentrations as a function of pH for an H_2S saturated aqueous solution at $p_{H_2S} = 1$ mbar, 25 °C, 1 wt% NaCl.

$$K_{bs} = \frac{c_{H^+} c_{S^{2-}}}{c_{HS^-}} \quad [69]$$

There is a very large discrepancy in the reported values for K_{bs}, varying from 1.0×10^{-19} to 1.1×10^{-12} kmol m^{-3} at room temperature (seven orders of magnitude). In addition, these values are very small compared with other equilibrium constants, all suggesting that using K_{bs} to calculate the concentration of sulfide species, $c_{S^{2-}}$ and further to predict the solubility product constants for ferrous sulfides should be avoided.

Given the same gaseous concentrations of H_2S and CO_2, one obtains a similar aqueous concentration of dissolved H_2S and CO_2 (see **Figure 11**) and the resulting pH is within 0.1 pH unit, therefore, values shown in **Figure 1** for CO_2 can be used for H_2S as the first approximation. The equilibrium distribution of sulfide species as a function of pH for an open system is shown in **Figure 12**. The concentration of bisulfide ion, c_{HS^-}, becomes significant only above pH 4, while the concentration of the sulfide ion, $c_{S^{2-}}$, is not even shown as it is very low and unreliable to calculate.

Many types of iron sulfides, such as amorphous ferrous sulfide (FeS), mackinawite ($Fe_{1+x}S$), cubic ferrous sulfide (FeS), troilite (FeS), pyrrhotite ($Fe_{1-x}S$ or FeS_{1+x}), smythite ($Fe_{3+x}S_4$), greigite (Fe_3S_4), and pyrite (FeS_2) occur. Studies have suggested that some of these are stoichiometric such as cubic ferrous sulfide, troilite, greigite, and pyrite, while others such as mackinawite, pyrrhotite, and smythite are not. Some are electrically nonconductive, others apparently behave as semiconductors. However, there is no consensus on these issues and the interested reader is directed to the vast literature on iron sulfides for a more in-depth treatment. The thermodynamics of these systems is very complicated; depending on environmental conditions and time, transformation from one type of ferrous sulfide into the other occurs. Limited information exists on aqueous solubility of the various sulfides. Avoiding the usage of the sulfide ion concentration, $c_{S^{2-}}$, one can write a general equation for precipitation of ferrous sulfide as

$$Fe^{2+} + H_2S \overset{K_{sp(FeS)}}{\Leftrightarrow} FeS_{(s)} + 2H^+ \quad [70]$$

where the solubility constant for one type of ferrous sulfide – mackinawite is known as a function of temperature[36,37]

$$K_{sp(FeS)}^{mackin} = 10^{(2848.779)/T_k - 6.347} \quad [71]$$

For other ferrous sulfides, only the values at room temperature are known, as listed in **Table 1** below. It is convenient to show various ferrous sulfide solubilities in terms of an equilibrium concentration of the Fe^{2+} as a function of pH at a given H_2S partial pressure (concentration). An example is presented in **Figure 13** where it can be seen that the much less soluble pyrrhotite and troilite are thermodynamically more stable forms compared to mackinawite and amorphous ferrous sulfide. For a typical ferrous ion concentration of $c_{Fe^{2+}} = 1$ ppm, the saturation with respect to troilite and pyrrhotite is reached already at pH 5.4, while for mackinawite it is pH 6 and for amorphous ferrous sulfide pH 6.7. Keeping in mind that the concentration of Fe^{2+} at a corroding steel surface can easily be much higher than in the bulk (e.g., 10 ppm or even higher) and that the pH is also higher at the surface than in the bulk (typically above pH 6), using **Figure 13** one can expect a whole range of different ferrous sulfides to form on a corroding steel surface at this H_2S concentration at different points in time.

SEM images of a ferrous sulfide surface layer formed on mild steel after a week long exposure are shown in **Figure 14**. The layered structure of the sulfide is prominent, and it can be identified as mackinawite. In longer exposures, the ferrous sulfide layer thickens and eventually becomes more protective. An image of a ferrous sulfide layer after a month long exposure is shown in **Figure 15**. The composition of the layer is a mixture of mackinawite and pyrrhotite. Another layered structure composed of a mixture of ferrous carbonate and ferrous sulfide is shown in **Figure 16**.

Table 1 Solubility product constants for various ferrous sulfides at 25 °C[38]

Type of ferrous sulfide	$-\log K_{sp(FeS)}$
Amorphous (FeS)	2.95
Mackinawite ($Fe_{1+x}S$)	3.6
Pyrrhotite ($Fe_{1-x}S$ or FeS_{1+x})	5.19
Troilite (FeS)	5.31

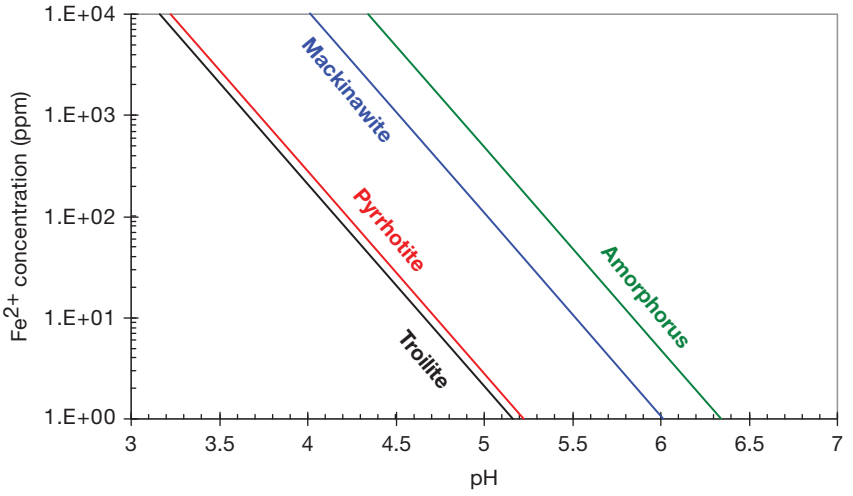


Figure 13 Calculated solubility of various iron sulfides as a function of pH shown in terms of the equilibrium concentration of Fe^{2+}, $p_{H_2S} = 1$ mbar, 25 °C, 1 wt% NaCl.

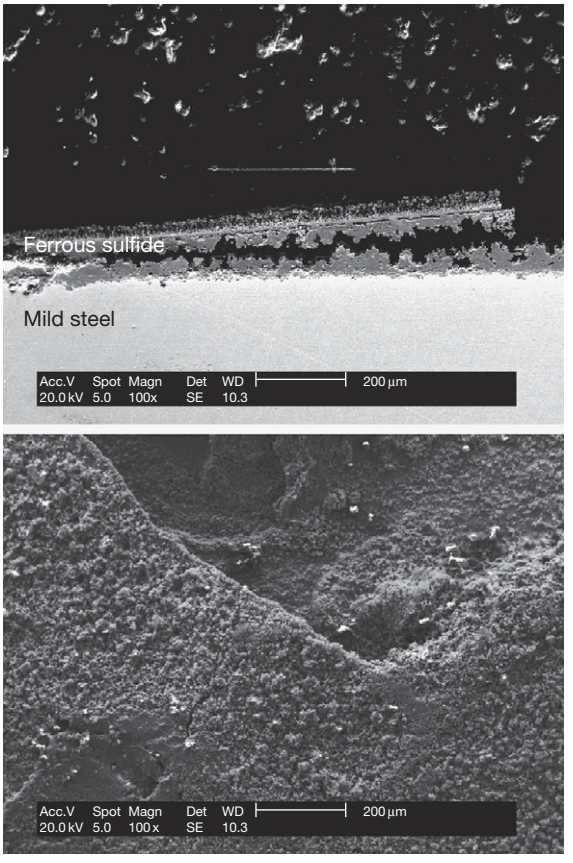


Figure 14 SEM images showing a cross-section and a top view of a ferrous sulfide layer formed on mild steel; 60 °C, pH 6, $p_{CO_2} = 7.7$ bar, $p_{H_2S} = 0.25$ mbar, 1 m s^{-1} single phase flow in a 100 mm ID pipe, 7 days exposure.

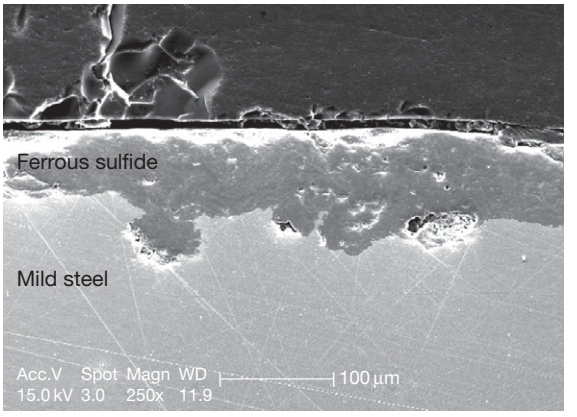


Figure 15 SEM images showing a cross-section view of a ferrous sulfide layer formed on mild steel; 60 °C, pH 6, $p_{CO_2} = 7.7$ bar, $p_{H_2S} = 0.25$ mbar, 1 m s^{-1} single phase flow in a 100 mm ID pipe, 30 day exposure.

2.25.3.2 Mild Steel Corrosion in H_2S and Mixed H_2S/CO_2 Saturated Aqueous Solutions

As aqueous H_2S is another weak acid, it can be seen as an additional reservoir of H^+ ions according to reaction [65], similar to H_2CO_3. Therefore, stimulation of the hydrogen evolution reaction could also be expected in the presence of H_2S. Using the analogy with CO_2 corrosion, one must also allow the possibility of direct reduction of H_2S, that is, that the H_2S molecule can be adsorbed at the steel surface,

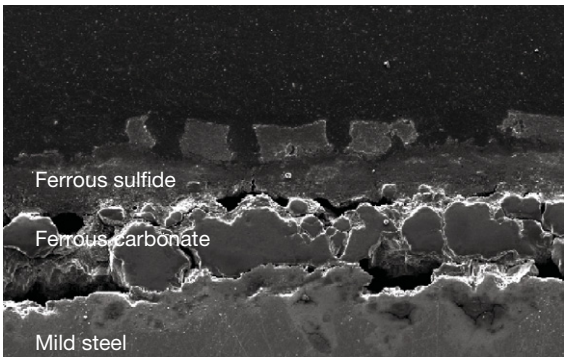


Figure 16 SEM images showing a cross-section view of a mixed ferrous carbonate and ferrous sulfide layer formed on mild steel; 60 °C, pH 6, $p_{CO_2} = 7.7$ bar, $p_{H_2S} = 1.2$ mbar, 1 m s^{-1} single phase flow in a 100 mm ID pipe, 25 day exposure.

followed by a reduction of the H^+ and oxidation of iron in the steel. One can write the overall corrosion reaction as

$$Fe_{(s)} + H_2S \rightarrow FeS_{(s)} + H_2 \qquad [72]$$

As solid ferrous sulfide (mackinawite) is always found on the corroding steel surface in the presence of H_2S, even below the solubility limit, this can been referred to as a direct 'solid state' reaction pathway as both the initial and final state of Fe are solid(s).[39]

Experimental evidence suggests that corrosion of mild steel by H_2S initially proceeds by adsorption of H_2S to the steel surface followed by a very fast redox reaction at the steel surface to form an adherent mackinawite film (much like a tarnish). This initial mackinawite film is very thin ($\ll$1 μm) but apparently rather dense and acts as a solid state diffusion barrier for the species involved in the corrosion reaction. Therefore, this thin mackinawite film is one of the most important factors governing the corrosion rate in H_2S corrosion. It also impedes the mobility of other species in reaching the steel surface and therefore corrosion rates due to CO_2 are affected even if very small amounts of H_2S are present in the gas phase (as little as 10^{-5} bar).

The thin mackinawite film continuously goes through a cyclic process of growth, internal stress growth, cracking, and delamination that generates an outer sulfide layer, which thickens over time (typically $\gg$1 μm) and forms an additional diffusion barrier. However, this outer sulfide layer is very porous and rather loosely attached to the steel surface. Over time it cracks, peels, and spalls, a process accelerated by turbulent flow. If the pH of the solution is below saturation level, the outer sulfide layer will undergo a process of chemical dissolution. Conversely, when the saturation is exceeded, ferrous sulfide precipitation from the bulk is possible. Eventually, the amount and protectiveness of the outer sulfide layer is determined by the balance of the various formation and removal processes.[39]

The transformation of mackinawite into other forms of less soluble and more stable ferrous sulfide (pyrrhotite and troilite, see **Figure 13**) may happen over time. Among the various ferrous sulfides, mackinawite is the prevalent ferrous sulfide that forms in the corrosion of mild steel at low H_2S concentration and low temperature. At increased levels of H_2S, mackinawite is less prevalent and pyrrhotite is the main corrosion product. At very high H_2S concentrations, pyrite and elemental sulfur appear. While thermodynamics of ferrous sulfides may favor other types of sulfide over mackinawite as the corrosion product, the rapid kinetics of mackinawite formation favors it as the initial corrosion product seen in most situations. Overall, however, there is currently no clearly defined relationship between the nature of the sulfide layer and the underlying corrosion process. It is generally thought that all types of ferrous sulfide layers offer some degree of corrosion protection for mild steel.

At very high H_2S concentrations, elemental sulfur can appear and lead to severe localized corrosion. Large amounts of elemental sulfur can precipitate out of the gas stream and can even block the line, due to the changes in pressure and temperature. Alternatively, when there is O_2 present, the most likely pathways for formation of elemental sulfur are as follows:

- ferrous sulfide reacts with O_2 and converts to iron oxide forming elemental sulfur probably via:

$$3FeS + 2O_2 \rightarrow Fe_3O_4 + 3S \qquad [73]$$

- at very high H_2S concentration, the following reaction can occur to yield elemental sulfur:

$$2H_2S + O_2 \rightarrow 2H_2O + 2S \qquad [74]$$

At very high temperatures, an alternative pathway is

$$H_2S \rightarrow H_2 + S \qquad [75]$$

Localized corrosion by elemental sulfur occurs via a reaction with the iron in the steel, represented by the

overall reaction

$$Fe + S \rightarrow FeS \quad [76]$$

It is not very clear at this stage what the detailed mechanism of this reaction is. It appears that rapid attack is seen only when direct contact of sulfur with the steel is achieved in the presence of water. A more in-depth discussion about the corrosion mechanisms of mild steel involving elemental sulfur exceeds the scope of this review.

2.25.3.3 Calculation of Mild Steel H_2S Corrosion Rate

Due to the complexity of the underlying processes and a lack of mechanistic understanding, predictive models of H_2S corrosion were not readily available until recently. One approach[40] which has the capability to address a few simple H_2S corrosion scenarios is presented below. A pure H_2S corrosion environment is described first followed by a mixed H_2S/CO_2 corrosion scenario.

2.25.3.3.1 Pure H_2S aqueous environment

Due to the presence of the inner mackinawite film and the outer porous sulfide layer, it is assumed that the corrosion rate of steel in H_2S solutions is always under mass transfer control. One can then write the flux of H_2S due to:

- convective diffusion through the mass transfer boundary layer as

$$\text{Flux}_{H_2S} = k_{m(H_2S)}\left(c_{H_2S} - c_{o(H_2S)}\right) \quad [77]$$

- molecular diffusion through the liquid in the porous outer sulfide layer as

$$\text{Flux}_{H_2S} = \frac{D_{H_2S}\varepsilon\psi}{\delta_{os}}\left(c_{o(H_2S)} - c_{i(H_2S)}\right) \quad [78]$$

- solid state diffusion through the inner mackinawite film as

$$\text{Flux}_{H_2S} = A_{H_2S}\exp\left(-\frac{B_{H_2S}}{RT_k}\right)\ln\left(\frac{c_{i(H_2S)}}{c_{s(H_2S)}}\right) \quad [79]$$

In a steady state, the three fluxes are equal to each other and are equivalent to the corrosion rate as

$$CR_{H_2S} = \text{Flux}_{H_2S} M_{Fe}/\rho_{Fe} \quad [80]$$

further corrected for appropriate corrosion rate unit.

By eliminating the unknown interfacial concentrations $c_{o(H_2S)}$ and $c_{i(H_2S)}$ from eqns [77] to [79], the following equation is obtained for the flux (corrosion rate) due to H_2S:

$$\text{Flux}_{H_2S} = A_{H_2S}\ln\frac{c_{H_2S} - \text{Flux}_{H_2S}\left(\dfrac{\delta_{os}}{D_{H_2S}\varepsilon\psi} + \dfrac{1}{k_{m(H_2S)}}\right)}{c_{s(H_2S)}} \quad [81]$$

This is an algebraic nonlinear equation with respect to Flux_{H_2S}, which does not have an explicit solution but can be solved by using a simple numerical algorithm such as the interval halving method or similar methods. These are available as ready-made routines in spreadsheet applications or in any common computer programming language. The prediction for Flux_{H_2S} depends on a number of constants used in the model which can be either found in handbooks (such as D_{H_2S}), calculated from the established theory (e.g., $k_{m(H_2S)}$) or are determined from experiments (e.g., $A_{H_2S}, c_{s(H_2S)}$). The unknown thickness of the outer sulfide layer change with time and need to be calculated as described below.

It is assumed that the amount of layer retained on the metal surface at any point in time depends on the balance of:

- layer formation kinetics (as the layer is generated by spalling of the thin mackinawite film underneath it and by the precipitation from the solution), and
- layer damage kinetics (as the layer is damaged by intrinsic or hydrodynamic stresses and/or by chemical dissolution):

$$\underbrace{\text{SRR}}_{\substack{\text{Sulfide layer}\\ \text{retension rate}}} = \underbrace{\text{SFR}}_{\substack{\text{Sulfide layer}\\ \text{formation rate}}} - \underbrace{\text{SDR}}_{\substack{\text{Sulfide layer}\\ \text{damage rate}}} \quad [82]$$

where all the terms are expressed in kmol m^{-2} s^{-1}. In order to simplify the calculations, it can be assumed that in the typical range of application ($4 < pH < 7$), precipitation and dissolution of ferrous sulfide layer do not play a significant role and so it can be written

$$\text{SRR} = \text{CR} - \text{SDR}_m \quad [83]$$

Some experiments involving mackinawite have shown that even in stagnant conditions about half of the outer sulfide layer that forms is lost from the steel surface due to intrinsic growth stresses by internal cracking and spalling, that is, $\text{SDR}_m \approx 0.5\text{CR}$, so one obtains:

$$\text{SRR} = 0.5\text{CR} \quad [84]$$

that is, about half of the iron corroded is found on the steel surface in the form of mackinawite. It is not known if and how this ratio is different when other types of ferrous sulfide layers form, for example, the more adherent and protective pyrrhotite. Moreover, additional experimentation is required to determine how the mechanical layer damage is affected by hydrodynamic forces.

Once the layer retention rate SRR is known, the change in mass of the outer sulfide layer can be easily calculated as

$$\Delta m_{os} = \mathrm{SRR} M_{FeS} A \Delta t \qquad [85]$$

The porosity of the outer sulfide layer was determined to be very high ($\varepsilon \approx 0.9$) by comparing the weight of the layer with the cross-sectional SEM images showing its thickness. On the other hand, this layer has proven to be rather protective (i.e., impermeable to diffusion) which can only be explained by its low tortuosity arising from its layered structure. By comparing the measured and calculated corrosion rates in the presence of the outer sulfide layer, the tortuosity factor was calculated to be $\psi = 0.003$.

A time-marching explicit solution procedure could now be established where

1. the corrosion rate Flux_{H_2S} in the absence of outer sulfide layer can be calculated by using eqn [81], and assuming $\delta_{os} = 0$;
2. the amount of sulfide layer δm_{os} formed over a time interval Δt is calculated by using eqn [85];
3. the new corrosion rate Flux_{H_2S} in the presence of sulfide layer can be recalculated by using eqn [81];
4. a new time interval Δt is set and steps 2 and 3 repeated.

At very low H_2S gas concentrations (ppm_w range), there is very little dissolved H_2S and the corrosion rate is directly affected by pH. A mackinawite layer still forms and controls the corrosion rate; however, the corrosion process is largely driven by the reduction of H^+ ions, rather than of H_2S. By analogy with the approach laid out above, the following expression is obtained for the flux of H^+ ions controlled by the presence of the ferrous sulfide layers:

$$\mathrm{Flux}_{H^+} = A_{H^+} \ln \frac{c_{H^+} - \mathrm{Flux}_{H^+}\left(\frac{\delta_{os}}{D_{H^+}\varepsilon\psi} + \frac{1}{k_{m(H^+)}}\right)}{c_{s(H^+)}} \qquad [86]$$

The flux Flux_{H^+} is directly related to the corrosion rate by H^+ ions:

$$\mathrm{CR}_{H^+} = \frac{\mathrm{Flux}_{H^+}}{2}\frac{M_{Fe}}{\rho_{Fe}} \qquad [87]$$

further adjusted for the appropriate corrosion rate unit.

By solving eqns [81] and [86] sequentially in time, the total corrosion rate in mixed pure H_2S aqueous environments can be calculated as

$$\mathrm{CR} = \mathrm{CR}_{H_2S} + \mathrm{CR}_{H^+} \qquad [88]$$

2.25.3.3.2 Mixed CO_2/H_2S environments

For mild steel corrosion in mixed CO_2/H_2S containing environments, one can account for the effect of CO_2 by assuming that the rate controlling step in this additional process is the diffusion of CO_2 through the ferrous sulfide layers. Then a similar expression can be obtained for the corrosion rate due to CO_2:

$$\mathrm{Flux}_{CO_2} = A_{CO_2} \ln \frac{c_{CO_2} - \mathrm{Flux}_{CO_2}\left(\frac{\delta_{os}}{D_{CO_2}\varepsilon\psi} + \frac{1}{k_{m(CO_2)}}\right)}{c_{s(CO_2)}} \qquad [89]$$

The flux Flux_{CO_2} is equivalent to the corrosion rate by CO_2:

$$\mathrm{CR}_{CO_2} = \frac{\mathrm{Flux}_{CO_2}}{2}\frac{M_{Fe}}{\rho_{Fe}} \qquad [90]$$

further adjusted for appropriate corrosion rate unit.

By solving eqns [81], [86], and [89], the total corrosion rate in mixed CO_2/H_2S environments can be calculated as

$$\mathrm{CR} = \mathrm{CR}_{H_2S} + \mathrm{CR}_{H^+} + \mathrm{CR}_{CO_2} \qquad [91]$$

2.25.3.4 Limitations of Modeling of Aqueous H_2S Corrosion of Mild Steel

The calculation model presented above covers uniform H_2S and CO_2/H_2S corrosion. There are numerous limitations:

- It does not predict localized corrosion in either environment.
- While it covers a very broad range of H_2S partial pressures, it is not recommended to use this model below $p_{H_2S} = 0.01$ mbar or above $p_{H_2S} = 10$ bar. Similar limits apply to the CO_2 partial pressure.

This leaves a very broad area of applicability for the present model.

- This H_2S model does not account for any precipitation of ferrous sulfide, ferrous carbonate, or any other scale; therefore, in cases where this is deemed important for corrosion, the model should be used with caution. The model also does not account for various transformations of sulfide layer from one type to another which are known to happen over time.
- The present model does not account for dissolution of the sulfide layer that may occur at very low pH. Therefore, the use of this model at pH < 3 is not recommended. Similarly, the model should be used with caution for pH > 7 where it has not been tested.
- The model in its present state does not cover the effect of organic acids on mixed H_2S and CO_2/H_2S corrosion, and therefore it should not be used when organic acids are present in the system. A practical threshold for the validity of the present model is <1 ppm of organic acids in the brine.
- The model does not account for the effect of high chloride concentrations, oxygen, elemental sulfur or any other unspecified condition, which is known to affect the corrosion rate and is not explicitly covered in the theoretical underpinnings discussed above.

While this calculation model is clearly not inclusive of all the important processes in aqueous H_2S corrosion of mild steel, it is believed that the main underlying assumptions about the formation and protective nature of a mackinawite layer are correct. The comparison of the performance of this model with experimental data is given in the following section, which covers the main factors affecting CO_2/H_2S corrosion of mild steel.

2.25.3.5 Key Factors Affecting Aqueous H_2S Corrosion of Mild Steel

Some of the key factors affecting aqueous H_2S corrosion are discussed in these sections, and the experimental results are compared with the model described above.

2.25.3.5.1 Effect of H_2S partial pressure

Corrosion rate of mild steel at extremely low H_2S partial pressures is seen in Figure 17 wherein atmospheric glass cell experiments p_{H_2S} ranged from 0.0013–0.32 mbar, corresponding to 1–250 ppm_m in the gas phase at 1 bar CO_2. Clearly, this is a CO_2 dominated corrosion scenario (p_{CO_2}/p_{H_2S} ratio is in the range 10^3–10^6); however, the presence of H_2S controls the corrosion rate. Even when present in such minute amounts, H_2S reduces the pure CO_2 (H_2S-free) corrosion rate by 3–10 times due to the formation of a thin mackinawite film. The model presented above successfully captures this effect as shown in Figure 17.

At higher H_2S partial pressures, the same effect is observed as shown in Figure 18, which shows results from autoclave experiments conducted at a very high

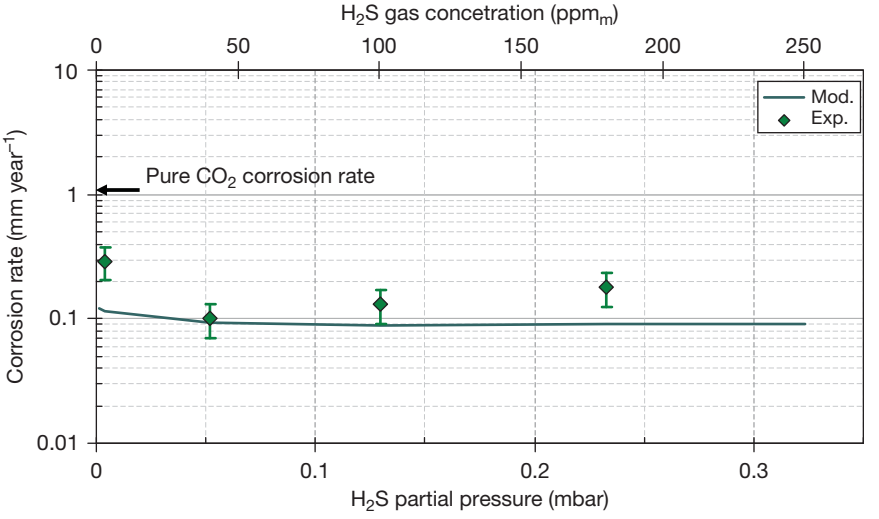


Figure 17 The corrosion rate vs. partial pressure of H_2S; experimental data (exp.) shown as points, model predictions (mod.) shown as lines; conditions: total pressure $p = 1$ bar, $p_{CO_2} = 1$ bar, $p_{H_2S} =$ 0.0013–0.32 mbar, $T = 20\,^{\circ}C$, reaction time 24 h, pH 5, 1000 rpm. For reference: pure CO_2 corrosion rate is measured to be 1 mm $year^{-1}$. Data taken from Lee.[41]

total pressure ($p = 138$ bar) and a high CO_2 partial pressure ($p_{CO_2} = 13.8$ bar). When comparing the predictions with the experimental results, it can be seen that the model underpredicts the observed rate of steel corrosion by approximately a factor of 2. However, when this is compared with a pure CO_2 (H_2S-free) corrosion rate under the same conditions (which is not reported but can be predicted to be almost 20 mm year^{-1}), the accuracy of the model can be considered as reasonable. At the highest p_{CO_2}/p_{H_2S} ratio of 3500 ($p_{CO_2} = 13.8$ bar, $p_{H_2S} = 40$ mbar), CO_2 accounts for ~70% of the corrosion rate and 30% can be ascribed to H_2S. At the lowest p_{CO_2}/p_{H_2S} ratio of 1180 ($p_{CO_2} = 13.8$ bar, $p_{H_2S} = 116$ mbar), CO_2 accounts for ~57% of the corrosion rate and 43% can be ascribed to H_2S.

Corrosion rates of mild steel at very high partial pressures of H_2S ($p_{H_2S} = 3$–20 bar) and CO_2 ($p_{CO_2} = 3$–12.8 bar) for exposures lasting up to 4 days are shown in **Figure 19**. This is a situation where the

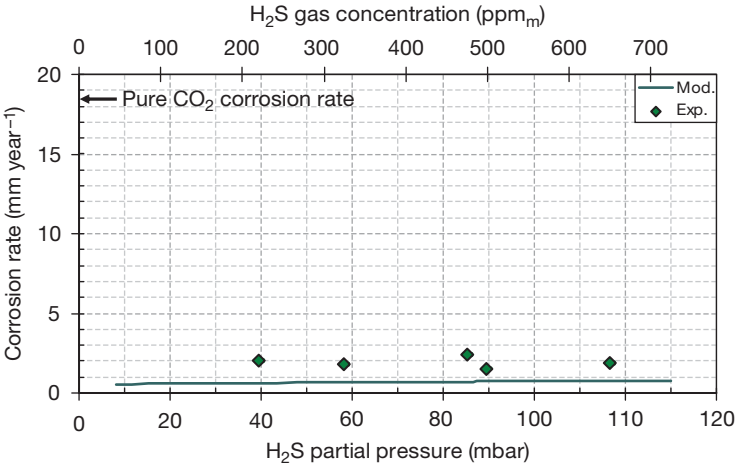


Figure 18 The corrosion rate vs. H_2S partial pressure; experimental data (exp.) shown as points, model predictions (mod.) shown as lines; conditions: total pressure $p = 137.9$ bar, $p_{CO_2} = 13.8$ bar, $p_{H_2S} = 40$–120 mbar, $T = 50\,^{\circ}C$, experiment duration 3 days, pH 4.0–6.2, stagnant. Experimental data taken from Smith and Pacheco *et al.*[31]

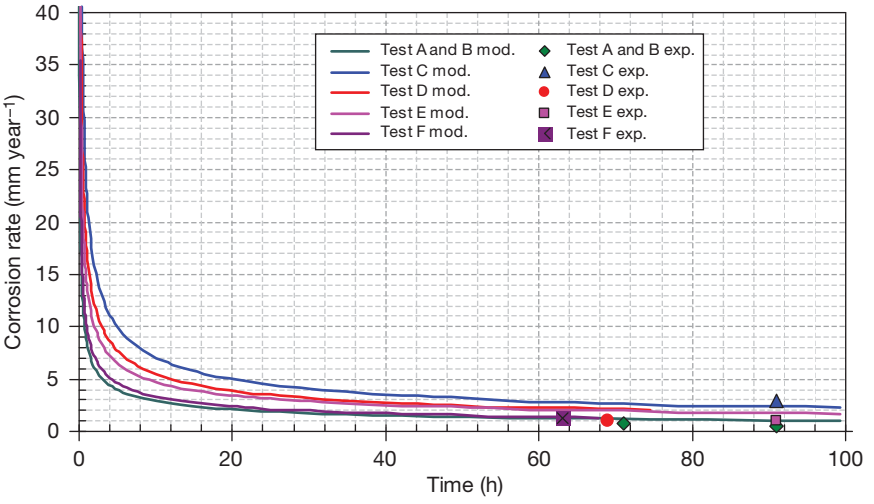


Figure 19 The corrosion rate vs. time; experimental data (exp.) shown as points, model predictions (mod.) shown as lines; Test A and B: $p = 8.3$ bar, $p_{CO_2} = 5.3$ bar, $p_{H_2S} = 3$ bar, $T = 60\,^{\circ}C$, 71 h (a) and 91 h (b); Test C: $p = 24$ bar, $p_{CO_2} = 4$ bar, $p_{H_2S} = 20$ bar, $T = 70\,^{\circ}C$, 91 h; Test D: $p = 15.7$ bar, $p_{CO_2} = 3.5$ bar, $p_{H_2S} = 12.2$ bar, $T = 65\,^{\circ}C$, 69 h; Test E: $p = 20.8$ bar, $p_{CO_2} = 12.8$ bar, $p_{H_2S} = 8$ bar, $T = 65\,^{\circ}C$, 91 h; Test F: $p = 7.2$ bar, $p_{CO_2} = 3$ bar, $p_{H_2S} = 4.2$ bar, $T = 65\,^{\circ}C$, 63 h; experimental data taken from Bich and Goerz.[43]

H_2S was the dominant corrosive species. At the highest p_{CO_2}/p_{H_2S} ratio of 1.8 ($p_{CO_2} = 5.3$ bar, $p_{H_2S} = 3$ bar), H_2S generated ~86% of the corrosion rate. At the lowest p_{CO_2}/p_{H_2S} ratio of 0.2 ($p_{CO_2} = 4$ bar, $p_{H_2S} = 20$ bar), H_2S generated 97% of the overall corrosion rate. It is also noted that the model predictions show that the corrosion rate in the first reaction hour is on average 20 mm year^{-1} with an initial corrosion rate of 60 mm year^{-1} and a final corrosion rate of 10 mm year^{-1}. The pitting corrosion rate was reported to be 30 mm year^{-1} in a field case with similar conditions, which is related to the very high, H_2S-driven corrosion seen at the beginning of experiments before a thick protective ferrous sulfide film forms.

2.25.3.5.2 Effect of flow

The effect of flow velocity in H_2S corrosion is shown in **Figure 20** for the three long-term experiments reported by Omar *et al.*[44] Flow loop experiments lasting 15–21 days were conducted at severe conditions: high partial pressure of H_2S (p_{H_2S}=10–30 bar), high partial pressure of CO_2 (p_{CO_2}=3.3–10 bar) and low pH 2.9–3.2. No effect of velocity on the uniform corrosion rate could be observed in these long-term exposures, which is due to the build-up of a thick protective sulfide layer. The model predictions also shown in **Figure 20** confirm this trend and show a remarkable agreement with the experimental results in the less extreme experiments 1 and 2 ($p_{CO_2} = 3.3$ bar; $p_{H_2S} = 10$ bar) both at low (25 °C) and high temperature (80 °C). In experiment 3 which was conducted at the most extreme set of conditions ($p_{CO_2} = 10$ bar; $p_{H_2S} = 30$ bar) and high temperature (80 °C) the model overpredicts the corrosion rate by a factor of 2.5. In all three experiments reported by Omar *et al.*,[44] the p_{CO_2}/p_{H_2S} ratio was about 0.3, that is, the corrosion process and corrosion rate were completely dominated by H_2S, which contributed ~95% of the corrosion rate.

2.25.3.5.3 Effect of time

A marked decrease of corrosion rate with time was seen in autoclave tests as reported in **Figure 19** above; the same was observed in stratified pipe flow experiments where pure CO_2 corrosion rate decreased with time due to the presence of H_2S, as shown in **Figure 21** below. The latter is also a mixed CO_2/H_2S corrosion scenario. At a p_{CO_2}/p_{H_2S} ratio of 200 (p_{CO_2} = 2 bar, p_{H_2S} = 4 mbar), the CO_2 contribution to the corrosion rate is 75% with most of the balance provided by H_2S. At the p_{CO_2}/p_{H_2S} ratio of 28 ($p_{CO_2} = 2$ bar, p_{H_2S}=70 mbar), both CO_2 and H_2S account for ~50% of the overall corrosion rate.

Corrosion experiments at high temperature (120 °C), high partial pressures of CO_2 ($p_{CO_2} = 6.9$ bar), and H_2S ($p_{H_2S} = 1.38$–4.14 bar) in exposures

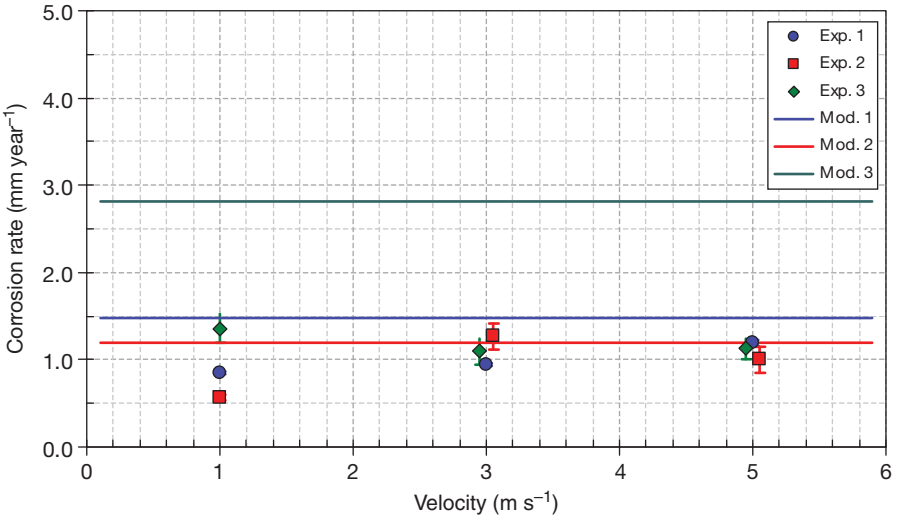


Figure 20 The corrosion rate vs. velocity; experimental data (exp.) shown as points, model predictions (mod.) shown as lines; exp 1.: 19 days, $p = 40$ bar, $p_{CO_2} = 3.3$ bar, $p_{H_2S} = 10$ bar, $T = 80$ °C, pH 3.1, $v = 1$–5 m s^{-1}; exp 2.: 21 days, $p = 40$ bar, $p_{CO_2} = 3.3$ bar, p_{H_2S}=10 bar, $T = 25$ °C, pH 3.2, $v = 1$–5 m s^{-1}; exp 3.: 10 days, $p = 40$ bar, $p_{CO_2} = 10$ bar, $p_{H_2S} = 30$ bar, $T = 80$ °C, pH 2.9, $v = 1$–5 m s^{-1}; experimental data taken from Omar *et al.*[44]

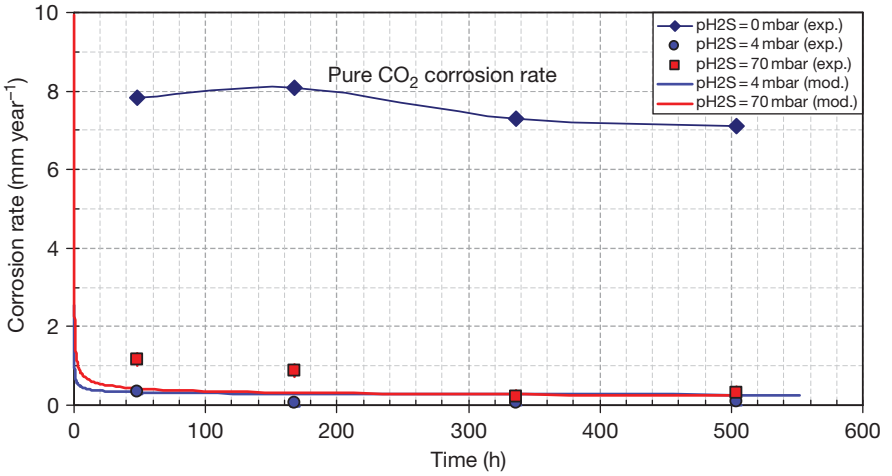


Figure 21 The corrosion rate vs. time; experimental data (exp.) shown as points, model predictions (mod.) shown as lines; conditions: total pressure $p = 3$ bar, $p_{CO_2} = 2$ bar, $p_{H_2S} = 3$–70 mbar, $T = 70°C$, experiment duration 2–21 days, pH 4.2–4.9, liquid velocity 0.3 m s^{-1}. Experimental data taken from Singer *et al.*[42]

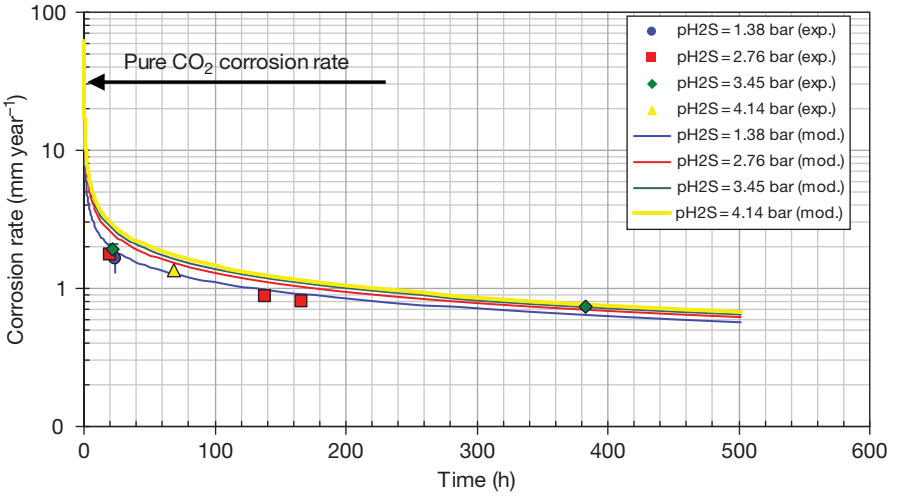


Figure 22 The corrosion rate vs. time; experimental data (exp.) shown as points, model predictions (mod.) shown as lines; conditions: total pressure $p = 7$ bar, $p_{CO_2} = 6.9$ bar, $p_{H_2S} = 1.38$–4.14 bar, $T = 120\,°C$, experiment duration 1–16 days, pH 3.95–4.96, liquid velocity 10 m s^{-1}. Experimental data taken from Kvarekval *et al.*[45]

lasting up to 16 days are shown in Figure 22. A steadily decreasing corrosion rate was observed due to build-up of a protective ferrous sulfide layer. The effect of p_{H_2S} increase on corrosion rate was very small and practically vanished over time. Both these effects were readily captured by the model with very good accuracy as seen in Figure 22. In this case, the H_2S is the dominant corrosive species. At the highest p_{CO_2}/p_{H_2S} ratio of 5 ($p_{CO_2} = 6.9$ bar, $p_{H_2S} = 1.38$ bar), H_2S generated ~70% of the corrosion rate. At the lowest p_{CO_2}/p_{H_2S} ratio of 1.67 ($p_{CO_2} = 6.9$ bar, $p_{H_2S} = 4.14$ bar), H_2S generated 82% of the overall corrosion rate.

The longest H_2S containing corrosion experiments which are practically achievable in the lab are of the order of a few weeks or at best a few months, while predictions are meant to cover a period of at least a decade, in order to be meaningful. With this in mind, it is interesting to take the experimental conditions above ($p_{CO_2} = 6.9$ bar, $p_{H_2S} = 3.45$ bar,

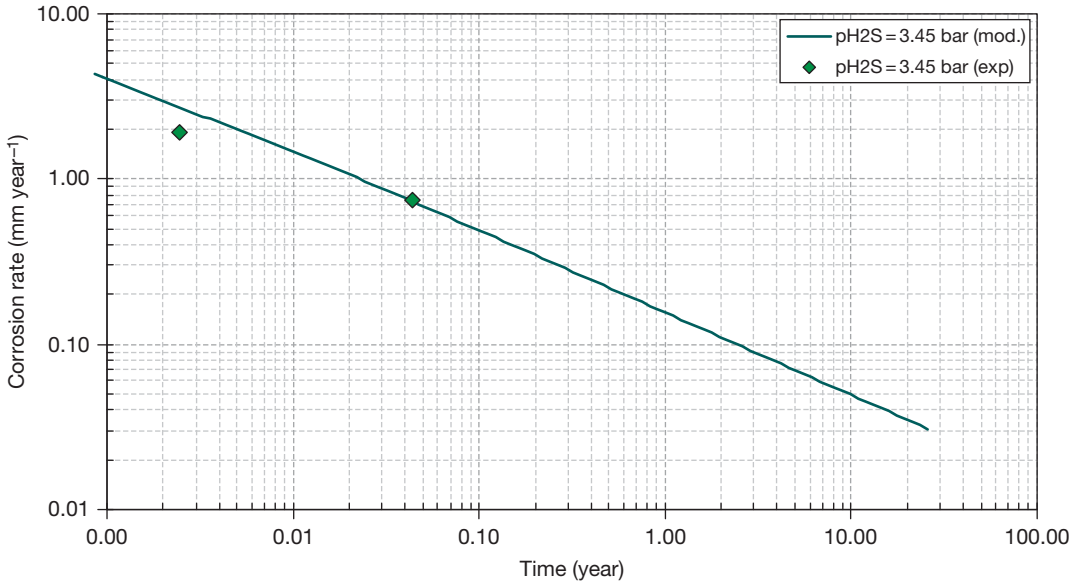


Figure 23 Extension of corrosion prediction to a 25-year lifetime; experimental (points), predicted (lines); conditions: $p_{CO_2} = 6.9$ bar, $p_{H_2S} = 3.45$ bar, $T = 120\,^{\circ}C$, pH 4, liquid velocity 10 m s^{-1}. Data from Kvarekval *et al.*[45]

$T = 120\,^{\circ}C$, pH 4, $v = 10$ m s^{-1}) and extend the simulation to 25 years. The result is shown in **Figure 23**. The corrosion rate was predicted to start out rather high as observed in the experiments; however, it was reduced to below 0.1 mm year^{-1} after 2 years and was as low as 0.03 mm year^{-1} after 25 years. The average corrosion rate over this period was only 0.06 mm year^{-1}, which amounts to a wall thickness loss of only 1.5 mm over the 25 years, an acceptable amount by any practical account. Actually, most of the other conditions simulated have shown that rather low H_2S uniform corrosion rates are obtained for very long exposures, which agrees with general field experience as recently discussed by Bonis *et al.*[33] Nevertheless, no quantitative long-term lab data are currently available to back-up these long-term predictions, and therefore they should be used with caution.

2.25.3.6 Localized H_2S Corrosion of Mild Steel in Aqueous Solutions

Localized H_2S corrosion of mild steel is even less understood than its uniform counterpart. While it is not very common, anecdotal evidence exists that has linked localized H_2S corrosion in aqueous environments to other factors such as high chloride content, the presence of elemental sulfur and the transformation of one type of sulfide into another. Intense research of these topics is ongoing with a breakthrough in understanding expected in the decade to come.

References

1. Bonis, M. R.; Crolet, J. L. In *Corrosion/89*; NACE International: Houston, TX, 1989; Paper no. 466.
2. Oddo, J.; Tomson, M. In *SPE of AIME*; Society of Petroleum Engineers: Richardson, TX, 1982; pp 1583–1590.
3. Palmer, D. A.; van Eldik, R. *Chem. Rev.* **1983**, *83*, 651.
4. Sun, W.; Nešić, S.; Woollam, R. C. *Corros. Sci.* **2009**, *51*(6), 1273–1276.
5. Sun, W.; Nešić, S. *Corrosion* **2008**, *64*, 334.
6. Drazic, D. M. In *Aspects of Electrochemistry*; Plenum Press, 1989; Vol. 19, p 79.
7. Lorenz, W.; Heusler, K. In *Corrosion Mechanisms*; Mansfeld, F., Ed.; Marcel Dekker: New York, 1987.
8. Nešić, S.; Thevenot, N.; Crolet, J. L. In *Corrosion/96*; NACE International: Houston, TX, 1996; Paper no. 3.
9. de Waard, C.; Milliams, D. E. *Corrosion* **1975**, *31*, 131.
10. Gray, L. G. S.; Anderson, B. G.; Danysh, M. J.; Tremaine, P. G. In *Corrosion/89*; NACE International: Houston, TX, 1989; Paper no. 464.
11. Eriksrud, E.; Søntvedt, T. In *Advances in CO_2 Corrosion*, Proceedings of the Corrosion/83 Symposium on CO_2 Corrosion in the Oil and Gas Industry; Hausler, R. H., Goddard, H. P., Eds.; NACE: Houston, TX, 1984; Vol. 1, p 20.
12. Schmitt, G.; Rothman, B. *Werkst. Korros.* **1977**, *28*, 816.
13. Gray, L. G. S.; Anderson, B. G.; Danysh, M. J.; Tremaine, P. R. In *Corrosion/90*; NACE International: Houston, TX, 1990; Paper no. 40.
14. Delahay, P. *J. Am. Chem. Soc.* **1952**, *74*, 3497.
15. Nešić, S.; Postlethwaite, J.; Olsen, S. In *Corrosion/95*; NACE International: Houston, TX, 1995; Paper no. 131.
16. Nešić, S.; Pots, B. F. M.; Postlethwaite, J.; Thevenot, N. *J. Corros. Sci. Eng.* **1995**, 1 Paper no. 3. Available at http://www.cp.umist.ac.uk/JCSE/Vol1/PAPER3/V1_p3int.htm
17. Chokshi, K.; Sun, W.; Nešić, S. In *Corrosion/05*; NACE International: Houston, TX, 2005; Paper no. 285.
18. Sun, Y.; Nešić, S. *Corrosion/2004*; NACE International: Houston, TX, 2004; Paper no. 380.

19. Nešić, S.; G.Solvi, T.; Enerhaug, J. *Corrosion* **1995**, *51*, p 773.
20. Sun, Y.; George, K.; Nešić, S. In *Corrosion/2003*; NACE International: Houston, TX, 2003; Paper no. 3327.
21. Gunaltun, Y. M.; Larrey, D. In *Corrosion/2000*; NACE International: Houston, TX, 2000; Paper no. 71.
22. Vitse, F.; Nešić, S.; Gunaltun, Y.; Larrey de Torreben, D.; Duchet-Suchaux, P. *Corrosion* **2003**, *59*, 1075.
23. Anderko, A.; Young, R. In *Corrosion/99*; NACE International: Houston, TX, 1999; Paper no. 31.
24. de Waard, C.; Lotz, U. *Corrosion/93*; NACE International: Houston, TX, 1993; Paper no. 69.
25. Berger, F. P.; Hau, K.-F. F.-L. *Int. J. Heat Mass Transfer* **1977**, *20*, 1185.
26. Eisenberg, M.; Tobias, C. W.; Wilke, C. R. *J. Electrochem. Soc.* **1954**, *101*, 306.
27. Shoesmith, D. W.; Taylor, P.; Bailey, M. G.; Owen, D. G. *J. Electrochem. Soc.* **1980**, *125*, 1007–1015.
28. Shoesmith, D. W. Formation, transformation and dissolution of phases formed on surfaces, In Lash Miller Award Address, Electrochemical Society Meeting, Ottawa, 27 November 1981.
29. Smith, S. N. In Twelfth International Corrosion Congress, Houston, TX, 19–24 September 1993; Paper no. 385.
30. Smith, S. N.; Wright, E. J. In *Corrosion/94*; NACE International: Houston, TX, 1994; Paper no. 11.
31. Smith, S. N.; Pacheco, J. L. In *Corrosion/2002*; NACE International: Houston, TX, 2002; Paper no. 02241.
32. Smith, S. N.; Joosten, M. In *Corrosion/2006*; NACE International: Houston, TX, 2006; Paper no. 06115.
33. Bonis, M.; Girgis, M.; Goerz, K.; MacDonald, R. In *Corrosion/2006*; NACE International: Houston, TX, 2006; Paper no. 06122.
34. Suleimenov, O. M.; Krupp, R. E. *Geochim. Cosmochim. Acta* **1994**, *58*, 2433–2444.
35. Suleimenov, O. M.; Seward, T. M. *Geochim. Cosmochim. Acta* **1997**, *61*, 5187–5198.
36. Benning, L. G.; Wilkin, R. T.; Barnes, H. L. *Chem. Geol.* **2000**, *167*, 25–51.
37. Sun, W.; Nesic, S.; Young, D.; Woollam, R. *Ind. Eng. Chem. Res.* **2008**, *47*(5), 1738–1742.
38. Criaud, A.; Fouillac, C.; Marty, B. *Geothermics* **1989**, *18*(5–6), 711–727.
39. Sun, W.; Nesic, S.; Papavinasam, S. *Corrosion* **2008**, *64*(7), 586–599.
40. Sun, W.; Nesic, S. *Corrosion* **2009**, *65*(5), 291–307.
41. Lee, K. J. Ph.D. dissertation. Ohio University, 2004.
42. Singer, M.; Brown, B.; Camacho, A.; Nešić, S. In *Corrosion/07*; NACE International: Houston, TX, 2007; Paper no. 07661.
43. Bich, N. N.; Goerz, K. In *Corrosion/1996*; NACE International: Houston, TX, 1996; Paper no. 26.
44. Omar, I. H.; Gunaltun, Y. M.; Kvarekval, J.; Dugstad, A. In *Corrosion/05*; NACE International: Houston, TX, 2005; Paper no. 05300.
45. Kvarekval, J.; Nyborg, R.; Choi, H. In *Corrosion/03*; NACE International: Houston, TX, 2003; Paper no. 03339.

2.26 Corrosion in Lubricants/Fuels

R. W. Wilson
Capcis House, 1 Echo Street, Manchester, M1 7DP, UK
S. B. Lyon
Corrosion and Protection Centre, School of Materials, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK

This article is a revision of the Third Edition article 2.11 by R.W. Wilson, volume 1, pp 2:143–2:154, © 2010 Elsevier B.V.

Glossary

Aromatic A hydrocarbon containing one or more conjugated (i.e., repeated single and double) bonds in the form of a ring (e.g., benzene – C_6H_6).

Asphaltenes Molecular substances found in crude oil that have some of the properties of asphalt, with high viscosity and melting point. Practically they are defined as crude oil fractions that are generally insoluble in *n*-heptane (C_7H_{14}) but soluble in toluene ($CH_3{\cdot}C_6H_5$).

Grease A solid or semi-solid lubricant material that consists of calcium, sodium or lithium soap that is emulsified with mineral or vegetable oil.

Oligomer A compound containing a limited (generally small) number of monomer units (as opposed to a polymer, which in principle contains a large number of monomer units).

Olefin A hydrocarbon that contains unsaturated bonds (e.g., ethene – C_2H_4).

Paraffin A hydrocarbon that contains no unsaturated bonds (e.g., ethane – C_2H_6).

Soap A metallic salt of a fatty acid.

Abbreviations

ASTM American Society for Testing and Materials
BS British Standard
EP Extreme pressure
VI Viscosity index (L, M, or H – low, medium, or high)
TAN Total acid number
TBN Total base number
SOA Spectrographic oil analysis

2.26.1 The Nature of Lubricating Oils and Greases

2.26.1.1 Introduction

Lubricants and fuels are not generally regarded as being corrosive, and therefore, in order to appreciate how corrosion can occur in lubricant systems, it is necessary to understand something of their nature. Once lubricants were almost exclusively derived from animal (e.g., whale) or vegetable (e.g., castor) oils or fats but modern requirements by way of volume and special properties have made petroleum the main source of supply. The production of lubricants represents between 1 and 2% of all petroleum products by volume but considerably more by value and type.

The main function of most lubricants is to reduce friction and wear between moving surfaces and to abstract heat. They also have to remove debris from the contact area (e.g., combustion products in an engine cylinder) or swarf in metal-cutting operations (e.g., in metal-working fluids). They may also be required to protect the lubricated or adjacent parts against corrosion, but this is not a prime function of most lubricants. On the other hand, many lubricants do contain corrosion inhibitors and some lubricating oils, greases, mineral fluids, and compounds are specially formulated to prevent the corrosion of machinery or machine parts, particularly when these components are in storage or in transit (e.g., temporary protectives).

2.26.1.2 Lubricating Oils

There are many hundreds of different varieties of lubricants, many of them tailored to meet specific requirements. For some applications, vegetable oils (especially those derived from the castor oil plant) are still preferred. However, due to the volume requirements petroleum-derived (mineral) lubricant base oil stocks predominate.[1]

Mineral base oils may be distillates or from residues and are derived from the vacuum distillation of a primary distillate with a boiling point range generally above that of gas oils (i.e., above about 200 °C). They are mixtures of hydrocarbons, generally containing more than about 20 carbon atoms per molecule, and range from thin, easily flowing long-chain 'spindle' molecules to thick, branched, 'cylinder' oils. For hydrocarbons having the same number of carbon atoms per molecule, generally, the higher the proportion of carbon to hydrogen, the more viscous the oil and the lower the viscosity index.

Distillate lubricating oils can be conveniently divided into three groups: low viscosity index oils (LVI oils), medium viscosity index oils (MVI oils), and high viscosity index oils (HVI oils). LVI oils are made from naphthenic distillates with low wax contents so that costly dewaxing is not required. MVI oils are produced from both naphthenic and paraffinic distillates; generally, the paraffinic distillates have to be dewaxed (i.e., those fractions that are solid at room temperature need to be removed from the base oil). The residues from the vacuum distillation can also be refined to provide very viscous lubricants; thus, residues from paraffinic base oils are generally solvent extracted and dewaxed. The main use of these products (bright stocks) is as blending components for heavy lubricants.

The mineral base oil fractions may be further categorized as predominantly long-chain aliphatic (paraffinic) hydrocarbons or aromatic hydrocarbons, depending on the crude feedstock. For higher performance, base oils may be synthesized to produce a more tightly specified range of compounds such as olefin oligomers and polyolefinic aromatics that are not present in the original crude oil feedstock. Although such materials can cost 5–10 times more than the conventional materials, their properties are more predictable and their performance (e.g., thermal and chemical stability) is significantly improved.

The primary component of a lubricant, which commonly comprises over 70% and generally up to 98% of the volume of the material, is the base oil stock. To this are added further components that modify and improve the overall performance of the lubricant for specific purposes. Additives that are present in most formulations include: viscosity modifiers, detergents and dispersants, antifoaming agents, oxidation inhibitors, and corrosion inhibitors (often called 'rust inhibitors'). Additives for specific purposes include biocides and antiwear and extreme pressure (EP) components, which are formulated to react with metal bearing surfaces under extreme conditions to produce a low friction surface.

2.26.1.3 Greases

Greases, in contrast to oils, are solid or semisolid lubricants that are made by thickening lubricating base oils with gelling agents such as metal ion salts of long-chain fatty acids (i.e., soaps), lubricious solids such as talc, mica, graphite, and molybdenum disulfide, as well as the conventional additives mentioned above. Typical greases are formulated from approximately a 1:1 ratio of lubricating base oil with emulsifying agent (usually calcium, sodium, or lithium soaps of long-chain fatty acids) plus additives. Greases are used where lubrication may be infrequent and/or where the lubricant is required to remain in place over a significant length of time. Greases are essentially emulsions of the base oil with the soap as the emulsifying agent and second phase and are thixotropic; that is under high shear the viscosity falls to values more representative of the base oil used in the formulation. Thus, they are ideal for lubrication of bearings and similar components. Since many vegetable oils are long-chain hydrocarbons and contain reactive functionality (i.e., double bonds or hydroxyl groups) they may be easily processed into fatty acids. Hence, many greases use fatty acid soaps derived from vegetable oils (e.g., castor oil).

2.26.2 Deterioration of Lubricants

2.26.2.1 Deterioration by Combustion Products

Lubricating oils deteriorate in service in two primary ways: they become contaminated and they undergo physical and chemical changes due to oxidation.[2] In engines the common contaminants are airborne dust and wear products, unburnt fuel, fuel combustion products, and water. The oxidation products are mainly acidic materials and asphaltenes. Asphaltenes in association with fuel contaminants and water tend to form solids such as sludges and lacquers that can coat component parts and become entrained in the lubricant. The acidic materials resulting from oxidation of the lubricant oil are generally weak organic acids. However, contamination by fuel combustion products (i.e., oxides of sulfur and nitrogen) is the source of almost all strong-acid contamination in lubricants. Recent trends in fuel compositions have resulted in the amount of sulfur, particularly in diesel fuels for automotive use, being reduced considerably. However, heavy fuel oils and residual diesel fuels can contain several percent of sulfur by weight. This sulfur is oxidized to sulfur acids, and sulfuric acid condensate may be encountered on the cooler surfaces. In petrol (gasoline) the sulfur content is now generally very small to negligible. Historically, halogen compounds added to gasoline as scavengers for lead-based antiknock compounds could be oxidized to halogen acids; however, with the general elimination of lead in this role, this is no longer of significant concern.

Water, absorbed from the environment or as a product of combustion, exacerbates the problem of acids in oils as the presence of water both promotes corrosion and solvates any ions present, including hydrogen ions from combustion acids. For this reason, lubricant oils for internal combustion engines frequently contain additives to combat acidity. Generally however, provided the water content in lubricating oil remains below about 0.1%, few problems occur.

2.26.2.2 Oxidative Degradation

Apart from unusual and highly specialized materials (i.e., lubricants based on silicone chemistry or fully fluorinated hydrocarbons) lubricants, being hydrocarbon based, are relatively susceptible to deterioration by oxidation.[3] Oxidation, being a thermally activated process, will increase in rate with temperature although this will depend on the nature of the compound being oxidized. Thus, oils with a high aromatic content (LVI and some MVI oils) tend to oxidize to give sludge-forming compounds, although some naphthenic oils give organic acids. Paraffinic (long-chain) oils (HVI and some MVI oils) oxidize more slowly to give weak acids.[4] In a plentiful supply of oxygen, oxidation proceeds at a significant rate at temperatures above about 60–130 °C depending on the composition of the lubricant. This oxidation is a complex process that typically involves the formation of organic peroxides as intermediates, which facilitate further hydrocarbon chain scission (breaking) or polymerization to form tars, sludges, and varnishes. Oxidation of hydrocarbons is catalyzed by the presence of transition metals with multiple oxidation states, such as copper, iron, and lead.

Oxidative degradation of hydrocarbons will, of course, lead eventually to the formation of carbon dioxide and water. Generally, however, lubricants form intermediate oxidation products, including aldehydes, ketones, and alcohols, with carboxylic acids as the common final oxidation products. Thus, the acidity of lubricating oils during use tends to increase as a function of time. Also, the intermediates, especially ketones and aldehydes, can polymerize, especially in the presence of strong acids from the

products of combustion (nitric and sulfuric). The presence of acidity is clearly corrosive to components and will form metal ion carboxylates (soaps) that are catalytic to oxidation. Thus, the oxidation of octane is accelerated from 5–10 times in the presence of 500 ppm iron octanoate[3] and excessive copper or iron will generally significantly enhance oxidation of lubricants at higher temperatures.[5] Lubricant (and fuel) additives may be formulated to control oxidation, corrosion, acidity, and radical formation in order to limit the degradation of the lubricant. However, such additives will themselves degrade (or be consumed) with time, hence the importance of regular changes of engine lubricants.[6]

2.26.2.3 Biodeterioration

Hydrocarbons provide a potential nutrient source of carbon for microorganisms and, consequently, lubricants and fuels are susceptible to biodeterioration by a range of fungal and microbial species.[7] Indeed, selection of specific microbial strains is of increasing interest as a mechanism of bioremediation of waste hydrocarbons. Lubricants are not normally susceptible to biodegradation unless they contain a significant separated water phase. However, the presence of 10–100 ppm water is sufficient to permit microbial growth on interfaces; thus, apparently water-free – (<0.1%) lubricants can contain significant microbial contamination. Lubricants can be particularly susceptible to biodeterioration since they contain essential elements for microbial activity (e.g., phosphorus, sulfur, nitrogen, etc.) in significant quantities. Engine and machinery lubricants that operate at a sufficiently high temperature in a closed environment will be essentially sterile. However, significant problems occur where lubricants are used at lower temperatures. Thus, metal-working lubricants (cutting fluids, etc.) that are left, for periods of time, open to the air at room temperature have increased susceptibility to microbial attack[8] unless a suitable biocide is present in the formulation.

2.26.3 Lubricant Additives

2.26.3.1 Types of Additive

The progressive development of engines and general machinery resulting in more arduous operating conditions, and particularly the use of longer oil-change periods, means that neither straight mineral oils nor compounded oils (mineral oils to which a proportion of an animal or vegetable oil has been added) are adequate for modern service requirements.

Despite the introduction of new, improved methods of refining, it has been necessary to enhance the performance of lubricants by the use of additives, either to reinforce existing qualities or to confer additional properties. Thus, almost all quality lubricants on sale today contain one or more additives. An enormous range of additives are available for use in lubricants, some produced by the oil companies and others provided by specialist manufacturers.[9] Additives are usually named after their particular function, but many additives are multifunctional. Thus, an antiwear additive may also protect a surface against corrosion. The main types of additive that can enhance the behavior of lubricants are listed in **Table 1**.

As can be seen from **Table 1**, many compounds are multipurpose and can act in a number of ways, some beneficial, some detrimental. The selection of additives thus involves a careful balance of properties for the required application.

2.26.3.2 Extreme Pressure Additives

Many additives, essential to the performance of the lubricant, provide no corrosion protection and some additives may become corrosive in certain circumstances. Thus, extreme pressure (EP) antiwear additives contain chemical groups which are designed to react chemically with metal surfaces when normal lubrication fails, forming easily sheared layers of metal oxides, sulfides, chlorides, or phosphates, thereby preventing catastrophic wear and seizure. Reaction between EP compounds and metal surfaces should only occur at local hot spots and the layers formed are extremely thin.[10] However, if the operating conditions are very severe these layers are continually generated and removed as they fulfill their antiwear function. A process of this nature is sometimes called 'chemical' wear, and if sliding surfaces operate continually under these conditions loss of metal from the rubbing surfaces can ultimately result in failure. Alternatively, all the EP additive may be used up (depleted) and then failure by seizure will occur. EP agents are intended to cater for the occasional overload condition and it must be emphasized that machinery should be designed so that it does not require the continual action of EP agents to function satisfactorily. Obviously, the selection of an EP additive requires great care; if it is too active, it may give rise to excessive metal removal (effectively corrosion) under normal operating conditions. Also, if a component is prone to 'fatigue pitting' (a form of contact fatigue in which fatigue cracks grow in such a way as

Table 1 Lubricant oil additives that can enhance anticorrosive function

Additive	*Function*	*Chemical types*
Antioxidant	To increase oxidation resistance of lubricants by interfering with the reactions that give rise to acid and asphaltene (polymer) formation	Oil-soluble amine and phenol derivatives (<120°C), dialklydithiophosphates and compounds listed below as metal deactivators (>120°C)
Metal deactivator	To form inactive protective films on metal surfaces which would otherwise catalyze oxidation and corrosion reactions	Trialkyl and triaryl phosphites, organic dihydroxyphosphines, some active sulfur compounds, diamines; for greases mercaptobenzothiazole and phosphites
Corrosion/rust inhibitors	To protect metal surfaces, particularly bearing surfaces, against corrosion/to eliminate rusting in the presence of moisture	Zinc dialklydithiophosphates; esters and derivatives of dibasic acids (e.g., barium and calcium sulphonates)
Water repellents	To impart water-resistant properties, particularly in greases	Aliphatic amines, hydroxyl fatty acids and organosilicone polymers
Basic compounds	To neutralize acids	Barium and calcium salts of sulphonic acids and alkyl salicylic acid
Dispersants/detergents	To keep insoluble combustion and oxidation products in suspension and dispersed	Salts of phenolic derivatives, polymers containing barium, sulfur and phosphorus; calcium or barium salts of sulphonic acids
Antiwear/extreme pressure	To protect metal surfaces from binding together (seizure) by formation of a surface reaction product that is easily sheared under local high pressures and temperatures	Sulfur or phosphate esters/amines, halogenated compounds (the latter are decreasing in use)

to cause 'chips' of metal to detach from the surface) in service the presence of an overactive EP agent may result in corrosion fatigue.

2.26.3.3 Interactions Between Additives

Modern high-performance lubricants contain a number of additives, each with a particular, special function. Thus, a turbine lubricant may contain an oxidation inhibitor, a rust inhibitor, an EP agent, and an antifoam compound. A high-grade diesel-engine lubricant may contain a viscosity improver, a dispersant, an antioxidant, a corrosion inhibitor, a basic compound, a pour-point depressant, and an antifoam compound. Sometimes, these additives may have undesirable side effects or interact adversely; in turbine oil, the rust inhibitor may act as an emulsifier, interfering with demulsification; in a diesel lubricant, the dispersant may promote oil oxidation. Frequently anticorrosion additives may not be able to exert their maximum effect because they are competing for sites on metal surfaces. The development of successful new lubricating oils requires skill and experience and always necessitates considerable laboratory and field testing in order to strike the right balance between the various additives.

2.26.3.4 Additives in Greases

All classes of additive mentioned earlier may be incorporated in the oil phase of greases and have similar function. However, with greases there is also the opportunity to employ oil-insoluble species that will partition to the emulsifying (soap) phase. The complete range of conventional water-soluble corrosion inhibitors, for example, nitrites and molybdates, can thus be easily incorporated into greases. However, care is required to avoid unwanted antagonistic interactions with other additives.

2.26.3.5 Sulfur-Containing Additives

Sulfur compounds occur naturally in most lubricants and many oil additives contain sulfur. In a properly formulated lubricant, these sulfur compounds should be inactive at ambient temperature. At elevated temperatures, they may decompose to give more active materials which can stain and corrode metals, particularly silver and copper. However, these same sulfur compounds have many beneficial qualities; this is why they are not removed completely in refining and why they are used as additives.

Thus, sulfur compounds in lubricants generally act as antioxidants, preventing acid and sludge formation. They can also adsorb or react to form thin films on metal surfaces protecting them from acid or peroxide attack. In addition, sulfur compounds are often used as EP agents. The oil chemist must try to strike a balance; the activity of the sulfur must be high enough for it to exert a beneficial effect and yet not so high as to stimulate corrosion.

All too frequently lubricants containing sulfur are exposed to more severe operating conditions than intended, and staining and corrosion may result. With traditional bearing materials, such as those containing silver and copper, significant corrosion can result, particularly where the sulfur-containing species have partially degraded. A more widespread problem is the corrosion of phosphor–bronze alloys (containing about 10% tin) particularly where temperatures can exceed 200°C. Two important metallurgical factors affecting the corrosion resistance of phosphor–bronze alloys are the amount of alloying element in solution in the copper-rich phase and the porosity of the alloy. For example, if the amount of tin in the solution can be increased by special casting techniques or heat treatments, the corrosion resistance is greatly increased. Similarly, zinc or silicon in solution also increases the resistance of copper to sulfur corrosion. If the alloy is porous the lubricant is drawn into the pores where it stagnates, and, at high temperatures, becomes very corrosive. Thus, as noted above, copper catalyzes oil oxidation with the consequent formation of corrosive sulfur compounds.

The most satisfactory solution to such problems is to employ corrosion-resistant bearing alloys. Traditional alloys with relatively high levels of tin and alloys of the gunmetal type, containing 2–4% zinc, have proved completely satisfactory. The substitution of zinc for phosphorus gives sounder castings and improves the corrosion resistance of the copper-rich matrix. More modern bearing materials such as thermally sprayed aluminum–copper–tin–(indium) have much improved resistance to such corrosion.[11]

2.26.4 Corrosion in Specific Lubricant Systems

2.26.4.1 Internal Combustion Engine Lubricants

Engine lubricants are exposed to severe operating conditions, being subjected to high temperatures, the products of combustion and a plentiful supply of oxygen. Consequently, unless the oil is changed at appropriate intervals, strong mineral acids and weak organic acids may accumulate. In addition, droplets of water may be formed and these can contain strong mineral acids derived from the fuel combustion gases. These droplets sometimes give rise to emulsions which deposit in the colder portions of an engine, for example, on the rocker-box covers; ferrous surfaces are most affected by this condensed moisture. A special dynamic corrosion test has been developed to study corrosion in these two-phase (water-in-oil) systems.[12] Problems associated with the retention of water in engine lubricants are likely to become more acute as antipollution devices are fitted to engines. The harmful effects are best countered by antirust and basic additives, and in diesel engines burning high-sulfur fuels, for example, marine diesel engines, very high levels of lubricant basicity are required.

Cast or sintered copper–lead or lead–bronze alloys are still widely used for engine bearings. The lead phase in such bearings is readily attacked by weak organic acids and almost all the lead can be leached out unless preventive measures are employed. However, the lead may be protected by an overlay of a lead–tin or lead–indium alloy.[13] About 3% tin or 5% indium in lead will render the lead resistant to attack by oil-oxidation acids. One reason why leaded bearings are protected by an overlay, and not by incorporating the protective alloying elements in the underlying lead, is that both tin and indium dissolve preferentially in copper.[14] In a cast or sintered bearing, therefore, any tin or indium will be found in solution in the copper-rich phase, leaving the lead-rich phase susceptible to attack. Aluminum–copper–tin alloys, which are increasingly used in modern automotive applications, are significantly more resistant to corrosion than conventional copper and lead-based alloys.

2.26.4.2 Corrosive Wear

Modern efficient engines that require extended service intervals and minimum maintenance, and that are able to accept a wide range of fuels whilst meeting strict exhaust emission standards, present special lubrication problems. In particular it is essential that the lubricant can prevent corrosion from fuel combustion products retained in the engine. Corrosive wear can be reduced or prevented by alkaline additives that can neutralize the acids responsible, by additives that prevent the acids reaching metal surfaces and by antirust additives that are adsorbed on metal surfaces and prevent the access

of water and oxygen.[15] Often all three additive types will be used in association although their relative concentrations will depend on the nature of the corrosive agents. In spark-ignition engines the use of nonconventional fuels such as compressed natural gas (methane), liquefied petroleum gas (propane and butane), methanol, ethanol and lead-free gasoline can give specific corrosion problems.[16] Broadly, organic acids of shorter chain length and greater acidity (lower pK_a) are generated from such fuels compared with normal gasoline, so higher additive levels are required.

Diesel-fuel sulfur levels can range from about 0.2% to 5.0%. The higher levels only occur in low- and medium-speed engines operating on residual fuels, but because of the high combustion pressures, sulfuric acid condensate forms at temperatures up to 200°C. On cross-head engines, which have independent cylinder lubrication systems the neutralizing power of the fresh oil depends on both oil feed rate and alkalinity.[17] The alkalinity of the lubricant is measured as the total base number (TBN) and lubricants with a TBN exceeding 100 are now available. This means that the lubricant has an acid neutralizing capacity equivalent to a 10% aqueous solution of KOH in oil-soluble form. Experience shows that when fuels with sulfur contents of 2.0% or more are used, the TBN of oils draining from the cylinders should not fall below 10 if excessive wear rates are to be avoided. However, the use of highly alkaline cylinder oils with low sulfur fuels is not recommended; not only is it needlessly expensive, it can also give rise to high wear rates.[15,17] In splash-lubricated piston engines running on high sulfur fuels, the copious quantities of oil splashed up from the crank case provide a greater reservoir of alkalinity and an initial TBN of 25–30 is generally adequate. Nevertheless, the TBN of the oil should not fall below about 3 times the sulfur content of the fuel.

2.26.4.3 Steam Turbine Lubricants

Lubricants in steam turbines are not exposed to such arduous conditions as those in engines. Thus, their main requirement is for high oxidation stability. However, they may be exposed to aqueous condensate or, in the case of marine installations, to sea water contamination, so they have to be able to separate from water easily and to form a rust-preventing film on ferrous surfaces, and it is usual to employ rust inhibitors. The problem of tin oxide formation on white-metal bearings is associated with the presence of water condensate in lubricants and can be overcome by keeping the lubricant dry or by the incorporation of appropriate inhibitors.[18]

2.26.4.4 Gear Lubricants

In addition to the usual oxidation and corrosion inhibitors, lubricants for heavily loaded gears almost always contain EP additives containing sulfur, chlorine, or phosphorus. In order to function, these additives must react locally with the metal surfaces, and yet the extent of the reaction should not be such that it could be described as corrosive, or promote fatigue pitting. These EP additives may be quite safe with ferrous surfaces, but may cause severe corrosion on copper alloys, for example, on bronze worm wheels, if for any reason excessive temperatures arise. Some turbine lubricants have to lubricate the turbine gears as well as the turbine; in these circumstances, any EP additives employed should not be corrosive in the presence of moisture.

2.26.4.5 Cavitation and Erosion

Cavitation and consequent erosion are an insidious cause of failure in plain bearings and hydraulic systems.[19] Vapor cavities form locally in reduced pressure areas of the lubricating fluid and then collapse rapidly in higher pressure regions. The shock waves generated by the vapor collapse can result in mechanical (impact) damage on surfaces. Common causes of cavitation are pressure fluctuations associated with the flow of the liquid and the vibration of a surface in contact with it. Unlike other kinds of damage, vapor cavitation is generally encountered on the unloaded areas of bearings. When the damage is due solely to cavitation, the damaged surfaces are rough: when foreign particles are present they are smooth (cavitation–erosion). Vapor cavitation can remove protective films, such as oxides, from metals and so initiate corrosion. In addition, the very high local pressures and temperatures associated with the final stage of cavity collapse can induce chemical reactions that would not normally occur. Thus certain additives are damaged by cavitation and their decomposition products can be corrosive.

2.26.4.6 Metal-Working Lubricants

Metal-working lubricants can be divided into two categories: metal cutting lubricants and metal forming lubricants.[20] Three types of cutting fluids are widely

used: soluble oils, water base fluids and straight cutting oils. Soluble oils are low-viscosity mineral oils (containing emulsifying agents) that, when added to water, form stable oil-in-water emulsions. Their main function is to cool the workpiece rather than to lubricate, but they should contain rust inhibitors to prevent corrosion of both the workpiece and the machine. Water base fluids contain no oil and employ the normal corrosion inhibitors used in coolants, for example, sodium nitrite, sodium benzoate, and triethanolamine phosphate. Straight cutting oils are used in severe operations where good lubrication is essential, and generally contain EP additives that should be carefully selected so that they do not corrode (stain) the workpiece or the machine tool. For example, very active sulfur compounds should not be employed with copper alloys. Chlorinated additives should be stable in moist air to avoid the risk of hydrochloric acid formation.

Metal-forming lubricants include formulations for rolling, drawing, extruding and forging. A vast range of compounds including fatty oils and compounded oils are used in these operations, and a major requirement is that they should not stain the workpiece during the forming operation or during subsequent annealing or in storage. Consequently, all oils and additives employed should be completely volatile, in addition to affording protection against rusting.

2.26.4.7 Water Base Lubricants

In some applications, for example, mining machinery, nonflammable lubricants are specified and water base or water-containing fluids are used. Standard corrosion inhibitors are used to combat corrosion but even a small amount of water in a mineral oil lubricant can adversely affect the fatigue life of components such as rolling bearings and gears.[21] Worthwhile improvements in fatigue life can be achieved by using additives of a completely different nature and by special heat-treatment techniques.[22]

2.26.4.8 Lubricants for Other Applications

Specially formulated lubricants are required for steam engines, compressors and exhausters, refrigerators, hydraulic equipment, textile machinery, transformers and switchgear, nuclear power plants, and many other diverse applications. Most of these lubricants will contain a carefully balanced set of additives, including some to prevent corrosion either directly by protecting the metal surface, or indirectly by preventing deterioration of the lubricant and combating the action of contaminants. When corrosion is encountered with a specialized lubricant, the cause is not likely to be any weakness on the part of the lubricant; it is probable that the lubricant is being exposed to extreme operating conditions, far more than those for which it was designed.

2.26.5 Monitoring and Testing

2.26.5.1 Oil Condition Monitoring

Historically, regular changes in lubricant ensured that they remained in optimum condition. However, lubricants are expensive and users have no wish to change them until necessary. Thus, much greater attention is now given to intermittent or continuous (online) monitoring of the condition of the oil and associated machinery.[23] Oil condition monitoring is often extended to include techniques that give advance warning of the deterioration or impending breakdown of the machinery, whether or not this is associated with loss of performance of the lubricant. Conventional condition monitoring collects particulate debris, for example, by magnetic filtration, and subjects these to *ex situ* analysis, typically for elemental components. Such methods include ferrography and spectrographic oil analysis SOA.[24] More recent techniques have been developed for online (*in situ*) use and these include particle counting, infrared analysis,[25] electrical methods[26] as well as online viscosity[27] measurement.

2.26.5.2 Testing

Methods for the testing and analysis of lubricants are the subject of many Company, National, and International Standards.[28] Typical tests for used diesel engine oils include viscosity, fuel dilution, flash point, water content, ash level, insoluble species, and neutralization value. Neutralization value is a measure of the acidity (total acid number, TAN), or alkalinity (total base number, TBN) of the oil. Laboratory examination of used oil samples and reporting results takes time and the simpler tests can be used in the field. Extensive tests for the performance (oxidation resistance) as well as for the corrosivity of oils and greases also exist and lie within the BS2000 series (now generally superseded) and the ASTM 'D' series. Thus, the most commonly known test for corrosivity of lubricants is the copper strip corrosion test specified in

BS 2000-112:2005; "Methods of test for petroleum and its products: Determination of corrosiveness to copper of lubricating grease (copper strip method)," which is broadly equivalent to ASTM D 130.

2.26.5.3 Health and Safety

The solvent action of mineral oil base stocks can cause skin problems and prolonged exposure has been found to be carcinogenic.[29] Thus, the use of additives that might be in any way harmful to health, for example orthotricresyl phosphate (antiwear) and sodium mercaptobenzothiazole (anticorrosion), is discouraged and discontinued where skin contact is likely. The presence of fungi and bacteria in water base or water-contaminated lubricants such as machining fluids and marine diesel crankcase lubricants may promote corrosion and are potentially of concern to health.[30] More broadly, increasing environmental concerns in recent years is driving research on environmentally friendly disposal of used lubricants and substitution of nonbiodegradable mineral oils with fully biodegradable oils that are often plant derived.[31]

References

1. Pisso, D. M.; Wessel, A. A. *Lubrication Fundamentals*, 2nd ed.; *Dekker Mechanical Engineering Series,* Vol. 137, Marcel Dekker, 2001.
2. Fox, M. F. In *Handbook of Lubrication and Tribology*, 2nd ed.; Totten, G. E. Ed.; CRC Press, 2006; Chapter 29.
3. Cochrac, G. J.; Rizvi, S. Q. A. In *Fuels and Lubricants Handbook: Technical Properties, Performance and Testing*; Totten, G. E., Westbrooke, S. R., Shah, R. J., Eds.; ASTM, 2003; Chapter 30.
4. Murray, D. W.; McDonald, J. M.; Wright, P. G. *Petroleum Rev.* **1982**, *36*, 36.
5. Clark, D. B.; Klaus, E. E.; Hsu, S. M. *Lubric. Eng.* **1985**, *41*, 280.
6. Barcroft, F. T.; Park, D. *Wear* **1986**, *108*, 213.
7. Allsopp, D.; Seal, K.; Gaylarde, C. *Introduction to Biodeterioration*, 2nd ed.; Cambridge University Press, 2003; Chapter 3.
8. Taylor, G. T.; Knee, N. D.; Keep, M. J. H.; Freestone, V. C. M. In Proceedings of the ''Waste Management '99'' Conference, Tucson, Arizona, USA,1999.
9. Rudnick, L. R. Ed. *Lubricant Additives: Chemistry and Applications*; CRC Press, 2003.
10. Farng, L. O. In *Lubricant Additives: Chemistry and Applications*; Rudnick, L. R. Ed.; CRC Press, 2003; Chapter 8.
11. Kato, K.; Adachi, K. In *Modern Tribology Handbook*; Bhusan, B. Ed.; CRC Press, 2001; Vol. 2, Chapter 21.
12. Hughes, R. I. *Corros. Sci.* **1969**, *9*, 535.
13. Forrester, P. G. *Trans. Inst. Metal Finishing* **1961**, *38*, 52.
14. Wilson, R. W.; Shone, E. B. *Anti-Corros. Meth. Mater.* **1970**, *17*(8), 9.
15. Wilson, R. W. In Proceedings 6th International Congress on Metallic Corrosion; Australasian Corrosion Association, 1981; Vol. 2 p 1479.
16. Thring, R. H. *SAE J. Automot. Eng.* **1984**, *92*(1), 60.
17. Golothan, D. W. K. *Trans. Inst. Mar. Engrs: Tech. Rep.* **1978**, *90A*(Pt 3), 137.
18. Hiley, R. W. *Trans. Inst. Mar. Engrs.* **1979**, *91*, 52.
19. Dowson, D.; Godet, M.; Taylor, C. M. In 1st Leeds–Lyon Symposium on Tribology, Section 7 Institution of Mechanical Engineers, 1975; p 177.
20. Savoit, R. E. *Lubrication – Texaco* **1969**, *55*(7), 65.
21. Grunberg, L.; Jamieson, D. T.; Scott, D.; Lloyd, R. A. *Nature* **1960**, *188*, 1182.
22. Hollox, G. E.; Hobbs, R. A.; Hampshire, J. M. *Wear* **1981**, *68*, 229.
23. Collacott, R. A. In *Industrial Tribology*; Jones, M. H., Scott, D. Eds.; Elsevier, 1983; Chapter 18.
24. Hoffmann, W. *Wear* **1981**, *65*, 307.
25. Agoston, A.; Ötsch, C.; Zhuravleva, J.; Jakoby, B. In Proceedings of IEEE Sensors; **2004**; Vol.1, P. 463 art. no. M4L-5.
26. Turner, J. D.; Austin, L. *Meas. Sci. Technol.* **2003**, *14*, 1794.
27. Agoston, A.; Ötsch, C.; Jakoby, B. *Sensors Actuators A: Phys.* **2005**, *121*, 327.
28. Hunter, M. E.; Baker, R. F. In *Fuels and Lubricants Handbook: Technical Properties, Performance and Testing*; Totten, G. E., Westbrooke, S. R., Shah, R. J., Eds.; ASTM, 2003; Chapter 31.
29. Warne, T. M.; Halder, C. A. *Lubric. Eng.* **1986**, *42*, 97.
30. Bennett, E. O. *Tribology Int.* **1983**, *16*, 133.
31. Chen, B.; Fang, J.; Dong, L.; Xia, S.; Wang, J. *Lubric. Sci.* **2008**, *20*, 311.

Further Reading

Lubrication; Texaco/Chevron, since 1925. Back issues available for download from Chevron: www.lubricantsuniversity.com.

Totten, G. E. Ed. *Handbook of Lubrication and Tribology*, 2nd ed.; CRC Press, 2006.

Totten, G. E.; Westbrooke, S. R.; Shah, R. J. Eds.; *Fuels and Lubricants Handbook: Technical Properties, Performance and Testing*; ASTM, 2003.

Bhusan, B. Ed.; *Modern Tribology Handbook;* CRC Press, 2001; Vol. 2.

2.27 Corrosion in Body Fluids

D. J. Blackwood
Department of Materials Science and Engineering, National University of Singapore, 9 Engineering Drive 1, Singapore 117576, Singapore

Glossary

Biocompatibility The extent to which an implanted material elicits an immune response in a host.

Carcinogen Cancer causing.

Cardiac Related to the function of the heart.

Cytotoxicity Toxic towards cells.

Extracellular The space outside the plasma membranes of cells and occupied by fluid; literally means outside the cell as opposed to intracellular which is inside the cell.

Hematoma A collection of blood, usually partially clotted, that results from the breakage of veins or blood vessels.

Inflammatory response Part of the human body's initial response to injury or infection.

In vitro Experiments conducted outside the body; literally means in-glass.

In vivo Data collected from materials implanted within a live subject, human or animal.

Orthodontic An area of dentistry concerned with treatment of the inability to bite correctly, due to the misalignment of teeth, irregular tooth growth or disproportionate relationship between the upper and lower jaws.

Oral lesions Abnormal tissue in or around the mouth, usually caused by disease or trauma.

Oral mucosa Mucous membranes of the mouth.

Osteoporosis Thin or porous bones caused by lack of calcium or stress shielding by load-bearing implants.

Prosthesis An artificial device that replaces a missing body part.

Pathological changes The human body's initial response to injury or infection.

Pulmonary Related to the function of the lungs.

Serum Blood plasma from which clotting factors have been removed.
Shape-memory alloy A material that, after it has been deformed, regains its original geometry by heating, due to temperature-dependent martensitic phase transformation.
Sherman plates Metal plates used to hold fractured bones in place to allow them to heal; typically the plate is screwed to the bone.

Abbreviations

ASTM American Society for Testing and Materials
FCC Face Centered Cubic
ISO International Standards Organization
L Low carbon
PREN Pitting resistance equivalent number $= \%Cr + (3.3\% \times Mo) + (16 \times \%N)$
SCC Stress corrosion cracking
SHE Standard hydrogen electrode
VM Vacuum melted

2.27.1 Introduction

Corrosion is of concern to the biomedical industry for two reasons. The first is common to all industries, being 'how will a device's lifetime/performance be impacted by corrosion?' However, the second concern is specific to the biomedical industry, being 'will the metallic ions that leach out of the device build up to levels sufficient to harm the patient, for example, by causing tumors to develop?' There is little known about what represents longterm safe limits for metallic ion concentrations in the body, and therefore, these could well be exceeded at corrosion rates that are insignificant with respect to the physical performance of the implant. Corrosion problems in dental applications are more common, mainly due to the high acidity and chloride contents of many foodstuffs. Fortunately, fixtures in the oral cavity are readily accessible for repair or replacement, but the toxicity of the metals leaching out is a major concern.

In most industries, corrosion is typically controlled by either coatings or altering the local environment. Unfortunately, until recently, coatings were of limited use for protecting surgical implants, since many of these (especially orthopedic devices) are subjected to wearing and abrasion processes. Likewise, the environment within the human body is essentially fixed, and therefore, it is not possible to lower the temperature or raise the pH. As a result, the only available method of reducing corrosion rates within the human body has been to fabricate the biomedical devices from a corrosion-resistant material. Nevertheless, since the late 1970s, this approach has been extremely successful, at least with respect to extending the lifetime of biomedical devices, thanks to the development of a range of corrosion-resistant alloys that have reduced the number of failures to extremely low levels. Besides, many of the few corrosion failures that still occur can be traced either to poor quality control or to an unexpected and unusually aggressive local environment around the implant, due to pathological changes in the surrounding tissue as this reacts to the surgical procedure. The remaining early failures of surgical implants are due to either fatigue or fretting, which may or may not be accelerated by corrosion.

Nevertheless, the lack of consensus over what represents safe levels for metallic ion concentrations within the body, or even which metals are toxic, means that the concerns about extended exposure to even very low levels of corrosion products resulting in medical complications remain. Furthermore, average life expectancies are increasing and the average age of patients receiving implants is decreasing. Ironically, this is both in part due to the modern popularity of physical sports, which place a large strain on joints, and therefore, the required performance lifetime of devices, and with it, the likelihood of both the corrosion-related failures and health problems instigated by elevated metallic levels is increasing. Additionally, a number of advanced materials, such as shape-memory alloys, are being introduced into the biomedical industry often with very little prior thought given to corrosion protection.

2.27.2 Historical Development

The earliest use of metallic materials in dental applications was over 4000 years ago in ancient Egypt; although initially for aesthetic purposes, by about 700 BC, this had developed to the level of using gold bridgeworks to secure false teeth formed out of ivory and bone. However, the modern era of using metallic surgical implants to fix damaged bones only began at the beginning of the twentieth century. Of the metals and alloys available at that time, vanadium

steel, developed in 1912, offered the best combination of corrosion resistance and mechanical strength. Although this alloy had poor tissue compatibility, it was not until the 1930s that austenitic stainless steels became sufficiently available as a viable alternative, first with compositions similar to that of grade 304 and later with molybdenum additions, that is, evolving towards grade 316L, which was endorsed by the American College of Surgeons in 1946.[1] In 1981, a 22Cr13Ni5Mo stainless steel was introduced for hip implants, but this was rapidly superseded by the introduction of high-nitrogen austenitic stainless steel, which offered the same corrosion resistance but without the need for a high-nickel content to maintain the FCC structure, dramatically reducing costs. In response to fears of nickel toxicity, the beginning of this millennium saw the introduction of nickel-free or low-nickel austenitic stainless steels implants, such as the nitrogen-strengthened 23Mn21Cr1Mo (ASTM F2229-02).

Cobalt–chromium–molybdenum alloys were first used in dentistry in the late 1920s and then for surgical implants in the late 1930s. The high corrosion resistance and apparent good biocompatibility made these alloys a favored material for orthopedic devices, but during the 1950s, the better strength-to-weight ratio of titanium meant that it started to increase in popularity; CoCrMo alloys remained the preferred choice in the United States up to the late 1960s. In 1974, the Ti6Al4V alloy, which has a much higher ultimate tensile strength than the commercial purity titanium, was introduced for trauma implants, and by the end of that decade, it had become the material of choice in orthopedic surgery, although CoCrMo alloys were still regularly used.

The late 1980s and 1990s saw the beginning of the introduction of advanced materials for biomedical applications, including rare earth magnets and NiTi shape-memory alloys. In the present decade, zirconium alloy joint prostheses were introduced and important advances have been made in the application of wear-resistant titanium nitride coatings.[1] Undoubtedly, the nature of surgical and dental implant materials will continue to evolve in the coming decades.

2.27.3 Health Effects Related to Corrosion in Body Fluids

There is no doubt that surgical implants do raise the level of metallic elements in the body, as illustrated by the data displayed in **Table 1**.[2–34] However, the form of the metal released into the body is often uncharacterized, for example, particulate matter or soluble ions, with the role of serum protein binding being virtually unknown. Furthermore, there is still no agreement over what constitutes safe levels for metals within human body fluids and tissues. Although cobalt,

Table 1 Approximate concentrations of metals in the human body with and without a total-joint replacement implant[2–4]

		Metal ion concentrations: body fluids (μM); tissues (ppm)						
		Co	*Cr*	*Mo*	*Ni*	*Ti*	*Al*	*V*
Serum	Normal	0.003	0.001	–	0.007	0.06	0.08	<0.02
	Implant	0.007	0.006	–	<0.16	0.09	0.09	0.03
Blood	Normal	0.002	0.058	0.009	0.078	0.35	0.48	0.12
	Implant	0.33	2.1	0.104	0.50	1.4	8.1	0.45
Liver	Normal	120	<14	–	–	100	890	14
	Implant	15 200	1130	–	–	560	680	22
Lung	Normal	–	–	–	–	710	9830	26
	Implant	–	–	–	–	980	8740	23
Spleen	Normal	30	10	–	–	70	800	<9
	Implant	16 000	180	–	–	1280	1070	12
Lymphatic	Normal	10	690	–	–	–	–	–
	Implant	390	690	–	–	–	–	–
Heart	Normal	30	30	–	–	–	–	–
	Implant	280	90	–	–	–	–	–

Source: Hallab, N. J.; Jacobs, J. J.; Gilbert, J. L. In Joint Replacement and Bone Resorption; Shanbhag, A., Rubash, H. E., Jacobs, J. J., Eds.; Taylor & Francis: New York, 2006; pp 211–254.

chromium, and nickel are essential trace elements, in excess, all these lead to carcinogenesis as well as other ailments;[2] likewise, vanadium leads to cardiac dysfunction and hypotension,[2] while iron has been linked with Parkinson's disease.[5] Excess aluminum can cause anaemia,[2] but claims that it causes Alzheimer's disease have been proved to be completely unfounded.[6] Titanium is a nonessential element, which is usually considered inert, but pulmonary disease has been reported in titanium production workers as well as in rats exposed to TiO_2 dust.[2]

The potential to cause cancer is obviously a major concern with metallic implants. Memoli *et al.*[7] reported a slight increase in sarcoma cancers in rats when implanted with devices with high Cr, Co, or Ni contents, and a significant number of implant site tumors have been reported in cats and dogs with stainless steel implants, but fortunately to date, malignant tumors associated with implants in humans are rare.[8] Nevertheless, the number of reported cases is increasing and is likely to continue to do so as the patient age decreases and the life-expectancy increases.[9]

Health concerns with orthodontic devices are also mainly associated with metal release, particularly nickel from stainless steels and the NiTi shape-memory alloy.[10] With respect to amalgams, apart from the release of toxic mercury, the main concern is the development of lesions of the oral mucosa that can lead to leukoplakia, lichen planus, and oral cancers. Early reports linked these lesions to the potential associated with galvanic cells that can exist between amalgams and precious metal restoration, and hence these have been termed galvanic lesions.[11,12] However, galvanic cells appear to be commonplace among healthy populations; Phillips *et al.* [13] did not find any evidence of leukoplakia in rats that were subjected to galvanic currents.[14] More recent studies have found that the lesions are usually caused by contact hypersensitivity to mercury[15]; the oral mucosa may be either in direct contact with an amalgam or contain some mercury deposits resulting from amalgam corrosion.[16] Clinical signs of amalgam corrosion have been reported to be significantly more frequent in patients suffering from oral lichen planus than in controlled groups, and of course, any galvanic cells between amalgams and precious metals would accelerate the corrosion and thus exasperate the problem.[17] Nevertheless, lesions of the oral mucosa caused by amalgam restorations are rare and can usually be solved by replacement of the amalgam.[15]

2.27.4 Environments Encountered in Biomedical Applications

Upon detecting the presence of a foreign entity, the body will produce enzymes and attempt to ingest the foreign object. If the foreign object is too large to be ingested, such as a surgical implant, the body will isolate and encapsulate the object in a fibrous tissue membrane. As such, the implant remains continuously exposed to extracellular tissue fluid. Although the actual compositions of body fluids are complicated, in terms of corrosiveness, the most important characteristics are the chloride, dissolved oxygen, and pH levels. The biological components in body fluids, for example, phosphates, cholesterols, and phospholipids, are usually considered either to play no role in the corrosion process or to exist at insignificant levels. **Figure 1(a)** shows the typical environmental conditions expected within a range of different body fluids, superimposed on the Pourbaix diagram for chromium in the presence of chloride ions.[18,19] From this diagram, it can be predicted that stainless steels and cobalt–chromium–molybdenum alloys are likely to suffer corrosion in many of the environments found within the body, but titanium would be in the passive state for virtually all physiological solutions (**Figure 1(b)**). Fortunately, the body fluid most likely to be encountered by an implant is blood, which contains ~0.9% NaCl and under normal conditions is at pH 7.4, with a redox potential in the vicinity of 0.0 V versus standard hydrogen electrode (SHE) and body temperature of ~37 °C. Under these conditions, many stainless steels and cobalt–chromium–molybdenum alloys can be expected to be in the passive state.

A review by Solar[20] in 1979 concluded that inorganic solutions based on 0.9% NaCl were satisfactory substitutes for human body fluids when studying the behavior of passive metals. As a result, the majority of *in vitro* corrosion experiments have been conducted in either 0.9% NaCl or standard isotonic solutions, such as SBF-K9, Ringer's, or Hank's solution, in which additions of bicarbonate and calcium chloride tend to be the main difference to a simple NaCl solution. Dissolved oxygen levels in blood are lower than in saline solutions exposed to air atmospheres, by factors of about 2 and 6 for arterial blood and for veinal blood, respectively. Conversely, bicarbonate levels are about 20 times higher in blood than in saline solutions (**Table 2**).[21–23] Usually, no attempt is made to lower the dissolved oxygen content of the isotonic NaCl solutions to that of the veinal blood, which may

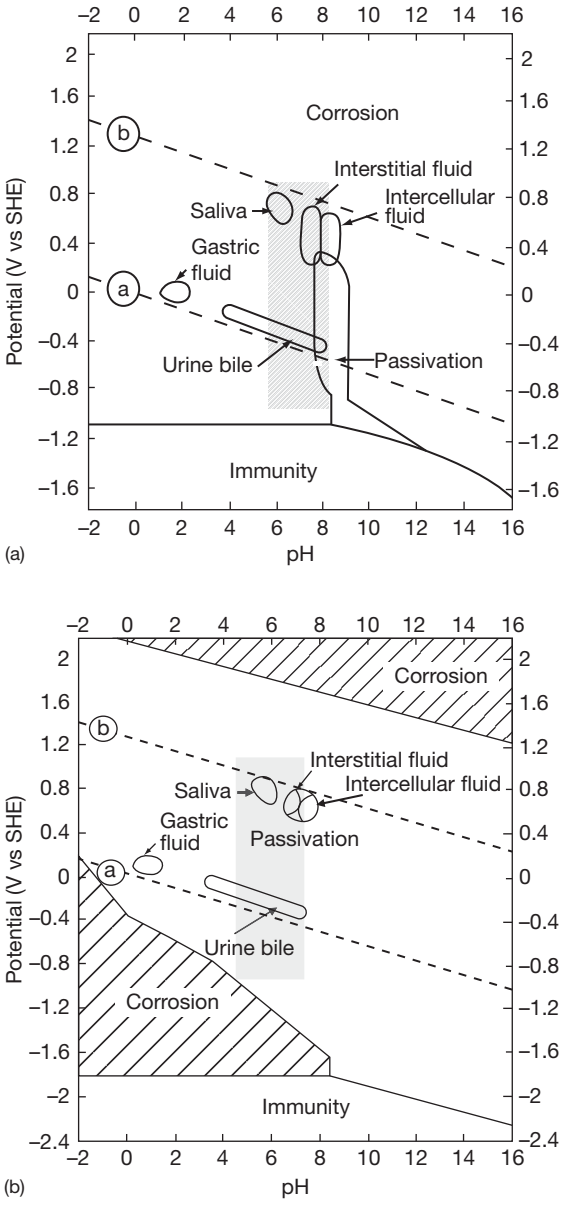


Figure 1 Representative environmental conditions for various body fluids superimposed on Pourbaix diagrams for (a) chromium in solutions containing chloride and (b) titanium. The shaded zones represent the conditions for physiological solutions as suggested by Schenk.[19] Adapted from Blackwood, D.J.; Seah, K.H.W.; Teoh, S.H. In *Engineering Materials for Biomedical Applications*; Teoh, S.H., Ed.; World Scientific: Singapore, 2004; pp 3.1–3.56, with permission from World Scientific.

explain some of the minor differences between the *in vitro* and *in vivo* corrosion behaviors of surgical implant materials.[22,24] In addition, the minor components in blood, which are usually ignored in *in vitro* studies, have occasionally been blamed for accelerated

Table 2 Comparison of oxygen and carbon dioxide levels in real and simulated human body fluids.[21–23]

	P_{O_2} (mmHg)	P_{CO_2} (mmHg)	HCO_3^- (mM)
Human artery blood	85–100	35–45	25
Human veinal blood	40	42–48	25
Rabbit artery blood	78	31	20
Rabbit veinal blood	28	39	23
NaCl solution	160	2	–

in vivo corrosion rates, for example, it has been reported that proteins can lower the pitting potential of 304L stainless steel.[25] Overall, it appears that typical body fluids are slightly less aggressive than is seawater, indeed Zitter[26] recommends that if a stainless steel is to be used to fabricate a biomedical device, it should have a pitting resistance equivalent number (PREN) not less than 26, somewhat below the value of 40 usually required for stagnant seawater.

Finally, it needs to be understood that the surgical operation plus the presence of the implant itself may cause the surrounding tissue to undergo severe pathological changes that can result in the development of a more corrosive environment. For example, Laing[27] reported that the buildup of hematomas, a condition that could last several weeks, can force the pH in the vicinity of a freshly inserted surgical implant down to as low as pH 4.0. Likewise, the initial stages of the inflammatory response can lead to the generation of hydrogen peroxide and elevated protein levels, both of which may favor the initiation of pitting and crevice corrosion.[21,28,29] Furthermore, the degree to which pathological changes occur depends not only on the biological activity of any released corrosion products, but also on the size and shape of the implant, which means that the extent of the pathological changes will vary across the surface of an implant. This could lead to the development of electrochemical cells and the potential gradients necessary to drive localized corrosion.

The environment within the oral cavity is not well defined. There are a number of recipes for artificial saliva, of which the most common is that of Fusayama[30] (NaCl, 0.400 g dm^{-3}; KCl, 0.400 g dm^{-3}; $CaCl_2 \cdot H_2O$, 0.795 g dm^{-3}; $NaH_2PO_4 \cdot H_2O$, 0.69 g dm^{-3}; $Na_2S \cdot 9H_2O$, 0.005 g dm^{-3}; pH 5.5), which may be acidified with citric acid to pH 4.0. However, in reality, the makeup of human saliva varies considerably between individuals, especially in the sulfide

content, which can cause tarnishing of both silver- and gold-based amalgams. In any case, many foodstuffs are acidic with high-chloride levels, and are thus far more corrosive than saliva. Moreover, oral hygiene has a strong affect on the corrosiveness of the oral environment. Many dental products also contain fluoride, some of the specialist varnishes used by dentists being over 2 wt% fluoride, which can cause corrosion even in commercial purity titanium.[31] An example of the problem of defining the environment for the oral cavity is illustrated in the work of Schiff *et al.*,[32] who investigated the corrosion performance of different orthodontic wires. All showed excellent resistance in artificial saliva, but in monofluorophosphate-containing mouthwashes, NiTi-based alloys were found to suffer strong corrosion, while TiNb alloys and TiMo-based alloys gave satisfactory performance. However, when a mouthwash containing stannous fluoride was used, the NiTi alloys outperformed the other two alloys, the TiMo alloys corroding badly.

2.27.5 Metals and Alloys Used in Biomedical Applications

In addition to an alloy's mechanical and corrosion resistance properties, its surface finish is also an important aspect for biomedical devices. The type of surface finish required on an implant is very much dependent on its final use. For example, on the one hand, for prostheses implants usually a rough surface is desirable, as this promotes attachment to the surrounding bone or cement. On the other hand, implants that come into contact with the blood stream, such as vials or reservoirs used to hold the chemicals required in chemotherapy, need to have smooth surfaces to prevent provoking thrombus or crystal formation of the applied chemotherapeutics. As a result, the compositions, mechanical properties, and surface finish of the major types of metals and alloys available are governed by a range of international and national standards, most notably, the various parts of ISO 5832.

2.27.5.1 Titanium and Titanium Alloys

Titanium has excellent corrosion resistance to most environments likely to be found *in vivo*, with the possible exception of acid anoxic regions (e.g., gastric fluids) where the protective passive oxide may not form. The ability of the passive oxide film to provide corrosion protection can be improved by anodizing; note that earlier concerns that anodizing may reduce titanium's resistance to stress corrosion cracking (SCC) and corrosion fatigue appear to be unfounded.[33]

Titanium alloys have even better strength-to-weight ratios than does pure titanium, but not quite as high a resistance to pitting corrosion as the parent metal has, and problems can be encountered if the local redox potential is high and the pH low. Theoretically, these conditions could exist during the initial stages of the inflammatory response following surgery, when both a reduction in pH and hydrogen peroxide production can occur, but the author is not aware of any such clinical cases.[21,27–29] Overall, titanium alloys have better corrosion resistance than do cobalt–chromium–molybdenum alloys and stainless steels.

In the previous decade, concerns about the cytotoxicity of vanadium have led to the development of a number of new alloys as potential replacements for the dual phase ($\alpha + \beta$ phases) Ti6Al4V alloy; there have been reports of elevated metal levels in soft tissues and bones surrounding implants fabricated from this alloy.[34] Many of the potential replacements are β stabilized, which allows a lower elastic modulus and thus should aid bone growth, but at the cost of lower wear and fatigue resistances. However, two replacement alloys, Ti15Mo5Zr3Al and Ti6Al2Nb1Ta0.8Mo, have already been used in hip prostheses in cemented and noncemented applications, respectively, with apparently excellent clinical results.[35]

Despite their excellent corrosion resistance and biocompatibility, titanium and its alloys are still not the perfect biomedical materials, as their poor shear strength makes them unsuitable for screws and other forms of fastener devices. This can lead to fixation problems if galvanic corrosion is to be avoided between the titanium alloys and attaching screws. In addition, titanium alloys also have a high coefficient of friction, which means that wear particles may form if rubbing against bone or another implant surface occurs; the latter case has led to failures due to fretting corrosion, in which the passive oxide film is worn away.[36] The poor fretting resistance of the Ti6Al4V alloy represents its most serious drawback for use in load-bearing prosthesis, and therefore, considerable effort has been made to find possible solutions.[37–40] Encouraging results have been reported for both anodizing and titanium nitride coatings.[37,41] Anodization has the advantages of low cost and ease of operation, while the nitride-coated Ti6Al4V has the better fretting resistance[38]; both techniques have demonstrated

biocompatibility.[37,42] Recently, a zirconium alloy containing 2.5% niobium has been introduced for joint prosthesis. A thick oxide layer formed on the Zr2.5Nb by heating to 535 °C gives it a far superior wear resistance to that of titanium alloys.[43] Zr alloys have similar corrosion resistance properties to titanium, but are difficult to machine as they show pyrotechnic tendencies.

2.27.5.2 Cobalt–Chromium–Molybdenum Alloys

The main alloys used are based on either CoCrMo alloy, which has been used extensively in dentistry and more recently for artificial joints, or CoCrNiMo, which has a very high ultimate tensile strength and thus is used for making the stems of prostheses for heavily loaded joints. Molybdenum is added to improve the mechanical properties by decreasing the grain size, rather than to improve the corrosion resistance. These alloys have excellent resistance to most forms of corrosion, including crevice corrosion and corrosion fatigue, but fretting corrosion can cause failures. The nickel-containing alloys have the better corrosion resistance, and ASTM F 1058 (40Co20Cr15Ni7Mo) has a long track record as a permanent implant alloy.[44] However, concerns about the release of toxic Ni^{2+} ions have resulted in the CoCrMo alloy, ASTM F 75 being the dominant cobalt–chromium–molybdenum alloy in use today. A final concern with using CoCrMo alloys is the potential release of chromate, a known carcinogen, into the body. This worry also applies to stainless steels, since these too contain chromium, although at lower levels.

2.27.5.3 Stainless Steels

Early attempts to use 12% Cr Sherman plates with type 304 stainless steel screws led to predictable galvanic corrosion problems; nonetheless, some of these plates remained in patients for over 30 years, only being removed when tumors (nonmalignant) develop over the corroding implants.[45] The galvanic corrosion problem was solved by using an all type 304 stainless steel construction, but this did not prevent problems associated with pitting and crevice corrosion. Nevertheless, small type 304 stainless steel Sherman plates have been removed from patients after ~30 years of service without any sign of corrosion; possibly the low oxygen content of body fluids prevented the pitting potential from being exceeded.[21] In contrast, a number of cases of larger type 304 stainless steel implants developing localized corrosion problems shortly after implantation have been reported.[46] These apparently contradictory results bear testament to the importance of the pathological changes that accompany the healing process in the first few weeks after surgery, and the extent of the healing process can be expected to be dependent on the size of the implant.[21]

The development of the molybdenum-containing type 316L stainless steel led to a significant decrease in the number of failures related to localized corrosion. However, from a review of failures during the years 1980–1989, Zitter[26] suggested that a PREN greater than 26 was required to prevent *in vivo* pitting corrosion, which is slightly above the PREN of most type 316L stainless steels produced, making this grade more suitable for temporary implant devices. The PREN value can be pushed above the recommended threshold of 26 by the addition of nitrogen, which also has the advantage of increasing the ultimate tensile strength of the material, but at the expense of a lower elongation at fracture. During the last two decades, the importance of sulfide and phosphide inclusions in pit initiation processes has been recognized and this has led to the development of type 316LVM stainless steel, in which the nonmetallic inclusion content is reduced by vacuum melting. The composition of the 316LVM grade is usually slightly above the normal 316 specification, typically being 18Cr14Ni3Mo, such that its PREN value of ~28 is above Zitter's threshold limit. The recently developed low-nickel stainless steels such as 23Mn21Cr1Mo (ASTM F 2229-02) are nitrogen strengthened, which can also push their PREN beyond 26.

The superior mechanical and formability properties of stainless steels over titanium alloys are a great advantage in orthodontic applications. However, the consumption of acidic food and beverages means that there is a much greater risk of localized corrosion in the oral cavity than inside the body. To counter this threat, ultraclean high-nitrogen austenitic stainless steels have been developed, for example, 21Cr10Ni3Mo0.3Nb0.4N.[47]

2.27.5.4 Nickel–Titanium Alloy

NiTi is a shape-memory alloy with super-elasticity, a property that is of interest in both surgical and dental applications. Mantovani[48] reviewed the possible uses of NiTi in medical appliances, typical applications

including dental braces, medical staples, and nails. The most likely form of corrosion on NiTi is pitting; its resistance to pitting in body fluids appears to be similar to that of 316L stainless steel type.[49] However, NiTi fares worse than stainless steel orthodontic wires in acidified artificial saliva.[50] Given that there is some evidence that nickel initiates cancer, the high nickel content of NiTi is a matter of concern.[2] To date, studies indicate that the amount of nickel released into the body from the corrosion of NiTi depends on the local environment and the tests in artificial saliva showed similar Ni release rates to 316L stainless steel,[51] while in simulated body fluids, the NiTi released three times as much nickel as in artificial saliva.[52] Rather alarmingly, Heintz *et al.*[53] found that when fibroblast cells adhere to the NiTi alloy, they damage the oxide film, leading to rapid localized corrosion and that this has caused failures in stent wires of explanted endovascular grafts. Hashimoto and Morita[54] have also reported significantly higher nickel ion release rates in the presence of living fibroblast cells. The mechanism by which this accelerated corrosion occurs is not yet known, possibly the attached fibroblasts lead to the development of occluded electrochemical cells. Nevertheless, the biocompatibility of NiTi, as determined in short-term *in vivo* tests on guinea pigs, was reported to be comparable with type 316LVM stainless steel.[55] This view appears to be supported by the largely favorable surgical clinical evidence available to date, including excellent biocompatibility in tendon tissue, minimal corrosion on retrieved implants, and nickel concentrations in the major organs at levels similar to those for stainless steel devices.[56] However, the concerns about potential Ni release and its poor performance in dental applications have led to a number of efforts aimed at improving the NiTi alloys' resistance to pitting corrosion, the most promising techniques being diamond-like carbon coatings and nitriding.[57–59] Recently, a new nickel-free shape-memory alloy has been developed, Ti18Nb4Sn, with a corrosion resistance in simulated body fluids apparently comparable with that of commercial titanium.[60] Once acceptable biocompatibility of this alloy has been demonstrated, there is a good chance of it eventually replacing NiTi, at least in orthodontics.

2.27.5.5 Porous Materials and Metallic Foams

As demonstrated earlier, titanium alloys and CoCrMo alloys have excellent corrosion characteristics for the construction of surgical implants. However, the elastic moduli of these solid metallic alloys are much higher than those of human bone, which means that stresses are not transferred to the surrounding bone effectively, leading to irregular bone growth and osteoporosis. One possible solution to this problem is to use porous implant materials that have lower elastic moduli closer to those of human bone. Such porous materials not only have the advantage of a more suitable elastic modulus, but also open up the possibility of allowing bone growth into the implant itself, thereby improving the adhesion and thus reducing the likelihood of fatigue or fretting failures. Unfortunately, the corrosion rates of porous titanium, porous CoCrMo, and porous NiTi in simulated body fluids have all been reported as being significantly higher than those of their solid counterparts, possibly due to crevice corrosion within the pore matrix.[61–64]

Very recently, interest has been shown in metallic foams, in which a foaming agent helps to control the porosity that can be up to 80%.[65] It has been reported that, for such titanium foams, corrosion rate as low as 0.07 μm year^{-1} can be obtained after aging for 6 days in a simulated body fluid (0.01 M phosphate buffer pH 7.4, 0.027M KCl + 0.137 M NaCl), which is comparable with or even better than that of solid titanium.[66] However, since the wetted surface area of a foam is more than 1000 times its geometric area, the concentration of metallic ions released is still much higher from foams than from solid materials, increasing the risk of health problems related to elevated metal levels in body tissues.

2.27.5.6 Magnesium Alloys

At first sight, most readers will be surprised at the suggestion of fabricating biomedical devices out of a material that corrodes as readily as magnesium; however, corrosion can in fact be a desirable feature! Preliminary investigations are being conducting into the use of magnesium alloys as degradable implants for musculoskeletal surgery,[67] the idea being that magnesium ions encourage bone cell activation so that the bone slowly regenerates as the metal corrodes. This represents an interesting new challenge to corrosion scientist: how to tailor the corrosion rate of the Mg alloy to match the bone regrowth rate? However, it should be noted that Witte *et al.*[67] reported that *in vivo* corrosion rates were four orders of magnitude lower than *in vitro* rates, and that the order of relative performance between the various Mg alloys tested changed with environment, leading these

authors to conclude that the present ASTM standard *in vitro* corrosion tests are unable to predict the *in vivo* corrosion rates of magnesium alloys.

2.27.5.7 Rare Earth Magnets

There are a number of ternary alloys containing rare earth elements that have remarkably strong magnetic properties, such as the samarium–cobalt family. Although originally developed for the magnetic storage industry, these alloys are used in a number of specialized medical applications, for example, as dental keepers; the strong magnetization is used to keep dental fixtures in place. Unfortunately, these rare earth magnets have very poor corrosion resistances and cannot be directly inserted into any body fluid. One solution is to completely seal the magnet inside a stainless steel cladding, but this must not reduce the effectiveness of the magnetization, which rules out the austenitic steels. In the case of dental keepers, ferritic stainless steels with chromium levels as high as 55% have been used, sufficient to withstand the most corrosive foodstuffs. The magnets hold the dental fixture to a ferromagnetic material usually cemented into a residual tooth root; typically this is a Pd–Co alloy offering good corrosion resistance, but there are concerns about the leaching of cobalt ions, which are cytotoxic at high concentrations.[68]

2.27.5.8 Dental Amalgams

Dental amalgams are high-strength multiphase alloys, which makes them vulnerable to localized galvanic or intergranular corrosion between the different phases. The majority of modern dental amalgams are prepared from two types of alloys: conventional silver tin amalgam and high-copper amalgams. The high-copper amalgams have superior clinical properties with a higher resistance to corrosion.[69] The corrosion of any amalgam is of concern, as it leads to the release of mercury into the body, and in rare cases, can cause oral lesions if this redeposits in the oral mucosa.[15] In conventional silver tin amalgams, the most base phase is γ_2 (Sn_7Hg), which releases mercury when it corrodes.[70] Conversely, the most corrosion-prone phase in high-copper amalgams is η' (Cu_6Sn_5), preferential corrosion of which does not release mercury into the body and thus these alloys have recently been favored.[71] However, Joska *et al.*[72] found that mercury release rates from conventional silver amalgams and high-copper amalgams were very similar, the method of preparation being critical. Furthermore, it has been shown that for most people, the major route for mercury to be taken into the body is via food, less than 10% coming from dental amalgams.[73]

2.27.5.9 Titanium Nitride Coatings

Very hard yet smooth, low friction coatings of titanium nitride, typically a few microns thick, can now be produced by a variety of methods, including ion implantation,[38] nitriding,[39,40] physical vapor deposition,[41] and magnetron sputtering.[74] These coatings have been shown to have excellent biocompatibility and have been successfully used on Ti6Al4V to protect against fretting corrosion and other forms of wear, allowing the world's first collar bone replacement to be preformed in 2003,[41] as well as on NiTi alloy and stainless steels to provide protection against pitting and crevice corrosion.[59,75] The excellent combination of biocompatibility with corrosion and wear resistance means that TiN coatings will undoubtedly find increased usage in biomedical applications.

One potential concern about nitride coatings that has yet to be fully investigated is the possibility of galvanic corrosion of the underlying metal at holidays or damaged areas of the coating. This has already been shown to be a problem at microstructural defects in TiN coatings on type 304 stainless steel in saline solutions.[76] Variations on titanium nitride coatings are also being developed, such as titanium aluminum nitride, that reportedly provide better protection to stainless steels without any loss of biocompatibility.[76,77]

2.27.6 Corrosion Types Encountered in Biomedical Applications

2.27.6.1 General Corrosion

For a successful implant material, the longterm general corrosion rate should certainly fall to much less than $1\,\mu m\,year^{-1}$; for almost any other application, such low corrosion rates would be considered insignificant. Nevertheless, even at these rates, it has been reported that after implantation the nickel, chromium, and cobalt levels in surrounding tissues can be significantly higher than normal values (**Table 1**).[2] It has also been shown that the presence of metallic ions released from nonmolybdenum-containing high-nitrogen stainless steels suppresses cell growth of human gingival fibroblasts.[78]

2.27.6.2 Pitting Corrosion

Although cases of pitting corrosion were common with the early stainless steel implants fabricated from grade 304, the addition of Mo (2–3%) to form 316L grade stainless steel has greatly reduced the number of failures. Zitter[26] has suggested that a PREN greater than 26 is required to prevent *in vivo* pitting corrosion. Stainless steel biomedical devices should thus be manufactured from at least high-quality 316L, preferably from ultraclean grades such as 316LVM or grades containing nitrogen additions. The risk of pitting corrosion of stainless steels in the oral cavity is higher than that in implants because of the number of chloride-containing acidic foodstuffs regularly introduced into the mouth. Although this is partly compensated by the ease by which orthodontic devices can be retrieved, it is still recommended that ultraclean high-nitrogen austenitic stainless steels be used rather than the standard 316L grade.

One of the major concerns about the use of surgical implants based on cobalt–chromium–molybdenum alloys is that pitting corrosion could lead to carcinogens being released into the body. This has resulted in numerous *in vitro* investigations into the pitting behavior of these alloys in pseudo-body fluids, all of which have reported excellent resistance to pitting as long as static conditions are maintained.[64,79] However, pitting corrosion has been observed when the CoCrMo alloys were subjected to either cyclic loads or severe cold working.[80]

Titanium metal is immune to pitting corrosion in any *in vivo* environment likely to be encountered. Although Ti alloys are less resistant, no *in vivo* pitting-related failures have been reported. The pitting behavior of nickel–titanium shape-memory alloys has already been discussed earlier.

2.27.6.3 Crevice Corrosion

The most common location of crevice corrosion found in biomedical applications is beneath the heads of fixing screws and it is a very serious problem with stainless steel devices, even when fabricated from high-grade molybdenum and nitrogen-containing alloys. In a survey conducted in 1959, Scales *et al.*[81] found that 24% of grade 316 stainless steel bone plates and screws removed from patients revealed signs of crevice corrosion. Although the introduction of the low nonmetallic inclusion type 316LVM stainless steel and the use of an austenitic microstructure free of any δ-ferrite phase have reduced the level of crevice corrosion problems, these have not been eliminated.

Crevice corrosion on CoCrMo alloys appears to be less of a problem than on stainless steels. Syrett and Davis[82] found no crevice corrosion on specimens removed from dogs and rhesus monkeys after 2 years of implantation. Similarly, Galante and Rostoker[83] found no crevice corrosion on CoCrMo alloy implants removed from rabbits after 1 year, although the latter authors did find single pits in the crevice regions that might have eventually developed into crevice corrosion if given sufficient time.

Crevice corrosion of titanium in neutral chloride environments has only been reported at temperatures in excess of 70 °C, that is, it is not expected *in vivo*. As with pitting corrosion, titanium alloys are less resistant to crevice corrosion than is pure titanium, and Galante and Rostoker[83] have reported single pits in the crevice regions of Ti6Al4V specimens implanted in rabbits for 1 year, but no actual crevice corrosion. Despite titanium's excellent resistance to crevice corrosion, it is not the answer to the problem, since its poor shear strength makes it unsuitable for screws and other fasteners, which are the main crevice formers.

Finally, some total-joint implant designs contain metal-on-metal press-fit conical tapers, which are subjected to stress and motion. Retrieval studies have found crevice corrosion at taper connections consisting of CoCrMo alloys heads with both CoCrMo alloy and Ti6Al4V stems.[2,84] Occasionally, titanium alloy stems have been attacked, which suggests that the wearing away of the passive film plays a role, that is, the mechanism is a combination of fretting and crevice corrosion.[36]

2.27.6.4 SCC and Hydrogen Embrittlement

To the best of the author's knowledge, neither SCC nor hydrogen embrittlement has been observed on recovered surgical implants. Although retrieved implants may show evidence of cracks, they do not show the physical characteristics associated with SCC, and thus almost certainly, result from mechanical damage either during manufacturing or during the recovery process. The laboratory experiments conducted to date support the presumption that the three classes of common implant alloys (Ti, CoCrMo alloys, and stainless steels) are not susceptible to SCC in *in vivo* environments. The only counter evidence comes from *in vitro* tests conducted under experimental conditions that are highly unlikely to ever exist in any true *in vivo* situation, such as extreme

negative potentials or acidic $MgCl_2$ solutions. However, Rodrigues *et al.* have very recently reported the first evidence of hydrogen absorption being implicated in the corrosion of retrieved titanium alloy hip implants.[106]

2.27.6.5 Corrosion Fatigue

One of the features that distinguish corrosion fatigue from mechanical fatigue is that the former is strongly dependent on the frequency of the applied loaded, with low-frequency cycles causing the most damage. Unfortunately, many medical devices are subjected to low-frequency loads, for example, simply walking represents a cyclic loading at about 1 Hz on a hip implant, so the threat of corrosion fatigue might be expected to be high. However, in 1975, Bechtol[85] reviewed all types of clinical fatigue-related failures and claimed that the most common root cause of failure was not related to corrosion, but rather to a breakdown of the bone–cement support interface. This leads to a widening of the separation between the metal prosthesis and bone–cement and finally to the deformity of the metal stem. This mechanism was recently supported by von Knock *et al.*,[86] who found no evidence of corrosion on 11 CoCrMo alloy femoral components retrieved after 2–15 years of service and suggested that the majority of the micromotion between the prosthesis and bone occurs at the bone–cement/bone interface.

Likewise, reviews of the literature related directly to corrosion fatigue of prostheses implants by Leclerc[87] in 1982 and Zitter[26] in 1991 both concluded that as long as the manufacture and metallurgical condition of the device conformed to international standards (e.g., ISO 5832 or ASTM F 138), corrosion played only a minor role in most fatigue failures. However, in contrast, Morita *et al.*[22] reported that the fatigue strengths of 316 stainless steel and a CoCrNiFe alloy were considerably less *in vivo* (rabbits) than in air and proposed that this was due to the low dissolved oxygen concentration in body fluids, causing a corrosive action on the alloys. Furthermore, in his 1982 review, Leclerc did note that the longer the prosthesis was implanted in the patient, the greater the role of corrosion, which is significant giving that, as explained in the introduction, it is anticipated that the required service life of surgical prosthesis will increase in the coming decades.

If corrosion fatigue does become an important issue in the future, the importance of material selection and design were emphasized by Piehler *et al.*,[88] who tested hip nail plates and found that large plates had better corrosion fatigue resistance than did small ones and that Ti6Al4V outperformed 316L stainless steel. The good performance of titanium alloys was also highlighted by Hughes *et al.*,[24] who reported that the corrosion fatigue resistance of titanium was virtually independent of pH over the range 2–7, whereas that of stainless steel declines rapidly below pH 4. These observations are consistent with the findings of Yu *et al.*[89] that corrosion fatigue in stainless steels can be initiated by pitting corrosion. The latter authors also reported that the corrosion fatigue resistance of the common implant alloy Ti6Al4V can be enhanced by nitrogen implantation and heat treatments to produce fine prior-β grain sizes.

2.27.6.6 Fretting Corrosion and Wear

Because all the successful surgical implant alloys are based on passive metals, any process that wears away the protective oxide film is of major concern. As a result, fretting corrosion represents the most important form of attack on load-bearing prosthesis and all three major classes of alloys used, namely, Ti alloys, CoCrMo alloys, and stainless steels, suffer fretting corrosion, sometimes in combination with crevice corrosion.[90] Fretting corrosion not only results in metal loss but also alters the dimensions of the prosthesis, causing fixation problems and allowing additional micromotions. This in turn can increase mechanical wear and lead to the loss of the surrounding bone–cement or bone, which, besides being a serious problem in itself, increases further the amount of movement of the implant, thereby increasing the likelihood of fatigue-related failures.[85] The situation is made worse by the fact that the corrosion products collect locally as particles that can cause further abrasion of the implant; for example, black titanium oxide debris is often found in the vicinity of implants.[91] The main cause of the shearing micromovements that eventually lead to the fretting corrosion is believed to be the large difference between the elastic moduli of solid metallic implants and the surrounding bone or bone–cement.

Morita *et al.*[92] investigated the wear resistances of a high-nickel version of 316L stainless steel (Fe17Cr14Ni2Mo), a cobalt–chromium–molybdenum alloy (Co28Cr6Mo), and the titanium alloy Ti6Al4V and found that rubbing between metal and ultra-high-molecular weight polyethylene, commonly used as a liner for the sockets into which the ball of hip implants are inserted, only caused the oxide on the Ti6Al4V alloy to suffer damage, that is, the titanium alloy was

the least resistant. Worse, when metal rubbed metal, the oxides on all the three implant alloys were destroyed, even at low loads. Furthermore, the addition of calcium chloride and/or hydrogen peroxide to saline solutions has also been shown to lead to increased fretting corrosion of Ti6Al4V.[93] Since H_2O_2 can be produced as part of the inflammatory response of damaged body tissues, it can be postulated that new load bearing prosthesis are at the most risk of developing fretting corrosion, although the corrosion may subside as the H_2O_2 production ceases, sufficient loosening of the device might occur to induce mechanical fatigue. Efforts to reduce the effects of fretting are based mainly on coatings, usually titanium nitride or aluminum oxide.[38–41,74,94]

Besides wear-resistant coatings, endeavors have been made to reduce the threat of fretting corrosion by improving the binding between the implant and its surroundings, be it bone or bone–cement. Proposals of how to achieve this include engineering the shape, topography, porosity, or composition of the implant to provide either in-growth of tissue or enhanced on-growth of mineralized bone[61,95]; plasma spraying a titanium coating with a specific surface roughness on the surface of the Ti6Al4V[96]; or depositing strongly adhered hydroxyapatite coatings that can fuse with the growing bone. Encouraging results have recently been obtained for silicon-doped hydroxyapatite coatings.[97] Although some reports have suggested that hydroxyapatite coatings do not provide any longterm improvement in fixation,[98] overall clinical results indicate that coated implants perform well, especially in young patients.[2,99] Occasionally, erosion–corrosion problems are encountered on implanted valves and pumps, as with nonbiomedical application, the solution to this problem is to use a more resistant material, such as titanium.[100]

2.27.6.7 Galvanic Corrosion

Galvanic corrosion has certainly caused the failure of a number of biomedical devices. However, the vast majority of these cases were caused by poor quality control or a lack of appreciation of the existence of galvanic couple between apparently similar materials. There have even been examples of galvanic corrosion-related failures arising, because just a single grade 304L screw was used in what was otherwise an all grade 316L construction[101]; since it is not possible to visually distinguish one grade of stainless steel from another, the solution to the problem is careful quality control. When the correct materials have been used, galvanic corrosion is not normally a problem; however, the poor shear strength of titanium alloys means that it is not suitable for use as fasteners, so there may be times when the production of a galvanic couple is unavoidable. In the event that titanium and any cobalt–chromium–molybdenum alloys or stainless steels are coupled together, it is likely that the former will become the cathode and thus accelerated corrosion of the latter alloys may be anticipated. However, titanium and its alloys are easily polarized and their passive films make them poor cathodes, which in practice means that the extent of accelerated corrosion caused to any metal from coupling to a titanium alloy can be expected to be small. This argument has been confirmed in a literature review by Mears[102] and also for titanium/cobalt–chromium–molybdenum alloy combinations by the *in vitro* experiments of Lucas *et al.*[80] and in clinical use as reported by Jackson-Burrows *et al.*[103] Nevertheless, Rostoker *et al.*[104] found that type 316L stainless steel suffered pitting corrosion in 1% NaCl solution at 37 °C when it was coupled to either Ti6Al4V, CoCrMo alloy, or graphite, but no pitting corrosion was found when any two of the other three materials were coupled together.

There is also one further aspect of galvanic corrosion to be considered, that is, any bimetallic couple is of course a small battery in which a current flows between the anode and cathode. Even if this current is too small to cause any significant corrosion problems, these could be sufficient to cause the patient pain, indeed persistent pain resulting from such situations has led to the need to retrieve some biomedical devices.[105] In dental applications, galvanic corrosion of amalgams has been linked to causing oral lesions, which can develop into cancer, particularly in patients with a hyposensitivity to mercury.[15]

2.27.7 Conclusions

The knowledge of corrosion and mechanical properties of materials has allowed the development of a number of extremely successful biomedical alloys. As a result, as long as the chosen materials match the requirements of national and international standards, the likelihood of a surgical implant suffering a corrosion-related failure is very low. The most important remaining areas of concern are fretting and corrosion fatigue. Even here, recent advances in titanium nitride coatings and fixation techniques are extremely encouraging, suggesting that a solution to

at least some of the problems is close at hand. The TiN coatings probably represent the most important advance in the protection of biomedical devices since Leclerc[87] wrote his chapter on surgical implants for the third edition of *Shrier's Corrosion* in 1994. However, TiN coatings have yet to be demonstrated in the tapered joints of some total-joint replacements where a combination of fretting and crevice corrosion can cause failures even in titanium alloys. Moreover, device failure is not the only concern; now that younger patients are receiving implants, the concern that extended exposure to even very low levels of corrosion products could result in medical complications, including cancer, is becoming more and more important. So there is still a continuing need to further reduce corrosion rates.

Furthermore, in recent years advanced materials have been developed that possess properties considered highly desirable for biomedical devices, for instance, porosity, shape-memory, and high magnetism. This trend is likely to continue, as the recent upsurge in funding on life science-related research will produce a number of new materials and devices specifically designed for biomedical applications, such as implanted sensors, automatic drug dispensers, and even micromachines. Techniques to protect these materials without interfering with the very functional properties that make them so desirable have been, and will have to be continually, developed, and occasionally, the biomedical industry will continue to throw out more unusual challenges to the corrosion scientist, as in the case of the sacrificial magnesium alloys being proposed to stimulate bone growth where a controlled corrosion rate is deemed desirable.

References

1. Blanchard, C. R.; Medlin, D. J.; Shetty, R. In *Joint Replacement and Bone Resorption*; Shanbhag, A., Rubash, H. E., Jacobs, J. J., Eds.; Taylor & Francis: New York, 2006; pp 559–592.
2. Hallab, N. J.; Jacobs, J. J.; Gilbert, J. L. In *Joint Replacement and Bone Resorption*; Shanbhag, A., Rubash, H. E., Jacobs, J. J., Eds.; Taylor & Francis: New York, 2006; pp 211–254.
3. Urban, R. M.; Jacobs, J. J.; Gilbert, J. L.; Galante, J. O. *J. Bone Joint Surg. Am.* **1994**, *76*, 1345–1349.
4. Jacobs, J. J.; Skipor, A. K.; Patterson, L. M.; Paprosky, W. G.; Black, J.; Galante, J. O. *J. Bone Joint Surg. Am.* **1998**, *80*, 1447–1458.
5. Thong, P. S. P.; Watt, F.; Ponraj, D.; Leong, S. K.; He, Y.; Lee, T. K. Y. *Nucl. Instrum. Methods Phys. Res. B* **1999**, *158*, 349–355.
6. Landsberg, J. P.; McDonald, B.; Watt, F. *Nature* **1992**, *360*, 65–68.
7. Memoli, V. A.; Urban, R. M.; Alroy, J.; Galante, J. O. *J. Orthop. Res.* **1986**, *4*, 346–355.
8. Jacobs, J. J.; Rosenbaum, D. H.; Hay, R. M.; Gitelis, S.; Black, J. *J. Bone Joint Surg. Br.* **1992**, *74*, 740–744.
9. Goodfellow, J. *J. Bone Joint Surg. Br.* **1992**, *74*, 645.
10. Wataha, J. C. *J. Prosthet. Dent.* **2000**, *83*, 223–234.
11. Solomon, H. A.; Reinhard, M. C. *J. Cancer* **1934**, *22*, 606–610.
12. Schiodt, M. *Oral Surg. Oral Med. O.* **1984**, *57*, 281–293.
13. Phillips, R. W.; Schnell, R. J.; Shafer, W. G. *J. Dent. Res.* **1968**, *47*, 666.
14. Muller, A. W.; Van Loon, L. A.; Davidson, C. L. *J. Oral Rehabil.* **1990**, *17*, 419–424.
15. Holmstrup, P. *J. Oral Pathol. Med.* **1991**, *20*, 1–7.
16. Ostman, P. O.; Anneroth, G.; Skoglund, A. *Scand. J. Dent. Res.* **1994**, *102*, 172–179.
17. Lundstrom, I. M. *Int. J. Oral Surg.* **1982**, *12*, 1–9.
18. Blackwood, D. J.; Seah, K. H. W.; Teoh, S. H. In *Engineering Materials for Biomedical Applications*; Teoh, S. H., Ed.; World Scientific: Singapore, 2004; pp 3.1–3.56.
19. Schenk, R. In *Titanium in Medicine*; Brunette, D. M., Tengvall, P., Textor, M., Thomsen, P., Eds.; Springer-Verlag: Berlin, 2001; pp 145–170.
20. Solar, R. J. In *Corrosion and Degradation of Implant Materials*; Syrett, B. C., Acharya, A., Eds.; ASTM Special Technical Publication, ASTM International: Philadelphia, 1979; Vol. 684, pp 259–273.
21. Blackwood, D. J.; Pereira, B. P. *J. Mater. Sci. Mater. Med.* **2004**, *15*, 755–758.
22. Morita, M.; Sasada, T.; Hayashi, H.; Tsukamoto, Y. *J. Biomed. Mater. Res.* **1988**, *22*, 529–540.
23. Moxham, J.; Costello, J. In *Textbook of Medicine*, 2nd ed.; Souhami, R. L., Moxham, J., Eds.; Churchill Livingstone: Edinburgh, UK, 1994; pp 444–534.
24. Hughes, A. N.; Jordan, B. A.; Orman, S. *Eng. Med.* **1978**, *7*, 135–141.
25. Kocijan, A.; Milosev, I.; Pihlar, B. *J. Mater. Sci. Mater. Med.* **2003**, *14*, 69–77.
26. Zitter, H. *Werkst. Korros.* **1991**, *42*, 455–466.
27. Liang, P. G. *Orthop. Clin. North Am.* **1973**, *4*, 249–273.
28. Hadjiargyrou, M.; Ahrens, W.; Rubin, C. T. *J. Bone Miner. Res.* **2000**, *15*, 1014–1023.
29. Tengvall, P.; Lundstrom, I. *Clin. Mater.* **1992**, *9*, 115–134.
30. Fusayama, T.; Katayori, T.; Nomoto, S. *J. Dent. Res.* **1963**, *42*, 1183–1197.
31. Joyston-Bechal, S.; Kidd, E. A. M. *Dent. Update* **1994**, *21*, 366–371.
32. Schiff, N.; Grosgogeat, B.; Lissac, M.; Dalard, F. *Biomaterials* **2004**, *25*, 4535–4542.
33. Zardiackas, L. D.; Roach, L. D.; Williamson, R. S. In *Titanium, Niobium, Zirconium, and Tantalum for Medical and Surgical Applications*; Zardiackas, L. D., Kraay, M., Freese, H., Eds.; ASTM Special Technical Publication, ASTM International: Philadelphia, 2006; Vol. 1471, pp 166–182.
34. Zaffe, D.; Bertoldi, C.; Consolo, U. *Biomaterials* **2004**, *25*, 3837–3844.
35. Maehara, K.; Doi, K.; Matsushita, T.; Sasaki, Y. *Mater. Trans.* **2002**, *43*, 2936–2942.
36. Heimgartner, P.; Schenk, R. In *Critical Factors in Localized Corrosion IV*, Proceedings of the 202nd Meeting of the Electrochemical Society, Salt Lake City, USA, Oct. 20–25, 2002; Virtanen, S., Schmuki, P., Frankel, G. S., Eds.; Electrochemical Society: New Jersey, 2003; Vol. 2002–24, pp 631–639.
37. Disegi, J. A. In Proceedings of the 16th Southern Biomedical Engineering Conference, Biloxi, Mississippi, USA, Apr. 4–6, 1997; Bumgardner, J. D., Puckett, A. D.,

Eds.; Institute of Electrical and Electronics Engineers: New York, 1997; pp 129–132.
38. Buchanan, R. A.; Rigney, E. D.; Williams, J. M. *J. Biomed. Mater. Res.* **1987**, *21*, 355–366.
39. Shenhar, A.; Gotman, I.; Radin, S.; Ducheyne, P.; Gutmanas, E. Y. *Surf. Coat. Technol.* **2000**, *126*, 210–218.
40. Starosvetsky, D.; Shenhar, A.; Gotman, I. *J. Mater. Sci. Mater. Med.* **2001**, *12*, 145–150.
41. Burslem, R. *Mater. World* **2004**, *12*, 31–32.
42. Manso-Silvan, M.; Martinez-Duart, J. M.; Ogueta, S.; Garcia-Ruiz, P.; Perez-Rigueiro, J. *J. Mater. Sci. Mater. Med.* **2002**, *13*, 289–293.
43. Mishra, A. K.; Davidson, J. A. *Mater. Tech.* **1993**, *8*, 16–21.
44. Clerc, C. O.; Jedwab, M. R.; Mayer, D. W.; Thompson, P. J.; Stinson, J. S. *J. Biomed. Mater. Res.* **1997**, *38*, 229–234.
45. Blackwood, D. J. *Corros. Rev.* **2003**, *21*, 97–124.
46. Zitter, H.; Schaschl-Outschar, D. *Werst. Korros.* **1981**, *32*, 324–331.
47. Pan, J.; Karlen, C.; Ulfvin, C. *J. Electrochem. Soc.* **2000**, *147*, 1021.
48. Mantovani, D. *J. Mater.* **2000**, *52*, 36–44.
49. Rondelli, G.; Torricelli, P.; Fini, M.; Rimondini, L.; Giardino, R. *J. Biomed. Mater. Res. B* **2006**, *79*, 320–324.
50. Huang, H.-H. *J. Biomed. Mater. Res. A* **2003**, *66*, 829–839.
51. Barrett, R. D.; Bishara, S. E.; Quinn, J. K. *Am. J. Orthod. Dentofacial Orthop.* **1993**, *103*, 8–14.
52. Rondelli, G. *Biomaterials* **1996**, *17*, 2003–2008.
53. Heintz, C.; Riepe, G.; Birken, L.; Kaiser, E.; Chakfe, N.; Morlock, N.; Delling, G.; Imig, H. *J. Endovasc. Ther.* **2001**, *8*, 248–253.
54. Hashimoto, T.; Morita, M. *Mater. Sci. Forum* **2005**, *475–479*, 2075–2078.
55. Wever, D. J.; Veldhuizen, A. G.; Sanders, M. M.; Schakenraad, J. M.; van Horn, J. R. *Biomaterials* **1997**, *18*, 1115–1120.
56. Ryhanen, J.; Shabalovskaya, S.; Yahia, L. *Mater. Sci. Forum* **2002**, *394–395*, 139–144.
57. Kobayashi, S.; Ohgoe, Y.; Ozeki, K.; Gei, L.; Hirakuri, K. K.; Aoki, H. *Key Eng. Mater.* **2005**, *284–286*, 783–786.
58. Starosvetsky, D.; Gotman, I. *Biomaterials* **2001**, *22*, 1853–1859.
59. Yeung, K. W. K.; Poon, R. W. Y.; Liu, X. M.; Chu, P. K.; Chung, C. Y.; Liu, X. Y.; Chan, S.; Lu, W. W.; Chan, D.; Luk, K. D. K.; Cheung, K. M. C. *Surf. Coat. Technol.* **2007**, *201*, 5607–5612.
60. Kawashima, A.; Watanabe, S.; Asami, K.; Hanada, S. *Mater. Trans.* **2003**, *44*, 1405–1411.
61. Seah, K. H. W.; Thampuran, R.; Chen, X.; Teoh, S. H. *Corros. Sci.* **1995**, *37*, 1333–1340.
62. Blackwood, D. J.; Chua, A. W. C.; Seah, K. W. H.; Thampuran, R.; Teoh, S. H. *Corros. Sci.* **2000**, *42*, 481–503.
63. Becker, B. S.; Bolton, J. D. *Powder Metall.* **1995**, *38*, 305–313.
64. Li, Y.-H.; Rao, G.-B.; Rong, L.-J.; Li, Y.-Y.; Ke, W. *Mater. Sci. Eng. A* **2003**, *363*, 356–359.
65. Wen, C. E.; Yamada, Y.; Shimojima, K.; Chino, Y.; Asahina, T.; Mabuchi, M. *J. Mater. Sci. Mater. Med.* **2002**, *13*, 397–401.
66. Menini, R.; Dion, M. J.; So, S. K. V.; Gauthier, M.; Lefebvre, L. P. *J. Electrochem. Soc.* **2006**, *153*, B13–B21.
67. Witte, F.; Fischer, J.; Nellesen, J.; Crostack, H. A.; Kaese, V.; Pisch, A.; Beckmann, F.; Windhagen, H. *Biomaterials* **2006**, *27*, 1013–1018.
68. Angelini, E.; Pezzoli, M.; Zucchi, F. *J. Prosthet. Dent.* **1991**, *65*, 848–853.
69. Sarker, N. K.; Eyer, C. S. *J. Oral. Rehabil.* **1987**, *14*, 27–33.
70. von Fraunhofer, J. A.; Staheli, P. J. *Nature* **1972**, *240*, 304–306.
71. Eley, B. M. *Br. Dent. J.* **1997**, *182*, 247–249.
72. Joska, L.; Bystrainsky, L.; Novak, P. *Mater. Corros.* **2003**, *54*, 152–156.
73. Newton, T. *Chem. Br.* **2002**, *38*(10), 24–27.
74. Hubler, R. *Surf. Coat. Technol.* **1999**, *116–119*, 1111–1115.
75. Neumann, H. G.; Beck, U.; Drawe, M.; Steinback, J.; Rostock, H. *Surf. Coat. Technol.* **1998**, *98*, 1157–1161.
76. Ibrahim, M. A. M.; Korablov, S. F.; Yoshimura, M. *Corros. Sci.* **2002**, *44*, 815–828.
77. Freeman, C. O.; Brook, I. M. *J. Mater. Sci. Mater. Med.* **2006**, *17*, 465–470.
78. Endo, K.; Abiko, Y.; Suzuki, M.; Ohno, H.; Kaku, T. *Zairyo Kankyo* **1998**, *47*, 570–576.
79. Syrett, B. C.; Wing, S. S. *Corrosion* **1978**, *34*, 138–145.
80. Lucus, L. C.; Buchanan, R. A.; Lemons, J. E.; Griffin, C. D. *J. Biomed. Mater. Res.* **1982**, *16*, 799–810.
81. Scales, J. T.; Winter, G. D.; Shirley, H. T. *J. Bone Joint Surg.* **1959**, *41B*, 810–820.
82. Syrett, B. C.; Davis, E. E. In *Corrosion and Degradation of Implant Materials*; Syrett, B. C., Acharya, A., Eds.; ASTM Special Technical Publication, ASTM International: Philadelphia, 1979; Vol. 684, pp 229–244.
83. Galante, J.; Rostoker, W. *Clin. Orthop. Relat. Res.* **1972**, *86*, 237–244.
84. Gilbert, J. L.; Buckley, C. A.; Jacobs, J. J. *J. Biomed. Mater. Res.* **1993**, *27*, 1533–1544.
85. Bechtol, C. O. *Orthop. Rev.* **1975**, *4*, 23–29.
86. von Knoch, M.; Bluhm, A.; Morlock, M.; von Förster, G. *J. Arthroplasty* **2003**, *18*, 471–477.
87. LeClerc, M. F. In *Corrosion*, 3rd rd.; Shrier, L. L., Jarman, R. A., Burstein, G. T., Eds.; Butterworth Heinemann: Oxford, 1994; Vol. 1, pp 2:164–2:180.
88. Piehler, H. R.; Portnoff, M. A.; Sloter, L. E.; Vegdahl, E. J.; Gilbert, J. L.; Weber, M. J. In *Corrosion and Degradation of Implant Materials: 2nd Symposium*; Fraker, A. C., Griffin, C. D., Eds.; ASTM Special Technical Publication, ASTM International: Philadelphia, 1985; Vol. 859, pp 93–104.
89. Yu, J.; Zhao, Z. J.; Li, L. X. *Corros. Sci.* **1993**, *35*, 587–597.
90. Syrett, B. C.; Wing, S. S. *Corrosion* **1978**, *34*, 379–386.
91. Engh, C. A., Jr.; Moore, K. D.; Vinh, T. N.; Engh, G. N. *J. Bone Joint Surg. Am.* **1997**, *79*, 1721–1725.
92. Morita, M.; Inoue, Y.; Sasada, T. *Toraiborojisuto* **1998**, *43*, 429–435.
93. Montague, A.; Merritt, K.; Brown, S.; Payer, J. *J. Biomed. Mater. Res.* **1996**, *32*, 519–526.
94. Sella, C.; Martin, J. C.; Lecoeur, J.; Bellier, J. P.; Davidas, J. P. *Adv. Biomater.* **1987**, *7*, 119–124.
95. Lemons, J. E. *Surf. Coat. Technol.* **1998**, *103–104*, 135–137.
96. Normand, B.; Renaud, F.; Coddet, C.; Tourenne, F. In *Thermal Spray: Practical Solutions for Engineering Problems*, Proceedings of the 9th National Thermal Spraying Conference, Cincinnati, USA, Oct. 7–11, 1996; Berndt, C. C., Ed.; ASM International: Materials Park, OH, 1996; pp 73–78.
97. Thian, E. S.; Huang, J.; Best, S. M.; Barber, Z. H.; Bonfield, W. *Mater. Sci. Eng. C* **2007**, *27*, 251–256.
98. Parvizi, J.; Sharkey, P. F.; Hozack, W. J.; Orzoco, F.; Bissett, G. A.; Rothman, R. H. *J. Bone Joint Surg. Am.* **2004**, *86A*, 783–786.

99. Dumbleton, J.; Manley, M. T. *J. Bone Joint Surg. Am.* **2004**, *86A*, 2526–2540.
100. Andersen, T. S.; Johansen, P.; Paulsen, P. K.; Nygaard, H.; Hasenkam, J. M. *J. Heart Valve Dis.* **2003**, *12*, 790–796.
101. Jedwab, J.; Burny, F.; Wollast, R.; Naessens, G.; Opdecam, P. *Acta Orthop. Belg.* **1974**, *40*, 877–886.
102. Mears, D. C. *J. Biomed. Mater. Res. (Symp.)* **1975**, *6*, 133–148.
103. Jackson-Burrows, H.; Wilson, J. N.; Scales, J. T. *J. Bone Joint Surg.* **1975**, *57B*, 148–159.
104. Rostoker, W.; Pretzel, C. W.; Galante, J. O. *J. Biomed. Mater. Res.* **1974**, *8*, 407–419.
105. Park, J. B.; Lakes, R. S. *Biomaterials: An Introduction*; Plenum: New York, 1992; pp 108–110.
106. Rodrigues, D. C.; Urban, R. M.; Jacobs, J. J.; Gilbert, J. L. *J. Biomed. Mater. Res. B* **2009**, *88*, 206–219.

2.28 Corrosion by Wood

M. J. Schofield
CAPCIS Ltd, 1 Echo Street, Manchester M1 7DP, UK

Glossary

Cellulose A complex polymeric carbohydrate, the chief constituent of the cell walls of all plants.

Creosote Oily liquid containing phenols and creosols, obtained from coal tar or by the destructive distillation of wood tar.

Heartwood The inner region of wood in a tree, it is dead.

Hemicellulose Plant cell-wall polysaccharides.

Lignin A polymeric organic compound consisting of various aromatic alcohols.

Lumber Wood that has been cut and surfaced for construction use.

Lumen The cavity bounded by a plant cell wall.

Moisture content The weight of water in wood expressed as a percentage of the weight of wood fibrous material (which is considered to be the oven dry weight of the sample).

Polyphenols Alcohols containing two or more benzene rings, each of which have at least one hydroxyl group (OH) attached.

Sapwood The outer layer of wood in a tree, next to the bark.

Tannins Soluble astringent complex phenolic substances of plant origin.

Timber Lumber whose smallest dimension is not less than 5 in.

Wood The hard, fibrous substance that forms the major part of the trunk and branches of a tree.

Abbreviations

ACQ Alkaline copper quat
AWPA American Wood Preservers' Association
CCA Copper–chromium–arsenic
DOT Disodium octaborate

2.28.1 Corrosion of Metals by Wood

Wood can cause corrosion of metals with which it is in contact, and nearby metals by emission of acidic vapors. The most important practical consequence of contact corrosion is the failure of fasteners in wood. Vapor corrosion affects mainly items that are packaged, stored, or displayed in wooden containers.

2.28.2 Structure and Composition of Wood

Wood is an organic material with a cellular structure and open spaces (lumens) between the cells. There

are many constituents in wood; the structural constituents are cellulose and hemicellulose fibers and lignin. Woods show great variability in terms of mechanical properties and chemical composition between species and within individual samples. Wood is very anisotropic in respect of most of its properties.

In dry wood, water is present only in the cell walls and within the cells; this water is bound and hence not available to support a corrosion process. However, dry wood is hygroscopic and it readily absorbs further water, becoming damp or wet. This water is present in the cell walls and in the lumens; it is free water and is usually present in sufficient quantities to support corrosion of metals in contact with the wood.

Woods are more or less resistant to biological attack; preservative chemicals are used to treat it to control fungal decay and insect attack. It is also flammable and therefore often treated with fire-retardant chemicals.

Woods contain organic acids, principally acetic acid with smaller quantities of formic, propionic, and butyric acids. Indeed, in the early twentieth century, most acetic acid was obtained by distillation from wood.[1] Wood also contains some natural corrosion inhibitors such as tannins. The water, acid, inhibitor, and treatment chemical contents of wood influence its corrosivity toward metallic materials.

2.28.2.1 Mechanisms of Corrosion by Wood

Two distinct mechanisms have been identified for the corrosion of metals by wood. These are conventionally called contact corrosion and corrosion by acidic vapor. This discussion applies primarily to the corrosion of iron and steel; corrosion of other metals is addressed in Section 2.28.2.1.3. Indirectly, some of the constituents of wood contribute to corrosion in pulping plants, but this subject is not addressed in this chapter.

2.28.2.1.1 Contact corrosion

Damp wood can be considered as an electrolyte or a slightly acidic solution,[2] hence metals in contact with wood are subject to conventional electrochemical corrosion in which electrons are released at anodic sites and consumed at cathodic sites. Pinion[3] demonstrated the production of alkali (OH^-) at the exposed end of a steel nail in wet wood. This is formed as a result of oxygen reduction in the cathodic reaction of the corrosion process

$$O_2 + 2H_2O + 4e \rightarrow 4OH^-$$

This process occurs most readily at areas of greatest oxygen access, for example, on the heads of fasteners embedded in wood.

The anodic reaction is metal dissolution, which in the case of iron and steel is

$$Fe \rightarrow Fe^{++} + 2e$$

This process occurs on the shanks of fasteners embedded in wood, resulting in damage that is not visible from the outside.

The corrosion of fasteners in wood is often described as a form of crevice corrosion, since metal loss occurs primarily in the area that is shielded from the higher oxygen content environment on the external surface of the wood (and the head of the fastener). However, differential aeration cell corrosion is a better description, because the term crevice corrosion is most commonly applied to corrosion-resistant alloys, for example, stainless steels. Differential aeration cells are reported to occur also on fasteners that connect two different woods,[3,4] resulting in corrosion within the wood of lower oxygen diffusivity.

Moisture content

The moisture content of wood is the dominant factor in controlling contact corrosion. As explained by Zelinka *et al.*,[5] in wood with moisture content below ~15–18%, the water in wood is bound to the hydroxyl sites within the cell walls; there is no free water, and therefore, the ionic conductivity of wood is very low and corrosion is minimal. Increasing the moisture content of wood results in free water within the cell walls; this supports the ionic conductivity, permitting corrosion to occur. This critical moisture content of the wood is equivalent to a relative humidity of the atmosphere around the wood of approximately 80–85%, depending on temperature.[6] Note that the critical moisture content varies between woods and a wider range of 10–20% is given in some references.[2,7] This critical moisture content is often described as the threshold for corrosion to occur, but the transition between dry and wet wood, or no corrosion and corrosion, is not so marked as the term implies.

The threshold moisture content for corrosion is also the minimum moisture content required for fungal decay (for which there is an upper limit of ~60% and oxygen must also be present), hence wood that is to be exposed to wet conditions is typically treated with preservatives to prevent such decay.

Acid content

The organic acid content of wood is an important factor in contributing to contact corrosion and vapor

corrosion of metals by wood. Acetic acid is the main acid involved, since it is present at significantly higher levels than are the other acids in wood. Acetic acid is formed by the hydrolysis of acetylated polysaccharides in the wood. As little as 0.5 ppmw is sufficient to influence corrosion,[8] with corrosivity increasing roughly in line with acid content, but it is to be noted that pH and other measured properties of wood can vary significantly between samples, between heartwood and sapwood, and across small distances in a single sample.

Table 1 shows the variable nature of the acidity of woods;[9–11] the pH of aqueous extracts from eucalyptus woods have been found to be 2.5 or even lower.[12] As a rough guide, woods of pH below 4 are very corrosive and woods of pH greater than 5 are low in corrosivity, but other factors are also relevant.

Salts

Woods naturally have a small salt content, notably sulfate and chloride, but this can increase by a number of means, including marine aerosol deposition, the floating of logs in seawater,[13] the introduction of chemical treatments such as preservatives and fire retardants, and even by the process of salt seasoning.[13] Salts contribute to corrosion by increasing the conductivity of the wood. The role of copper ions, which are an important component of many preservative treatments, is discussed in Section 2.28.2.4.1.

Polyphenols

Polyphenolic compounds that possess two or three adjacent hydroxyl groups are reported to be capable of forming iron chelates. This chelation, along with the acidity of the wood, is considered to be a contributing factor in the corrosion of steel saw blades.[14]

Bimetallic corrosion

Contact between dissimilar metals in damp wood results in bimetallic corrosion.[2,15] An example is that of steel-framed boats with wooden planking attached by copper alloy fasteners.[11] This is easily avoided by preventing such contact or selecting compatible fasteners.

Effect of corrosion on wood

The alkali (OH^-) that is formed at the cathodic site during the corrosion process degrades the wood. Hemicelluloses and lignins are readily attacked by alkali, making the wood spongy. This effect is commonly known as nail sickness[3,16] and results in a loss of holding power of nails and other fasteners in the wood.

At anodic sites, iron corrosion products react with tannin to form blue-black staining. This is a common problem in wooden railway sleepers and it affects other materials of cellulosic structure, including textiles.[17] It also occurs where iron corrosion products from steel wool (used to finish wood) and handling chains come into contact with the wood. Tannate

Table 1 Relative corrosivity of woods by vapor corrosion

Wood	*Classification in Defence Guide-3A*[10]	*Typical pH values*
Oak	Most corrosive	3.35, 3.45, 3.85, 3.9
Sweet chestnut	Most corrosive	3.4, 3.45, 3.65
Steamed European beech	Moderately corrosive	3.85, 4.2
Birch	Moderately corrosive	4.85, 5.05, 5.35
Douglas fir	Moderately corrosive	3.45, 3.55, 4.15, 4.2
Gaboon	Moderately corrosive	4.2, 4.45, 5.05, 5.2
Teak	Moderately corrosive	4.65, 5.45
Western red cedar	Moderately corrosive	3.45
Parana pine	Least corrosive	5.2–8.8
Spruce	Least corrosive	4.0, 4.45
Elm	Least corrosive	6.45, 7.15
African mahogany	Least corrosive	5.1, 5.4, 5.55, 6.65
Walnut	Least corrosive	4.4, 4.55, 4.85, 5.2
Iroko	Least corrosive	5.4, 6.2, 7.25
Ramin	Least corrosive	5.25, 5.35
Obeche	Least corrosive	4.75, 6.75

Source: Rance, V. E.; Cole, H. G. *Corrosion of Metals by Vapours from Organic Materials*, 3rd edn.; HMSO: London, 1958.
DG-3A *Defence Guide for the Prevention of Corrosion of Cadmium and Zinc Coatings by Vapours from Organic Materials*; HMSO: London, 1966.
Gray, V. R. *J. Inst. Wood Sci.* **1958**, *1*, 58.

staining is generally restricted to a depth of 0.4 mm[18] and hence is not a significant form of degradation.

Some sources suggest that acidic conditions develop at the site of the anodic reaction, by reaction between chloride ions and ferrous ions (though more accurately the acidity will be produced by metal ion hydrolysis),[19] but the significance of this is not well documented.

2.28.2.1.2 Corrosion by acidic vapor

The organic acids in wood are volatile and can cause corrosion of some metals in confined spaces. It is principally acetic acid that is responsible for this form of corrosion. The mechanical properties of woods can be enhanced by thermal treatment, but this increases the volatility of organic acids, which increases the risk of vapor corrosion at least in the short term. This form of corrosion was traditionally associated with the use of wooden packaging and pallets, but more recently, it has become a serious issue in conservation.

Aoyama *et al.*[20] attributed the corrosion of nickel-plated brass to acetic acid released from wooden storage boxes. In extreme conditions of heat and humidity, sufficient acetic acid vapor is produced from wood to cause stress corrosion cracking of quenched and tempered AISI 1035 carbon steel nails.[21] The imposition of hardness and tensile strength limits was found to be necessary to prevent such cracking. Volatile corrosion inhibitors such as *m*-dinitrobenzene and anthraquinone have been shown to be effective in suppressing acetic acid vapor corrosion of mild steel in wooden boxes in hot, humid conditions.[22]

A common problem in conservation is the corrosion of artifacts stored in wooden boxes or displayed in wooden showcases. As summarized by Berndt,[23] corrosion is due mostly to the presence of acetic acid vapor, although other chemicals, including formic acid, formaldehyde (which is oxidized by atmospheric oxygen to formic acid), and esters (which hydrolyze to form alcohols and acids) can also contribute. Work conducted at the British Museum has shown that Egyptian bronze artifacts corroded when stored in wooden cupboards due to the presence of acetic acid vapor at levels in the range 1000–2500 mg m^{-3}. Subsequent work found that corrosion did not occur on a range of lead and copper alloys when the acid vapor was reduced to below ~350 mg m^{-3} in showcases, but this limit is a function of humidity and temperature.[24] Humidity control and acid-scavenging chemicals are used to control this form of corrosion, but the preferred solution is to select materials of construction that release less acidic vapor.

2.28.2.1.3 Corrosion of other metals

Aluminum

Aluminum alloys have reasonable resistance to corrosion in wood of low moisture content, but it is recommended that they should be protected by coatings or other technique if the wood becomes wet in service.[25] Specific corrosion risks are posed by wood that contains chloride ions (e.g., by absorption from salt water) and by copper-containing and mercury-containing preservative treatment chemicals. This corrosion risk extends to flashing that is exposed to water runoff from treated wood. Some preservatives such as pentachlorophenol and chromates are considered to be innocuous to aluminum, but note that chromate treatments have been largely phased out for environmental reasons. Coal tar creosote is considered to have corrosion-inhibiting properties toward aluminum, but some corrosion testing that forms the basis of some reports has failed to separate out other effects, such as water repellency, associated with creosote.

Brass and other copper alloys

Copper and its alloys are susceptible to corrosion by acetic acid, as noted in **Section 2.28.2.1.2**, and are therefore subject to corrosion in most woods. Brass screws were reported to suffer dezincification in damp wood that had been treated with zinc chloride and sodium dichromate.[26] Copper and its alloys are susceptible to corrosion by ammonia, which is released from ammonium salt fire-retardant treatments in some circumstances.

Lead

Lead roofs have been corroded by acetic and formic acid released from wood that has been wet by condensation in poorly ventilated conditions.[27]

2.28.2.2 Corrosion Testing

A number of corrosion test methods are applicable to contact corrosion; these are reviewed by Zelinka *et al.*[28]

Simple exposure tests have been standardized by the American Wood Preservers' Association (AWPA). AWPA standard method E12-94 'Standard Method of Determining Corrosion of Metal in Contact with Treated Wood,' is an accelerated test method, in which a metal coupon is placed between two pieces of preservative-treated wood and exposed for 240 h at 49 ± 1 °C, 90 ± 1% relative humidity. The assessment of the test coupons is based on mass loss

and visual appearance. Because the test is accelerated, it cannot be used to predict a service life for a metal component in contact, rather, it is used to rank the corrosivity of preservative treatments. The industry is currently working to develop new tests that will be more directly applicable to the corrosion of fasteners in treated wood. ASTM subcommittee G01.14.01 'Corrosion of Metals in Treated Lumber' is also addressing this issue.

A further AWPA test is E17-99 'Standard Method for Determining Corrosion Rates of Metals in Contact With Treating Solutions,' in which metal coupons are exposed in the preservative treatment chemicals. This method is applicable to corrosion in treatment plant equipment, and not to corrosion on contact with treated wood.

Electrochemical corrosion test methods have also been employed. Linear polarization resistance measurement has been used successfully to study the corrosion behavior of carbon steel, stainless steel, and zinc in solutions of preservative treatment chemicals[29] and in aqueous extracts prepared from wood.[5] However, it is not a suitable technique for corrosion in contact with wood because of the relatively poor ionic conductivity of wood and it is not capable of measuring localized corrosion effects such as pitting. Alternating current impedance testing is more appropriate for the study of contact corrosion; it has been used with some success on preservative-treated wood.[30] However, Zelinka *et al.*[31] reported that results obtained in their tests could not be related to corrosion rate 'because of gaps in understanding the impedance properties of wood.'

2.28.2.3 Corrosion Modeling

The modeling of the corrosivity of wood toward embedded fasteners has been undertaken in Australia,[32] taking into account several factors, including the moisture content and acidity of the wood, temperature, local sheltering, and weathering, in addition to climatic and atmospheric factors. The objective of this modeling is the prediction of the service life of fasteners in wooden structures.

2.28.2.4 Industrial Significance of Corrosion of Metals by Wood

Wood is the oldest construction material known to man. It has long been used to construct dwellings, boats, and structures such as bridges, and to manufacture weapons. In the first stages of the Industrial Revolution, most machines were made of wood.[33] Although its industrial use has declined in recent years, it is expected to increase again now that is has been recognized[34] that wood is an environmentally friendly material with significantly lower energy cost and carbon footprint than steel and concrete in many building applications.

2.28.2.4.1 Building

The most important use of wood within the purview of this book is its use as a construction material for buildings. Wood in ground contact, or exposed to rain or to humid air, has a moisture content sufficient to sustain contact corrosion of metallic fasteners. Corrosion in this condition is often dominated by the effects of preservative or fire-retardant chemicals. The following discussion also applies to other wood in ground contact such as railway sleepers and utility poles, which are routinely treated with these chemicals.

Preservative treatments

A range of preservative chemicals are used for the treatment of wood against fungal decay and insect attack. Traditional preservatives include creosotes. Creosotes are low in corrosivity toward most metallic materials except lead,[8] and are often reported as having corrosion inhibiting properties. Coal tar creosote is reported to be less corrosive than is wood tar creosote toward carbon steel.[35] Other organic chemicals such as pentachlorophenol are considered to be noncorrosive.

Many modern preservatives are based on copper compounds, since copper is a potent fungicide. Copper ions introduced into wood are a major contributing factor to corrosion in treated wood, but their role has not been established with certainty. Baker[2] proposed a form of galvanic corrosion between the copper and less noble fastener materials. However, work by Simm *et al.*[36] failed to find metallic copper on the surface of steel fasteners that had corroded in copper–chromium–arsenic (CCA) treated wood, implying that the possible cathodic reaction of Cu^{2+} ion reduction was not operative. Copper salts in common use include copper naphthenate, copper azoles, CCA, and alkaline copper quat (ACQ).

CCA preservatives are now less commonly used, since the U.S. Environmental Protection Agency has concluded that their arsenic content may threaten human health. They have been largely replaced by ACQ formulations, which are more corrosive toward steel and galvanized steel. This is considered to be due to the higher level of copper ions in the newer formulations, and because the chromium and arsenic

in CCA acted as corrosion inhibitors. Tests conducted on galvanized steel in accordance with AWPA E12 standard (see Section 2.28.2.2) showed corrosion rates < 0.03 mm year^{-1} for DOT, $\sim$0.03–0.08 mm year^{-1} for CCA, and 0.08–0.16 mm year^{-1} for ACQ types. Types 304 (UNS S30400) and 316 (UNS S31600) stainless steels performed well in all treatments.[37]

Recommended fasteners[36] for use in wood treated with modern copper-containing preservatives are galvanized steel, with at least 0.9 oz ft^{-2} (45 μm) of zinc, or stainless steels. Aluminum alloy fasteners are not recommended for this application. Previous work[2] showed that the corrosion resistance of copper and Monel (UNS N04400) in CCA-treated wood was intermediate between that of stainless steels and galvanized (and bare) steel.

Copper-free formulations such as zinc naphthenate[11] and disodium octaborate (DOT, $Na_2B_8O_{13}{\cdot}4H_2O$) are reported to be noncorrosive.

Fire-retardant chemicals

Common fire-retardant treatment chemicals are based on boron compounds (boric acid or sodium tetraborate), zinc chloride, or ammonium salts (sulfate or phosphate). Corrosion inhibitors such as sodium dichromate and ammonium thiocyanate are present in some formulations.

Boron-based treatments are considered not to pose a risk of corrosion to metallic materials in contact with treated wood.[8] Zinc chloride is corrosive, but reports for the ammonium-based treatments are variable.[11]

2.28.2.4.2 Water tanks

Wood is still used for tank construction and may have "the best life cycle cost of any material in handling water, acidic wastewaters, and weak acid solutions up to temperatures of 60 °C."[38] Corrosion-resistant materials should be used for fasteners and fittings in such tanks.

2.28.2.4.3 Water cooling towers

Wood is the most economic material of construction for some components of water-cooling towers of certain design and capacity. Wood degrades in this service, but fastener failure by corrosion does not appear to be an issue.

2.28.2.4.4 Wood cutting tools

In use, saw blades[11] and other wood cutting tools[39] suffer corrosion and corrosive wear that has an electrochemical component that 'is not negligible and can be very important.'

2.28.2.4.5 Conservation

In addition to the issues discussed in Section 2.28.2.1.2, corrosion of metals by wood endangers some of our cultural heritage. Sánchez-Gómez *et al.* report that steel bars embedded in a wooden door to a belfry tower dating back to 1724 corroded, damaging the door.[40]

References

1. http://en.wikipedia.org/wiki/Acetic_acid (viewed on 13/07/2009).
2. Baker, A. J. *Durability of Building Materials and Components*; ASTM STP 691; American Society for Testing and Materials: West Conshohocken, 1980; pp 981–993.
3. Pinion, L. C. The degradation of wood by metal fastenings and fittings; Timberlab 174–183, 1968–1972, Paper No. 27, Forest Products. Research Laboratory: Princes Risborough, Aylesbury, Bucks, England, 1970; pp 61–66.
4. Causes of Metal Corrosion in Timber Fixings; British Stainless Steel Association. http://www.bssa.org.uk (viewed on 13/07/2009).
5. Zelinka, S. L.; Rammer, D. R.; Stone, D. S. *Corrosion of Metals in Contact with Treated Wood: Developing Test Methods*; Proceedings of Corrosion 2008, New Orleans, March 16–20, 2008; NACE International: Houston, 2006.
6. *Technical Notes*. *Treated Wood News*; Western Wood Preservers Institute, 2001.
7. Hof, T.; Van der Elburg, J. *TNO Nieuws* **1967**, *22*(4), 217–218.
8. Davis, R. I. *Timber and Metal – The Connection*; Proceedings of the Pacific Timber Engineering Conference; Timber Research and Development and Advisory Council: Gold Coast, Australia, 1994; pp 439–448.
9. Rance, V. E.; Cole, H. G. *Corrosion of Metals by Vapours from Organic Materials*; HMSO: London, 1958.
10. DG-3A Defence Guide for the prevention of corrosion of cadmium and zinc coatings by vapours from organic materials; HMSO: London, 1966.
11. Gray, V. R. *J. Inst. Wood Sci.* **1958**, *1*, 58.
12. Hillis, J. A. F. *Wood Extractives and Their Significance to the Pulp and Paper Industries*; Academic Press: New York, 1962; pp 389–392.
13. Corrosion of Metals by Wood. Guides to Practice in Corrosion Control; National Physical Laboratory: Teddington, UK, 1979 (available online at http://www.npl.co.uk).
14. Krilov, A.; Gref, R. *Wood Sci. Technol.* **1986**, *20*(4), 363–370.
15. Packman, D. F. *Wood* **1954**, *19*(9), 358–359.
16. Baker, A. J. Degradation of wood by products of metal corrosion, U.S.D.A. Forest Service Research Paper FPL 229; U.S. Department of Agriculture Forest Service Forest Products Laboratory: Madison, WI, 1974.
17. Bell, W. A.; Gibson, J. M. *Nature* **1957**, *180*, 1056.
18. Wengert, E. M. Causes and Cures for Stains in Dried Lumber. http://www.woodweb.com/knowledge_base/

Causes_and_Cures_for_Stains_In_Dried_Lumber.html (viewed on 13/07/2009).
19. Crossman, M.; Simm, J. *Manual on the Use of Timber in Coastal and River Engineering*; Thomas Telford Publishing: London, 2004.
20. Aoyama, Y.; Kamioka, M. *Corros. Eng.* **1965**, *14*(9), 381–401. (Article in Japanese with English abstract.)
21. Fox, S. P.; Barton, G. M.; Hawbolt, E. B. *Mater. Perform.* **1977**, *16*(10), 36–40.
22. Srivastava, K.; Dwivedi, S. K. *J. Corrosion Sci. Eng.* **2007**, *10*, Paper 21. http://www.jcse.org (viewed on 13/07/2009).
23. Berndt, H. Assessing the detrimental effects of wood and wood products on the environment inside display cases. American Institute for Conservation 15th Annual Meeting, Vancouver, Washington, DC.
24. Bradley, S.; Thickett, D. *The Pollutant Problem in Perspective*; IAP, 1998.
25. Hatch, J.E. *Aluminum – Properties and Physical Metallurgy*; ASM, 1984.
26. Baechler, R. H. *Proc. Am. Wood-Preservers' Assoc.* **1939**, *35*, 56–63.
27. Edwards, R.; Bordass, W.; Farrell, D. *Analyst* **1997**, *122*, 1517–1520.
28. Zelinka, S. L.; Rammer, D. R. Review of test methods used to determine the corrosion rate of metals in contact with treated wood. United States Department of Agriculture Forest Service, Forest Products Laboratory, General Technical Report, FPL–GTR–156, 2005.
29. Zelinka, S. L.; Rammer, D. R.; Stone, D. S.; Gilbertson, J. T. *Corros. Sci.* **2007**, *49*(3), 1673–1685.
30. Bailey, G.; Schofield, M. J. *J. Inst. Wood Sci.* **1984**, *10*(1), 14–18.
31. Zelinka, S. L.; Rammer, D. R. Proceedings 9th World Conference on Timber Engineering, Portland, USA, August 6–10, 2006.
32. Nguyen, M. N.; Leicester, R. H.; Wang, C. Embedded corrosion of fasteners in timber structures. Manual No. 6. Forest and Wood Products Australia, 2008.
33. http://www.puhsd.k12.ca.us/chana/staffpages/eichman/Adult_School/us/fall/industrialization/1/industrial_revolution.htm (viewed on 13/07/2009).
34. Bowyer, J. L.; Shmulsky, R.; Haygreen, J. G. *Forest Products and Wood Science*, 4th ed.; Blackwell: Oxford, UK, 2003; p 506.
35. Paes, J. B.; Vital, B. R.; Della Lucia, R. M.; Della Lucia, T. M. C. *R. Arvore, Viçosa-MG* **2002**, *26*(5), 621–627.
36. Simm, D. W.; Button, H. E. *Corros. Prevent. Control* **1985**, *32*(2), 25–35.
37. TechnicalBulletin. Preservative treated wood. http://www.strongtie.com (viewed on 13/07/2009).
38. Smallwood, R. E. *Life Cycle Maintenance Considerations of Non-Metallic Process Equipment Compared to Metallic Equipment*; Proceedings of Corrosion 2006, San Diego, USA, March 12–16, 2006; NACE International: Houston, 2006.
39. Gauvent, M.; Rocca, E.; Meausoone, P. J.; Brenot, P. *Wear* **2006**, *261*(9), 1051–1055.
40. Sánchez-Gómez, M. A.; Esteso, M. A. *Corrosion Effects on the Cultural Heritage*; Proceedings of Corrosion 2002, Denver, USA, April 7–11, 2002; NACE International: Houston, 2002.

Further Reading

Nguyen, M. N.; Leicester, R. H.; Wang, C. *Atmospheric Corrosion of Fasteners in Timber Structures.* Manual No. 5. Forest and Wood Products Australia, 2008.

2.29 Corrosion in Radiolysis Induced Environments

G. O. H. Whillock
National Nuclear Laboratory, B170 Sellafield, Cumbria CA20 1PG, UK

Glossary

Gray SI unit of absorbed radiation; 1 Gy = 1 J kg^{-1}.

***G*-value** Radiolytic yield in terms of the number of particles formed per 100 eV of absorbed energy.

LET Linear energy transfer; measures the energy deposited by ionizing radiation per unit path length in the absorbing medium.

Rad Non-SI unit of absorbed radiation dose; 100 rad = 1 Gy.

Radiolysis Chemical decomposition brought about by the absorbance of ionizing radiation.

Abbreviations

IASCC Irradiation-assisted stress corrosion cracking

LET Linear energy transfer

LWR Light water reactor

PWR Pressurized water reactor

Symbols

V_{SCE} Electrode potential, volts with respect to the saturated calomel electrode

V_{NHE} Electrode potential, volts with respect to the normal hydrogen electrode

***X*(*Y*)** Element *X* present in valency state *Y*; for example, Cr(VI) is $Cr_2O_7^{2-}$ or CrO_4^{2-}

2.29.1 The Nature of Ionizing Radiation

Ionizing radiation is produced by a number of processes, such as radioactive decay, nuclear fission, and nuclear fusion. There are many types of ionizing radiation, including neutrons, protons, α-particles (helium nuclei), β-particles (electrons), γ-radiation, X-rays, and cosmic rays. Ionizing radiation imparts energy to matter with which it interacts, sufficient, by definition, to displace electrons from the matter. The amount of energy absorbed, ϕ_{ads}, obeys the exponential decay law, depending on the energy of the incident radiation flux, the nature of the radiation, the physical properties of the matter (such as its density), and the depth of the matter through which the radiation passes:

$$\phi_{ads} = \phi_0 - \phi_0 e^{-\mu x} \quad [1]$$

where ϕ_0 is the incident flux, μ is the attenuation coefficient (which is specific to the type of radiation and depends on the density of the absorbing matter), and x is the depth of the absorbing matter.[1] It is

immediately clear why thick sections of dense materials such as lead are used to shield ionizing radiation in order to limit occupational exposures to acceptable levels. The range of charged particles, that is, electrons, protons, α-particles, and heavier ions, is much shorter and decreases with particle mass. For example, α-particles will not penetrate through a sheet of paper, and β-radiation is typically completely absorbed by a few millimeters of a light metal such as aluminum. However, heavy charged particles deposit significantly more energy per unit path length than do lighter uncharged particles. This is termed the linear energy transfer (LET) value[2]; LET values for 3 MeV β-radiation, protons, and α-radiation in water are 0.2, 20, and 150 keV μm^{-1} respectively.[1]

The SI unit of absorbed radiation is the gray (Gy), which is defined to be the absorbance of 1 J kg^{-1} of matter. However, a unit termed 'rad' is still frequently used (1 Gy = 100 rad). Absorbed dose rates are measured in grays per second or, more conveniently, grays per hour. A related unit is the sievert (Sv), which is used to measure absorbed dose equivalents in biological samples, using weighting factors to take into account the different damage effects of the various types of ionizing radiation.

2.29.2 Corrosion Effects Arising from Radiation Absorbance

2.29.2.1 The Effect on Metals and Surface Films

The majority of the effects of ionizing radiation on corrosion processes are associated with radiolytic alteration of the environment (see **Section 2.29.2.2**), but a brief mention is made here of effects on the metal substrate and surface films, which may be important, sometimes critically so, in some circumstances.

All types of ionizing radiation may affect the semiconducting properties of passive films (termed the photoradiation effect) by altering the number and type of charge carriers.[3,4] Cathodic processes may consequently be affected, leading to change in the corrosion rate. However, the photoradiation effect is small and can be confused with the much more significant effect on cathodic processes arising from solution radiolysis.

Heavy, energetic particles (i.e., neutrons and protons) promote atomic displacement and the formation of increased numbers of defects, both in metal lattices and oxide films. A number of important physical properties are affected as a result (e.g., electrical resistivity and hardness), but as far as corrosion processes are concerned, there are two principal effects, both deriving, in very simplistic terms, from increased defect density arising from atomic displacements:

- the thickness and porosity of oxide films forming on metals such as aluminum and zirconium increases[5,6];
- some alloying elements may preferentially segregate to grain boundaries at elevated temperatures (i.e., in nuclear reactors), leading, in the case of stainless steels, to radiation-induced sensitization.[7] This may either have significant *in situ* corrosion consequences in the form of susceptibility to irradiation-assisted stress corrosion cracking (IASCC),[8–10] or introduce a legacy effect when, for example, reactor-irradiated stainless steel is subsequently exposed to water containing very low chloride concentrations.

The other principal types of ionizing radiation, α, β, γ, X-rays, and cosmic rays, do not detectably affect the metal microstructure.

2.29.2.2 The Effect on the Environment – Radiolysis

Radiolysis is defined as chemical decomposition brought about by the absorbance of ionizing radiation. Significant corrosion consequences can arise from the products resulting from the decomposition of water, solutes, gases, and polymers and, in some circumstances, from the alteration of the valence state of certain solutes.

2.29.2.2.1 Water radiolysis

Water is fragmented by ionizing radiation, resulting in the formation of a range of short-lived intermediates, which react with themselves, water, and some solutes to generate hydrogen, hydrogen peroxide, and to a much lesser extent, oxygen, as stable molecular products[11–13]:

$$H_2O \rightarrow e_{aq}^-, H, OH, O_2^-, HO_2, H^+, OH^-, H_2, H_2O_2, O_2 \quad [1]$$

In a completely closed single-phase situation, almost never encountered in practice, the irradiation of pure water produces no overall oxidizing or reducing effect. When, however, the gaseous products are able to escape, a net oxidizing or reducing effect is produced depending on the composition of the cover gas; the former results in the presence of air, oxygen, or argon, whereas the latter results in the presence of excess hydrogen.

Radiolytic reaction schemes are complex, involving a host of interactions that limit the ultimate yield of the molecular products. For example,

$$e^-_{aq} + H_2O_2 \rightarrow OH + OH^- \quad [2a]$$

$$H + H_2O_2 \rightarrow H_2O + OH \quad [2b]$$

It therefore follows that any solute that preferentially reacts with the reducing radicals acts to increase the overall yield of hydrogen peroxide by suppressing reactions [2a] and [2b]; in such circumstances, the radiolytic yield of hydrogen is suppressed. Oxygen does this directly, reactions [3a] and [3b], and Cl^- does this indirectly, reactions [4a]–[4b].[11,12] The term radical scavenger is used to denote species that act like this.

$$e^-_{aq} + O_2 \rightarrow O_2^- \quad [3a]$$

$$H + O_2 \rightarrow HO_2 \quad [3b]$$

$$Cl^- + OH \rightarrow Cl + OH^- \quad [4a]$$

$$e^-_{aq} + Cl \rightarrow Cl^- \quad [4b]$$

$$H + Cl \rightarrow H^+ + Cl^- \quad [4c]$$

Cl^- is not consumed in the process, so low concentrations exert a pronounced effect.

The yield of radiolytic species is quantified by *G*-values, which measure the number of molecules decomposed or the number of molecules, ions, or radicals formed per 100 eV of absorbed radiation energy. *G*-values depend on the type of radiation (i.e., its LET value); fast neutrons and α-radiation produce higher yields of the molecular products than do β- and γ-radiation, the last two being essentially equivalent. **Table 1** illustrates these effects for high-purity, high temperature water.[14] The effect of radical scavengers also depends on the type of radiation, it being more pronounced for neutrons than for β/γ-radiation. Hence, considerable care must be taken in applying the findings of studies carried out with one type of radiation (γ-rays being frequently used) to situations where other types of radiation may dominate (e.g., reactor cores).

G-values are of fundamental importance in elucidating radiolytic mechanisms, usually done using pulse radiolysis techniques, but from a corrosion viewpoint, have little direct relevance, since it is the steady-state concentration of certain species, primarily hydrogen peroxide, that is important. The steady-state concentration depends not only on the net rate of radiolytic formation, which from the *G*-value concept can be seen to be dose rate dependent, but also on any nonradiolytic consumption mechanisms (including corrosion reactions). However, radiolytic production rates, along with kinetic data for relevant chemical and electrochemical (i.e., corrosion) reactions, do have an important application in forming the basis for mechanistic models predicting the corrosion potential in light water reactors (LWRs) (see **Section 2.29.3.1**).

2.29.3 The Effect of Radiolysis on the Corrosion of Metals in Aqueous Environments

A number of important effects of radiolysis on corrosion processes have been discovered as a consequence of the exploitation of nuclear energy, ranging from energy production in water-cooled power reactors, through reprocessing operations, to the disposal of nuclear wastes.

Upon initial irradiation, the potential of platinum and stainless steel electrodes initially falls toward the reversible hydrogen potential, but then rises with increasing dose uptake toward potentials dominated by hydrogen peroxide redox reactions.[15] This behavior reflects the faster diffusion rate of hydrogen to the electrode compared with hydrogen peroxide and the greater radiolytic yield of hydrogen peroxide compared with hydrogen. For most corrosion processes, the oxidizing effect due to the production of hydrogen peroxide dominates.

Table 1 Radiolytic yields in high-purity, high temperature water

Energy source	*LET (eV Å^{-1})*	*Yields of species (number per 100 eV)*							
		$-H_2O$	H_2	$H+e^-_{aq}$	H	e^-_{aq}	H_2O_2	OH	HO_2
β, γ	0.02	3.74	0.44	2.86	0.55	2.31	0.70	2.34	0.00
Neutrons	4.0	2.79	1.12	0.72	0.36	0.36	1.00	0.47	0.17
$^{10}B(n,a)^7Li$	24	3.30	1.70	0.20	0.16[a]	0.04[a]	1.30	0.10	0.30

[a]Estimated.

Source: Hart, E. J.; Boag, J. W. *Nuclear Chemistry – Theory and Applications*; Pergamon Press: Oxford, 1980; Chapter 15, pp 323–324.

2.29.3.1 High Temperature Water – LWRs

The major application here is the primary coolant circuit of LWRs (i.e., normal, as opposed to heavy, water), of which there are two principal types: boiling water reactor (BWR) and pressurized water reactor (PWR), differentiated by whether steam is raised directly in the core or indirectly via separate steam generators. The water temperature varies a little from reactor to reactor, but is ~300 °C. High-purity water is used in BWRs, whereas in PWRs, additions of boric acid are made to control the reactivity of the fuel as it burns (i.e., undergoes fission) with the acidity counteracted by lithium hydroxide.

Various corrosion processes affect reactor operations, the principal ones being

- fuel cladding corrosion, leading to the formation of adherent oxides, which impact on heat transfer;
- general corrosion of stainless steels and inconels, the latter are used in PWR steam generators, leading to the formation of deposits (termed cruds), which can impact on heat transfer rates and also pose a dose uptake issue in areas that require maintenance, since the corrosion products become radioactive as a result of the ingrowth of various isotopes; and
- stress corrosion cracking (SCC), intergranular or transgranular, of structural stainless steels and inconels.

All of the above are affected by water radiolysis. Corrosion in LWRs arguably constitutes the most important application of water radiolysis, and consequently, LWR chemistry has been extensively studied, and continues to be so as new problems emerge and new reactor designs are being considered. The interested reader is directed to the proceedings of recent international conferences on the subject.[16,17]

Within the core of a BWR operating under normal water chemistry (i.e., deoxygenated pure water), water radiolysis results in oxidizing conditions arising from the greater radiolytic yield of hydrogen peroxide compared with hydrogen. Complex models considering the various radiolytic and electrochemical processes and the flow conditions, coupled with direct measurement in a number of instances, indicate that the free corrosion potential of stainless steels is elevated significantly by as much as 250 mV compared with out-of-core exposure.[18,19] This renders 304 stainless steel susceptible to SCC, particularly if sensitized; if not already present as a result of welding operations, sensitization tends to grow in as a result of neutron damage. Irradiation-assisted SCC is therefore influenced both by microstructural changes and water radiolysis.[8–10] To counteract the radiolytic oxidizing effect, hydrogen–water chemistry has been introduced to BWRs. Figure 1 shows that small concentrations of hydrogen, of the order of 5 cm^3 of H_2 per kg, act to suppress almost completely the radiolytic production of oxidizers in the bulk water.[20] In PWRs, a hydrogen overpressure is similarly used, thus ensuring that conditions within the core are generally sufficiently reducing to prevent the attainment of potentials where SCC of stainless steel can occur, −230 mV_{NHE} being accepted commonly as the safe limit.

The effect of crevices (i.e., occluded regions) requires careful consideration, since local electrochemical conditions within the crevice can differ significantly from the bulk. Complex models considering the rates of radiolytic reactions, transport processes, fluid flow, and electrochemical reactions are

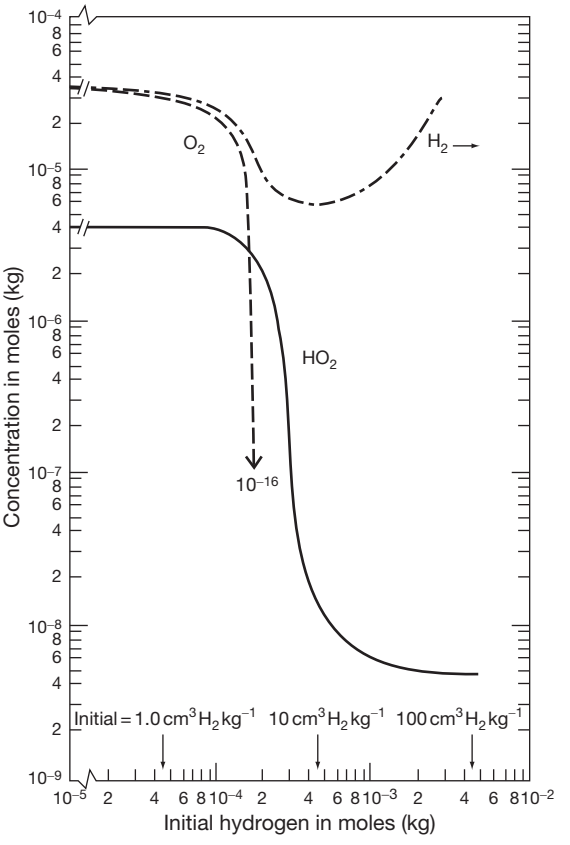


Figure 1 Computed steady-state concentrations of the indicated radiolytic products in PWR conditions as a function of initial hydrogen concentration. Reproduced from Solomon, Y. In *Water Chemistry of Nuclear Reactor Systems*, Bournemouth, UK, 24–27 October 1978; BNES, London, UK, 1979; pp 101–109, with permission from BNES.

used to estimate the potential distribution, since *in situ* direct measurement is not possible. Such models indicate that higher bulk hydrogen concentrations, that is, 20–25 cm^3 of H_2 per kg, are needed to ensure that the corrosion potential at the tip of long crevices does not exceed −230 mV_{NHE}; still higher hydrogen concentrations, that is, 50 cm^3 of H_2 per kg, are required to prevent significant acidification at the crevice tip.[21]

High hydrogen concentrations and slightly alkaline conditions also act to reduce the general corrosion rate of stainless steels and inconels, thus reducing activity build-up in the coolant circuit. However, increasing hydrogen concentration appears to increase the SCC susceptibility of inconel alloy 600. There is, thus, a drive to reduce hydrogen concentrations in PWRs slightly in order to limit SCC of alloy 600, while not compromising on either SCC control of stainless steels or crud deposition due to general corrosion within the reactor circuit.[22]

2.29.3.2 Chemical Plant Heating/Cooling Waters and Related Environments

Stainless steels, almost exclusively austenitic grades, are widely used in the construction of nuclear reprocessing plant and ancillary items associated with the storage and disposal of radioactive wastes. In these applications, stainless steels are exposed to a variety of environments, including irradiated waters containing low to moderately high concentrations of chloride at ambient to moderately high temperatures. A strong susceptibility to localized corrosion propagating in the form of pitting corrosion has been noted in some circumstances where such tendency would otherwise not be expected, that is, at <10 mg l^{-1} Cl^- and temperatures as low as 30 °C.

All localized corrosion processes consist of discrete initiation and propagation phases, both of which are influenced by the adopted corrosion potential, and hence the redox potential exerted by the environment. **Table 2** shows that the corrosion potential of passive stainless steel in water is strongly elevated by exposure to γ-radiation, but that the corrosion potential is insensitive to the dose rate.[23] This effect is known to be attributable to water radiolysis rather than subtle modification of the metal's passive film, since the corrosion potential shift is reproduced in nonirradiated stainless steels exposed to previously irradiated water. In addition, in contradiction to early interpretation of the radiolytic corrosive effect,[24] it is clear that the observed oxidizing

Table 2 Effect of γ-radiation on the corrosion potential of 304L stainless steel in 10 mg l^{-1} Cl^- at 30 °C

Dose rate (Gy h^{-1})	*Potential (mV_{SCE})*
0	−150 to 80
10	250–300
200	250–300
1000–2000	260–380

Source: Ferreira, M. G. S.; Simoes, A. M. P. Eds.; *Mater. Sci. Forum* **1995**, *192–194*, 469–476; Work carried out by Marsh, G. P.; and Taylor, K. J. AERE Harwell, under contract to BNFL, reported by Herbert, D.; Whillock, G. O. H.; Worthington, S. E. In *Electrochemical Methods in Corrosion Research V*; Proceedings of the 5th International Symposium, Sesimbra, Portugal, September 5–8, 1994.

effect is due to stable molecular products, rather than the short-lived intermediates, since the latter do not persist outside of the directly irradiated region. Chemical additions of hydrogen peroxide to nonirradiated waters, in fact, produce essentially the same effect on the corrosion potential, including independence on concentration, as illustrated in **Figure 2**.[23] The behavior of hydrogen peroxide in fixing the potential of (passive) stainless steel is consistent with mechanistic studies, which indicate that the potential of inert (i.e., noble or passive) surfaces is determined by the discharge of hydroxyl radicals, OH, resulting from the decomposition of hydrogen peroxide; the hydrogen peroxide concentration in solution is unimportant, provided it is sufficient to give full surface coverage of adsorbed OH.[25] **Figure 2** indicates that for stainless steel, the critical hydrogen peroxide concentration is less than 0.5 mg l^{-1}. This mechanism also explains the reported independence of stainless steel corrosion potential on ozone concentration,[26] since OH radicals also arise from the decomposition of ozone in water.[27]

At Cl^- concentrations of the order of 10 mg l^{-1} and temperatures up to 80 °C at least, the elevation of corrosion potential promoted by water radiolysis (or hydrogen peroxide additions) is not, however, sufficient to initiate true pitting corrosion, since the pitting breakdown potential is far too high. The breakdown potential of commercial grades of stainless steel is, in fact, elevated a little, on average, by ionizing radiations due to the elimination of low breakdown potentials from the distribution.[24] The origin of this effect is not known for certain; it may possibly reflect slight passive film thickening or the removal of surface-breaking inclusions, which might otherwise act as pitting sites, or it might somehow be related to the photoinhibition effect.[28] However, if a suitable crevice

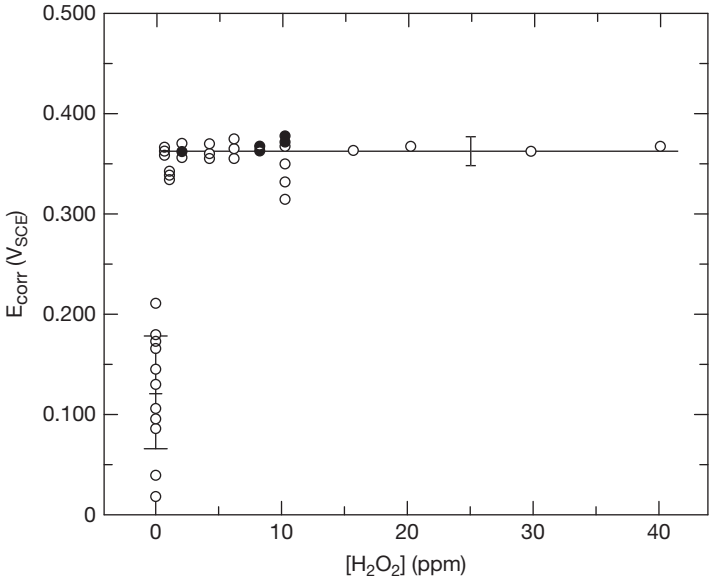


Figure 2 Effect of hydrogen peroxide on the free corrosion potential of 304L stainless steel in 300 mg L^{-1} Cl^- solution at ambient temperature. Adapted from Ferreira, M. G. S.; Simoes, A. M. P. Eds.; *Mater. Sci. Forum* **1995**, *192–194*, 469–476; Herbert, D.; Whillock, G. O. H.; Worthington, S. E. In *Electrochemical Methods in Corrosion Research V*; Proceedings of the 5th International Symposium, Sesimbra, Portugal, September 5–8, 1994.

is present, the elevation in corrosion potential promoted by water radiolysis is sufficient to initiate breakdown within the crevice. Table 3 demonstrates the effect of γ-radiation on the ease of crevice initiation. These data refer to engineered crevices manufactured from 304L stainless steel, tested in water containing 10 mg l^{-1} Cl^- at 30 °C.[23] Although the dataset is incomplete, hampering unequivocal interpretation, a wealth of other data collected on different stainless steels at different temperatures and dose rates confirms a dose rate effect. The dependence on dose rate, despite the independence of the initial potential elevation, is due to the metastable nature of crevice initiation; without sufficient maintenance of the potential provided by the cathode, an incipient crevice can be switched off by the potential fall accompanying the burst of anodic dissolution.

In the propagation phase of crevice/pitting corrosion, the additional cathodic reaction provided by the reduction of hydrogen peroxide, generated either radiolytically or due to deliberate addition, increases the propagation rate. This is illustrated in Figure 3 by recording the current flow from a large passive stainless steel surface to an engineered crevice in the presence and absence of γ-radiation.

Once initiated, crevice corrosion propagates in essentially the same manner as pitting corrosion. However, the existence of the crevice is crucial; without it, the deleterious effect of the ionizing radiation on stainless steel exposed to low Cl^- concentrations does not happen. Experience indicates that even very minor defects such as a slight rolling lap or microporosity at a weld, which would not be expected to be detectable at the manufacturing stage, can act as the required crevice, provided the radiation dose rate is high enough. Figure 4 presents an example. At lower radiation dose rates, a more

Table 3 Effect of γ-radiation on the ease of crevice initiation: 304L stainless steel engineered crevices, 10 mg l^{-1} Cl^- at 30 °C

Dose rate (Gy h^{-1})	*Initiation frequency*[a]
0	2/16
2.8	1/2
25	1/2
200	2/2
1000	3/3
1600	1/1

[a]n/m indicates *n* corroded out of *m* tested.
Source: Ferreira, M. G. S.; Simoes, A. M. P. Eds.; *Mater. Sci. Forum* **1995**, *192–194*, 469–476; Work carried out by Marsh, G. P.; and Taylor, K. J. AERE Harwell, under contract to BNFL, reported by Herbert, D.; Whillock, G. O. H.; Worthington, S. E. In *Electrochemical Methods in Corrosion Research V*; Proceedings of the 5th International Symposium, Sesimbra, Portugal, September 5–8, 1994.

severe crevice would be expected to be required to initiate sustained propagation (all other factors such as Cl^- concentration and temperature being equal). There are insufficient data and the interactions are too complex to be definitive as to whether there is a dose rate threshold, below which crevice effects cease to be of practical significance. However, as noted earlier, the transition from initiation to stable propagation requires a vigorous cathode, otherwise, the potential drops too far and the corrosion switches off.

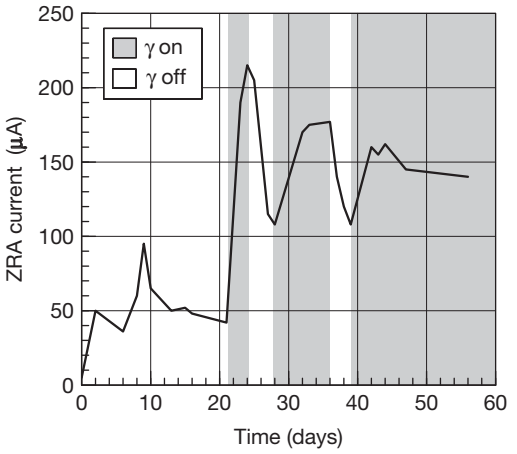


Figure 3 Effect of γ-radiation (1 kGy h^{-1}) on the current flowing to an AISI 347 crevice in 100 mg L^{-1} Cl^- solution at ambient temperature. Reproduced from Ferreira, M. G. S., Simoes, A. M. P., Eds.; *Mater. Sci. Forum* **1995**, *192–194*, 469–476; Work carried out by Marsh, G. P.; and Taylor, K. J. AERE Harwell, under contract to BNFL, reported by Herbert, D.; Whillock, G. O. H.; Worthington, S. E. In *Electrochemical Methods in Corrosion Research V*; Proceedings of the 5th International Symposium, Sesimbra, Portugal, September 5–8, 1994.

It is in fact the case that all examples, known to the author, of a clear effect on crevice initiation/propagation come from high dose rate environments (>100 Gy h^{-1}).

Mitigation options are essentially the same as those that pertain to nonirradiated situations. Reducing the Cl^- concentration and temperature have strong beneficial effects in reducing the likelihood of crevice initiation; the former has a secondary effect on radiolyzed waters if reduced low enough, since the radical scavenging action of Cl^- is eliminated. Inhibitors such as hydroxide and nitrate can also be effective if applied at the outset, but for the latter, at least, consideration needs to be given to its radiation stability (which also precludes organic inhibitors from serious consideration). The selection of a more resistant alloy, such as 316L stainless steel, would also eliminate irradiation-assisted crevice corrosion except at high Cl^- concentration.

Although the earlier discussion focuses on stainless steels, there is no reason to suppose that similar radiation effects would not be manifest in other alloy systems where the elevation of the corrosion potential by a few hundred millivolts is required to initiate crevicing. However, there appear to be no reported occurrences owing, presumably, to an absence of industrial applications. Likewise, a deleterious effect of radiolysis on other localized corrosion mechanisms, notably SCC and corrosion fatigue, would be expected, but appears not to be reported except in high temperature, high-purity waters (see **Section 2.29.3.1**) and for the pond-storage of some types of spent nuclear fuel where the stainless steel cladding is highly sensitized as a consequence of in-reactor damage.[29]

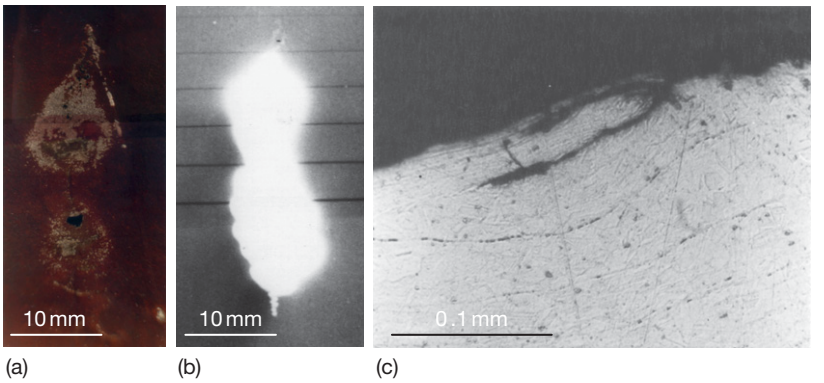


Figure 4 Corrosion pit resulting from crevice corrosion at a rolling lap on an AISI 321 stainless steel plate exposed for 13 Gy to irradiated water (~500 Gy h^{-1}) containing ~10 mg l^{-1} Cl^- at ~40 °C: (a) external appearance (the bright scratch-like feature is the initiating defect); (b) radiograph showing the extent of pit growth; (c) cross-section through the initiating defect. The plate is question was only 4.7 mm thick and the pit did not penetrate to the outside; instead, a relatively large exit hole formed on the same side as the corrosion started from about half way down the pit.

2.29.3.3 Nitric Acid Solutions

Nitric acid solutions are used in nuclear fuel reprocessing, exploiting the propensity of uranium and plutonium to form nitrate complexes that dissolve in nucleophilic solvents such as tri-*n*-butyl phosphate. Stainless steels, mainly austenitic grades, are commonly used in the construction of reprocessing plant and are either passive or suffer intergranular corrosion (leading to grain dropping, and hence general wastage rather than cracking), depending on the oxidizing power of the solution, which is strongly increased at high temperatures (>70 °C) and in the presence of certain dissolved species, which act as oxidants. A more detailed exposition is given elsewhere (see **Chapter 2.24, Corrosion in Nitric Acid**).

Radiolysis of aqueous nitric acid, and indeed, neutral nitrate solutions, produces the nitrite ion as the principal dissolved product,[30] along with varying amounts of hydrogen peroxide, depending on the acidity, since nitrous acid is oxidized by hydrogen peroxide (at least, at room temperature; in hot nitric acid solutions, hydrogen peroxide decomposes with the production of NO_x gases, indicating the reduction of the nitrate ion, that is, reversal of the redox couple). The reaction schemes are complex, involving various linked chemical reactions as well as radiolytic ones, and the yields depend strongly on the LET of the radiation.[30]

The presence of significant concentrations of nitrous acid modifies the corrosive behavior of nitric acid considerably, the overall effect depending on the circumstances. In pure aqueous nitric acid solutions, nitrous acid catalyzes nitrate reduction and, therefore, acts to elevate the corrosion potential of stainless steel, increasing its corrosion rate a little if the temperature is high enough to support intergranular corrosion, even though the redox potential of the solution falls, since nitrous acid is less oxidizing than nitric acid (see **Figure 12** in **Chapter 2.24, Corrosion in Nitric Acid**). In more complex liquors containing certain dissolved species such as Cr(VI) and Ce(IV), the effect of nitrous acid production can be profound, sharply reducing the corrosion rate if a complete reduction to Cr(III) and Ce(III) occurs, since the corrosion potential moves back toward passivity.[31] In such liquors, the corrosion potential and hence the corrosion rate attained depends on the overall redox balance; whether oxidation by the acid or reduction by irradiation predominates depends on the nitric acid concentration and temperature, the increase of either of which favors oxidation, and radiation dose rate, the increase of which favors reduction.

Owing to the fact that the corrosion resistance of metals such as zirconium and tantalum is unaffected by oxidizing species in solution, no effect of radiolysis is apparent. No effect of radiolysis on the corrosion resistance of titanium has been reported, although a deleterious effect might be expected if reliance were placed on oxidizing ions, rather than dissolved titanium ions, for the maintenance of passivity (see **Chapter 2.24, Corrosion in Nitric Acid**).

2.29.3.4 Polymer Degradation Products

Polymeric materials are susceptible to radiolytic degradation and some, notably halogenated ones, release species capable of exerting strong corrosive effects in some circumstances. For example, the radiolysis of polyvinyl chloride (PVC) produces hydrogen and chlorine free radicals as primary products, which combine producing HCl, Cl_2, and H_2.[32] Likewise, fluoropolymers produce F_2 and HF (if either hydrogen is present within the polymer or water is present) as potential corrodants. In general, owing to the fact that the release rates are relatively low, significant corrosive effects arise only in situations where the corrodants are able to accumulate. For example, the use of PVC tape to attach temporary signage during the construction of stainless steel plant, which will subsequently be irradiated, is usually banned owing to concerns over the possibility of chloride-induced SCC.

2.29.4 The Effect of Radiolysis on the Corrosion of Metals in Atmospheric Environments

In addition to hydrogen peroxide production due to radiolysis of condensed water films, atmospheric corrosion processes may be affected by the production of ozone, nitrogen oxides, and nitric acid from radiolysis of air. In air, most of the observed effects are attributable to nitric acid production, hence, the corrosion rate of metals, which are attacked by dilute nitric acid (e.g., iron, copper, lead, aluminum, zinc, nickel), is increased, whereas metals passivated by nitric acid, such as stainless steels (but see later text), are unaffected.[33] The radiolytic reaction scheme is complex, but nitric acid results from interactions such as

$$O + NO \rightarrow NO_2 \qquad [5a]$$

$$NO_2 + OH \rightarrow HNO_3 \qquad [5b]$$

Nitric acid is produced only if water, nitrogen, and oxygen are simultaneously present. Hence, no corrosion acceleration results in irradiated dry air[33,34] or in irradiated moist oxygen-free nitrogen.[33] Radiation shifts the critical relative humidity required to support atmospheric corrosion to much lower value, 10–20%, compared with 60–70% in the absence of radiation.[35] In general, increasing dose rate increases the corrosive effect, although at very high dose rates, surface temperatures can increase sufficiently to prevent the formation of condensed water films, in which case, no corrosion results.[33] It is also reported that there is a lower dose rate threshold ($\sim$50 Gy h^{-1}), below which negligible corrosive effect results on iron,[33] although the evidence presented to support that assertion appears weak.

Subtle effects arising from the radiolysis of air have been noted, such as an inhibiting effect (due to the nitrate ion) on the initiation of crevice corrosion in stainless steels tested in low Cl^- water. However, the principal application relates to potential, deleterious effects during the storage and disposal of spent nuclear fuels and wastes.[36] Stainless steel fuel cladding, if severely sensitized as a consequence of neutron damage, is known to be susceptible to intergranular attack in moist air; the corrosion mechanism appears to be SCC, constituting the only practical example, known to the author, of the occurrence of this failure mechanism in the stainless steel–nitric acid system. In the absence of severe sensitization (i.e., in excess of that produced by normal welding operations), no corrosion occurs.

2.29.5 Corrosion Test Considerations

Careful consideration needs to be given to the design of tests employing radiation if representative radiolytic effects are to be adequately produced. This includes the selection of the dose rate and the type of radiation, since the steady-state radiolytic yield of hydrogen peroxide (which is the principal active agent modifying aqueous corrosion) depends on both. However, it appears to be generally the case that sufficiently representative effects are produced by radiation dose rates of the same order of magnitude as those that pertain in the environment being simulated. Likewise, it is the case that different types of radiation generally produce similar radiolytic effects, although the exact effect may be more or less pronounced depending on LET values.

Consideration also needs to be given to the selection of suitably radiation-resistant materials for the construction of the test apparatus to avoid inadvertent contamination of test liquors by released polymeric degradation products or, in the case of electrochemical tests, to avoid problems arising from the breakdown of electrical insulation. Where necessary, local shielding can be employed to protect items that cannot be made intrinsically more robust. With regard to electrochemical testing, it appears to be the case that γ-radiolysis exerts negligible effect on the stability of saturated calomel or silver/silver chloride reference electrodes.

References

1. Choppin, G. R.; Rydberg, J. *Nuclear Chemistry – Theory and Applications*; Pergamon Press: Oxford, 1980; Chapter 14, pp 265–307.
2. Choppin, G. R.; Rydberg, J. *Nuclear Chemistry – Theory and Applications*; Pergamon Press: Oxford, 1980; Chapter 15, pp 308–331.
3. Gorse, D.; Rondot, B.; da Cunha Belo, M. *Corr. Sci.* **1990**, *30*, 23–36.
4. Byalobzheskii, A. V. *Radiation Corrosion*; Israel Program for Scientific Translations: Jerusalem, 1970; pp 79–93.
5. Byalobzheskii, A. V. *Radiation Corrosion*; Israel Program for Scientific Translations: Jerusalem, 1970; pp 70–79.
6. Simnad, M. T. In *The Effects of Radiation on Materials*; Harwood, J. J. Ed.; Reinhold: New York, 1958; pp 129–133.
7. Norris, D. I. R. Ed. Radiation-Induced Sensitisation of Stainless Steels, Proceedings of the Symposium held at Berkeley Nuclear Laboratories on 23 September 1986, CEGB.
8. Andresen, P. L. In *Stress Corrosion Cracking – Material Performance and Evaluation*; Jones, R. H. Ed.; ASM International: Materials Park, OH, 1992; pp 181–210.
9. Scott, P. *J. Nucl. Mat.* **1994**, *211*, 101–122.
10. Was, G. S. In *Radiation Effects in Solids*; Sickafus, K. E., Kotomin, E. A., Uberuaga, B. P., Eds.; Proceedings of the NATO Advanced Study Institute on Radiation Effects in Solids, held in Erice, Sicily, Italy, 17–29 July 2004; Springer: Netherlands, 2007; pp 421–447.
11. Byalobzheskii, A. V. *Radiation Corrosion*; Israel Program for Scientific Translations: Jerusalem, 1970; pp 11–15.
12. Spinks, J. W. T.; Woods, R. J. *An Introduction to Radiation Chemistry*, 3rd ed.; John Wiley & Sons: New York, 1990.
13. Burns, W. G.; Moore, P. B. *Rad. Effects* **1976**, *30*, 233–242.
14. Hart, E. J.; Boag, J. W. Reported by Choppin, G. R.; Rydberg, J. *Nuclear Chemistry – Theory and Applications*; Pergamon Press: Oxford, 1980; Chapter 15, pp 323–324.
15. Airey, P. L. *Radiat. Res. Rev.* **1973**, *5*, 341–371.
16. Swann, T. Ed. Water Chemistry of Nuclear Reactor Systems 7: Proceedings of the Conference Organized by the British Nuclear Energy Society and Held in Bournemouth on 13–17 October 1996; Thomas Telford: London, 1998.
17. 5th International Workshop on LWR Coolant Water Radiolysis and Electrochemistry, San Francisco, 15 October 2004.
18. Indig, M. E.; Nelson, J. L.; Wozadlo, G. P. Proceedings of the 5th International Symposium on Environmental Degradation of Materials in Nuclear Power Systems – Water Reactors, Monteray, California, 25–29 August 1991;

Cubicciotti, D., Simonen, E. P., Gold, R., Eds.; ANS: La Grange Park, IL, 1992; pp 941–952.
19. Indig, M. E.; Nelson, J. L. *Corrosion* **1991**, *47*, 202–209.
20. Solomon, Y. *Water Chemistry of Nuclear Reactor Systems,* Bournemouth, UK, 24–27 October 1978, BNES, London, UK, 1979; pp 101–109.
21. Vankeerberghen, M.; Weyns, G.; Gavrilov, S.; Henshaw, J.; Deconinck, J. *J. Nucl. Mater.* **2009**, *385*(3), 517–526.
22. Feron, D.; Corbel, C.; Hickel, B. 5th International Workshop on LWR Coolant Water Radiolysis and Electrochemistry, San Francisco; 15 October 2004.
23. Herbert, D.; Whillock, G. O. H.; Worthington, S. E. In *Electrochemical Methods in Corrosion Research V*, Proceedings of the 5th International Symposium, Sesimbra, Portugal, September 5–8, 1994; Ferreira, M. G. S., Simoes, A. M. P. Eds. *Mater. Sci. Forum* **1995**, *192–194*, 469–476.
24. Marsh, G. P.; Taylor, K. J.; Bryan, G.; Worthington, S. E. *Corr. Sci.* **1986**, *26*, 971–982.
25. Bockris, J. O'M.; Oldfield, L. F. *Trans. Faraday Soc.* **1955**, *51*, 249–259.
26. Lu, H. H.; Duquette, D. J. *Corrosion* **1990**, *46*, 843–852.
27. Rice, R. G.; Wilkes, J. F. Fundamental Aspects of Ozone Chemistry in Recirculating Cooling Water Systems, paper 205 in Corrosion '91, 11–15 March 1991, Cincinnati, NACE, Houston, TX, 1991.
28. Macdonald, D. D.; Heaney, D. F. *Corr. Sci.* **2000**, *42*, 1779–1799.
29. Hands, B. J. *Conference on Materials and Nuclear Power,* Proceedings of EuroMat 96, Bournemouth, UK, 21–23 October, 1996; Institute of Materials: London, 1997; pp 459–468.
30. Kazanjian, A. R.; Miner, F. J.; Brown, A. K.; Hagan, P. G.; Berry, J. W. *Trans. Faraday Soc.* **1970**, *66*, 2192–2198.
31. McIntosh, A. B.; Evans, T. E. *Peaceful Uses of Atomic Energy,* Proceedings of the 2nd International Conference, Geneva, Switzerland, September, 1958; UN: Geneva, Switzerland, 1958; pp 206–225.
32. Clough, R. L.; Gillen, K. T. *Radiat. Phys. Chem.* **1981**, *18*, 661–669.
33. Byalobzheskii, A. V. *Radiation Corrosion*; Israel Program for Scientific Translations: Jerusalem, 1970; Chapter 5, pp 141–162.
34. Shatalov, I. V.; Nikitina, V. A. *Peaceful Uses of Atomic Energy,* Proceedings of the 2nd International Conference, Geneva, Switzerland, September, 1958; UN: Geneva, Switzerland, 1958; pp 284–296.
35. Vernon, W. H. J. *Trans. Faraday Soc.* **1935**, *31*, 1668–1700.
36. Reed, D. T.; Van Konynenburg, R. A. *Effect of ionizing radiation on moist air systems, Lawrence Livermore Laboratory report UCRL 97936*, 1987.

2.30 Electrochemical Methods

R. A. Cottis
University of Manchester, Manchester, UK

Abbreviations

ADC Analog to digital converter
CE Counter electrode
CSE Copper sulfate reference electrode
EIS Electrochemical impedance spectroscopy
EN Electrochemical noise
IR Voltage produced by current *I* passing through resistance *R*
K–K Kramers-Kronig
OCPD Open circuit potential decay
PSD Power spectral density
RE Reference electrode
SCE Saturated calomel reference electrode
WE Working electrode

Symbols

A Area (m^2), gain of operational amplifier
b Bandwidth of measurement (s^{-1})
B Stern-Geary coefficient (V)
C Capacitance (F)
C_{dl} Double layer capacitance (F m^{-2} or F)
d Separation between conductors (m)
e_- Inverting input to operational amplifier
e_+ Non-inverting input to operational amplifier
e_o Output from operational amplifier
E_n Electrochemical potential noise (V)
f Frequency (s^{-1})
f_n Characteristic frequency (s^{-1})
G Amplifier gain
i Current density (A m^{-2})
i_{corr} Corrosion current density (A m^{-2})
I_n Electrochemical current noise (A)
I Current (A)
L Inductance (H)
q Characteristic charge (C)
Q Charge (C)
R Resistance (Ω)
R_n Electrochemical noise resistance (Ω m^2 or Ω)
R_p Polarization resistance (Ω m^2 or Ω)
R_s Solution resistance (Ω)
V Voltage (V)
V_C Control voltage (V)
V_{CE} Voltage of counter electrode (V)
V_i Input voltage (V)
V_o Output voltage (V)
V_{RE} Voltage of reference electrode (V)
Z Impedance (complex)
ΔE Small change in potential (V)
Δi Small change in current density (A m^{-2})
ε Relative permittivity
ε_0 Permittivity of free space ($8.854187817\times10^{-12}$ F m^{-1})
ρ Resistivity (Ω m)
σ_E Standard deviation of potential (V)

σ_I Standard deviation of current (A)
Ψ_E Power spectral density of potential (V^2 Hz^{-1})
Ψ_I Power spectral density of current (A^2 Hz^{-1})
ω Angular frequency (s^{-1})

2.30.1 Introduction

This chapter commences by providing a brief summary of the basics of electricity and electrical measurements. It then considers the various measurements involved in the study of electrochemistry and how these can be applied to the study of corrosion. Finally the design of test electrodes and electrochemical cells is discussed. For a more comprehensive introduction to electrochemical methods for corrosion studies, the readers are referred to Kelly *et al.*[1]

2.30.2 Basics of Electricity

2.30.2.1 Charge

Charge is the most fundamental electrical property. It derives from the charge inherent in the electrons and protons that make up the atoms and molecules of metals, solutions, and gases that we are concerned with. The unit of charge is the Coulomb, named after Charles Augustin de Coulomb, with the symbol C. The charge on the electron is $1.602176487 \times 10^{-19}$ C, and therefore, one coulomb corresponds to 6.24151×10^{18} electrons (strictly, the charge on the electron is negative, and therefore, that should really be −1 C).

2.30.2.2 Current

Current is the flow of charge and a current of one Ampere (named after André Marie Ampère, and usually abbreviated to Amp) corresponds to 1 C passing a given point in an electrical circuit in 1 s. Currents in metals invariably consist of movement of the mobile valence electrons. The movement is surprisingly slow (typically of the order of 10^{-4} m s^{-1}), since there are a lot of valence electrons available. Note that the electrons move in the opposite direction to the 'conventional current,' which is treated as being due to the movement of positive charges. Currents in solutions consist of the flow of ions, both positively and negatively charged, with positive ions moving in the direction of the conventional current and negative ions moving in the opposite direction.

2.30.2.3 Current Density

In electrochemical systems in which reactions are occurring at the same rate over the entire electrode surface, it is useful to normalize measured currents by dividing by the area to give the current density (units A m^{-2}). In this chapter, we use the convention that I stands for current, while i stands for current density. Note that it is not always reasonable to quote current densities; for example, in electrochemical noise measurements (see the following section), current is not expected to be proportional to area, and hence current density is area-dependent. Similarly, it is misleading to refer to current density in relation to pitting or crevice corrosion (except when referring to true current densities, such as the local current density at the corroding surface within the pit or crevice, or the cathodic current density at the external surface).

2.30.2.4 Potential

The potential at a point is defined as the work done in moving unit positive charge to that point from infinity. It is clearly a theoretical concept rather than anything that can really be used! The unit of potential is volt (named after Alessandro Volta), abbreviated as V, and the potential at a point is 1 V if 1 W is required to move a positive charge of 1 C to that point.

2.30.2.5 Potential Difference and Voltage

The potential difference or voltage between two points is the difference between the potentials at the two points. The basic principles of thermodynamics imply that this is identical to the work done in moving unit charge from one point to the other (which we can measure, at least in principle).

2.30.2.6 Resistance and Resistivity

It is commonly found that the voltage across a component is proportional to the current passing through the component. This behavior is known as Ohm's Law (named after Georg Simon Ohm), given by the well-known formula $V = IR$, where V is the voltage (V), I the current (A), and R is the conventional symbol for the *resistance* of the component, with the unit Ohm and the symbol Ω. It should be appreciated that, like many laws, Ohm's Law is an

approximation to reality, and many components (including semiconductors and metal–solution interfaces) do not exhibit a linear relationship between current and potential. When Ohm's Law is not obeyed, we can still talk about the resistance of a component, defined as the slope of the voltage-versus-current curve, but the value of the resistance will vary according to the applied voltage or current.

The resistance of a uniform conductive solid decreases as the cross-sectional area increases and as the length decreases. We can define the *resistivity* of the solid as the resistance across opposite faces of a unit cube of the solid. Then the resistance of a section of wire will be $l\rho/A$, where l is the length of the wire (unit m), A the cross-sectional area (m^2), and ρ the *resistivity* (Ω m).

2.30.2.7 Conductivity and Conductance

In some situations, it is convenient to work in terms of the conductance of a component or the conductivity of a metal or solution. These are simply the inverse of the resistance and resistivity respectively. The unit of conductance is Ω^{-1}, also known as Siemens (after Ernst Werner von Siemens), with the symbol S, while conductivity has the unit $\Omega^{-1}m^{-1}$ or Sm^{-1}.

A particular application of conductance is found in the relationship of the conductivity of a solution to the composition. For dilute solutions, each ion contributes a specific amount to the conductivity, and therefore, the conductivity can be derived by summing the conductivities due to the individual ions making up the solution (see **Chapter 2.38, Modeling of Aqueous Corrosion** for further details).

2.30.2.8 Capacitance

If two conductors are placed close together, a change in the voltage between the conductors will lead to a separation of charge, and the amount of charge will be proportional to the voltage. This leads to the concept of a capacitor, a device that stores charge, the governing equation being $Q = CV$, where Q is the stored charge (C), V the voltage across the capacitor (V), and C the capacitance (it is unfortunate but inescapable that C is commonly used as both the symbol for capacitance and for the unit of charge). The unit of capacitance is Farads with the symbol F (after Michael Faraday), and a 1 F capacitor will store 1 C when 1 V is applied across it.

Note that if we differentiate the governing equation for a capacitor with respect to time, we obtain

$$\frac{dQ}{dt} = I = C\frac{dV}{dt}$$

that is, the current is proportional to the rate of change of voltage across the capacitor.

One of the simplest forms of capacitor consists of two parallel plates of conductor separated by an insulator, known as a 'parallel plate capacitor.' In this case, the capacitance is defined by $C = A\varepsilon\varepsilon_0/d$, where A is the area (m^2), ε is the relative permittivity of the insulator (dimensionless), ε_0 is the permittivity of free space ($8.854187817 \times 10^{-12}$ F/m), and d is the separation between the conductors (m) (strictly this formula is approximate, because it ignores edge effects).

An example of a parallel plate capacitor that is found in electrochemical systems is metals covered with insulating films immersed in solution. In this case, the metal and the solution form the two plates, while the coating provides the layer of insulator.

In the absence of an insulator, there is also a capacitance between a metal and a solution, known as the double-layer capacitance (see **Chapter 1.02, Electrochemistry** for further information).

2.30.2.9 Inductance

As current flows along a conductor, it induces a magnetic field. When the current changes, this magnetic field must also change and a voltage is required to cause this change. The inductance, L (unit Henry, symbol H, named after Joseph Henry), of a component is defined by

$$V = L\frac{dI}{dt}$$

While a capacitor can be thought of as storing charge, an inductor effectively stores current (i.e., once a current is flowing, the inductor tries to keep it flowing). This leads to problems when switching inductive loads, such as motors, as the inductor will develop a large voltage as it tries to keep the current flowing through the switch. This can lead to arcing and damage to the switch contact, together with the production of high-frequency electromagnetic radiation that can be picked up as interference.

2.30.2.10 Impedance

Resistance, capacitance, and inductance all represent the response of a component to applied voltage or current, and they can be combined into a single property, the impedance of the component. We can consider a simple example of impedance by thinking

about the properties of a paint-coated metal immersed in solution. An ideal paint will be a perfect insulator, but real paints will display some conductivity. Thus, a paint of thickness d and area A, together with the metal and solution either side of it, will exhibit a resistance given by $\rho d/A$, while the capacitance, which will be in parallel with the resistance, will be given by $A\varepsilon\varepsilon_0/d$. Note that the resistance decreases with increasing area, whereas the capacitance increases.

The impedance of a resistor is independent of frequency and has a magnitude that is equal to the resistance, and a phase shift of zero.

The magnitude of the impedance of a capacitor decreases with increasing frequency, being given by $1/\omega C$, where ω is the angular frequency, and the phase shift is $-\pi/2$ radians ($-90°$).

The magnitude of the impedance of an inductor increases with increasing frequency, being given by ωL, and the phase shift is $\pi/2$ radians ($90°$).

2.30.2.11 Admittance

Just as we can consider conductance to be the inverse of resistance, admittance is the inverse of impedance.

2.30.2.12 Measurement Fundamentals

All electrochemical studies involve the measurement and/or control of current and voltage. In this section, the fundamental requirements for accurate measurements are presented. Measurements come in a variety of forms:

- Analog or digital

Analog signals have a continuous range of possible values; almost all real variables (current, voltage, velocity, etc.) are analog. It is essentially impossible to measure the value of an analog variable completely, as this would require an infinite resolution. The term analog, when applied to measurements, therefore implies not that the measurement reproduces the measured property completely, but rather that it does not have clearly defined measurement limits. Thus, a conventional electromechanical chart recorder can be described as analog, in that the pen can adopt any position across the chart and the chart paper nominally moves continuously. In practice, the resolution is limited by many factors, including the size of the pen tip, the grain in the chart paper, noise in the measurement electronics, and steps in the chart motion (it is common for the chart drive to use a stepper motor which moves in small discrete steps).

Digital signals have a well-defined set of possible values. The number of values is frequently defined by the number of bits (binary digits) used to represent all of the possible values (since digital instruments often use binary arithmetic internally). One bit can have two possible values (0 and 1), eight bits can have 256 (2^8) possible values (00000000 to 11111111), and in general n-bit numbers provide 2^n possible values. Common resolutions that are available in measuring systems are indicated in Table 1. Not all instruments use binary arithmetic, and many devices such as multimeters that are primarily intended for human interpretation use decimal arithmetic. In this case, the resolution is defined in terms of the number of decimal digits.

- Continuous or sampled

Besides having a finite set of possible values, real measurements also have limited resolution in time. In the case of analog measurements, this usually corresponds to a drop in sensitivity to rapid changes, while for digital measurements, it corresponds to a finite sampling rate. Note that the process of sampling data has important consequences on their quality. In particular, all signal frequencies at or above half the sampling frequency will be converted to false signals below the sampling frequency by a process known as aliasing, and these must be removed by analog filtering (Section 2.30.2.12.4) before sampling.

Table 1 Common digital measurement resolutions (as is common, for binary measurements 1 bit is assumed to be used to signify the sign of the measured voltage, while for decimal instruments the sign is additional to the number of digits)

Binary		
Number of bits	Number of values	Resolution for ±1 V full scale
8	256	7.8 mV
10	1024	2 mV
12	4096	400 μV
16	65 536	30 μV
20	1 048 576	2 μV
24	16 777 216	120 nV
Decimal		
Number of Digits	Number of values	Resolution for ±1 V full scale
3	1000	1 mV
3.5	1999	500 μV
4	10 000	100 μV
4.5	19 999	50 μV
5	100 000	10 μV
6	1 000 000	1 μV
7	10 000 000	100 nV
8	100 000 000	10 nV

2.30.2.12.1 Time and frequency domains

We experience the world as a sequence of events that follow on in time, and it is natural to think in the 'time domain,' that is, the variation of the measured parameters with time. However, the properties of many systems can be understood and interpreted more easily when expressed as the behavior at different frequencies, that is, in the 'frequency domain' (the brain works in the frequency domain when processing sound – we hear the different frequencies in the sound, rather than the moment to moment fluctuations in air pressure).

2.30.2.12.2 Instrumentation fundamentals

In this section, we shall briefly review the various components that make up modern electronic instruments. This is not intended to be a full coverage of the subject, for which the readers are referred to specialist works, but rather to give a general understanding of the basis.

2.30.2.12.3 Amplifiers

A common requirement in instruments is to change the amplitude of a signal or to convert from a current to a voltage or vice versa. This function is performed by amplifiers. In general, the relationship between the input and the output of an amplifier can be defined by its 'transfer function.' To simplify the discussion, we shall just consider voltage amplifiers (i.e., devices that take an input voltage V_i and produce an output voltage V_o). Then the transfer function of an amplifier can simply be defined by $V_o = G(V_i)$, where G is the gain of the amplifier. We have written G as a function in order to emphasize that G may not be a constant multiplier. Thus the amplifier may be nonlinear over voltage and/or frequency. However, amplifiers are often designed to be linear for voltages and frequencies within their design range, and we can therefore write $V_o = GV_i$, where G is a constant. Thus, an amplifier with a gain of 1000 will convert a 1 mV signal to a 1 V signal.

There is no reason why G must be positive or greater than 1; thus, it is perfectly feasible to design an amplifier with a gain of -0.1, which would convert a 1 mV signal to -0.1 mV.

Other important properties of an amplifier are the input and output impedance. The input impedance is the impedance that the input presents to the circuit to which it is connected – for voltage amplifiers we often want this to be very high so that the connection of the amplifier does not affect the system to which it is connected. The output impedance is an indication of the way in which the output voltage is affected by the impedance of the circuit to which it is connected. We usually want amplifiers to have a low output impedance so that the output voltage is unaffected by the circuit to which it is connected. An important type of amplifier is the 'unity-gain buffer'; this has a gain of 1 (so the output is the same as the input), but the input impedance is very high, and the output impedance very low. This allows us to connect high-impedance voltage sources (such as the potential measured on a painted sample) to low-impedance measuring devices. A practical realization of a unity-gain buffer is described in the following section.

In contrast to voltage amplifiers, a current amplifier takes an input current, and we want it to have a very low input impedance, so that connecting the amplifier does not affect the current, while a current output should have a very high output impedance, so that the current is not affected by the impedance of the system to which it is connected. Note that there is no reason why an amplifier should not have a current input and a voltage output; indeed this is exactly the characteristics of a 'zero-resistance amplifier' (in a later section). Similarly, an amplifier can be designed with a voltage input and a current output, producing a device that provides a current that is controlled by an input voltage (i.e., a voltage-controlled galvanostat).

2.30.2.12.4 Filters

Another important building block of modern instruments is the filter. This is a device that responds differently to different frequencies. The simplest filter consists of a resistor and a capacitor (**Figure 1**). At high frequencies (above $1/2\pi RC$ Hz), the capacitor has a low impedance compared with the resistor, while at low frequencies it has a high impedance. Then a low pass filter (one that allows low frequencies to pass but removes high frequencies) can be obtained by placing the resistor in the signal path and the capacitor across it (**Figure 1(a)**). Swapping the two components produces a high pass filter (**Figure 1(b)**). (The black triangle at the bottom of each figure is a conventional symbol for the earth or ground connection, voltages being assumed to be measured with respect to ground.)

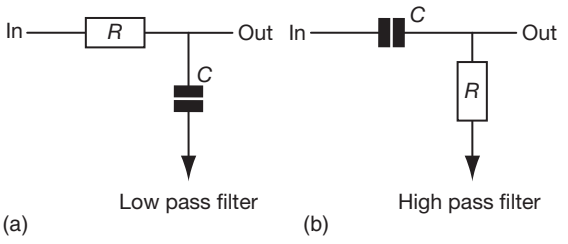


Figure 1 RC filter circuits.

Simple filters using only passive components (resistors, capacitors, or inductors) are rather limited, and many practical filters will use either analog amplifier circuits or digital techniques, which permit a wide range of transfer functions to be produced (see Horowitz and Hill[2] for further information).

2.30.3 Measurement and Control Methods

2.30.3.1 Measurement of Voltage

2.30.3.1.1 Fundamentals

An ideal voltage-measuring instrument will determine the voltage between two points without drawing any current (since this could modify the measurement). In practice, all voltage-measuring instruments will draw a small amount of current, and this is usually defined in terms of the input impedance of the instrument. Many lower-cost multimeters will have an input impedance of 10 MΩ, and if such a meter is used to measure a voltage of 1 V, it will draw a current of 0.1 μA. Whether or not this is acceptable will depend on the system being measured, but for many corrosion applications, the impedance would be too low, and a value of 1 GΩ or above is recommended (an even higher input impedance may be necessary for some systems, such as intact paint-coated metals or passive alloys).

A second consideration in the selection of voltage measurement instruments is the measurement quality required. There are several aspects to this, and we can define three terms with slightly different meaning:

- Resolution or precision is the ability to detect small changes in a large value (e.g., if the meter can measure a maximum of 1 V and can resolve 1 mV the resolution is one part in one thousand).
- Sensitivity is the ability to measure small values, irrespective of the range being used (it is relatively easy to measure changes of 1 mV in a signal with a maximum of 1 V, but it is much more difficult if the maximum is 1 kV).
- Accuracy is the ability to measure the 'true' value (this is often much less well-defined than resolution, as it depends on how well the instrument is maintained and calibrated as well as its initial specification).

An important aspect of voltage measuring devices is the speed with which they can adjust to and measure changes. For 'manual' instruments (i.e., instruments where the only output is viewing by a human, who may or may not choose to write down the result), changes occurring over periods of less than a second cannot be resolved by the observer, and speed is a minor limitation. However, when the output is being recorded, whether by a computer or on a chart recorder, then the measurement speed is also important. For corrosion studies most information tends to be at rather low frequencies (rarely more than 1 kHz), although impedance measurements may go to higher frequencies.

2.30.3.1.2 Single-ended and differential measurements

Voltage-measuring devices may have one input terminal connected to earth, known as a 'single-ended input,' or both input terminals may be independent of ground, known as a 'differential input.' In general, differential inputs are to be preferred, as they help to avoid problems associated with earth loops (see the following section).

2.30.3.1.3 Instruments

A range of instruments is available to measure voltage:

- Moving coil meters are actuated by current but can be adapted to measure voltage by adjusting the resistance of the coil (plus a series resistor if necessary) to give an appropriate calibration. However, they generally have too low an impedance to be useful for corrosion measurements, and they have limited resolution (usually about 1% of full scale).
- Electronic voltmeters use a semiconductor circuit to amplify and convert the voltage to a suitable form for display, either on a moving coil meter or on a digital display (in which case the term digital voltmeter is normally used). Whichever form of display is used, the input impedance is a function of the amplifier used and can be very high (values of 10^{14} Ω are commercially available, usually termed electrometers).
- Chart recorders provide a means of recording the variation of voltage with time. They are gradually becoming less widely used, both because they are relatively expensive, and, unlike computer-based data acquisition systems, the recorded data cannot easily be further processed. As with electronic voltmeters, the input impedance is a function of the electronics used, a value of 10 MΩ being common.
- Oscilloscopes can be thought of as specialized chart recorders that are optimized for viewing high-speed signals. Electrochemical measurements are only rarely concerned with high-speed measurements, and hence oscilloscopes are not widely used,

especially as their input impedance is usually rather low (1–10 MΩ being common). However, they can be useful for checking for high-frequency oscillation or viewing transient behavior (but note that modern computer-based data acquisition systems may provide similar or enhanced functionality for lower cost).

- Computer data-acquisition systems can be thought of as being similar to digital voltmeters, but with the display components being replaced by a connection to the computer, which can take over the display functions (as well as doing many other things). Unfortunately, the majority of computer data acquisition systems are optimized for rapid data acquisition (10^5 or 10^6 samples per second being common), rather than high input impedance or high precision, and consequently they are not ideal for corrosion measurements.

2.30.3.1.4 Earthing and shielding

Induced interference is a common problem in electrical measurements. This arises because the signal circuits act as aerials for the pick-up of electromagnetic radiation. Much of this interference can be eliminated by appropriate earthing (also known as grounding) and screening of the signal circuits. An electrical earth (or ground) is a connection to a stable voltage at approximately the potential of the local soil. It is common to use the safety earth associated with mains power sockets to earth an appropriate part of the measuring circuit. In electrochemical measurements, it is common to earth the working electrode, although there may be circumstances in which this is not desirable. When configuring earth connections, it is very important to avoid an 'earth loop,' which involves two separate connections to earth. This results in a loop of wire, with earth forming part of the loop, and this loop can act as a very effective aerial for the pickup of mains frequency interference.

Interference pickup can be significantly reduced by screening the signal circuits. This involves surrounding the circuits with a metallic conductor that is connected to earth. Electromagnetic radiation cannot penetrate this conductor, and consequently cannot interact with circuits inside the screen. It is common to screen cables by using a coaxial cable, in which the inner conductor is screened by an outer braided screen, which is earthed (note that the comments above about earth loops also apply to the screen connection of coaxial cables). Larger circuits, such as electrochemical cells, can be enclosed in an earthed, conductive box, known as a Faraday cage.

2.30.3.2 Measurement of Current

2.30.3.2.1 Fundamentals

An ideal ammeter will measure the current between the two input terminals while maintaining the voltage between the terminals at zero. In many modern instruments, current is measured by measuring the voltage across a resistor (known as a current shunt) that is inserted in the current path. This results in a small but significant voltage being developed (this is know as the 'voltage burden' and is typically of the order of 100 mV for a full-scale reading to allow the voltage to be measured with an acceptable accuracy), and such meters are unsuitable for many electrochemical measurements. The capabilities of an ammeter can be defined in terms of the input impedance (essentially the resistance across which the voltage is measured) and ideally we want the input impedance to be zero. An alternative current measurement method uses a form of amplifier known as a current amplifier (essentially an amplifier with a current input and a voltage output) that can measure current with a very low input impedance, and this is preferred for corrosion studies. In the corrosion field the term 'zero-resistance ammeter' is used to describe a current amplifier coupled with a suitable voltage measuring device.

2.30.4 Instruments

Many current-measuring methods use similar instruments to those described earlier for voltage measurement. Such instruments normally use the method of measuring the voltage drop across a resistor to measure currents, but even if the instrument offers such a capability internally (as in the case of cheaper digital multimeters), it may be better to use it in voltage measurement mode with an external resistor, as this allows a greater control of the compromise between sensitivity and input impedance (i.e., a smaller resistor, and hence a smaller voltage, may be used, giving a less sensitive measurement, but one that is less affected by the input impedance of the measuring system).

There are three main classes of instrument that can measure current with negligible voltage drop. All of these use essentially the same current amplifier circuit, the only real difference being the packaging and primary function of the device:

- Electrometers normally include a very high input impedance voltage amplifier and a very low input impedance current amplifier capable of measuring

very small currents with a very low voltage burden (typically less than 1 mV).
- Zero-resistance ammeters are normally designed specifically to measure currents in the range of concern for electrochemistry.
- Potentiostats (see Section 2.30.4.3) can be configured to measure current with negligible voltage burden.

A limitation of all devices based on the current amplifier circuit is the limited maximum current that can be measured, as the full measured current must be supplied by an internal amplifier, and once the current exceeds the capability of the amplifier, the current amplifier will behave as a very high resistance.

2.30.4.1 Measurement of Resistance

2.30.4.1.1 Fundamentals

In general, resistance is measured by applying a current to the circuit in question and measuring the voltage developed across it. Then the resistance is computed from Ohm's Law, $V = IR$, making the tacit assumption that the behavior of the component is correctly described by this law. Note that Ohm's Law does not apply to electrochemical interfaces, and somewhat more complex methods must be used.

2.30.4.1.2 Instruments

Most digital multimeters can measure resistance, normally by applying a constant current, which will vary according to the resistance range selected, and measuring the voltage developed. This is generally satisfactory for larger resistances, but for accurate measurements, especially for small resistances, it is necessary to compensate for the resistance of the connecting cables. This is done using a 4-terminal measurement (Figure 2), in which the current is applied to the component using one set of cables, and the voltage is measured using a second set of cables. More sophisticated multimeters provide the separate current and voltage connections that are required for this method. When measuring the resistance of electrochemical interfaces, a variant of the 4-terminal measurement method must be used (see the discussion on the measurement of polarization resistance).

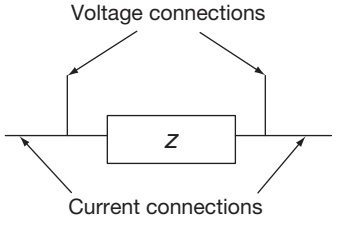


Figure 2 Four-terminal method for resistance or impedance measurement.

2.30.4.2 Measurement of Impedance

2.30.4.2.1 Fundamentals

Impedance measurement can be seen as an extended form of resistance measurement in that a current is applied to the component in question and the resulting voltage is measured. In the case of impedance measurement, however, the current must be a sine wave and the voltage measurement must determine both the amplitude and the phase (with respect to the current) of the resulting voltage. In the general case, it is also necessary to vary the frequency of the sine wave and to account for nonlinearity and noise in the measurement. These aspects are discussed further in the following section.

2.30.4.3 Control of Potential – The Potentiostat

In the general case, the control of the potential of an electrode in an electrochemical cell can be achieved using a four terminal method that is similar to the 4-terminal method of measuring resistance, the current being supplied through one pair of connections and the voltage measured through a second pair. Then a control circuit can adjust the current to maintain a constant potential difference (or a potential difference that follows an independent control signal) between the two voltage measuring connections. However, one of the two connections (to the working electrode) is generally very low resistance and can use a single connecting wire in most cases, whereas the other (to the solution) has a high resistance (and other problems), and must use two connections, one for current (supplied to a separate conductor known as the counter electrode or secondary electrode or auxiliary electrode) and the other to measure the potential using a reference electrode. This leads to the three-terminal device (though it may have more connections to provide additional facilities) known as a potentiostat. In essence, a potentiostat is a very simple voltage amplifier that amplifies the difference between a control potential and the potential of the reference electrode by a very large amount and then applies it to the counter electrode (with the working electrode being connected to ground). Then if the control potential

differs from the reference electrode potential, the counter electrode potential will be changed in the opposite direction, tending to bring the reference electrode potential back to being the same as the control potential. This process is known as 'negative feedback.' A very basic potentiostat can be constructed for a few pounds (Steinberg and Lowe[3] provides an example circuit, although the method of current measurement is more complex than is necessary for most applications, and the circuit could be even simpler). Commercial potentiostats typically provide higher currents, better facilities for the measurement of current and potential, more controlled frequency response, etc.

2.30.4.3.1 Operational amplifier model

The very simplest potentiostat can be constructed from a single operational amplifier (op-amp). (Nowadays the operational amplifier is a standard integrated circuit component that is available with a wide variety of specifications, costing less than £1 for lower specification devices.) The ideal op-amp is a device with three terminals (plus power supply connections – typical op-amp designs use power supplies with three outputs, −15, 0, and +15 V, the 0 V output providing the reference for voltage measurements) as shown in Figure 3.

The relationship between the output of the amplifier (e_o) and the noninverting (e_+) and inverting (e_-) inputs is given by

$$e_o = A(e_+ - e_-)$$

where A, the gain of the amplifier, is very large (typically of the order of 10^6 or more, although it falls with increasing frequency). For the ideal amplifier, the input impedance is infinite and therefore no current will flow into the inputs of the amplifier; for real amplifiers, it is normal to specify the input leakage current rather than the input impedance, and typical values range from 100 nA down to less than 1 pA, depending on the technology used to construct the amplifier.

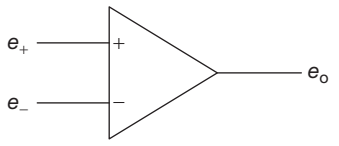


Figure 3 Operational amplifier (power supply connections omitted for simplicity).

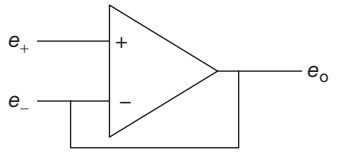


Figure 4 Unity gain buffer.

2.30.4.3.2 Principle of negative feedback

Most operational amplifier circuits rely on 'negative feedback' to provide the required function. In essence, this means that a signal derived from the amplifier output is fed back to the input such that the desired relationship between the input and the output is obtained.

The simplest negative feedback circuit is the unity gain buffer:

In this circuit, the input signal is connected to e_+ and e_o is connected back to e_- (Figure 4). We can analyze the behavior by substituting for e_- in the op-amp gain equation:

$$e_o = A(e_+ - e_o)$$

Hence

$$e_o = e_+[A/(1+A)]$$

As A is very large, $[A/(1+A)]$ is very close to 1, so $e_o \approx e_+$.

By design, the input impedance of the operational amplifier is very large, and the output impedance is very low (and made lower still by the action of the negative feedback), thus providing an amplifier with a gain of one, very high input impedance and very low output impedance.

2.30.4.3.3 Simple potentiostat circuit

The basic circuit to provide control of potential is also very simple (Figure 5).

Again, we can see how the circuit works by putting the various voltages into the gain equation for the op-amp (using V_{CE} for the potential at the counter electrode terminal and V_{RE} for the reference electrode,

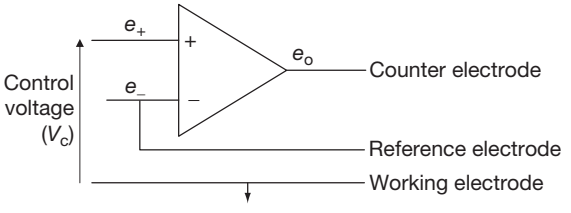


Figure 5 Operational amplifier circuit for potential control.

both relative to the working electrode, which is connected to the power supply ground):

$$V_{CE} = A(V_c - V_{RE})$$

Hence

$$V_{RE} = V_c - V_{CE}/A$$

As A is very large (usually $> 10^6$), this implies that $V_{RE} \approx V_c$. As the electrode potential of the working electrode is measured relative to V_{RE}, this means that the potential of the sample will be $-V_c$.

Somewhat more detailed descriptions of potentiostat circuits are presented in **Chapter 1.02, Electrochemistry**. Note that the circuits for the unity gain buffer and the potentiostat illustrate a general fact for operational amplifier designs, namely that when operating with correct negative feedback, the voltage at the two input terminals is the same, that is, $e_+ = e_-$, and this provides a simple way of working out how the circuit will behave.

2.30.4.3.4 Some potentiostat configurations

For simplicity, we shall draw the potentiostat as a box with three connections, labeled CE, RE, and WE (for Counter Electrode, Reference Electrode, and Working Electrode respectively), and a control input that permits control of the potential with an external input. Then the simplest potentiostat circuit controls potential, optionally following an external device, such as a sweep generator that provides a potential that varies with time (Figure 6).

The potentiostat may provide a means of monitoring the current, but if this is not available, then the current can be measured by inserting an ammeter in the lead from the potentiostat to the counter electrode or by inserting a resistor and measuring the voltage across it (note that the potentiostat automatically compensates for the voltage drop across the ammeter/resistor, and hence this is not a problem, providing it is not excessive; up to 1 V is reasonable).

2.30.4.3.5 Control of current

If a potentiostat is configured to control the potential across a fixed resistor, this will produce a controlled current from the CE terminal to the RE terminal (Figure 7). It should be noted that in this circuit, the RE is at a voltage of V_c relative to instrument ground (the WE terminal), while the CE terminal is at whatever potential is necessary (within the capability of the potentiostat of course) to supply the current. (The term 'instrument ground' is used to refer to the 0 V connection of the power supply. This may or may not be connected to a true local ground (e.g., the earth pin of the mains connector).) It is often required to measure the potential that results from the application of a controlled current; note that the reference electrode must not be connected to the RE terminal to achieve this; some potentiostats may provide a mechanism for measuring the potential in galvanostatic mode (check the manual); alternatively an independent voltmeter must be used.

2.30.4.3.6 Measurement of current

If a potentiostat is configured to control the potential between the RE and WE terminals at 0 V, then the current required to do this can be measured between the CE and RE terminals (Figure 8). The voltage burden is not particularly important, as the potentiostat will adjust the voltage at the CE terminal to supply the required current. An ammeter is shown in Figure 8, but this can, of course, be replaced by a resistor and voltmeter (providing the voltage across the resistor is within the capability of the potentiostat). Note that one current measuring connection is connected to instrument earth in this configuration, while the other (connected to the RE terminal) is a 'virtual earth' (i.e., it is at earth potential, but not actually connected to earth).

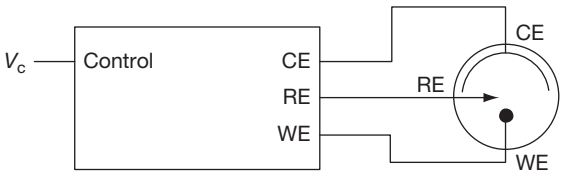


Figure 6 Potentiostat used for control of potential.

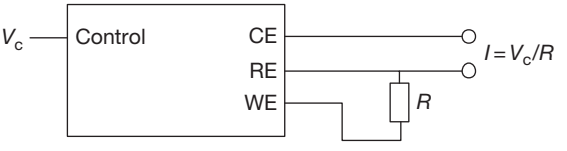


Figure 7 Potentiostat configuration for current control.

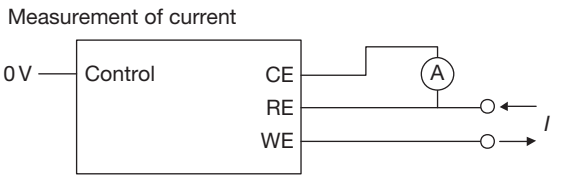


Figure 8 Configuration of potentiostat to measure current.

2.30.4.3.7 Problems in potentiostatic measurements

The potentiostat is a very useful device, but users should be aware of a number of possible problems:

- Overloading

The potential of the working electrode is controlled by current supplied from the amplifier in the potentiostat. The current and voltage available will be determined by the design of the amplifier and the power supply in the potentiostat. The voltage limit is very often ±15 V, as this is the normal maximum supply voltage range for integrated circuit operational amplifiers (though higher voltage potentiostats are available). The current limit will vary from ±10 mA for standard operational amplifiers, up to many amps for power devices. If the required potential demands more current or voltage than the potentiostat can supply, then the potential control will break down; this can usually be recognized by an unchanging current, corresponding to either the current limit of the potentiostat or the voltage limit in conjunction with the total circuit resistance.

- Instability

As already explained, the operation of the potentiostat depends on negative feedback. However, at higher frequencies, elements in the circuit (including the cable from the reference electrode to the potentiostat and circuits internal to the potentiostat) can introduce delays between a change in the potential being applied to counter electrode and the resultant potential appearing at the inverting amplifier input. At a frequency where the delay corresponds to half of the period of one cycle, the delay will convert negative feedback into positive feedback, which can result in strong oscillation of the potentiostat (this is essentially the same as the 'howl' when the microphone picks up the sound from the public address system). Very often the only measurements taken from a potentiostatic experiment are the potential and current – if these are made using dc instruments, the 'howl' will not be detected, and the potentiostat will just appear not to be working, as the dc value of the potential will typically not be what is intended. In this situation, it is important to check the current using an ac meter; alternatively check the cell voltage with an ac meter – do not try to measure the ac voltage between the reference electrode and the working electrode, as this will typically be very small (remember that it will be amplified by the gain of the amplifier in the potentiostat before it is applied to the counter electrode). Such oscillations can usually be stopped (at the expense of slowing down the response of the potentiostat) by connecting a capacitor from the counter electrode terminal to the reference electrode terminal, which provides strong negative feedback at higher frequencies (1 μF is usually sufficient). It may also be useful to reduce the effective impedance of the reference electrode at higher frequencies (especially when using high-resistivity solutions) by coupling it with a capacitor to a length of platinum wire wrapped around the tip of the Luggin probe (see **Section 2.30.8**).

- Frequency response

Many applications of potentiostats are concerned with 'near-dc' conditions, and the frequency response is not particularly important. For this reason and to avoid problems with instability, some potentiostats have a relatively low frequency response. Some applications, however, such as the measurement of electrochemical impedance or transient measurements with scratching electrodes, demand a response at high frequencies. Users should therefore take care to select an appropriate instrument for their application.

2.30.4.4 Control of Current

In principle, the control of current is simpler than the control of voltage, as we are not concerned with the source of any voltages developed in the external circuit. Thus, the simplest galvanostat is a two-terminal device, and many laboratory power supplies offer a controlled current mode of operation that may be suitable for some corrosion experiments (though they are typically not very easy to set accurately, especially for smaller currents). Alternatively, a potentiostat can be used very easily as a galvanostat by using it to control the voltage across a resistor (**Figure 7** and also in the Chapter **Electrochemistry** for an example circuit). Note that there can be earthing issues with galvanostats, as the simplest circuits (including potentiostats used as galvanostats) do not connect either of the output terminals to earth.

As with potentiostats, galvanostats have limits on the output current and voltage, but they are less susceptible to oscillation and similar problems (as the feedback is contained within the instrument, it is much easier to avoid the positive feedback that could give problems).

2.30.5 Electrochemical Methods

2.30.5.1 Potential Monitoring

The measurement of potential is very simple – a reference electrode is placed in the same solution as the working electrode and a voltmeter with suitable (high) input impedance is connected to measure the voltage difference between the two. Note that the negative terminal of the voltmeter should be connected to the reference electrode in order for the measured voltage to have the correct polarity. However, if the working electrode is connected to earth (as is quite likely in the case of plant measurements), this configuration may compromise the input impedance of the meter (since the negative input terminal of a voltmeter often has a much lower impedance to earth than the positive terminal), and in this case it is better to connect the meter 'the wrong way round' (i.e., connect the higher impedance positive terminal to the reference electrode) and reverse the sign of the measured potential.

Provided no current is being applied to the working electrode, it is not particularly important where the reference electrode is situated, and it is not necessary to use a Luggin probe (see Section 2.30.8). The only exception to this is for measurements of the distribution of potential over the surface of the sample, when the design and location of the reference electrode is critical.

While the measurement of potential is relatively easy, the interpretation may present problems. The corrosion potential is affected by many factors, and in the absence of additional information (such as a detailed understanding of the system in question), it is not possible to deduce anything about what is happening. As an example of a situation where the potential does provide useful information, the potential of carbon steel in concrete drops markedly when it starts to corrode, but even here, there is a wide range of potential over which the behavior is considered as uncertain, as indicated in Table 2.

Table 2 Determination of condition of steel in concrete from measured potential (from Tullmin *et al.*[4])

Potential/V(CSE)	*Condition*
> -0.20	Passive
-0.20 to -0.35	Active or passive
< -0.35	Active (i.e., corroding)

2.30.5.2 Galvanic Current

As with the potential, the coupling current between two electrodes is relatively easy to measure, but more difficult to interpret. With a well-defined galvanic couple, the current may provide a useful indication of some feature of the environment, but it is difficult to be sure in the general case whether it will provide useful information and how to extract that information from the measured data.

Note that galvanic currents must be measured with a low potential drop, and therefore an electrometer, zero-resistance ammeter, or potentiostat should be used as described earlier.

2.30.5.3 Polarization Curves

2.30.5.3.1 Introduction

One of the most common electrochemical methods in corrosion research and testing is the determination of the polarization curve – the relationship between the current and the potential, usually over a relatively wide range (of the order of a volt). It is normally the objective to determine polarization curves under near steady-state conditions (i.e., the measured current at each potential is essentially that that would be obtained after a long period at a constant potential). For at least two reasons, this is not really possible:

- When corrosion reactions are occurring, the electrode will change with time, especially in the anodic region of the curve, and hence there is really no such thing as steady-state.
- It is not usually feasible to use more than a few samples, and therefore measurements must be made at different potentials on the same sample; the prior history of the sample is thus likely to influence the behavior.

Thus, the measurement of a polarization curve involves an element of compromise, and the optimum method and parameters to use will depend on the system being studied. As a general rule, systems exhibiting a low corrosion rate benefit from slow measurement methods (since the electrodes will change only slowly by corrosion and the low currents involved will be more easily perturbed in the short term by factors such as double-layer charging currents and currents associated with thickening of passive oxide films), while systems exhibiting rapid corrosion can (and should) be measured more quickly, since the actively corroding surface will reach near-steady state quickly and will change more quickly as a result of corrosion.

2.30.5.3.2 Measurement methods

Polarization curves can be measured either by controlling the current and measuring the potential or by controlling the potential and measuring the current. Since control of current is somewhat easier than control of potential, early measurements tended to use controlled current, but it has a number of disadvantages; in particular, it is not possible to control the potential in the passive region of active–passive polarization curves using controlled current (Figure 9). Thus, modern approaches almost invariably use controlled potential.

The polarization curve is notionally a continuous function, in that the potential can be held at any value and the current measured. In practice, however, there are two approaches to the measurement – the potential can be swept smoothly through the range of values of interest (leading to a *potentiodynamic* polarization curve) or it can be stepped over a range of reasonably small steps (typically in the range of 10–50 mV and leading to a *potentiostep* curve). The advantage of the potentiodynamic method is that it provides an essentially continuous curve. The advantage of the potentiostep method is that the measured current can be taken at the end of the step hold period, and therefore it approaches steady-state a little more closely for the same overall measurement time. It also permits the use of new specimens for each potential, which minimizes effect due to the prior history of the specimen (though it also much more time-consuming and expensive, and introduces questions about sample-to-sample reproducibility).

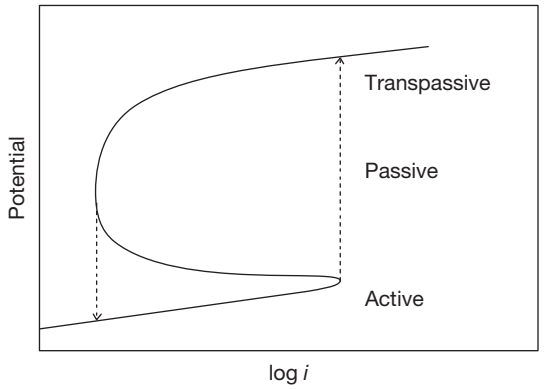


Figure 9 Potentiostatic and galvanostatic measurement of polarization curve (schematic). Note that there is only one current density for each potential, so potentiostatic measurement gives the complete curve. In contrast there are three possible potentials (only two of which are stable) corresponding to much of the range of current densities, so galvanostatic measurement cannot reach some parts of the curve, and flips from the active peak to transpassive corrosion on the positive-going scan direction, and from the passive current density to the active curve in the negative-going direction (denoted by the dashed arrows). Note also that it is impossible to hold the system in the passive region using galvanostatic polarization, as slight changes in condition will cause a transition to active corrosion.

The potentiodynamic method requires a potentiostat, a linear sweep generator (to produce the potential sweep) and a recording device of some kind to record the current and potential. In principle, the latter can be an X–Y recorder (possibly with a log converter on one axis to record log(current)), or a conventional recorder with a hard-working research student to extract and process the information. However, most modern systems will use a computer data acquisition system to manage the measurement and facilitate the processing of the data. The wide dynamic range of the measured current presents something of a problem here, and it is common for dedicated systems to use automatic range switching in order to permit the measurement of both large and small currents.

The sweep rate is an important parameter, as it controls the closeness with which steady-state is approached. Typical values of sweep rate are of the order of 1 mVs^{-1}; as indicated earlier, a slow sweep rate is required for low corrosion rates, while a faster sweep can be used (and may be necessary to avoid excessive modification of the working electrode by corrosion) for systems that are corroding rapidly.

2.30.5.3.3 Interpretation

The simplest approach to the interpretation of polarization curves is to model the curve as the summation of a series of curves corresponding to the individual electrochemical reactions that can occur. As activation-controlled reactions exhibit a linear relationship between potential and log(current density), it is normal to plot potential using linear axes and current density using log axes (strictly the absolute value of current density, since log(x) does not exist for $x < 0$). Since the potential is usually the controlled variable and current density the measured variable, polarization curves should strictly be plotted with potential on the x-axis. However, early workers controlled current and therefore plotted current density on the x-axis, and many workers still use this convention.

Figure 10 (taken from ASTM Standard G5:1998, which provides a standard method for determining polarization curves) shows an example polarization curve, in this case for a ferritic stainless steel in sulfuric acid (the curve displays as a band because

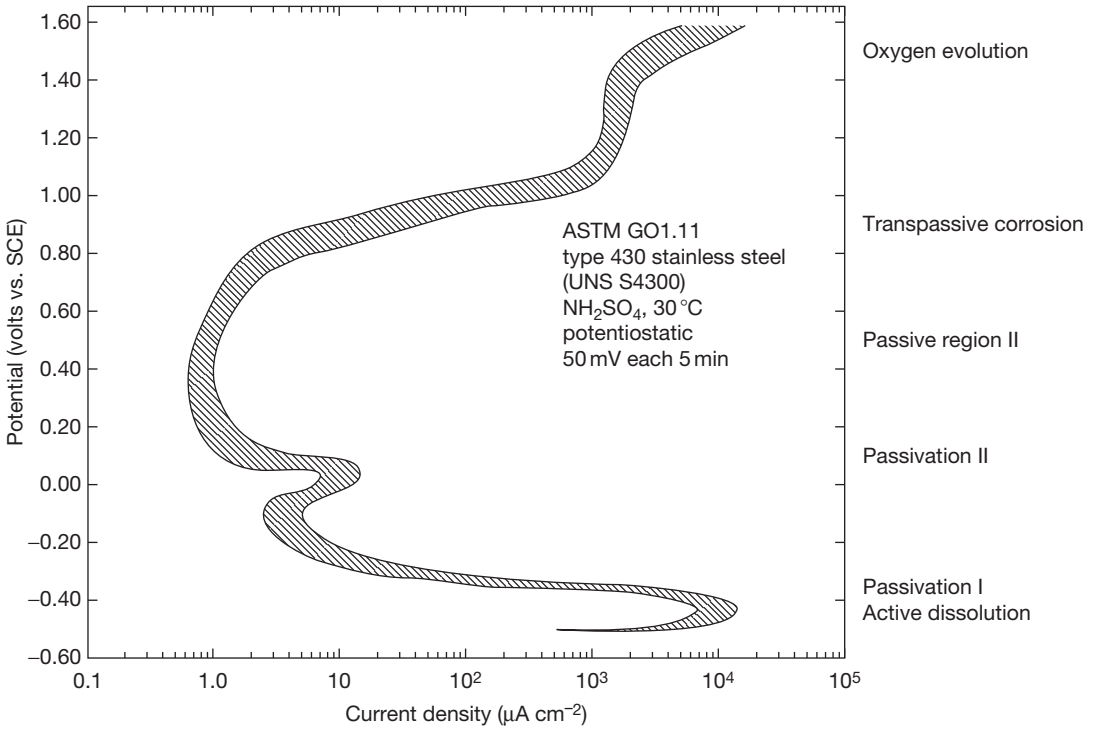


Figure 10 Polarization curve for ferritic stainless steel in sulfuric acid (from ASTM standard[5]).

this figure shows the range of curves obtained by different labs on a particular batch of steel). The labels on the right of **Figure 10** indicate what is happening in each region of the curve. Note that there are two passivation reactions; the first corresponds to passivation by Cr_2O_3, while the second corresponds to passivation by Fe_2O_3 as well as Cr_2O_3.

An important process in the analysis of polarization curves is the estimation of the Tafel slope (the slope of the linear region of the $E-\log|i|$ curve). For an ideal activation controlled reaction, this is straightforward (though easy for beginners to miscalculate the slope, for example by taking $\log(i_1-i_2)$ rather than $\log(i_1)-\log(i_2)$). However, real polarization curves tend not to show a clear Tafel region, and it can be impossible to determine the Tafel slope. As a guide, a Tafel slope should be considered to be reliable only if it meets the following conditions:

- The polarization curve is straight for at least one decade of current (i.e., the current changes by a factor of ten).
- The straight region of the polarization curve should start at about one Tafel slope away from the open circuit potential (i.e., if the Tafel slope is 60 mV, then the curve should become linear at about 60 mV from the corrosion potential). This is not an exact rule, as it depends on the relationship between the anodic and cathodic slopes; if in doubt, a good check is to overlay the calculated curve onto the measured curve.
- The Tafel slope should not 'undercut' the polarization curve between the open circuit potential and the start of the Tafel region. If this occurs, it suggests that the Tafel slope corresponds to a different reaction from the one that is dominant at the open circuit potential; a common student error is to fit a Tafel slope to the hydrogen evolution reaction when the dominant reaction at the open circuit potential is oxygen reduction (**Figure 11**). The Tafel slope may be correct in this situation, but it is not relevant to the corrosion process (therefore, it cannot be used, for example, to compute the Stern–Geary coefficient).

2.30.5.4 Linear Polarization Resistance

The linear polarization resistance is the slope of the $E-i$ relationship at the open circuit potential (OCP) (i.e., at zero current), and it provides a method of

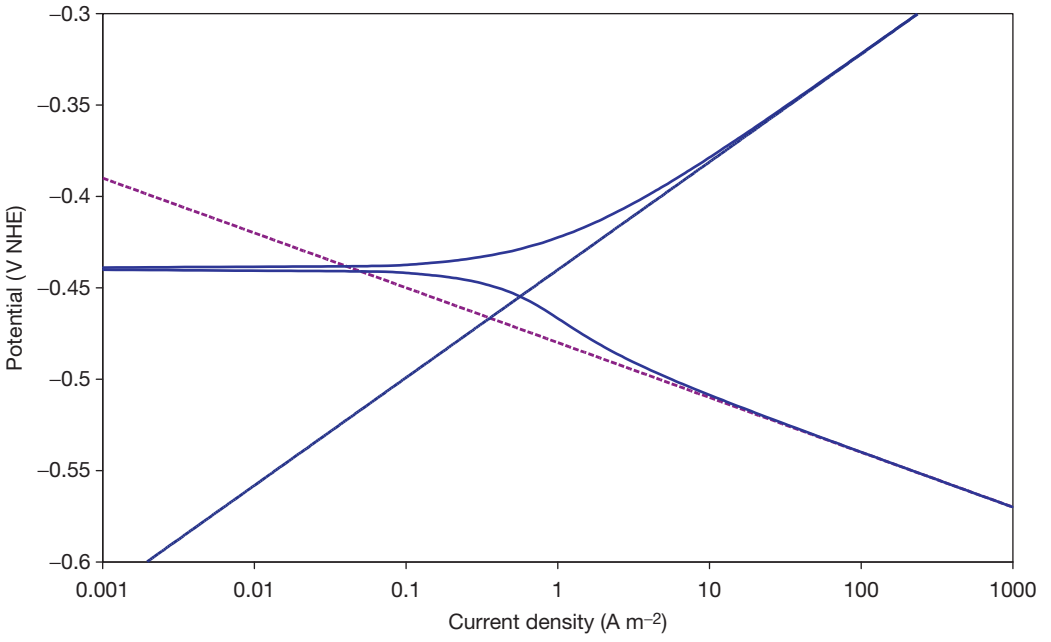


Figure 11 Calculated polarization curve exhibiting 'undercutting' of cathodic Tafel region due to multiple cathodic reactions. This graph is based on the idealized behavior of iron in a neutral solution, where the primary cathodic reaction is oxygen reduction at the OCP, which gives rise to the 'bump' in the cathodic curve. Note that the anodic and cathodic Tafel slopes do not intersect at the corrosion potential, and the current density at the intersection is less than the corrosion current density (which is defined by the limiting current density for oxygen reduction, set to 1 A m^{-2} for this model).

estimating the instantaneous corrosion rate. It can be measured in a number of ways:

- The potential can be swept through a narrow range either side of the OCP, the current recorded, and the slope determined. This method is commonly used in laboratory measurements, as it has a number of advantages – the $E-i$ plot produced provides an indication of measurement problems such as high noise levels (**Figure 12**), and it makes a true measurement of dE/di at $i=0$. However, it is relatively difficult to automate, as it is necessary both to compensate for changes in the OCP and to determine the slope of the curve at $i=0$.
- The current can be swept through a small range about zero and the potential measured. This has the advantage of automatically making a measurement that is centered on zero current, but the current range needs to be adjusted to compensate for changes in the measured resistance. This method also reveals measurement noise.
- The potential can be stepped between OCP $-\Delta E/2$ and OCP $+\Delta E/2$, where ΔE is a small potential difference (typically 10 to 20 mV). Then R_p can be determined as $\Delta E/\Delta i$, where Δi is the change in current density. This suffers from the difficulty of compensating for changes in the OCP, and errors can be introduced if the anodic and cathodic Tafel slopes are markedly different.
- An applied current can be stepped between $-\Delta i/2$ and $+\Delta i/2$ and the corresponding potential step, ΔE, measured. This is the simplest approach for corrosion monitoring, as it automatically makes the measurement centered on the OCP, and the polarization resistance is directly proportional to ΔE (since $R_p = \Delta E/\Delta i$ and Δi is constant). The current amplitude should ideally be adjusted to keep ΔE in a reasonable range, but this is not too problematic unless the corrosion rate varies over quite a wide range. This method is also subject to error if the anodic and cathodic Tafel slopes are markedly different.
- A more sophisticated approach to the measurement of polarization resistance is to apply a sine wave of current or potential and measure the amplitude of the potential or current respectively. Essentially, this consists in making a single frequency impedance measurement – see the following section for further information. This method has two main advantages: it is easier to account for extraneous noise, and the frequency at which the measurement is made is well-defined, which

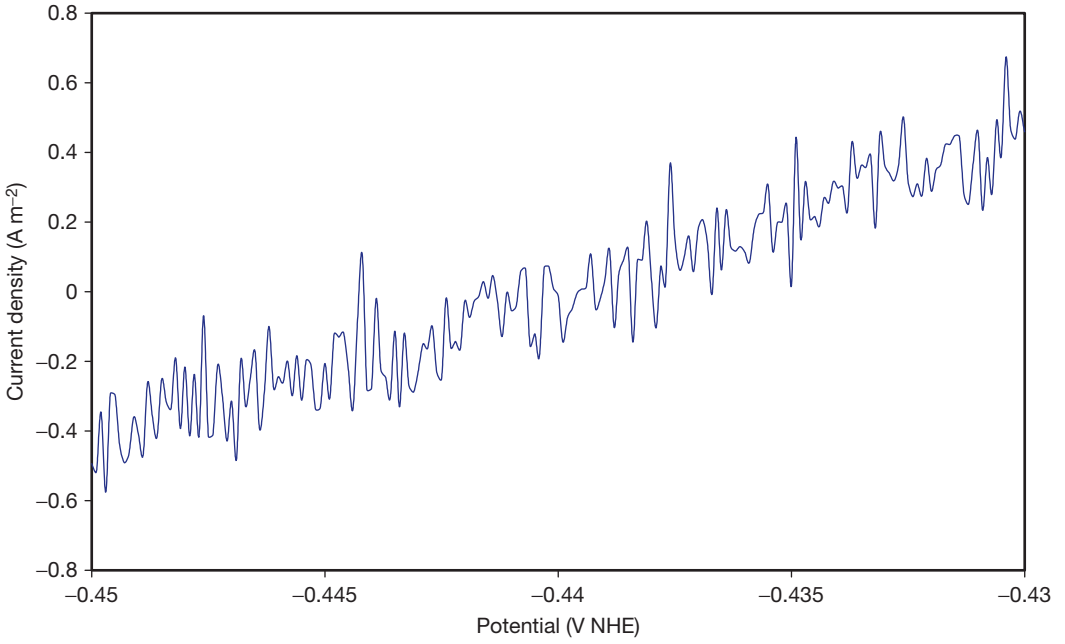


Figure 12 Linear polarization resistance measurement using a potential sweep and exhibiting measurement noise. In this case the amplitude of the noise is low enough that we can make a reasonable estimate of the polarization resistance.

facilitates interpretation. With the low cost of sophisticated electronics, this has become a good method for corrosion monitoring, especially as it can easily be combined with Harmonic Analysis for the estimation of the Stern–Geary coefficient (see **Section 2.30.5.6**).

In addition to the variety of methods of polarization that are available, polarization resistance can be measured either as the resistance between two equivalent electrodes (this has some practical advantages, notably the fact that the expected dc potential difference is zero) or as the resistance of a single electrode, using a reference electrode to connect to the solution in a reproducible fashion (**Figure 13**).

Note that the normal unit of R_p is Ωm^2, that is, it is not, strictly speaking, a resistance, but a resistance times area.

A question for all of the measurement methods is the optimum cycle time for the measurement. The basic theory on which the polarization resistance method is based assumes that R_p is measured at a low frequency, such that capacitive and other time varying currents can be neglected. However, in some situations diffusion processes can change very slowly (notably in environments such as soil and concrete), and it is not feasible to measure R_p at sufficiently low frequencies. In this case it has been found that relatively fast

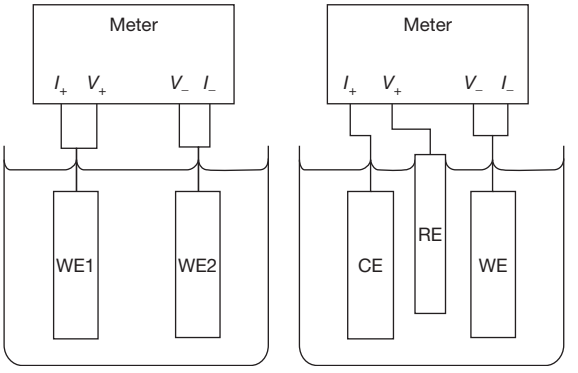


Figure 13 Two- and three-electrode methods for polarization resistance and impedance measurement. Generic connections have been indicated on the measuring device – if a potentiostat is used for the measurement, I_+ will correspond to the counter electrode terminal, V_+ the reference electrode and V_- and I_- the working electrode (this will be one terminal unless the potentiostat has separate current and voltage measuring connections for the working electrode).

measurements (with a typical measurement cycle time of 30 s) still give good correlation with corrosion rate. This is attributable to the measurement of the charge transfer resistance with only a small contribution from the diffusional impedance (see Cottis and Turgoose[6] for further explanation), but note that this does lead to a B-value different than for very low frequency measurements.

2.30.5.4.1 Interpretation

The interpretation of the polarization resistance is normally based on the Stern–Geary relationship:[7]

$$R_p = \frac{B}{i_{corr}}$$

where B is the Stern–Geary coefficient (it is often called the Stern–Geary constant, but it is not a true constant, being dependent on the metal and environment being studied), or simply the B-value; its unit is V.

Note that the B-value cannot be derived from the measured data (but see the sections on harmonic analysis and intermodulation distortion) and must be obtained from separate experiments, either by measuring the anodic and cathodic Tafel slope from a polarization curve or by using a practical calibration against corrosion rates measured by an alternative method, such as weight loss. In the absence of a suitable B-value, it is common to use 26 mV for activation controlled systems and 52 mV for systems with the cathodic reaction being limited by diffusion. These values correspond to typical Tafel slopes for these two cases, and the calculated corrosion rates will usually be within a factor of two or three of the true value, which is usually tolerable for corrosion monitoring, where the objective of the electrochemical corrosion monitoring is primarily to detect major changes in behavior.

2.30.5.4.2 Errors

The polarization resistance method makes a number of assumptions about the system under investigation, and errors arise if these assumptions are invalid.

- It is assumed that there is only one anodic reaction (metal dissolution) and one cathodic reaction (oxygen reduction or hydrogen evolution). Note that this implies that the rates of the reverse reactions are negligible. This is not always valid: relatively noble metals such as copper and nickel may be close to equilibrium at their corrosion potential, and it is possible for R_p to be dominated by the metal–metal ion exchange reaction; similarly, the kinetics of hydrogen–hydrogen ion reaction typically control the potential of nickel-base alloys in hydrogen-containing high temperature water, and electrochemical measurements give little information about the corrosion reaction.
- It is assumed that both anodic and cathodic reactions obey Tafel's Law, which implies that they must be under activation control. Truly diffusion limited reactions (i.e., reactions that are at the limiting current density) can also be regarded as obeying Tafel's Law, with a Tafel slope of infinity, but intermediate cases (where the current is at a significant fraction of the limiting current density) do not obey the Stern–Geary equation and will give errors. Similarly, passive alloys will have an anodic behavior that does not obey Tafel's Law. Fortunately in the case of passive systems, the cathodic reaction will often obey Tafel's Law, while the anodic reaction can be regarded as having a very high Tafel slope; consequently, they may obey the Stern–Geary relationship. This is not a major problem for corrosion monitoring, as the passive alloy will always have a high polarization resistance, and the exact corrosion rate is not usually of concern.
- The electrochemical reactions are uniformly distributed over the electrode (i.e., uniform corrosion is occurring). Polarization resistance measurements on electrode that are subject to localized corrosion are normally considered to estimate the average corrosion rate. This is justified if the corrosion process consists solely of a cathodic reaction (either obeying Tafel's Law or fully diffusion controlled) on the passive surface and an anodic reaction at the localized corrosion site (e.g., within the pit or crevice) that obeys Tafel's Law. However, this is not necessarily the case, and polarization resistance should be regarded as more of a qualitative measure if localized corrosion is occurring. Note that it is not possible to identify purely from the measured value whether or not the corrosion is localized or uniform.
- It is generally assumed that the relatively small current or potential perturbation that is applied as a result of the measurement of polarization resistance does not change the corrosion behavior. This is not necessarily correct, and the perturbation may change the behavior somewhat. This is more likely when using controlled potential methods, as these are more likely to apply a nonzero mean current to the electrodes.
- Polarization resistance measurements are necessarily performed on probe electrodes when used for corrosion monitoring. It is assumed that the probes have the same behavior as the plant, but there are several reasons why this may not be valid.[8]

2.30.5.5 Electrochemical Impedance Spectroscopy

Electrochemical impedance spectroscopy, EIS, extends the polarization resistance method by measuring how the impedance varies with frequency. There are a

number of ways in which this measurement can be made (see Turgoose and Cottis[9] and Orazem and Tribollet[10] for further information on the technique), but in essence the measurement consists in applying a series of ac currents and measuring the ac potential response or applying a series of ac potentials and measuring the current response. The impedance is then calculated by dividing the potential by the current at each frequency (using complex arithmetic in order to maintain the phase information). There are a number of criteria that must be satisfied in order to make a valid measurement,[11] including:

- Linearity: the response of the system must be proportional to the perturbation (this is rarely true for electrochemical systems, but we approach it by using small amplitude perturbations).
- Causality: the response of the system must be a direct result of the perturbation (power line noise and similar interference would result in a breach of this requirement, as would randomly occurring currents associated with localized corrosion).
- Stability: this requirement requires that the system does not exhibit characteristics such as multiple values for a given perturbation, or (equivalently) regions of negative resistance (electrochemical systems are often not stable, including any systems that include an active–passive transition).

It is possible to test EIS data for failure to meet these requirements, the most common method being to use the Kramers–Kronig (K–K) transform; this allows for the calculation of the phase response from the amplitude response and vice versa. If the calculated and measured data do not match up, this implies that the data are not valid (note that this is a necessary, but not sufficient, condition – valid data will not fail the K–K transform test, but it is possible for invalid data to pass it).

2.30.5.5.1 Presentation of EIS data

The EIS measurement produces a set of amplitude and phase values for a range of frequencies, and there are two main ways of presenting these:

- The Bode plot presents log(amplitude) and phase against log(frequency). Figure 15 presents a Bode plot for the equivalent circuit of Figure 14 (this is known as the Randles equivalent circuit and has three components: R_s is the resistance of the solution, R_p is the polarization resistance of the metal–solution interface, and C_{dl} is the double-layer capacitance of the metal–solution interface). On the Bode plot, a resistor produces a horizontal line on the amplitude plot with amplitude equal to the resistance and a constant phase of zero on the phase plot. A capacitor produces an amplitude that falls with a slope of -1 as the frequency increases (the amplitude of the impedance is $1/(2\pi fC)$, where f is the frequency and C the capacitance), and a constant phase of $-90°$; as we are normally dealing with resistors and capacitors, it is common to invert the phase axis (i.e., plot–phase) so that capacitive circuit elements give data above zero.
- The Nyquist plot normally plots the imaginary part of the impedance against the real part (Figure 16). The Nyquist plot invariably inverts the imaginary axis (i.e., it plots the imaginary component of impedance with increasingly negative values on the y-axis), so that capacitive circuit elements plot above the x-axis. One weakness of the Nyquist plot compared with the Bode plot is that it does not implicitly include the frequency of each measurement point, so at least some points should have their frequency indicated.

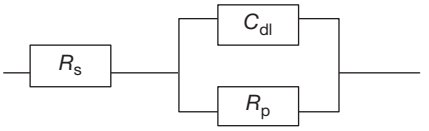


Figure 14 The Randles equivalent circuit for a corroding interface.

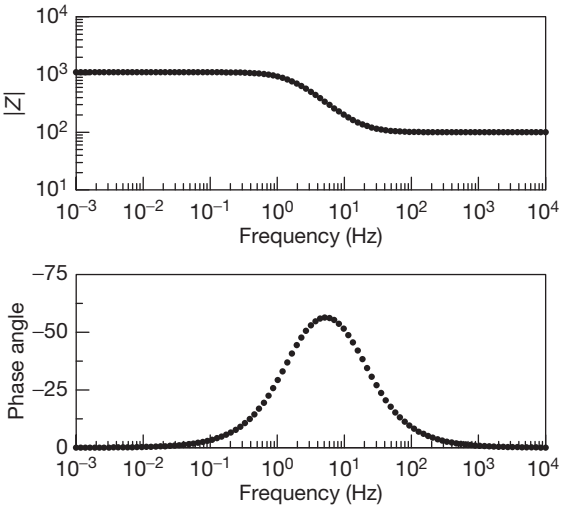


Figure 15 Example Bode plot for the equivalent circuit of Figure 14, with component values of $R_s = 100\ \Omega$, $R_p = 1000\ \Omega$, $C_{dl} = 0.0001$ F (from Cottis[8]). Note that |Z| tends to R_s at high frequency and to $R_s + R_p$ at low frequency.

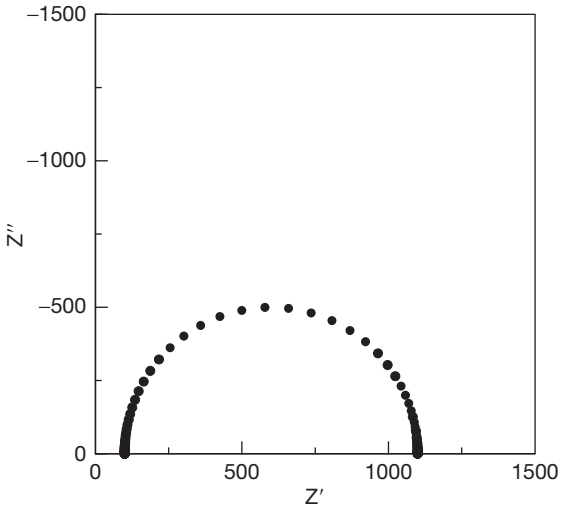


Figure 16 Example Nyquist plot (from Cottis[8]).

2.30.5.5.2 *Interpretation of EIS measurements*

The interpretation of EIS data is essentially concerned with determining the components inside a sealed 'black box' with just two leads connected to the outside. It is relatively easy to determine the values of internal components if we know what their configuration is, but if we do not know the configuration or the number of components there is no unique solution to the problem, as there are an infinite number of circuits that could produce the observed behavior. The simplest example of this is circuits that contain only resistors. Any resistor network will appear from the outside as a pure resistor (i.e., the measured impedance will have constant amplitude and zero phase). Consequently, in order to obtain a valid interpretation, it is important to use prior knowledge of the expected behavior in order to model the real physical system.

There are two basic approaches to model EIS performance, and hence to determine the properties of the elements of the electrochemical interface by adjusting the parameters of the model to match the measured data:

- The response of the chemical and electrochemical processes to a fluctuating potential is analyzed with a mathematical or numerical model in order to produce a simulated spectrum, and the parameters of the model are adjusted to fit the measured spectrum. This is technically an ideal approach, as it provides a direct relationship between the physical processes occurring and the observed spectrum, and can include the effects of nonlinearity in response, but it is very demanding of time and skill, and it is therefore used only infrequently.
- A much more common approach is for the components of the system under investigation to be modeled by electrical equivalent circuit elements that have similar characteristics to the actual processes concerned. Thus, a paint film can be modeled as a resistor (corresponding to ionic current passing through the film) in parallel with a capacitor (corresponding to the capacitance of the paint film acting as a parallel plate capacitor); underneath the paint film the metal–solution interface gives rise to a resistor (corresponding to the charge transfer resistance) in parallel with a capacitance (the double-layer capacitance). The response of the resultant electrical circuit is then modeled using conventional methods from electrical engineering, and the parameter values adjusted to optimize the fit between the model and real data. Note that the equivalent circuit is derived first, based on the physical processes occurring; it may then be necessary to modify the equivalent circuit to fit features of the measured data. It is important that the added elements are linked to a physical process; some less-experienced workers try different equivalent circuits to find the one that best fits the measured data and then try to work out what the elements correspond to; while this may provide a good fit to the data, the circuit may have the wrong configuration so that the values of circuit elements give no information about the real processes occurring.

2.30.5.5.3 *Advanced impedance measurements*

Significant advances have been made in the EIS technique in recent years. These are largely related to the detection and handling of errors in the measurement and enhancing the speed of measurement.

In a conventional EIS measurement, the impedance is acquired sequentially at each of a set of frequencies. The frequencies are normally arranged as a geometric series (i.e., each frequency in the series is the product of the immediately preceding sample and a multiplier (the latter is typically chosen to give around seven frequencies per decade). With this analysis scheme, the time taken to determine a full spectrum is dominated by the lowest frequency included in the measurement. It is not uncommon for the lowest frequency to be 1 mHz, implying that the measurement of an impedance spectrum takes

several hours. Besides leading to long measurement times, this leads to questions about the stationarity of the system. Van Gheem *et al.*[12,13] have developed a measurement method that uses multiple simultaneous sine waves, permitting the simultaneous measurement of a full spectrum (the use of multiple sinewaves is not new; Van Gheem *et al.*'s method is novel in the way that it allocates measurement frequencies in such a way that the measurement of the power at frequencies that are not present in the applied signal permits the estimation of the impact of noise and nonlinearity on the measurement).

In a similar vein, Orazem[14] has developed a method to use the error structure of EIS data to optimize the fitting of equivalent circuit models to the data.

Unfortunately, the methods developed by Van Gheem and Orazem were not available in commercial EIS measurement systems at the time of writing, but they provide an important advance in the development of reliable EIS analysis methods.

2.30.5.6 Harmonic Analysis

Harmonic analysis may be thought of as a single-frequency impedance measurement that is extended to measure the first and second harmonics of the input signal.

When a sine wave voltage is applied to a linear system (see the definition of linearity given earlier), the resultant current consists only of a sine wave at the same frequency. However, if the system is nonlinear, as is the case with electrochemical interfaces, this produces some distortion of the sine wave, which creates harmonics of the input signal. Using similar assumptions to those used by Stern and Geary,[7] Dévay and Mészáros[15,16] showed that the amplitudes of the first and second harmonics can be used to determine the Tafel slopes of the anodic and cathodic reactions, and hence determine the *B*-value to be used in the estimation of i_{corr}. Thanks to the low cost and power of modern electronics, this approach has been adopted as a corrosion monitoring technique with the claimed advantage (based on the theoretical analysis, but not yet really validated in practice) that it does not depend on prior knowledge of the *B*-value.

In his analysis Mészáros showed that the double layer capacitance could be allowed for by taking measurements at two frequencies. However, this requires the assumption that the equivalent circuit is a simple Randles circuit, which is often not the case, and as far as the author is aware, all practical implementations of this method depend on the measurement being made at a low frequency, such that the capacitive components of the equivalent circuit can be neglected.

2.30.5.7 Intermodulation Distortion

A further development of the harmonic analysis technique applies two sine waves to the corroding sample. These interact with the nonlinear interface to produce a rather more complex set of output frequencies (**Figures 17 and 18**). Besides the harmonics of the two input signals, intermodulation components are formed at the sum and difference frequencies. The method was first proposed by Mészáros and Dévay,[17] and the method was subsequently developed by Bosch *et al.*,[18] who termed it 'Electrochemical Frequency Modulation,' although the process by which the output signals are produced is termed 'intermodulation distortion' in the signal processing literature (and in Mészáros and Dévay's paper). (Intermodulation has been widely used in electronic systems for many years as a method for changing the frequency of a signal. Thus in a modern satellite TV system, the LNB (Low Noise Block downconverter) takes the microwave signal picked up by the satellite dish (with a frequency in the range 5–20 GHz) and mixes it through a nonlinear device with a locally generated signal to convert it to a lower frequency (which will be the difference between the incoming frequency and the local signal frequency) that can be handled more easily in the receiver itself, the original signals and all the other intermodulation components being removed by filtering.)

Analysis of the signals produced permits, in theory, the extraction of the Tafel slopes in much the same way as harmonic analysis, but in addition, certain pairs of frequencies can be shown to have constant ratios, and these have been proposed as a test of the reliability of the data (Bosch *et al.*, who first drew attention to these relationships, termed them 'causality factors' 2 and 3 on the basis that they should have values of 2 and 3 if the outputs are caused by the inputs, rather than being random noise). This is potentially a very attractive ability, but at the time of writing this method had been relatively little used, and it is unclear whether or not real systems follow the theoretical behavior well enough for the results to be relied upon.

The theoretical derivation of the analysis procedures for this method implicitly assumes that the measurement is made at a low enough frequency that capacitive effects can be ignored. The relevant

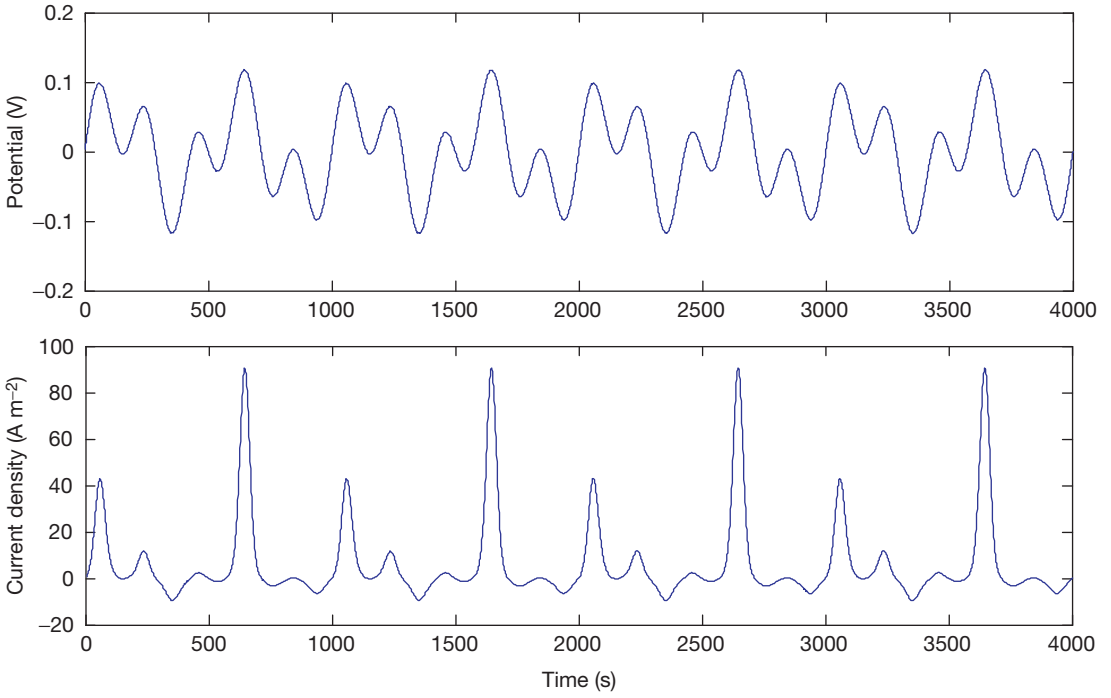


Figure 17 Potential and current versus time for a simulated intermodulation measurement (controlled potential, cathodic Tafel slope 120 mV anodic slope 60 mV). Note the much lower amplitude of the cathodic currents as a result of the larger Tafel slope, and the resultant distortion of the signal compared to the input signal.

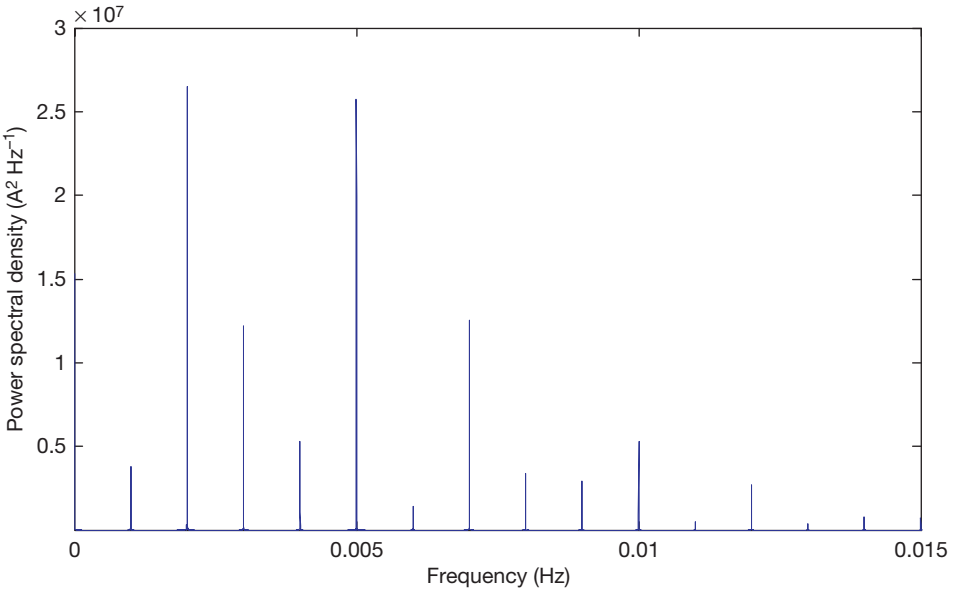


Figure 18 Illustration of intermodulation distortion (the two input frequencies used were 0.002 and 0.005 Hz (f_1 and f_2); the signal at 0.003 and 0.007 Hz are the difference ($f_2 - f_1$) and sum ($f_2 + f_1$) signals; the signals at 0.004 and 0.01 Hz are the first harmonics of f_1 and f_2; the remaining signals result from intermodulation between f_1, f_2, and harmonics and intermodulation components. The current was based on a 1 m^2 electrode, which is why the observed power spectral density is so high.

frequency is that of the highest component used in the analysis, and this typically requires the use of very low frequencies for the two input signals. It may be possible to analyze the phase and amplitude of the various output components in order to test for, and possibly to correct for, capacitive components, but this had not been done at the time of writing.

2.30.5.8 Transient Techniques

A number of transient techniques have been developed to probe the properties of electrochemical interfaces. Several of these are essentially attempting to determine the properties of elements of an assumed equivalent circuit. These have largely been superseded by EIS, since this obtains the same information in a more accurate and more general way. An example of such a technique is open circuit potential decay (OCPD). In this method, the metal–solution interface is charged by the application of a constant current until a steady potential is achieved. Then the current source is disconnected, and the potential allowed to decay back to its open circuit value, giving a potential–time curve similar to that shown in **Figure 18**.

Various methods can be used to interpret the resultant curve (see Bosch *et al.*[18] for details); most of these depend on the assumption that the interface can be described by a simple equivalent circuit and that the interface has been charged for long enough that the charging current into the double layer capacitance, C_{dl}, can be neglected.

2.30.5.9 Electrochemical Noise

Electrochemical noise (EN) can be defined as naturally occurring fluctuations in the current and/or potential of corroding electrodes.[20] Early workers regarded these fluctuations as merely a nuisance, to be removed by filtering or averaging if possible, but Iverson[21] realized that they may contain information about the rate and nature of the corrosion process. A number of techniques have been devised for the measurement of electrochemical noise,[8] but most measurements now determine the potential noise as either the fluctuation of the electrochemical potential relative to a low-noise reference electrode or the fluctuation of the potential between two nominally identical working electrodes, while the current noise is measured as the current between two nominally identical working electrodes. An extension of the method measures the potential and current noise simultaneously using three electrodes (**Figure 19**). This has a number of advantages, including the ability to determine the electrochemical noise resistance (see the following paragraph).

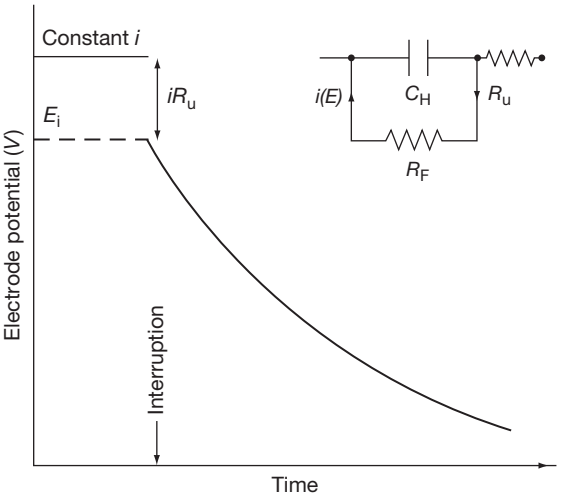


Figure 19 Typical OCPD plot (from Kelsall[19]).

2.30.5.9.1 *Interpretation of electrochemical noise*

Early work on electrochemical noise used largely heuristic methods to analyze the data, that is, a number of analysis procedures were tried, and the ones that seemed to work determined by experiment. More recently, more detailed theoretically-based analyzes have been developed.[19,22]

One of the most useful parameters to come out of the early heuristic work was the electrochemical noise resistance, R_n. This can be defined as $A\sigma_E/\sigma_I$, where A is the area of each electrode and σ_E and σ_I are the standard deviations of current and potential respectively, assuming the conventional three-electrode configuration. (Throughout this chapter we take R_p, R_n, and R_{ct} as being normalized to unit area (i.e., their unit is Ohm m^2.) This convention has not always been used (including by the author) and care should be taken to ensure that area effects are properly handled.) This result was first suggested by Dawson *et al.*[23] The result can be proved theoretically with relatively few assumptions:

- The current noise sources on the two working electrodes have the same mean and standard deviation and are uncorrelated.
- The measurements are made at a low enough frequency that capacitive and other effects can be ignored.

- The response of the two working electrodes to an applied current is defined by AR_p, where R_p is the polarization resistance.

With these assumptions, it can be shown that $R_n = R_p$ (strictly we should say the expected value of R_n is R_p, since this is a statistical result that is only true on average). Thus, i_{corr} can be estimated as B/R_n, where B is the Stern–Geary coefficient.

The simplest theoretical analysis of the expected amplitude of the potential and current noise is based on the premise that the electrochemical process consists of a series of events that are characterized by a pulse of charge. At the lowest level, this pulse might be two electrons produced by the oxidation of one iron atom to Fe^{2+}or four electrons consumed by the reduction of one molecule of oxygen. (Note that corrosion texts often state that the anodic and cathodic processes associated with corrosion are tightly coupled, such that every oxygen molecule reduced requires two iron atoms to be dissolved. However, this is not strictly correct, and the coupling only occurs indirectly through the change in electrochemical potential that occurs as the double-layer capacitance is discharged. Thus the anodic and cathodic events are uncorrelated in the short term, and we can therefore treat them as independent noise sources.) Larger events would correspond to such processes as the anodic charge produced by the initiation, growth, and death of a metastable pit.

If we assume that individual events are uncorrelated (i.e., the occurrence of one event does not change the probability of a subsequent event occurring), we can use the shot noise analysis produced by Shottky[24] to describe noise production in thermionic valves. Individual events are treated as pulses of zero duration, but nonzero charge, q. Then if the mean current is I, the average number of events occurring each second is I/q. (Note that we use I, I_n, I_{corr}, etc. to indicate total current, rather than current density, because current noise is not expected to be proportional to specimen area, and therefore it is more meaningful to talk about the current for a given specimen area, rather than considering the current density. R_n on the other hand is expected to be inversely proportional to specimen area and it is therefore appropriate to normalize it to unit area (giving the unit of $\Omega\ m^2$).) Analysis of the statistics of this process leads to the shot noise formula $\overline{I_n^2} = 2qIb$, where $\overline{I_n^2}$ is the variance of the current noise signal and b the bandwidth of the measurement (the range of frequencies included in the measurement).

For a corroding electrode two processes must be occurring simultaneously (e.g., metal dissolution and oxygen reduction). In principle, we can estimate the noise due to each source and sum them to derive the total noise amplitude. However, it will commonly be the case that one of the processes produces significantly more noise than the other, in which case we only need to consider that process. Thus, for pitting corrosion driven by oxygen reduction, the pitting process is expected to produce much more noise than oxygen reduction under normal circumstances, and therefore, we can simply treat oxygen reduction as providing a constant current that brings the net current to zero. With this assumption, and assuming that we can apply the Stern–Geary equation to estimate I_{corr} from R_n, it can be shown[19] that we can estimate I_{corr}, the charge, q, in the transient events, and the frequency, f_n of those events:

$$I_{corr} = \frac{B}{R_n} = \frac{B\sigma_I}{\sigma_E}$$

$$q = \frac{\sigma_I \sigma_E}{Bb}$$

$$f_n = \frac{I_{corr}}{q} = \frac{B^2 b}{\sigma_E^2}$$

where B is the Stern-Geary coefficient; σ_I, the standard deviation of current; σ_E, standard deviation of potential, and b is the bandwidth of measurement.

Note that these equations assume that only low frequencies are included in the calculation of standard deviation, and it is arguably better to use an estimate of the power spectral density (PSD) at a low frequency:

$$I_{corr} = \frac{B}{R_n} = B\sqrt{\frac{\Psi_I}{\Psi_E}}$$

$$q = \frac{\sqrt{\Psi_I \Psi_E}}{B}$$

$$f_n = \frac{I_{corr}}{q} = \frac{B^2}{\Psi_E}$$

where Ψ_I is the PSD of current and Ψ_E is the PSD of potential.

2.30.5.9.2 Measurement

Electrochemical noise is almost invariably measured as digitized time records. Thus, the measurement process consists of a number of steps, illustrated in **Figure 20**. Potential noise typically has a low amplitude, and it is therefore common to amplify it – this is done as the first step in order to minimize the effects of other processes on the instrument noise. If the potential is measured with respect to a reference

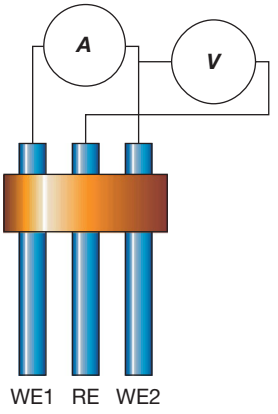


Figure 20 Three electrode EN measurement configuration.

electrode, it will typically have a relatively large dc value, and the voltage amplifier may also incorporate a dc offset to allow a larger gain to be used without exceeding the output voltage range of the amplifier.

Current noise is normally converted to a voltage signal using a current amplifier. It is not usually necessary to provide for a current offset as the current noise normally has an expected mean of zero.

Following amplification, the voltage and current noise signal should be filtered to remove unwanted frequency components (although the amplifier and filter circuits may be combined in a practical instrument). In particular, a low-pass antialiasing filter should be included to remove frequency components that could give rise to aliasing in the sampling process. A high-pass filter may also be incorporated to remove the dc level (as an alternative to the use of a fixed dc offset), although it is difficult to produce good analog filters with a very low frequency, and even good quality filters present 'issues' associated with the filter settling time when first switching on.

The continuous analog signal must then be sampled to convert it to a regular sequence of samples, and these sampled values must be converted from analog to digital values. These two processes may be combined if certain types of analog to digital convertor are used. Most electrochemical noise information is present at low frequencies (typically below about 10 Hz), and hence EN measurement systems do not have particularly stringent sampling rate requirements, and it is more important to have a high resolution. Most computer data acquisition systems are optimized for high-speed, relatively low-resolution sampling, and they are therefore unsuitable for EN measurements. Conventional digital multimeters or electrometers usually provide more suitable measurement capabilities, and they may also provide some of the required signal conditioning (although for the best results separate signal conditioning systems are normally required).

Finally the sampled digital time series is either recorded for subsequent processing (this is typically done for research purposes) or processed online to produce summary parameters (typically used for corrosion monitoring).

The measurement of electrochemical noise requires care in order to avoid sources of noise other than true electrochemical noise. Some of the more likely sources of error are:

- Instrument noise – electrochemical noise often has a very low amplitude, and high-quality current and potential measuring systems are necessary in order to minimize the addition of electronic noise from the instrument. Instrument noise levels should be checked to ensure that the level of such noise is significantly below that of the measured signals. Instrument noise levels are influenced by the impedance of the system being measured, and a dummy cell with a similar impedance to that of the corroding system should be used in the calibration process (**Figure 21**, based on Ritter *et al.*[25]).
- Quantization noise – virtually all electrochemical noise analysis methods involve converting the analog noise signal to a sampled digital time record. If the resolution of the analog to digital convertor is inadequate, steps will be observed in the time record (**Figure 22**). These add a form of noise, known as quantization noise, to the data; providing the measured signal is changing sufficiently rapidly, the contribution of the quantization noise is to add white noise (noise with a constant PSD) to the signal. (If the signal is changing slowly compared with the quantization step size, the measured value may be constant for long periods, and in this case the effect of the quantization may be to reduce the apparent noise.)
- Aliasing – during the sampling process, signals with a frequency above half the sampling frequency are transformed into samples that appear to have a lower frequency (**Figure 23**). (Some digital-to-analog conversion methods reduce the sensitivity to aliasing by averaging the measured signal over the sample period. However, this does not completely prevent aliasing, and it is good practise to use antialiasing filters for all EN measurements.) It is important to use antialiasing filters to remove these frequencies

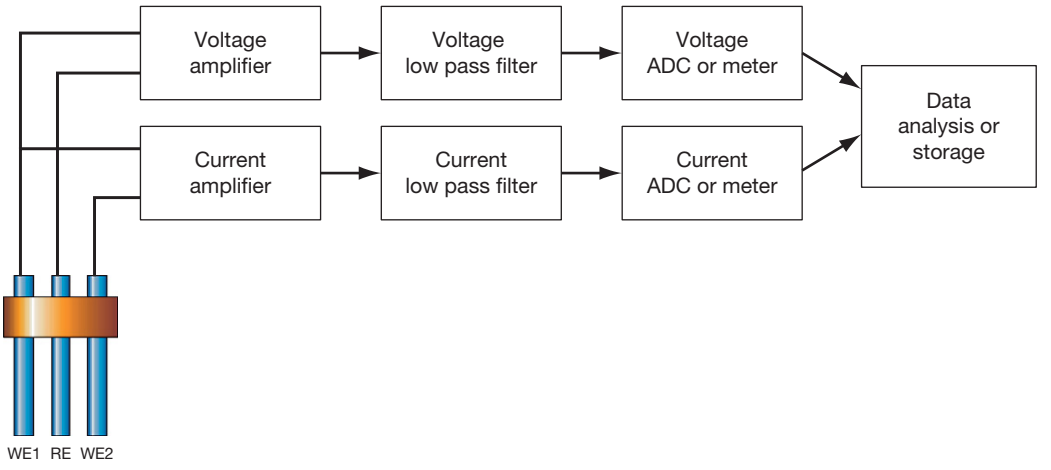


Figure 21 EN Measurement System (ADC = Analog to Digital Convertor) (from Cottis[8]).

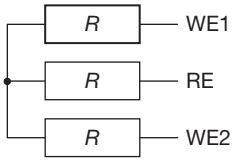


Figure 22 Dummy cell for determination of instrument noise levels (from Schottky[24]).

before sampling occurs, as it is impossible to distinguish the resultant sampled data from real low-frequency signals. It should also be appreciated that aliasing does not just occur within electronic devices; it can also be created in software. Thus, if the sampling frequency of a signal is reduced by decimation (selecting every nth sample), this will also produce aliasing, and a low-pass software filter should be applied to the original data before decimation. (As software filters can be constructed to have essentially ideal properties, this can be a useful approach. Indeed there are arguments that the most effective approach to the development of instrumentation for EN is to sample at a fixed, relatively high frequency and then deliver data at the required rate by filtering and decimation. Besides avoiding expensive, low-frequency analog filters, this tends to reduce the amplitude of quantization noise (this process is known as oversampling in the electronics literature).)

- Interference – electromagnetic interference from a number of sources may be picked up and amplified by the EN measurement system. This leads to two main types of error: mains frequency (power-line) noise (50 or 60 Hz depending on location) and relatively rapid transients (duration of the order of ms) due to inductive effects associated with the switching of high currents (the laboratory refrigerator is a common source of the latter). While both of these types of interference can be identified relatively easily, it may be more difficult to do this once the data have been sampled. Thus, mains frequency noise will typically be aliased to a much lower frequency, and transients will appear to be at least one sample period long. Thus, it is always best to prevent such interference from getting into the measurement system. This is normally achieved by careful shielding and earthing.

2.30.6 Electrode Design

Specimens for corrosion electrochemistry can take many forms, and much will depend on factors such as the form in which the material is available, the objective of the measurement, and the type of corrosion behavior that is expected. Typical electrode constructions are shown in **Figure 24**.

Factors that should be taken into account in the design of the electrode include:

- Surface to be studied
  Different planes in an alloy tend to behave differently as a result of segregation occurring during casting and subsequent thermo-mechanical treatment. Consequently, it may be important to study a particular surface, in which case only planar electrodes, **Figures 25(b), 25(d), 25(e), and 25(f)**, are suitable (**Figure 25(a)** is not suitable, as the edges of the specimen are necessarily exposed as well as

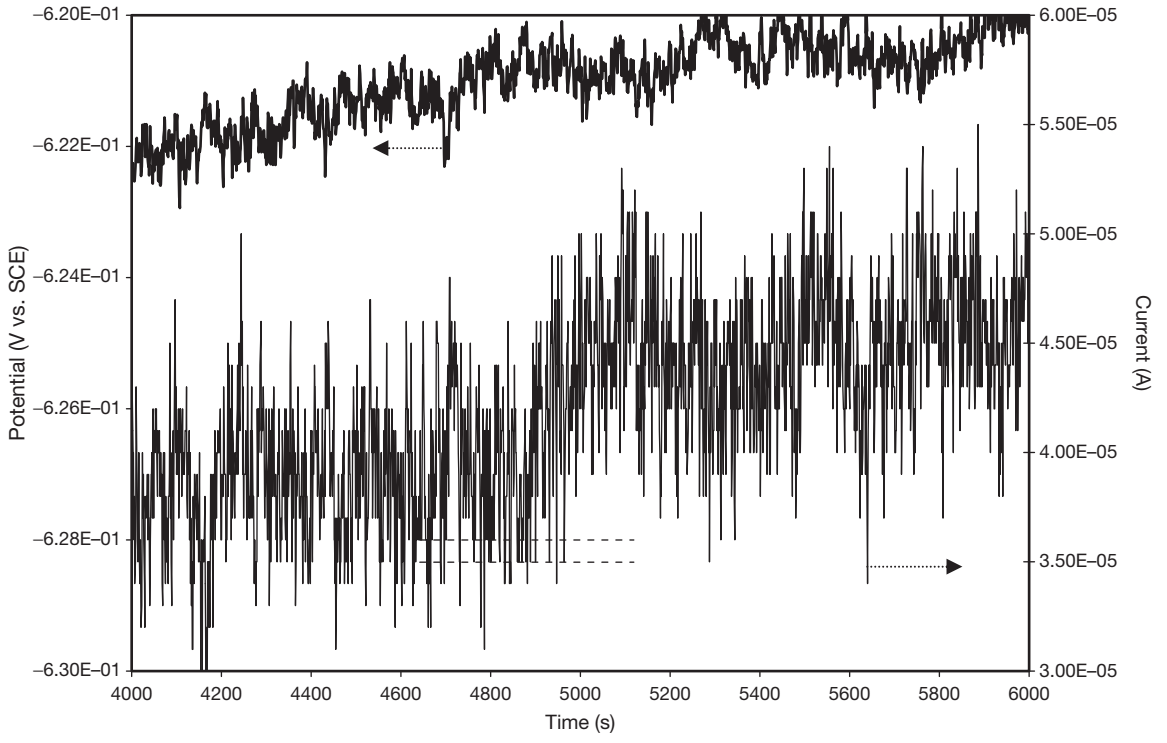


Figure 23 EN time record showing quantization noise in the current signal (from Cottis[8]). The quantization can be seen as the 'banding' in the plot, illustrated by the two dotted lines superimposed on the plot at two adjacent quantization levels. This level of quantization is tolerable, and is not going to have a very large impact on the analysis.

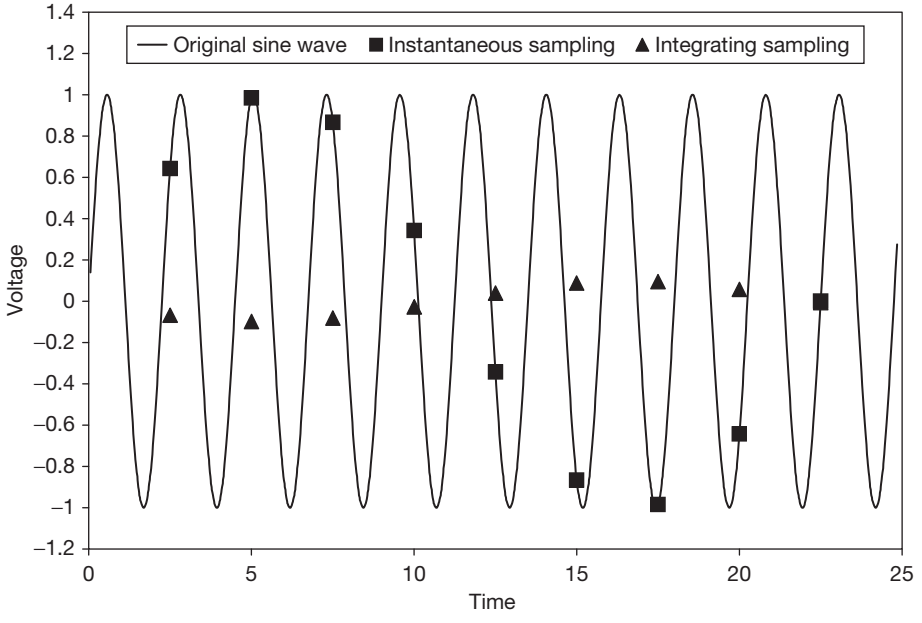


Figure 24 Illustration of aliasing (from Cottis[8]). Integrating sampling is used by many voltmeters, and averages the voltage over the sample period; instantaneous sampling just takes value of the voltage at the sample time, and is much more seriously affected by aliasing.

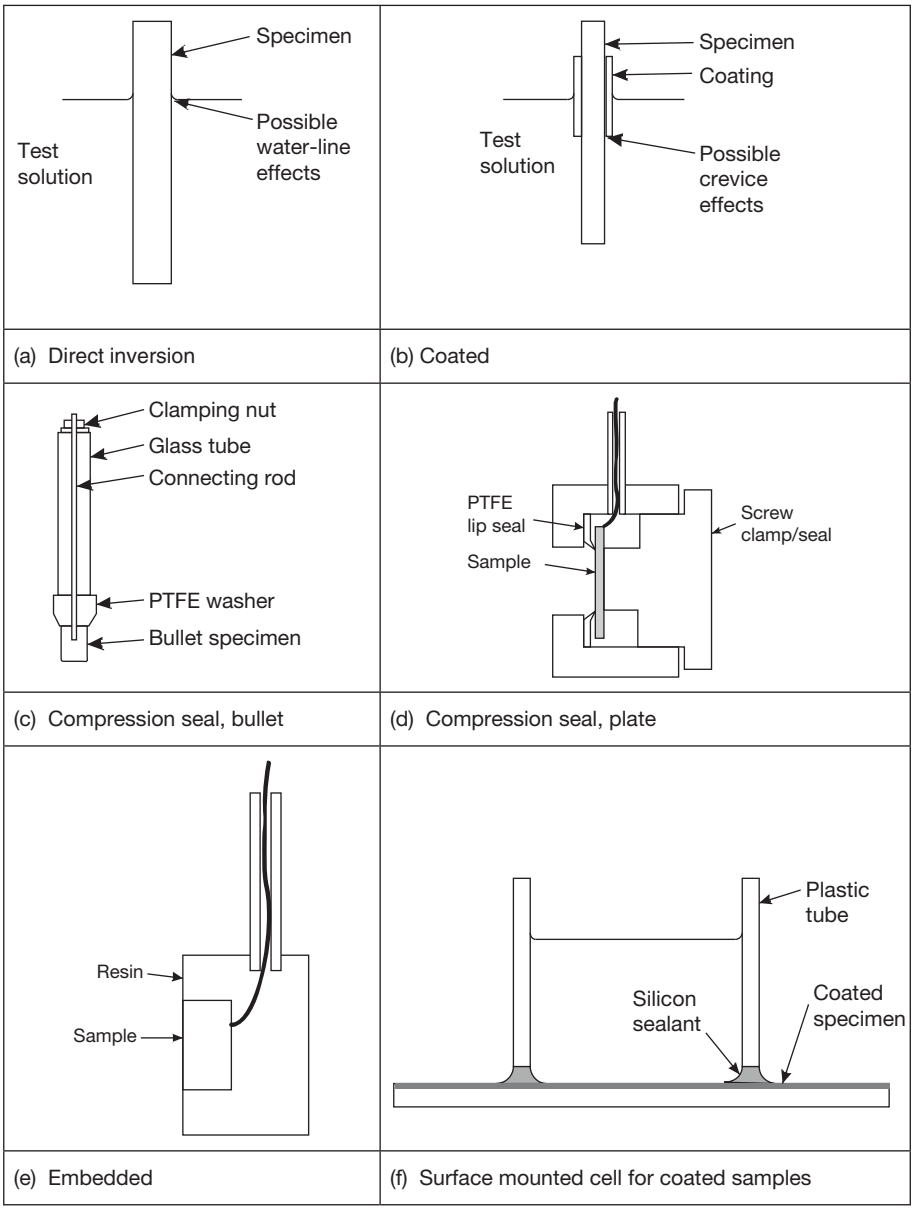


Figure 25 Typical electrode designs for corrosion studies.

the surface of interest). Thought should also be given to the state of the surface (ground, polished, as received, etc.) as this can have a marked effect on the electrochemical behavior (the embedded electrode of Figure 25(e) is suitable for studying a specific plane, but not a preexisting surface, as the electrode preparation must include grinding and/or polishing of the embedded electrode to remove resin from the surface).

- Susceptibility to crevice corrosion
  Passive alloys, including stainless steels, are susceptible to crevice corrosion. It is almost inevitable that electrode designs expose metal–insulator junctions to the test solution (Figure 25(a) is an exception, but it has its own problems associated with the waterline) and these are potential sites for crevice corrosion. Besides blatant damage to the specimen, this can lead to apparent passive current densities

that are much higher than the true values (Figure 26). A number of approaches have been used to minimize the risk of crevice corrosion:

1. The use of a PTFE-metal compression seal[26] can prevent crevice corrosion, but careful design and manufacture are necessary (PTFE-metal contacts have also been used as crevice-formers[27]).
2. Embedding a metal sample in epoxy or similar resin can lead to the formation of a crevice as the resin cracks away from the metal. The risk of this happening can be minimized by maximizing the adhesion of the resin to the metal by using an appropriate pretreatment (contact the resin manufacturer for advice), and by using a thin layer of resin (this minimizes the shrinkage stress between the resin and the metal; if a thick resin layer is required to facilitate surface preparation, this can be added as a second layer, when cracking will usually occur harmlessly between the two resin layers). Another approach that has been used with stainless steels is to run a trace of nitric acid round the edge of the specimen to passivate any crevice (although it may be difficult to do this reproducibly).
3. Some coatings are particularly liable to act as a crevice at the edge of the coating, while others, particularly those with a very good adhesion, may be less susceptible. Wax-based coatings may be useful, as the wax can usually be removed easily after the test.[28]
4. Crevice corrosion arises because of chemical changes (normally acidification) that occur in the crevice, and a method of avoiding crevice corrosion is to flush the crevice so that an aggressive solution cannot build up. (There are systems where the dominant factor involved in crevice corrosion is the potential drop down the crevice, but this is relatively unusual.) This is the basis of the Avesta Cell,[29] which uses a seal fabricated from filter paper that is flushed with distilled water.

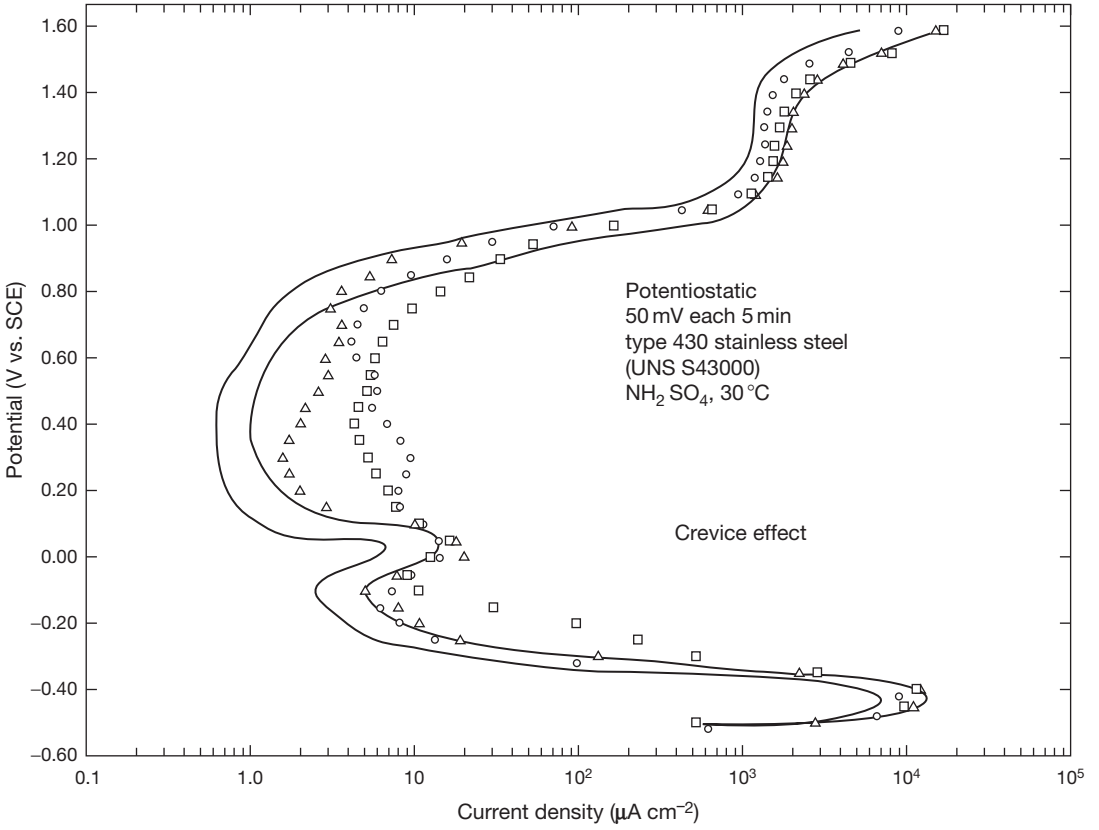


Figure 26 Effect of crevice corrosion on passive current density, revealed by the increase in passive current density (markers) compared to the standard curve (solid lines) (from Tullmin *et al.*[4]).

2.30.7 Cell Design

The design of cells for electrochemical corrosion measurements should take a number of factors into account:

- Current distribution
  In most cases, one objective for an electrochemical measurement will be to make the measurement at a uniform current density and potential over the surface of the specimen. In general, this is facilitated by ensuring a uniform current path between the WE and the CE (e.g., using by a cylindrical WE at the center of a cylindrical CE or by using two parallel plates). It is often suggested that the CE should be large compared with the WE, but this is not necessarily beneficial from the point of view of the current distribution. In general, it becomes more difficult to achieve a uniform current distribution as the solution conductivity decreases, since the IR drop will have a larger relative influence on the current distribution.
- Mass transport
  Besides the distribution of current, it is also important to consider mass transport effects, particularly when applying relatively large overpotentials, as in the case of the measurement of a polarization curve. If well-characterized flow conditions are required, a variety of configurations is available (see the Chapter on **Flow Assisted Corrosion** for further information). Where mass transport is less critical, stirring of the solution may be sufficient. Gas bubbling serves both to stir the solution and to control the gas concentration in the solution. Note that of the range of controlled flow configurations that are available, only a few offer mass transport conditions that are the same across the electrode, the rotating cylinder electrode probably being the best from this perspective.
- Influence of reaction products
  Any electrochemical reaction will induce a change in the local chemistry, and it is generally desirable to minimize the influence of such changes on the measurement. The nature and significance of such changes will vary according to the system under study, concentrated solutions typically being less seriously affected than dilute solutions. The local concentration at the metal surface will be influenced by mass transport, while the bulk concentration will be influenced by the solution volume and the applied current. The reaction products at the counter electrode should also be considered, and in critical cases, it may be necessary to separate the working electrode and the counter electrode compartments.
- Control of gas composition of solution
  Dissolved gases, notably oxygen, can be very important in corrosion reactions, and it is important to control the gas concentration of the solution. For aerated solutions this can be simply achieved by bubbling air through the solution. Deaeration is much more difficult, both because commercial gases typically contain traces of oxygen and because oxygen diffuses very rapidly (it even diffuses relatively rapidly through many polymers). Consequently, for rigorous deaeration, it is necessary to take great care to prevent oxygen access (e.g., by sealing the cell, bubbling the exhaust gas from the cell through water to restrict back diffusion, and possibly by purging the gas used to deaerate the solution by passing it though a suitable solution such as vanadous chloride). Any gases produced at the counter electrode must also be considered, as noted earlier.
- Temperature
  Chemical and electrochemical processes are generally rather sensitive to temperature (for typical activation energies corrosion rates double every 10 °C). Consequently for accurate measurements, the temperature should be controlled (or at least recorded). Temperature control is easier when the temperature is above ambient, and it is usual to perform tests somewhat above room temperature (e.g., at 30 °C). Tests at temperatures approaching boiling point become difficult, with respect to the maintenance of the solution composition (especially for dissolved gases) and because condensed moisture from the test can cause electrical leakage problems. For tests at above ambient pressure, it is necessary to use an autoclave, and electrochemical measurements become far more difficult and are beyond the scope of this article.

2.30.8 Reference Electrode

A range of reference electrodes is available (see **Table 3** and the Chapter **Electrochemistry**.

Note that the solution in the reference electrode plays an important part in controlling the potential of the electrode. Thus, a saturated calomel electrode must have saturated KCl as the solution in contact with the Hg/Hg_2Cl_2 paste. This means that the SCE

Table 3 Common reference electrodes

Common Name	*Electrode*	*V vs NHE*	*Notes*
Saturated* Calomel Electrode (SCE)	Hg/Hg_2Cl_2/sat. KCl	+0.241	Probably the most common electrode in the laboratory, but use of mercury presents safety hazards
Calomel	Hg/Hg_2Cl_2/1 M KCl	+0.280	Better temperature stability than SCE
Mercurous sulfate	Hg/Hg_2SO_4/sat. K_2SO_4	+0.640	Useful for avoiding Cl^- contamination of test solution
	Hg/Hg_2SO_4/0.5M H_2SO_4	+0.680	
Mercurous oxide	Hg/HgO/1M NaOH	+0.098	Good for alkaline solutions
Silver chloride	Ag/AgCl/sat. KCl	+0.197	Very easy to make, but light sensitive, tending to replace SCE as most commonly used
Copper sulfate	Cu/sat. $CuSO_4$	+0.316	Robust, commonly used for cathodic protection in soils (best avoided in chlorides because of precipitation of $CuCl_2$)
Zinc/seawater	Zn/seawater	−0.8	Not a true reference electrode, but robust and quite stable

*Note that the S in SCE stands for Saturated (not Standard) owing to the use of a saturated solution of KCl.

tends to leach chloride (and potassium of course, although this is usually less significant for corrosion) into the solution in contact with it. This leaching will contaminate the solution in contact with the reference electrode. Whether or not this leads to a serious problem depends on the test solution and the configuration of the reference electrode. In general, it is beneficial to use reference electrodes based on anions that are already present in the test solution. Thus, SCE is fine for seawater and other solutions containing high concentrations of chloride, but would be undesirable in chloride-free solutions.

It is a general objective of electrochemical measurements to measure the potential of the working electrode relative to the solution immediately adjacent to it. If there is no current applied to the working electrode, then this is not a problem, as there will not normally be significant currents in the test solution, and hence the potential will be essentially constant throughout the solution. However, when the working electrode is polarized, there will be potential drops in the solution (commonly called 'IR-drops') due to the interaction of the applied current and the resistance of the solution. There are several approaches to minimizing or correcting for the IR-drop:

- The reference electrode can be placed close to the surface of the working electrode. In practice this is usually undesirable due to the relatively large size of most reference electrodes (although micro electrodes are available) and the contamination of the working electrode by species leaching from the reference electrode. A way around both of these problems is the use of a Haber–Luggin probe (also known as a Luggin probe), which is an electrically insulating tube with a fine tip that is placed close to the surface of the working electrode. The other end of the tube is connected to an electrode holder into which the reference electrode is inserted, and the tube and electrode holder are filled with the test solution. (A tube filled with solution is known as a 'salt bridge' and is essentially the ionic conduction equivalent of an insulated metal wire (although the resistance is, of course, much higher).) This provides a small probe that can be placed close to the surface of the working electrode (but not too close, about 3 tip diameters is generally recommended to avoid shielding part of the surface), and it also reduces the significance of leaching of the reference electrode solution into the test solution.
- The potential drop can be compensated for in one of a number of ways. The simplest method is to correct the measured potential based on an estimate of the solution resistance, but this may have undesirable consequences for some measurements, such as the measurement of polarization curves, where the effective sweep rate may be very different from that intended as a result of changes in the IR-drop during the measurement. Alternatively the estimated resistance can be used by the potentiostat to compensate for the IR drop. In essence, the applied current is passed through a resistor with the same value as the estimated solution resistance, and the resultant voltage is added to the control voltage. This is a form of positive feedback, and will result in instability in the potentiostat if the estimated resistance is too large; consequently, it is difficult to compensate completely

for the IR-drop. Another approach is to use the fact that the metal–solution potential tends to decay relatively slowly when the current is disconnected (due to the effect of the double-layer capacitance), whereas the IR-drop disappears essentially instantaneously. This leads to the current-interruption method, whereby the true potential is estimated when the current is briefly switched off.

2.30.9 Counter Electrode

In some cases, typically when working with microelectrodes, or when the current is very small for some other reason, it is possible to use the reference electrode as the counter electrode. The use of a specific counter electrode can also be avoided by using two working electrodes, and this may be useful in the measurement of polarization resistance or electrochemical impedance (seen already). In all other cases where current is applied to the working electrode, a counter electrode is required as the second connection to the test solution.

In most cases, the counter electrode should not affect the measurement, and it should therefore have the following properties:

- It should have a relatively low polarization resistance so that the potential drop between the counter electrode and the solution does not limit the polarization that can be applied.
- It should not contaminate the solution. In practice, there will always be some electrochemical reaction at the counter electrode, and what we really need is for the products of that reaction to be harmless or easily removed. Inert electrodes, such as platinum or graphite are often used, in which case the reaction products are usually gases (oxygen or chlorine when anodic or hydrogen when cathodic) that can be removed by bubbling air or nitrogen past the counter electrode (although there may also be a pH change at the counter electrode). In closed systems (e.g., in autoclave studies), it is more difficult to dispose of gaseous reaction products, and it may be better to use a reactive electrode and trap the reaction products close to the counter electrode (e.g., by using an ion-exchange membrane between the counter electrode and the working electrode). It is also wise to check that supposedly inert electrodes are actually inert – some platinum dissolution can occur at high enough anodic current densities and graphite electrodes tend to release traces of impurities into the solution as the graphite is oxidized to CO_2. Note that many metals, including carbon steels and stainless steels, are essentially inert when cathodically polarized, and they can therefore be used as counter electrodes when only anodic polarization of the working electrode is required (e.g., when studying pitting breakdown potentials of stainless steel).
- It is often stated that the area of the counter electrode should be large compared with the working electrode, but this may be less important than ensuring that the overall cell configuration provides the required current density distribution. Thus, a counter electrode of the same size as the working electrode, mounted parallel to it, may optimize the current distribution without significant adverse effects from the size of the counter electrode.

References

1. Kelly, R. G.; Scully, J. R.; Shoesmith, D. W.; Buchnheit, R. G. Electrochemical Techniques in Corrosion Science and Engineering. Boca Raton FL: *CRC Press*, 2002.
2. Horowitz, P.; Hill, W. *The Art of Electronics*, 2nd edn. Cambridge: Cambridge University Press, 1989.
3. Steinberg, M. D.; Lowe, C. R. *Sens. Actuators B* **2004**, *97*, 284–289.
4. Tullmin, M. A. A.; Hansson, C. M.; Roberge, P. R. The Corrosion of Steel, and its Monitoring, in Concrete contribution to Intercorr/96, from www.corrosionsource.comviewed on 22nd April 2009.
5. ASTM Standard G5: 1998.
6. Cottis, R. A.; Turgoose, S. In *Electrochemical and Optical Techniques for the Monitoring of Metallic Corrosion*; Ferreira, M. G. S., Melendres, C. A., Eds.; Dordrecht: Kluwer, 1991; pp 123–133.
7. Stern, M.; Geary, A. L. *J. Electrochem. Soc.* **1957**, *104*(1), 56–63.
8. Cottis, R. A. ''Corrosion Monitoring – What's the Point'', in ''Corrosion Monitoring in Nuclear Systems'', ed. S. Ritter, European Federation of Corrosion, in press.
9. Turgoose, S.; Cottis, R. A. *Corrosion Testing Made Easy: Electrochemical Impedance and Noise.* NACE: Houston, 1999.
10. Orazem, M. E.; Tribollet, B. *Electrochemical Impedance Spectroscopy*, Wiley: New York, 2008.
11. Urquidi-Macdonald, M.; Real, S.; Macdonald, D. D. *Electrochim. Acta* **1990**, *35*(10), 1559–1566.
12. Van Gheem, E.; Pintelon, R.; Vereecken, J.; Schoukens, J.; Hubin, A.; Verboven, P.; Blajiev, O. *Electrochim. Acta* **2004**, *49*(26), 4753–4762.
13. Van Gheem, E.; Pintelon, R.; Hubin, A.; Schoukens, J.; Verboven, P.; Blajiev, O.; Vereecken, J. *Electrochim. Acta* **2006**, *51*(8–9), 1443–1452.
14. Orazem, M. E.; Tribollet, B. *Electrochim. Acta* **2008**, *53* (25), 7360–7366.
15. Dévay, J.; Mészáros, L. *Acta Chimica Academiae Scientiarum Hungaricae* **1979**, *100*(1–4), 183–202.
16. Mészáros, L.; Dévay, J. *Acta Chimica Academiae Scientiarum Hungaricae* **1980**, *104*(3), 311–316.

17. Mészáros, L.; Dévay, J. *Acta Chimica Academiae Scientiarum Hungaricae* **1980**, *105*(1), 1–17.
18. Bosch, R. W.; Hubrecht, J.; Bogaerts, W. F.; Syrett, B. C. *Corrosion* **2001**, *57*(1), 60–70.
19. Kelsall, G. H. In *Techniuques in Electrochemistry, Corrosion and Metal Finishing*; Kuhn, A. T., Ed.; New York: Wiley, 1987; pp 75–94.
20. Cottis, R. A. *Corrosion* **2001**, *27*(3), 265–285.
21. Iverson, W. P. *J. Electrochem. Soc.* **1968**, *115*(6), 617–618.
22. Gabrielli, C.; Huet, F.; Keddam, M. *J. Appli. Electrochem.* **1985**, *15*(4), 503–508.
23. Eden, D. A.; John, D. G.; Dawson, J. L. Corrosion Monitoring UK Patent 8611518 1986, US Patent 5139627 (filed 1987, granted 1992).
24. Schottky, W. *Ann. Phys.* **1918**, *57*, 541–567.
25. Ritter, S.; Huet, F.; Cottis, R. A. Reliability of electrochemical noise measurements – part 2: First ECG-COMON guideline for an assessment of electrochemical noise measurement devices to be presented to Eurocorr 2009.
26. Stern, M.; Makrides, A. C. *J. Electrochem. Soc.* **1960**, *107*, 782.
27. Carranza, R. M. In *Proc. MRS Fall Meeting*; Rebak, R. B., Hyatt, N. C., Pickett, D. A., Eds.; 2008, 1124; Paper 1124-Q09–01 2008.
28. Fonseca, C.; Barbosa, M. A. *Corrosion Sci.* **2001**, *43*(3), 547–559.
29. Qvarfort, R. The Avesta Cell – a New Tool for Studying Pitting **1988**, acom, no. 2–3, pp 2–5.

2.31 Spectroscopies, Scattering and Diffraction Techniques

H.-H. Strehblow
Institut für Physikalische Chemie und Elektrochemie, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany
D. Lützenkirchen-Hecht
Fachbereich C – Physik, Bergische Universität Wuppertal, Gauß-Str. 20, 42097 Wuppertal, Germany

Abbreviations

AES Auger electron spectroscopy
ATR Attenuated total reflection
EXAFS Extended X-ray absorption fine structure
GDOES Glow discharge optical emission spectroscopy
IR Infrared
IRAS, IRRAS IR reflection absorption spectroscopy
ISS Ion scattering spectrometry
MALDI Matrix assisted laser desorption
MEIS Medium energy ion scattering
NEXAFS Near edge X-ray absorption fine structure
RBS Rutherford back scattering
SEM Scanning electron microscopy
SIMS Secondary ion mass spectrometry
TOF Time of flight
UHV Ultrahigh vacuum
UPS UV-photoelectron spectroscopy
XANES X-ray absorption near edge structure
XAS X-ray absorption spectroscopy
XPS X-ray photoelectron spectroscopy
XRD X-ray diffraction
XRR X-ray reflectivity

Symbols

$\boldsymbol{a}_i$ Base vector of the elementary unit cell
$\boldsymbol{b}_i$ Base vector of the reciprocal lattice
$\boldsymbol{d}$ Lattice spacing, layer thickness
$\boldsymbol{D}$ Particle size, atomic density
E_B Binding energy
E_{kin} Kinetic energy
E_P Passivation potential
E_{Pass} Pass energy
$e\Phi_W$ Work function
$e\Phi_{Th}$ Threshold energy
F_{hkl} Structure factor
f_n Atomic form factor
h Planck's constant

$h\nu$ Photon energy
I Intensity of radiation
m Mass of projectile ions in ISS, MEIS, RBS
M Mass of a target atom
n Complex index of refraction
Q Reciprocal lattice vector
Q_z Transferred momentum
R_n Lattice vector of atom n in real space
T Spectrometer transmission function
β Imaginary part of the refractive index
δ Real part of the refractive index
$\Delta\Theta$ Width of a diffraction peak
ϕ Incidence angle
ϕ_c Critical angle of total reflection
λ Wavelength (X-ray, photon, photoelectron)
λ Photoelectron mean free path
μ Absorption coefficient
Θ Scattering angle in RBS, ISS, Bragg angle in a diffraction experiment
σ Photoionization cross section
ω Angle of acceptance of an analyzer

2.31.1 Introduction

Corrosion is the attack of solid material by the environment, which could be liquid or gas phase. The pertinent reactions involve oxidation or reduction of the material's surface which yield products that are transferred, at least temporarily, into the surrounding phase, that is, species into the gas phase, species dissolved into the electrolyte, or deposits on the surface. Corrosion studies usually involve weight loss or weight gain measurements, electrochemical studies, and the analysis of the electrolyte and the solid surface with spectroscopic methods. Most of these investigations start with assumptions about a corrosion process and its chemistry. The analysis of the data may be very accurate and sensitive like in the case of electrochemical methods. However, the nature of the process itself often remains uncertain. Thermodynamic data provide a reasonable basis for assumptions of the processes occurring at the surface, which may be deduced from potential–pH diagrams, the so-called Pourbaix diagrams.[1] They give the first insight into the most likely chemical and electrochemical processes, but they are uncertain because they are based on pure equilibrium data and thus do not account for the kinetic characteristics of the system which may dominate the process. As a result, any study on the basis of electrochemical and other physicochemical data requires the application of surface analytical methods to learn about the nature of the chemical processes which then may be used for their interpretation. Besides the knowledge of the chemical processes, one needs information on the structure of the surface and of surface compounds and their changes which may significantly influence surface reactions. Similar to the information on the surface and the surface compounds, one needs knowledge of the chemistry and the structure of soluble or gaseous corrosion products, which may be obtained by appropriate methods such as rotating ring disc studies or spectroscopic investigation of the electrolyte or the gas phase.

Spectroscopic methods, which are sensitive to the surface chemistry and structure, may be applied *ex situ* and/or *in situ*. Although *in situ* methods are often preferred because they allow one to follow the surface reaction in its natural environment, several *ex situ* methods give such rich and detailed chemical and structural information that they are widely used and well accepted for the interpretation of a corroding system. Of course, one has to be sure that the transfer of the specimen from its environment to the spectrometer does not introduce changes, which then are studied and lead to misinterpretation. The loss of contact with an aggressive atmosphere or an electrolyte solution, the loss of potential control after electrochemical treatment and sample transfer to a vacuum, which is necessary for several methods, should not introduce artifacts to the composition and structure of the specimen's surface. Experience to date indicates that there are methods and equipments that largely meet these requirements.

All spectroscopic methods involve excitation of the solid surface or the homogeneous medium in front (gas or liquid phase) which then induces a response from the system under study. Exciting radiation may consist of electrons, atoms, ions or electromagnetic radiation from the infrared (IR) to the hard X-ray range. Emission of electrons, ions, molecules, or electromagnetic radiation is the response. This chapter intends to give a brief summary of important spectroscopic methods which have been applied successfully to corrosion research and to the solution of practical corrosion problems. The principle, the strength, and the limits of these methods will be discussed and an example in each case will illustrate their application. **Table 1** presents a summary and a comparison of the various techniques. A more detailed description of surface analytical methods and their application to corrosion are discussed elsewhere.[2,3]

Table 1 Characteristics of some surface analytical methods for corrosion in research and practice

Method	*Depth resolution*	*Lateral resolution*	*Information*
XPS (X-ray photoelectron spectroscopy)	2 nm	100 nm	Composition of surface and surface films, binding, and oxidation state, *ex situ*
UPS (UV photoelectron spectroscopy)	1 nm	>1 mm	Work function, threshold energies, band structure, and binding orbitals, *ex situ*
AES, SAES ([scanning] auger electron spectroscopy)	2 nm	20 nm	Composition of surface and surface films, and high lateral resolution, *ex situ*
ISS (ion scattering spectroscopy)	0.3 nm	20 nm	Composition of surface and surface films, modest lateral information, and high depth resolution (monolayer), *ex situ*
RBS (Rutherford back scattering)	5 nm	0.1 mm	Quantitative quasi-nondestructive depth profile, thick films, modest depth resolution, *ex situ*
SEM (scanning electron microscopy)	2 nm	1 nm	Surface topography with high lateral resolution, *ex situ*
EMA (electron microprobe analysis)	2 μm	2 μm	Composition of surfaces and surface layers, *ex situ*
SIMS (secondary ion mass spectroscopy), TOF SIMS	0.5 nm	20 nm	Composition of surfaces and surface layers, high sensitivity and mass resolution, isotopes, large molecules detection, *ex situ*
X-ray reflectivity	0.1 nm	>1 mm	Surface roughness, *in situ*
XRD (X-ray diffraction)	2–3 nm	100 nm	Atomic structure, (micro) crystalline materials (>3 nm), surface sensitive for grazing incidence, *in situ*
XAS (X-ray absorption spectroscopy)	2–3 nm	100 nm	Near range order also of amorphous structures by EXAFS, analysis of electrode surfaces and electrolyte films, chemical information by XANES, *in situ*
GDOES (glow discharge optical emission spectroscopy)	~1 nm	1 mm	Fast elemental analysis of surfaces and thin layer depth profiling, *ex situ*
IR (infrared spectroscopy), ATR (attenuated total reflection) IR spectroscopy, and IRAS or IRRAS (IR reflection absorption spectroscopy)	~0.5 nm	10 μm	Chemical information via oscillation modes of adsorbates and thin layers, *in situ* by grazing incidence or total reflection from the rear side of thin vapor deposited films despite water absorption

2.31.2 Methods Working in the Vacuum and Sample Transfer

Many valuable spectroscopic methods function only in vacuum, for example, X-ray photoelectron spectroscopy (XPS), Auger electron spectroscopy (AES), secondary ion mass spectroscopy (SIMS), ion scattering spectroscopy (ISS) for low energy ions (1–5 keV) or high energy ions, and Rutherford backscattering (RBS). Indeed, typically, the evaluation of energy and mass of emitted electrons and ions requires ultrahigh vacuum (UHV) conditions. There has been a strong debate on the usefulness of these methods, with their application in corrosion research being compared to experiments in surface physics. In physics, mostly single crystals were investigated, and the UHV equipment was not flexible enough for a rapid specimen exchange. However, since the mid-1970s, when spectrometers were equipped with effective sample transfer systems, these methods have been applied very successfully to studies in surface chemistry, corrosion science, catalysis, and practical surface problems in industry. Any kind of surface, such as metals, semiconductors, insulators, and polymers may be studied as long as it is stable under vacuum conditions and does not contaminate the spectrometer. The wide application of these methods in various fields has shown that the achieved data are extremely useful, and one almost cannot avoid his or her use for the solution of many problems in research and industry. However, one has to keep in mind that they require sample transfer to UHV, and as a result, the specimen loses contact with a gaseous environment or the electrolyte and potential control during

sample transfer. Furthermore, exposure to the laboratory atmosphere may cause contamination of the surface and oxidation of surface compounds and corrosion products. In addition to this, the structure of the surface and surface layers may change due to dehydration and oxidation. These problems may be minimized or avoided in corrosion research by an electrochemical specimen preparation within a closed system and its transfer via a protecting inert gas atmosphere.

Several systems have been described in the literature with a shuttle taking specimens from the preparation stage to the UHV environment of the spectrometer, or a specimen preparation facility within a special vessel attached to the spectrometer and an appropriate transfer system to the analyzer chamber.[4,5] **Figure 1** presents an example of the equipment that has been used successfully for corrosion research and the investigation of passive layers in the laboratory of the authors during the last 25 years. In addition to the three chambers of a commercial spectrometer, the analyzer chamber, preparation chamber, and the fast entry lock, an electrochemical preparation chamber is attached which contains a small electrochemical vessel of $\sim 3\,cm^3$ volume for electrochemical specimen preparation under well controlled conditions and under the protection of a purified argon atmosphere.[4,5] Water and the electrolyte are degassed by purified argon and are introduced from the containers to the electrochemical vessel, which is equipped with a counter electrode and the reference electrode. The specimen is contacted to the electrolyte surface using a hanging meniscus geometry. Pulse programs given to the potentiostat allow a specimen preparation at the potential of choice in the range of milliseconds or less up to hours. A typical specimen preparation starts with cleaning by argon ion sputtering within the preparation chamber. Then the sample is transferred to the electrochemical chamber and contacted to the electrolyte at sufficiently negative potentials so that one starts with an oxide free clean surface. The potential is then pulsed to a value of interest and then returned to a value where the sample is not expected to change further. Finally, it is washed with degassed water and transferred to the analyzer chamber for its surface analytical investigation. This system has been applied to many studies of corrosion and passivity. It is even suited for the study of the electrochemical double layer by XPS when the electrode can be removed from the electrolyte under water repelling (i.e., hydrophobic) conditions.[4,7] In this case, the surface is not washed by water after its preparation. For these studies, one may follow with XPS, both qualitatively and quantitatively, changes in the chemical composition of the double layer with the electrode potential and may also investigate the changes in the energy of XPS features with the potential, which is related to the position of the species within the double layer.[8] The system also contains an additional port on the fast entry lock to allow for introduction from a transfer vessel of samples which have been prepared within a glove box or in an *ex situ* experimental stage, enabling exposure to aggressive gases, such as HCl. Together with a shuttle, the spectrometer has been applied successfully for high temperature corrosion research in aggressive gas phases.[9]

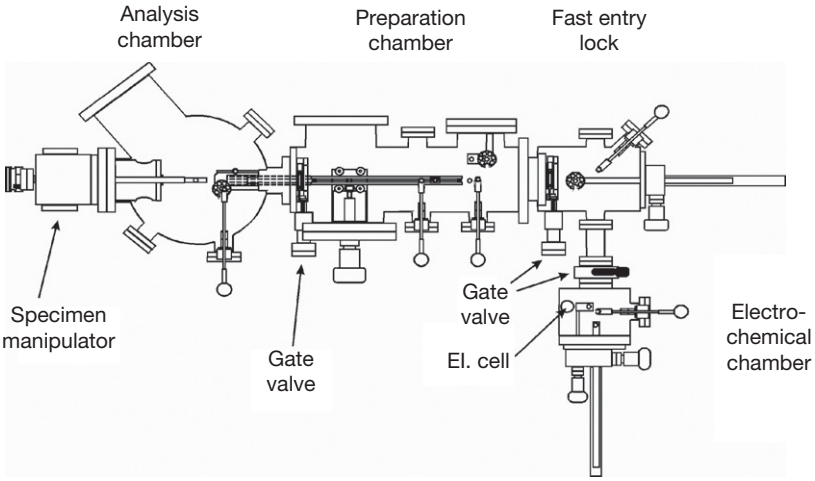


Figure 1 Schematic representation of a commercial XPS Spectrometer with three vacuum chambers, and an attached electrochemical chamber with an electrochemical vessel. A rack and pinion drive system permits a fast and flexible specimen transfer.[4–6]

2.31.3 X-ray Photoelectron Spectroscopy

XPS is a widely used soft surface analytical method, which usually does not introduce changes due to the analysis itself. An XP-spectrometer consists of an X-ray source, an energy analyzer, and a detector. A monochromatic X-ray beam leaves the X-ray source, strikes the specimen surface, and ejects electrons. Typically, a twin anode X-ray source is used, providing either Mg K_α ($E = 1253.6\,eV$) or Al K_α ($E = 1486.6\,eV$) radiation. The kinetic energy of the photoelectrons is measured by an electrostatic energy analyzer. Modern spectrometers use a spherical or a spherical sector analyzer and an electrostatic lens focusing the electrons to its entrance slit (see **Figure 2**). The standard detector is a channeltron, an electron multiplier of special design with a potential drop of 3–4 keV across its surface. In most cases, an array of several channeltrons is used, or a channel plate with a two-dimensional arrangement of many small channeltrons on a plate to increase the sensitivity. The spectral data emerging from the channeltrons are sorted out with respect to their energy and added to each other appropriately with the help of the computer. The analyzer consists of a spherical condenser supplied with a voltage which allows electrons with the appropriate pass energy to travel through it via its entrance and exit slits. It may be seen as a band pass filter which permits the transfer of only those electrons with the correct energy, the so-called pass energy E_{Pass}. A linearly changing retardation voltage between the specimen, usually at ground, and the entrance slit is applied, so that all electrons have the chance to get successively through the analyzer when a spectrum is taken. Due to the specific energetic situation between the specimen and the analyzer, the energy balance of the photoelectrons follows eqn [1].

$$h\nu = E_B + E_{kin} + e\Phi_A \qquad [1]$$

The energy $h\nu$ of an absorbed X-ray photon is used to ionize an electron with binding energy E_B from an atom at the surface. For solid specimens, E_B is measured relative to the Fermi level E_F. Any excess hν energy, beyond that required for ejection, provides the electron with its kinetic energy E_{kin}, which is measured relative to the energy analyzer. **Figure 3(a)** illustrates the energy balance of the photoelectron of eqn [1]. It contains the work function $e\Phi_A$ of the analyzer, which is a constant of the spectrometer that is compensated internally after its calibration. With all

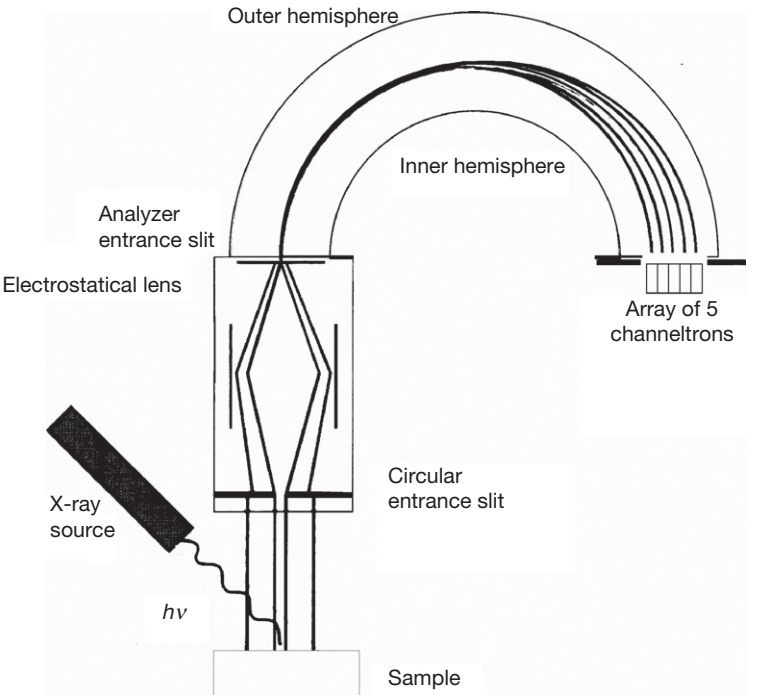


Figure 2 Schematic diagram of an XPS spectrometer with an X-ray source, focusing electrostatic lens, spherical sector analyzer, and a channeltron array as detector.[6]

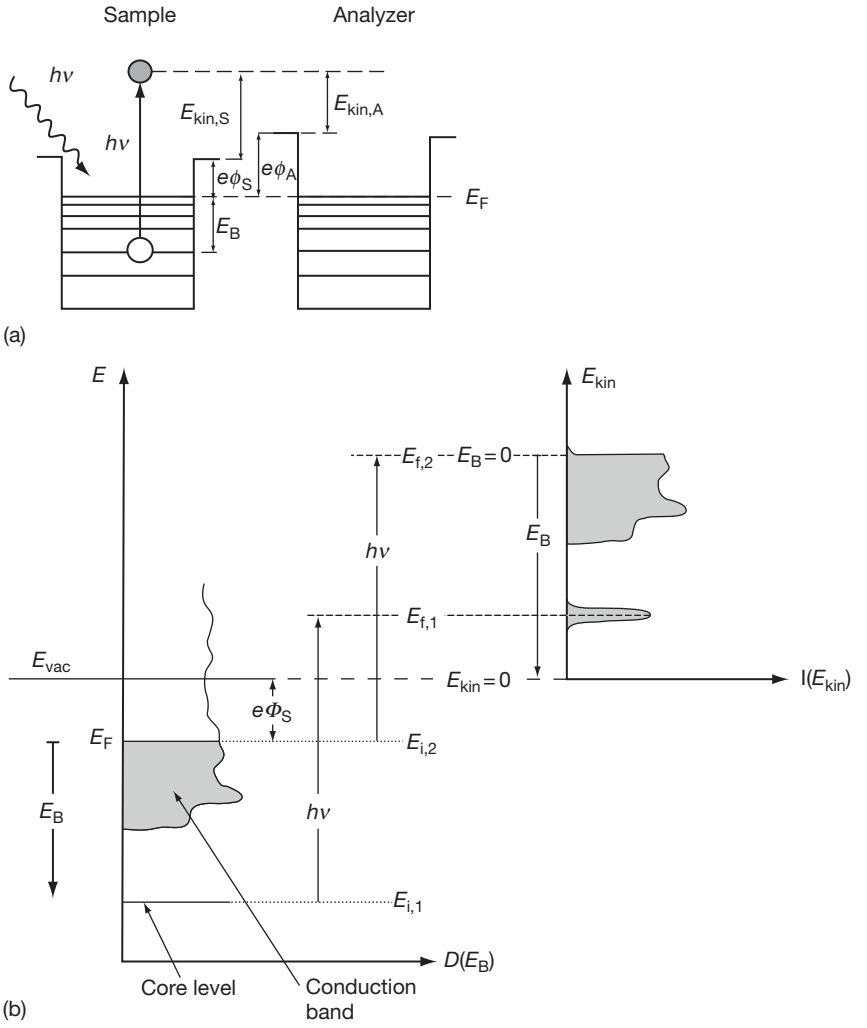


Figure 3 (a) Energy diagram for XPS showing the potential wells of the specimen and the analyzer along with electronic levels, Fermi energy E_F, the related work functions $e\Phi_S$ and $e\Phi_A$, contact potentials difference Φ_A–Φ_S, and the related kinetic energies $E_{kin,S}$ and $E_{kin,A}$. A contact situation between analyzer and the specimen is shown with equal Fermi energies E_F.[10] (b) Mechanism of XPS with exciting X-rays of energy $h\nu$ and electrons emitted from a core level (a) and from the conduction band (b) with initial energies E_i and final state energies of E_f. The resulting XP-spectrum is displayed, indicating the binding energy E_B and kinetic energy E_{kin} of the photoelectrons. For simplicity, the contact potential difference of the specimen and the analyzer as well as the retardation voltage are omitted.[10]

other quantities known, E_B is given by eqn [1]. Usually, a spectrometer is calibrated relative to the XPS-signals of well characterized standards mostly of pure metals such as E_B $(Au4f_{7/2}) = 84.00\,eV$, E_B $(Ag3d_{5/3}) = 368.26\,eV$, E_B $(Cu2p_{3/2}) = 932.67\,eV$.

In most cases, a spectrum contains several XP signals from each of the elements at the surface of a specimen due to their characteristic electronic levels, as indicated in **Figure 3(b)**. It also contains X-ray induced Auger features, which arise through filling of the ejected photoelectron's core hole with an electron from a higher energy level. As shown in **Figure 4**, the excess energy resulting from this transition may lead to the ejection of a third electron, the Auger electron. Another competing process is the emission of a photon, that is, X-ray fluorescence. Heavier elements ($Z > 35$) have a high Auger yield so that their XPS spectra contain strong Auger lines.

The energy resolution of the XP-spectra is limited by the width of the excitation lines, which is ~0.8 eV for Mg K_α and 1.0 eV for Al K_α radiation. The exact position of an XPS-signal contains a chemical shift

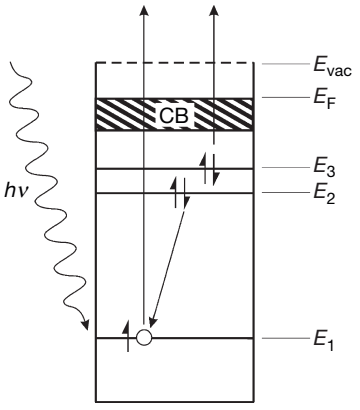


Figure 4 Auger process with photoionization at level E_1, transfer of an electron from E_2 to E_1 and ejection of an Auger electron from level E_3.[10]

which has been found and interpreted first by Kai Siegbahn. An increasing charge on the atom causes a slightly higher ionization energy for any core level and thus a smaller kinetic energy or a higher binding energy E_B. As a result, one may measure energy shifts of the XP-signal, the so-called chemical shifts, which informs one about the charge of an ion or the oxidation number of an element. Electronegative binding partners similarly induce a more positive charge and thus a higher E_B. The chemical shift is often in the range of the width of the signals. The natural width of the exciting X-rays may be improved with an X-ray monochromator to ~0.5 eV, *albeit* at the expense of intensity loss. Modern energy analyzers and sensitive detector arrays and channel plates will compensate for these losses. In many cases, one has to deconvolute a measured signal in order to separate it into the contributions of the different species. The integration of these contributing peaks yields a quantitative measure for the amount of the related species.

XPS analysis usually starts with a background subtraction, which is often achieved by a linearly increasing line or a correction according to Shirley,[11] that is, a curve with background contributions proportional to the preceding part of the signal. The background corrected signal is then synthesized with peaks of well characterized standards, which are usually described by Gauss–Lorentzians with appropriate parameters. The energy and shape of these peaks are usually kept constant with an adjustment of their height to get the best fit of the synthesized to the measured signal.[10] Usually, relative signal intensities are used (i.e., intensity ratios) in order to compensate for spectrometer characteristics such as sensitivity of the detector, the intensity of the X-ray source, or the intensity/energy performance of the energy analyzer. The intensity ratio of two elements A and B is given by eqn [2]:

$$\frac{I_A}{I_B} = \frac{\sigma_A T_A \lambda_A D_A}{\sigma_B T_B \lambda_B D_B} \quad [2]$$

D is the atomic density of components A and B. The photoionization cross-section σ takes care of the relative sensitivity of elements A and B to the photoionization process and is given by the data of Scofield.[12] The mean free path of the electrons λ within the specimen generally increases with kinetic energy and is given by the empirical relation of Seah and Dench[13] (eqn [3]).

$$\lambda = B\sqrt{E_{kin}} \quad [3]$$

Parameter $B = 0.054\,\text{nm}\,\text{eV}^{-0.5}$ for elements, $B = 0.097\,\text{nm}\,\text{eV}^{-0.5}$ for inorganic compounds, and $B = 0.087\,\text{nm}\,\text{eV}^{-0.5}$ for organic compounds. The transmission function T depends on the specific design of the spectrometer and changes with the pass energy E_{Pass} and the kinetic energy E_{kin} of the photoelectrons. Equation [4] gives the transmission function T for the VG ESCALB Mk2 analyzer as an example.

$$T = \omega^2 0.01 E_{Pass}^{1.5} E_{kin}^{-0.5} \quad [4]$$

ω is the angle of acceptance of the analyzer for the electrons. In this case, the square root dependence of T and λ on E_{kin} cancel each other for eqn [2].

A frequent problem in corrosion is the examination of thin passive layers, that is, some few nanometers thick anodic oxide films on a metal substrate. The XPS intensity ratio of the cations and the metal atoms provides a possibility to calculate the oxide thickness. The parameters σ, λ, and T in eqn [2] cancel, in this case due to the ratio of the same element. However the signal I_{ox}^{Me} from the cations of Me within the oxide are submitted to a self attenuation factor $[1 - \exp(-d/\lambda_{ox}^{Me}\cos\theta)]$, and the substrate signal I_{Me}^{Me} is attenuated by a factor $\exp(-d/\lambda_{ox}^{Me}\cos\theta)$ due to the oxidic layer above. This attenuation is a consequence of inelastic energy losses of the photoelectrons on their way through the surface layers to the vacuum. d is the layer thickness, and Θ the angle between the direction of the detector and the surface normal of the specimen. The intensity ratio of eqn [5] indicates an increase with increasing angle Θ. This is simply due to increased masking of the electrons from the substrate, because their path through the thin oxide layer increases with Θ. Equation [6] permits the calculation of the thickness d of the passive layer.

$$\frac{I_{ox}^{Me}}{I_{Me}^{Me}} = \frac{D_{ox}^{Me}}{D_{Me}^{Me}} \cdot \frac{[1 - \exp(-d/\lambda_{ox}^{Me}\cos\theta)]}{[\exp(-d/\lambda_{ox}^{Me}\cos\theta)]} \quad [5]$$

$$d = \lambda_{Me}^{ox}\cos\theta \ln\left[\frac{I_{ox}^{Me} D_{Me}^{Me}}{I_{Me}^{Me} D_{ox}^{Me}} + 1\right] \quad [6]$$

In the case of an oxide layer on a metal substrate, such a calculation is quite trivial, but in a multilayered system, one must decide where the various cations and anions are located relative to each other. Angular resolved XPS measurements (ARXPS) can be used to decide the distribution of species perpendicular to the surface plane. As an example, **Figure 5** presents ARXPS results from a passive layer on Fe formed at $E = -0.16$ V in 1 M NaOH for various times in the range of 100 ms to 100 s. It is clear that a majority of Fe(III) ions are located above the Fe(II) ions, as suggested on the basis of thermodynamics.[15] One may further see that the passive film develops over this period of time. At 100 ms, almost no Fe(III) is present, and the bilayer structure is still not developed. With increasing time, the intensity ratio and the upward bending increase due to the growth of Fe(III) on top of Fe(II), and the Fe(II) to Fe(III) oxidation.

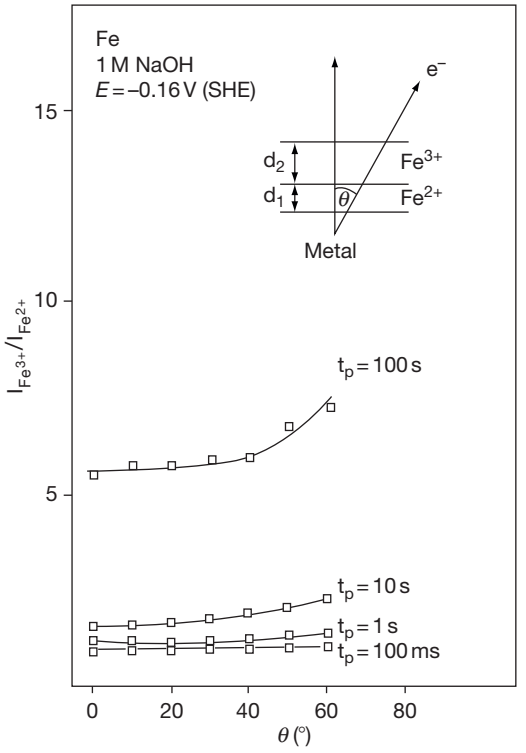


Figure 5 ARXPS intensity ratio of Fe(III)/Fe(II) intensities for Fe passivated at $E = -0.160$ V (SHE) in 1 M NaOH for different times t_P as indicated.[6,14]

At 100 s, the passive film has its final form and thickness. These data have been evaluated to obtain the thickness of the Fe(II) and Fe(III) layers for anodic oxidation at all potentials from the prepassive to the transpassive range as presented in **Figure 6**.[14] The film composition follows the characteristic peaks of the polarization curve of sputter cleaned iron in 1 M NaOH, which is also displayed in **Figure 6**. Film growth starts in this alkaline electrolyte at $E_{P1} = -0.80$ V. Below $E = -0.20$ V, the film contains a large fraction of Fe(II) which is oxidized at $E_{P2} = -0.20$ V to Fe(III). At $E > 0.20$ V, only a small residue of Fe(II) ions remain located below a growing Fe(III) film, which is at the detection limits of the method. These XPS results confirm the thermodynamic predictions of the composition of the passive layer. E_{P2} is the passivation potential of Fe in acidic

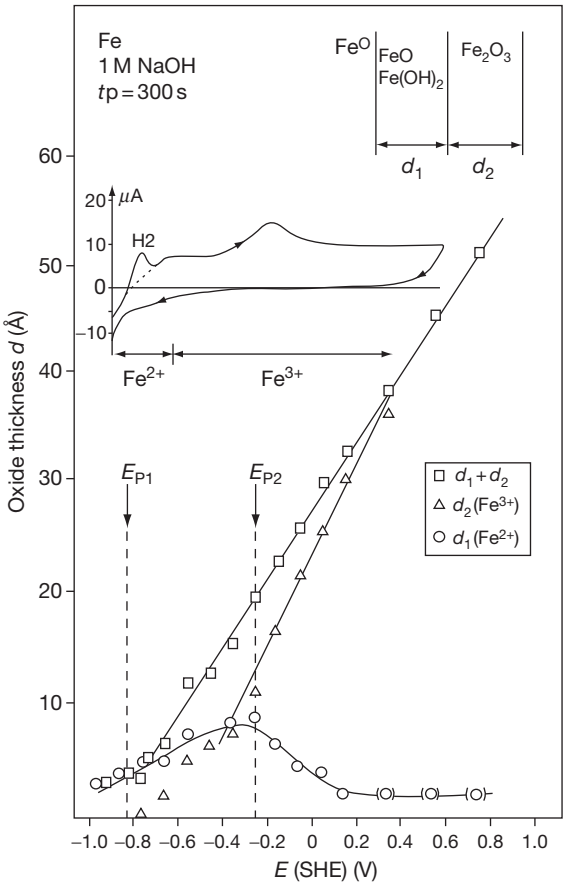


Figure 6 Evaluation of XPS studies of Fe passivated in 1 M NaOH for 300 s on the basis of a bilayer model with Fe(II) and Fe(III) layer thicknesses of d_1 and d_2, respectively. Polarization curve of sputter cleaned Fe within the spectrometer with indication of the passivation potentials E_{P1} and E_{P2}.[6,14]

solution, the so-called Flade potential extrapolated to pH = 13 with a −0.059 V/pH dependence. It was interpreted as the Fe_3O_4 to Fe_2O_3 oxidation with thermodynamic data, which is confirmed by the XPS results.[14,15] In acidic electrolytes, the Fe(II) rich layer between E_{P1} and E_{P2} is not protective and dissolves immediately. However, it is protective in alkaline solution. These results may be obtained by the exclusion of air only, that is, by an investigation in the closed XPS system as described above. Fe(II) would be oxidized immediately to Fe(III) by traces of oxygen so that this species otherwise could not be found. Many passive layers have been examined in detail on pure metals and binary alloys as a function of the potential, time of passivation and electrolyte composition, in particular pH. A compilation of data and literature can be found elsewhere.[6] **Figure 7** presents schematically the multilayered structure of passive layers on several different pure metals and binary alloys obtained by a detailed XPS analysis along with ion scattering measurements for completion. Higher valent species are found for sufficiently positive potentials only. Usually, hydroxides are outside and oxides inside. Higher valent cations are located in the outer part of the films.

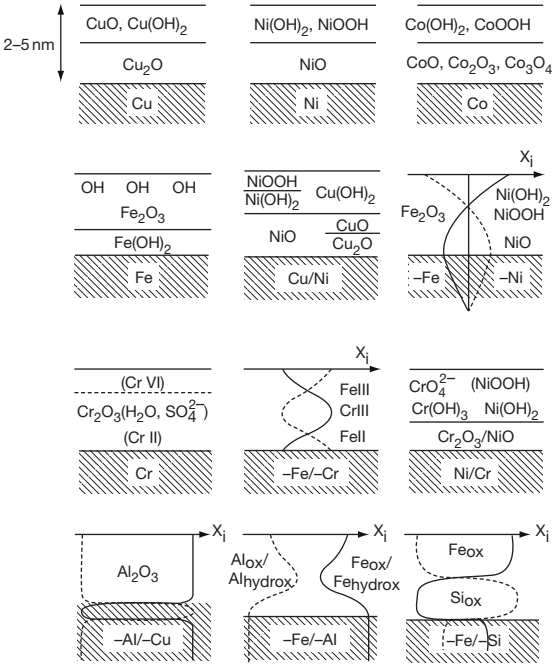


Figure 7 Schematic diagram of passive layers on various metals and binary alloys deduced from XPS and ion scattering measurements demonstrating their multilayer character.[6]

The lateral resolution of XPS is usually poor and corresponds to the diameter of the exciting X-ray beam (several mm). The inlet slit to the analyzer may reduce the accepted surface area of the sample to some extent. Furthermore, the beam may be focused, especially together with a monochromator, with an ellipsoidal diffracting crystal down to several micrometers only. New instruments also contain an imaging option. This can be achieved by rastering a focused X-ray beam across the sample surface. Another design uses lenses at the entrance and exit of the analyzer to focus an image of the whole specimen or a major part of it onto a channel plate, which registers an image of the elemental distribution at the surface. Modern spectrometers allow a switching from a spectroscopic mode registering XPS to an imaging mode registering the elemental distribution of any element, which is chosen by its binding energy at the spectrometer control unit.

In addition to the basic research, XPS may also be used for a wide variety of industry problems. Besides the investigation of corrosion problems and corrosion layers, it may be used to study polymers, semiconductors, and insulators. One may investigate contaminations and surface problems in the production of any kind. Many laboratories of surface analysis help the industry and other institutes to provide a fast solution for their problems. A serious component of such activity is, ensuring that the specimen is obtained and prepared without the contamination of its surface in order to get results which are relevant to the technical situation.

2.31.4 Ultraviolet Photoelectron Spectroscopy

In ultraviolet photoelectron spectroscopy (UPS), a surface is irradiated with UV light, usually the He I (21.2 eV) or the He II line (40.8 eV), from a lamp attached to the analysis chamber of a spectrometer. The UV-light of the helium discharge is transmitted via a capillary to the specimen's surface, which permits a sufficiently high transmission without ruining the vacuum. As the excitation energy is low compared to the XPS, one only gets contributions to the photoelectron spectrum from low binding energy levels, including the valence band. Thus, UPS mirrors the broad distribution of states of the conduction band of a metal and the valence band of a semiconductor or insulator. Consequently, interpretation of the data can rather be more involved than

for XPS. One straightforward, but important, application of UPS is the determination of the work function $e\Phi_W$ of a metal, or the threshold energy $e\Phi_{Th}$ of a semiconductor or insulator ($e\Phi_{Th}$ = difference between the upper valence band edge and the vacuum level). **Figures 8(a) and 9** show UP-spectra of Cu–metal without and with a thin passive layer of Cu_2O, respectively. One may see clearly the contribution of the surface oxide to the spectrum. **Figures 8(b) and 8(c)** show the potential wells of Cu metal and the analyzer with the related work functions similar to **Figure 3(a)**. Application of a negative bias to the specimen of $\sim U = -5$ V ensures that all photoelectrons reach the analyzer (**Figure 8(c)**). This bias negates a potential barrier which the electrons would otherwise have to overcome, as presented in **Figure 8(b)**, causing a drop off of the UV-spectrum at the low energy side close to $E_{kin} = 0$ eV. There exists the obvious relationship, indicated in eqn [7], between the work function $e\Phi_W$, the kinetic energy E_{kin}, and the width ΔE of the spectrum. The cutoff edge in the biased spectra corresponds to photoelectrons with $E_{kin} = 0$ eV, which have been excited from deeper levels within the valence band, or which have lost their kinetic energy due to inelastic interactions. Similarly, eqn [8] allows one to deduce the threshold energy $e\Phi_{Th}$ of a semiconductor or insulator. These equations and data have been used to determine $e\Phi_W$ of Cu and $e\Phi_{Th}$ of its anodic oxides and thus the absolute value of the Fermi energy of a bare Cu electrode immersed from the electrolyte at appropriate potentials and of the upper valence band edge of its anodic oxides.[6]

$$e\Phi_W = h\nu - \Delta E \quad [7]$$

$$e\Phi_{Th} = h\nu - \Delta E \quad [8]$$

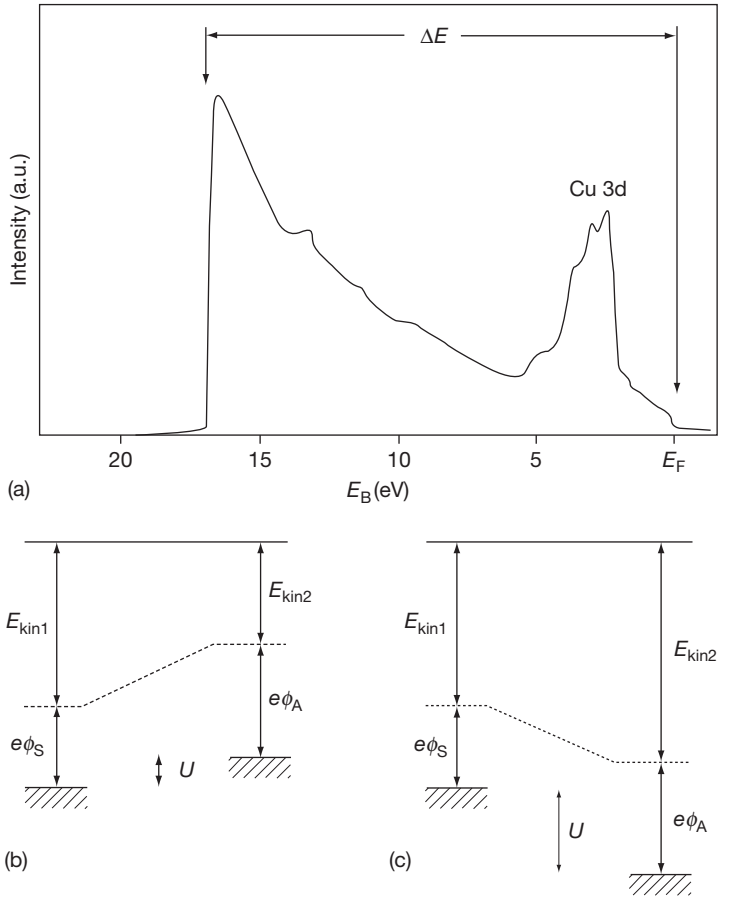


Figure 8 (a) UP-spectrum of sputter cleaned copper with a width ΔE. (b) and (c) Energy diagrams showing the work functions of the specimen $e\Phi_S$ and the analyzer $e\Phi_A$ and the related kinetic energies in direct contact, and with the application of a negative bias U to overcome the contact potential difference.[10]

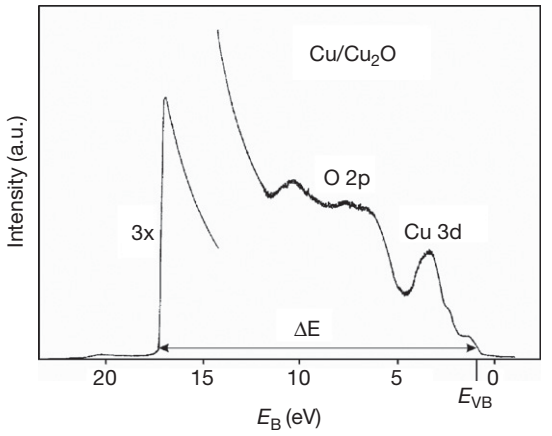


Figure 9 UP-spectrum of an anodic Cu_2O layer on Cu showing the O bands and a step related to the valence band edge.

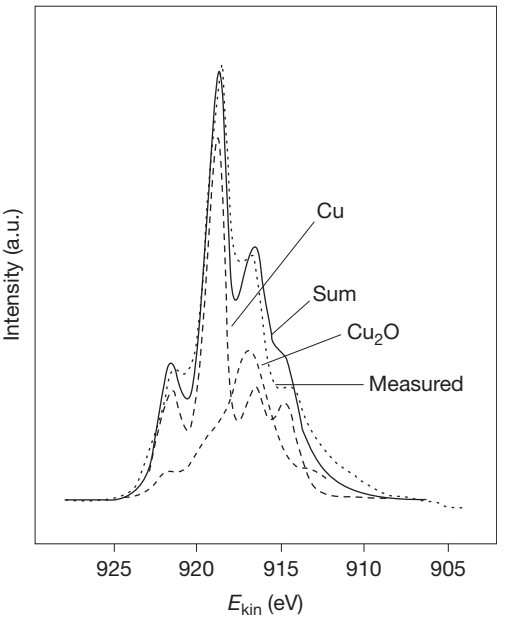


Figure 10 X-ray induced LMM Auger signal of a Cu specimen passivated in 1 M NaOH with a ~2 nm Cu_2O overlayer, and its separation into contributions from Cu and Cu_2O, respectively,[10,16] using reference compound data.

2.31.5 Auger Electron Spectroscopy

If an electron is emitted from its orbital, the remaining hole may be filled by a more energetic electron. The energy difference between the two electrons leads to characteristic X-ray fluorescence or to the ejection of a third electron, the Auger electron, as competing processes (**Figure 3**). Such emission may be initiated by X-rays or incoming electrons, corresponding to the X-ray and electron induced AES. In the XPS subsection, it has been mentioned briefly that Auger lines are present in XP-spectra. These Auger lines are sometimes relied upon for the determination of the chemical state, when the chemical shift of XP lines is so small that it cannot be resolved with XPS, as is the case for Cu metal and Cu(I) oxide. In **Figure 10**, the Cu LMM line acquired from a thin Cu_2O film on Cu has been deconvoluted with the aid of the Auger lines from appropriate Cu reference compounds. From these quantitative data for the two Cu species, the thickness of the Cu_2O layer has been determined in good agreement with electrochemical results.[16] However, the AES lines are usually more complicated due to the three electron process of AES which requires more effort for a reliable quantitative evaluation. Therefore, AES is used for a chemical analysis mainly if the chemical shift for XPS is too small to distinguish between the two species.

Electron induced AES is of greater importance for surface analysis. The beam from an electron source may be focused by a lens to a sufficiently small spot, so that one may get AE-spectra with very high lateral resolution of ~10 nm. Modern high intensity electron sources use thermoionic emission or field assisted emission of electrons as, for example, from LaB_6 filaments or pin shaped Schottky emitters. Given these source characteristics, AES has been used for local spot analysis, line profiles, or even an elemental mapping, that is, imaging with respect to the elements of interest by scanning the beam. In corrosion, many localized effects are of decisive importance like inclusions and local attack at these sites, pitting and crevice corrosion, and grain boundary segregation and grain boundary attack. In all these cases, information on the distribution of metal components, impurities, and corrosion products is important to understand the leading mechanisms which require surface analysis with high lateral resolution.

Conventional electron induced AES yields the first derivative of the Auger lines. This is a consequence of modulation of the pass energy of the energy analyzer, usually by 5–10 mV together with the application of lock-in techniques. The obtained signal is therefore the first derivative of the Auger line with respect to the energy scale. The advantage is removal of a large and especially seriously changing background, which is the case for Auger signals of elements like Si at low kinetic energy. The disadvantage of differentiated signals is the loss of some

information. In modern spectrometers, the sensitivity is high enough to measure the nondifferentiated signal, which provides the advantages mentioned above.

Auger lines are named using the electronic levels involved in the process, for example, KLL. The first letter describes the orbital which is ionized by the impinging electrons or X-rays, the second is associated with the level from which an electron falls into the hole, and the third is the level from where the Auger electron is emitted. A KLL process starts with the ionization of a K shell, followed by a transfer of a L-electron to the core hole, and finally transfer of the energy difference to a third electron from the L-shell. The energy of the Auger electron is thus determined by the three involved energy levels. Due to spin–orbit coupling, the p- and d-electrons of the L and M shells are split into different levels, for example, L_1–L_3 (s, $2p_{1/2}$ and $2p_{3/2}$), leading to the nomenclature $KL_1L_{2,3}$ for an Auger line. Thus, Auger lines are more complicated compared to XPS and therefore a quantitative evaluation including different oxidation states requires more effort. If electrons of the valence band are involved, their description contains V, for example, LVV for an electron ejected from the L shell with two additional electrons involved from the valence band.

2.31.6 Ion Spectrometry

The backscattering of noble gas ions has been developed into techniques for elemental analysis of surfaces and surface films. Two main approaches are described here: the backscattering of He- or Ne-ions with a primary energy of 1–5 keV (ISS, or low energy ion scattering (LEIS)) and the backscattering of He with a primary energy of 2 MeV (RBS). In addition, SIMS will also be discussed in this subsection.

2.31.6.1 Ion Scattering Spectrometry or Low Energy Ion Scattering

In ISS, only impinging ions, which are backscattered by the outermost atomic layer of a surface, have a small, but significant, probability of emerging charged. At the second or any deeper layer, essentially all noble gas ions pick up an electron and are scattered as neutral atoms. Given that only charged particles are detected, using an electrostatic energy analyzer as employed for XPS but with reversed polarity, ISS detects the composition of the outermost surface layer. A soft sputter process yields a depth profile of the surface region with a monolayer resolution. The energy loss suffered by scattered ions of mass m is described by eqn [9a], where E is the ion energy after scattering, and E_0 the primary energy. If the backscattering angle between the primary beam and the scattered ions is $\Theta = 90°$, eqn [9a] simplifies to eqn [9b]. In this geometry, each target mass M will lead to an energy loss with a characteristic E/E_0 ratio depending on the involved masses m and M only.

$$\frac{E}{E_0} = \frac{m^2}{(m+M)^2}\left\{\left[\cos\theta + \left(\frac{M^2}{m^2} - \sin^2\theta\right)\right]^{1/2}\right\}^2 \quad [9a]$$

$$\frac{E}{E_0} = \frac{M-m}{M+m} \quad [9b]$$

Thus, the energy spectrum of the backscattered ions leads to a peak for each target mass M. The integrated peak areas provide a measure of the amount of atoms within the surface. After calibration for the backscattering cross section, that is, the relative sensitivity of the elements, one may calculate quantitative results for surface concentrations. In combination with sputtering, one gets a quantitative depth profile. **Figure 11** shows the depth profiles of Cr for a Fe15Cr

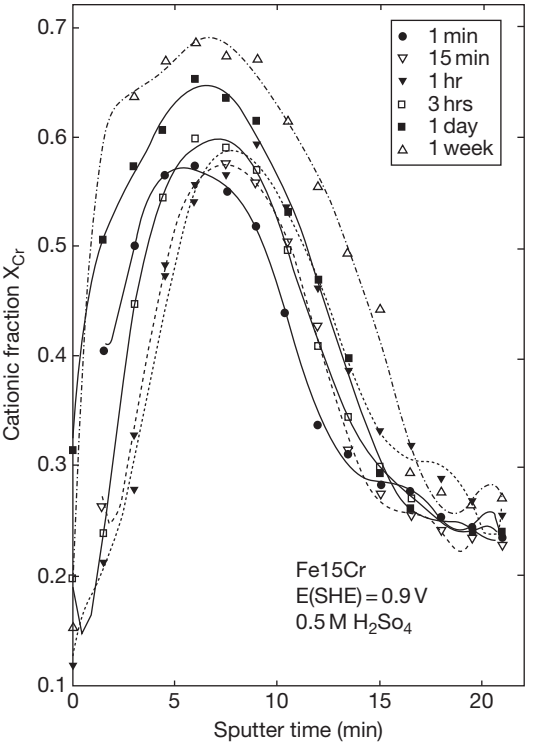


Figure 11 Quantitative ISS depth profiles of Fe15Cr passivated at $E = 0.90$ V in 0.5 M H_2SO_4 for different times in the range of 1 min to 1 week.[6]

electrode passivated in 0.5 M H_2SO_4 at $E = 0.90$ V for different times in the range of 1 min to 1 week.[6] The profile demonstrates that the passive layer is significantly enriched in Cr. At the right side, it levels off to 15% Cr for the bulk metal alloy. The outermost surface is only slightly enriched in Cr. Apparently, the Cr content accumulates on the surface region during passivation. Also, the maximum gets slightly higher with time from 58% for 1 min to 70% Cr for 1 week. These results agree well with those obtained by XPS. In contrast to XPS, which provides chemical information, the advantage of ISS is the high depth resolution for this 2 nm film. Therefore, both methods are complementary to each other. The high enrichment of Cr(III) ions within the layer is a consequence of their extremely slow dissolution rate, 1–2 orders of magnitude smaller than Fe(III). This property in turn is the reason for the excellent stability of FeCr alloys and FeCrNi stainless steel in comparison to iron.

2.31.6.2 Rutherford Back Scattering

For RBS He ions are accelerated by a high voltage to 2 MeV, provided by a Van der Graaf generator. These high energy ions may penetrate materials to a maximum depth of ~1 μm. They are backscattered at the nuclei of atoms at different depths and then get back to the surface where they are measured at a backscattering angle of $\Theta = 175°$ by a solid state detector. The ring shaped Si-detector contains a hole for the primary impinging beam. Two kinds of energy losses are obtained, the loss by the backscattering process which follows eqn [9], and the losses due to 'inelastic' processes of the ions on their way in and out. The latter depend on the kinetic energy and may be taken from tables for the stopping cross sections. As a result, one obtains broad signals whose leading edge is characteristic of the mass of the target atoms and whose width corresponds to the depth. The height is proportional to the concentration of the atoms when corrected for the element's sensitivity. The intensity increases with the square of the atomic number. **Figure 12** shows an example of RBS from an Al 1%Cu vapor deposited film, before and after anodization to 100 V in citrate buffer pH 6.0.[6,17] The leading edge of Cu is at high energy due to its larger mass well separated from the light metal Al. A step in the Al signal after anodization corresponds to the grown Al_2O_3 anodic film with a smaller Al-concentration compared to the metal substrate. Its width is a measure of the oxide thickness.

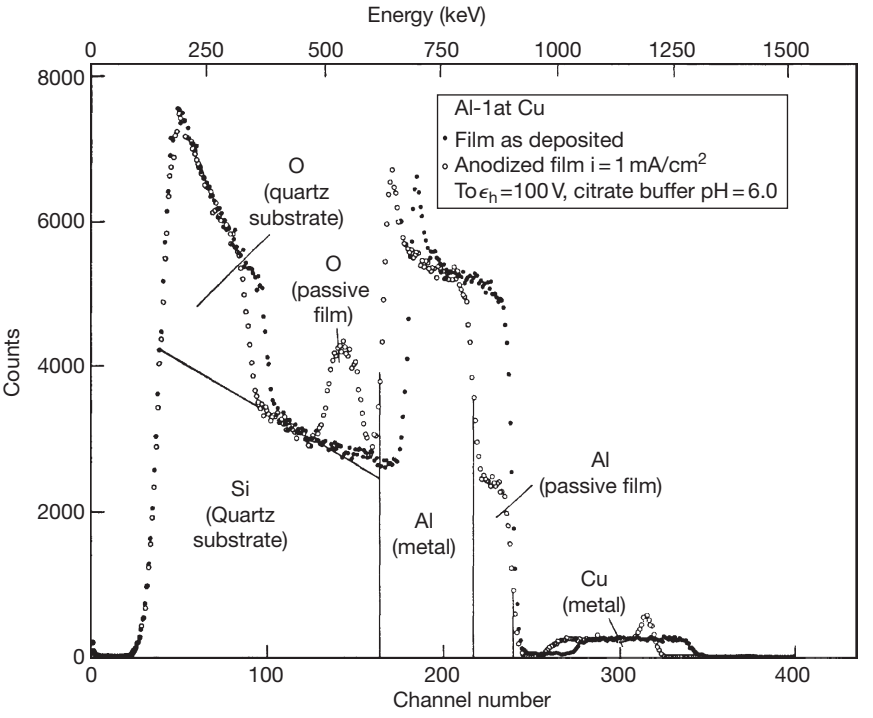


Figure 12 RBS spectrum of vapor deposited Al 1 at.% Cu film on quartz substrate before and after passivation with $I = 1\ \text{mA cm}^{-2}$ at 100 V in citrate buffer pH = 6.0.[17]

The oxygen signal of the anodic oxide is superimposed on the signal of Si from the quartz substrate for the vapor deposited metal film. The leading edge of the Cu signal shows a shift to smaller energies, due to the formation of the anodic oxide which is Cu free. A superimposed small peak close to its leading edge corresponds to Cu accumulation at the metal surface underneath the oxide. In the case of anodization of these Al/Cu films Cu accumulation at the metal surface finally causes a film breakthrough and a destruction of the vapor deposited metal film. This example demonstrates that one obtains nondestructive depth profiles of corroding surfaces by RBS which explain the details of elemental distribution and their changes during the reactions.

2.31.6.3 SIMS and Related Methods

If an ion beam strikes a surface, the energy of the impact is transmitted to atoms and molecules of the specimen, which causes their transfer to the vacuum. These ions may be submitted to mass analysis for their identification. For the mass separation, quadruple mass analyzers may be used instead of magnetic sector analyzers. Another possibility is time-of-flight mass spectrometers (TOF-SIMS). In this case, a pulsed ion beam creates bunches of secondary species, which are accelerated to a fixed kinetic energy. Given $E_{kin} = 1/2mv^2$, lighter ions travel faster, and so arrive at the detector quicker than the heavier ones. These ions are detected by a channeltron or a channel plate. Mass spectrometry is very sensitive, so traces of species may be detected. Furthermore, the method may distinguish between the various isotopes of one element and may also separate species with only slight mass changes, less than one unit due to the mass defect. It is also one of the few methods which may detect hydrogen which is not seen by methods such as XPS or AES. A frequent application is the determination of dopants in semiconductors and their depth profiles which may be measured by ion sputtering and SIMS. The method is also important for electrode kinetics and corrosion research. As isotopes may be separated by the method, one may study the incorporation of oxygen during anodic film growth on an already preexisting layer by SIMS depth profiling. The distribution of the O^{18}-isotope within anodic oxide grown in O^{18}-enriched water may help distinguish, where new oxide is formed at the metal surface or at the oxide–electrolyte interface. This enables the measurement of transfer rates of oxygen anions versus metal cations within the preexisting anodic oxide film. It has been found that oxygen ions can migrate inward[18] while metal cations are moving outward. These results have been confirmed by the study of the movement of Xe-implants in oxide films by RBS during further oxide growth. High intensity ion sources like liquid metal ion sources with Cs and Ga permit a very small focus of the beam from 5 μm to <20 nm. These ion guns facilitate the study of corrosion processes with high lateral resolution and allow imaging by SIMS showing the lateral distribution of elements with appropriate resolution.

SIMS may be used as a quantitative method. However, the SIMS signal depends strongly on the matrix, that is, the electronic conditions at the specimen's surface. A metal with a high concentration of electrons at the Fermi level causes a high probability of neutralization via tunnel processes for the ions immediately in front of the surface. Therefore, the SIMS signal of cations is very small for pure metal surfaces, whereas it gets much higher if the specimen is covered with a thin semiconducting or insulating oxide film, due to its band gap where no electrons are available. The SIMS signal may vary for the same concentration of species by orders of magnitude. However, it will be fixed if the interesting species is a dopant in a matrix which does not change during ion impact. This is the case for the analysis of dopant implants in semiconductors. One possibility to stabilize the SIMS efficiency is the inlet of small amounts of oxygen to the surface of the metal forming a thin layer of semiconducting oxide. Another is the ionization of the neutrals after their sputtering with an electron beam, or by the stabilized formation of secondary ion species by a glow discharge (GD).

An ion beam may also be used to remove large ionized molecules from a metal surface. These processes have been studied in detail by SIMS of large organic molecules deposited on Ag. In this case, one obtains the M^+, the $[M + H]^+$ and the AgM^+ signals, besides fractions of the molecule M. Apparently, the energy of the ion beam is distributed via a cascade to the site of the molecule, which then finally is responsible for its transfer to the vacuum. The chemical interaction of the molecules with the silver surface causes the AgM^+ peak. A very similar mass spectrometry of organic molecules is matrix assisted laser desorption (MALDI). For this analytical method, the organic molecules are codeposited with a matrix molecule at a metal surface, mostly a derivative of benzoic acid. The matrix is vaporized by a pulsed laser beam which ionizes and transfers the embedded molecules together with the matrix to the vacuum.

After acceleration by a voltage in the range of 10–30 keV, the molecules are analyzed and detected by a TOF-mass spectrometer. Very large masses have been detected by MALDI-TOF up to 10^6 mass units. TOF mass spectrometry is an excellent tool to study organic films, inhibitors and polymers due to the high mass range of up to 10^6 atomic mass units which is accessible to this method.

2.31.7 Methods Using Hard X-rays: X-ray Diffraction (XRD), X-ray Reflectivity (XRR), and X-ray Absorption Spectroscopy (XAS)

In the laboratory, sealed X-ray tubes are the most widely used X-ray sources, which emit continuous Bremsstrahlung-radiation as well as characteristic X-ray emission lines. The X-rays are excited by an accelerated electron beam impinging on a metal target within a vacuum tube. Although the maximum electrical input power of such an X-ray tube is typically about several kilowatts, the resulting X-ray intensities which are measured on a sample surface are relatively weak, because typically, only fractions of a percent of this electric input power is transformed into X-ray radiation leaving the tube through X-ray windows (typically beryllium). This is a serious drawback for *in situ* investigations of electrodes and related corrosion processes, since the parasitic absorption by the electrolyte further reduces the intensity which is available for the experiments. Assuming, for example, a rotating anode X-ray tube which provides $\sim 10^6$ monochromatic X-ray photons per second per square millimeter, a Bragg reflectivity of a real sample (such as a passive film on a metal substrate) of $\sim 10^{-3}$% and a transmission of the electrochemical cell of ~10%, we end up with a single photon per second only. In a practical diffraction experiment, thus, an extremely large accumulation time is required for each data point in order to achieve a useful signal to noise ratio even in the photon counting mode using sophisticated detector equipment. Nevertheless, some few *in situ* experiments with laboratory equipment have been reported in literature.

Much more intense X-ray beams are provided by synchrotron radiation (SR).[19] SR is generated when charged high energy particles (usually electrons or positrons with kinetic energies of ~2–8 GeV) are moving on a curved path which is induced by a magnetic structure. Compared to X-ray anodes, SR has a continuous emission spectrum ranging from the IR to hard X-rays. It is well defined and can be calculated from the knowledge of the local curvature of the electron path and the energy of the stored electrons or positrons.[19] The quality of the emitted radiation is characterized by its brightness and brilliance. The latter parameter is defined by the number of photons which is emitted in one second from a source area of 1 mm^2 into a cone defined by 1 $mrad^2$ normalized to a spectral band width of 0.1%.[19] In Figure 13, the brilliance of some synchrotron sources is illustrated as a function of the photon energy. For comparison, the radiation emitted by a rotating anode with a point focus is $\sim 10^{10}$ photons/(s mm^2 $mrad^2$ 0.1%), which is at the lower end of the scale of Figure 13. In contrast to the discrete line spectrum which is provided by the emission lines of an X-ray anode, the SR storage ring emits over a large range of energy from the IR to hard X-rays. Thus, the energy (wavelength) for the experiments can be selected freely from this continuum and allows an optimization depending on the actual sample or problem

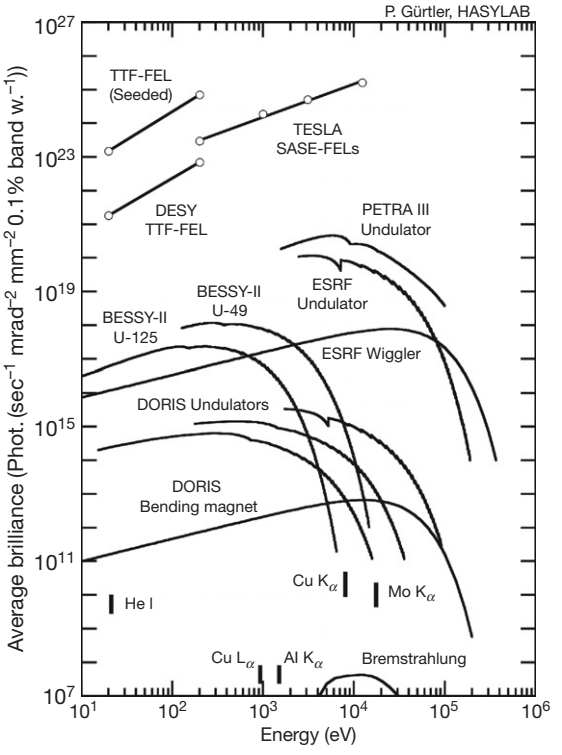


Figure 13 Average brilliance for different radiation sources at HASYLAB (DESY, Germany), BESSY (Berlin, Germany) and the ESRF (Grenoble, France). For comparison, the brilliance expected for future sources (free electron lasers) is also depicted. Courtesy of P. Gürtler, HASYLAB.

under investigation. For the application of extended X-ray absorption spectroscopy (EXAFS, see below), the use of SR is mandatory, since the photon energy is scanned over ~1 keV in these experiments. Here, even for a simple transmission mode experiment, the Bremsstrahlung-continuum of X-ray tubes is much too weak in intensity.

While second generation SR sources make use of the radiation emitted by bending magnets, sophisticated periodic magnet-structures (insertion devices, wigglers, and undulators) are installed at third generation synchrotron sources. The periodic field inside these devices forces the passing electrons and positrons on a sinusoidal path. Depending on the magnetic field strength, the magnetic gap and the number of antiparallel arranged magnets, the individual emission processes are without any phase relation (wiggler: strong field, small number of periods) or in phase (undulator: weak field, large number of undulator periods), improving the spectral brightness by at least two orders of magnitude compared to a bending magnet (see **Figure 13**). This means that SR is able to provide highly intense monochromatic radiation with extremely small lateral and horizontal divergence. These properties are an indispensable prerequisite for example, for surface X-ray scattering or diffraction experiments, especially in an electrochemical environment.

The charge carriers in a storage ring are generally accumulated in small bunches, the size of which is measured in terms of the time which is necessary to pass the focal point of the experiment. While the length of an individual bunch is of the order of several picoseconds, the distance between two bunches amounts to several nanoseconds. Therefore, SR enables time resolved experiments using this periodic bunch structure. Furthermore, due to the fact that the accelerated particles are kept in the orbit plane, the emitted synchrotron light is generally linearly polarized in this plane. This is important with regard to investigations of highly anisotropic systems such as adsorbates on single crystal surfaces. In addition, slightly below and above the orbit plane, SR-light is circularly polarized, which gives unique experimental opportunities for the investigation of magnetic materials and newly constructed insertion devices provide circularly polarized light with extremely high brilliance even in the orbit plane.

The increased availability of intense SR light sources, especially those of the new third generation, has its benefit for *in situ* investigations of corrosion processes, such as localized corrosion or passivation. *In situ* studies of electrode surfaces with X-rays generally suffer from the parasitic absorption by the electrolyte. While this effect can be tolerated at high X-ray energies of ~15 keV or more, the absorption of the electrolyte increases dramatically with decreasing photon energy. For example, for energies ~25 keV, a 10 mm electrolyte (water) layer has an X-ray transmission of ~65%, while at the energy of the Cu K_α-emission line (~8.0 keV), the transmission of an electrolyte layer of only 3 mm decreases to ~12%. While for X-ray scattering or diffraction experiments, this parasitic absorption by the electrolyte can be circumvented simply by using higher photon energies, it is a crucial drawback for XAS, since the electrode material with its characteristic absorption edges predefines the energy range which is necessary for the experiments. Depending on the scientific problem and the chosen experimental technique, different types of *in situ* cells have been applied, meeting both the experimental requirements for electrochemistry and the chosen analytical methods and techniques. A schematic representation of several different cell geometries is given in **Figure 14**. Simple transmission type cells (**Figure 14(a)**) may be used for XRD and XRR experiments, and they are also suited to the study of corrosion products within the electrolyte.[20] A number of studies of pitting corrosion applied this geometry for the investigation of artificial pits also.

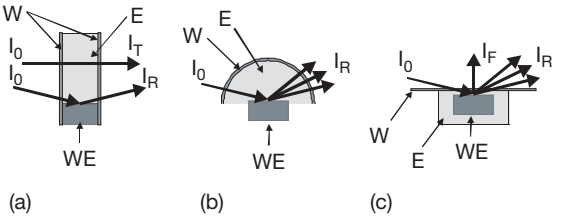


Figure 14 Schematic representations of different electrochemical cells for *in situ* studies of corrosion processes. (a) Transmission type cells which may be used for investigations of the electrode using X-ray diffraction (XRD), reflection mode X-ray spectroscopy as well as for the characterization of the electrolyte in front of the electrode, for example, for the spectroscopic identification of corrosion products. (b) Cell geometry which is optimized for diffraction experiments. (c) Thin layer cell favorable for fluorescence spectroscopy or diffraction experiments of electrode surfaces. While the electrochemical preparation is performed with an inflated window allowing an unhindered mass transport from and to the electrode surface, the thin elastic window is pressed onto the surface of the electrode for the X-ray measurements with a thin electrolyte layer remaining in order to retain potential control. WE – working electrode, W – X-ray window, E – electrolyte.

A second type of cell is optimized for XRD experiments (**Figure 14(b)**).[21] It mainly consists of a cylindrical body with a sample holder and a hemisperical X-ray window of fused silica. Due to its symmetry, all solid angles in the hemisphere above the sample surface are accessible for the incident and reflected beam. Since the X-ray path length in the electrolyte and the window material are angle independent, this cell design is ideal for diffraction, surface diffraction as well as for XRR studies.[21]

A third type of electrochemical cells reduces the electrolyte to a thin film with a typical thickness well below 10–20 μm, which is trapped between the sample (working electrode, WE) and an organic polymer film (polypropylene, Mylar, Kapton, etc.) of a few μm thickness.[22] Such a cell, presented schematically in **Figure 14(c)**, is well suited for fluorescence mode detection experiments and X-ray standing wave techniques[22,23] since the path length of the detected fluorescence photons is minimized and independent from the incidence angle Θ. However, for surface sensitive studies, the grazing incidence geometry with glancing angles of typically below 0.5° is mandatory, so that the path length of the X-rays in the electrolyte and the window material increase dramatically causing an enhanced parasitic absorption of the impinging and also the emerging radiation accordingly. As a result, *in situ* studies of electrode surfaces are currently limited to X-ray energies of above ~6000 eV, even if state of the art detector equipment such as multielement solid state detectors[22,24] and high flux insertion device beamlines at third generation storage rings are used. For investigations of dynamical processes consuming species from the solution or generating significant amounts of soluble reaction products, cells with a thin layer geometry cannot be used due to the strongly restrained mass transfer from and to the WE. In this case, cells without mass transfer limitations such as the transmission type cells (**Figure 14(a) and 14(b)**) have to be applied or the electrode has to reach a stationary state with low reaction rates prior to the measurement in order to avoid erroneous results or misinterpretation.

2.31.7.1 X-ray Diffraction

The XRD technique is based on the elastic scattering of X-rays from structures that have long range order as schematically shown in **Figure 15**. X-rays reflected at net planes parallel to the surface may lead to constructive interference (Bragg-peak) if the incidence angle Θ (Bragg-angle) between the impinging radiation and the lattice planes satisfies the relation

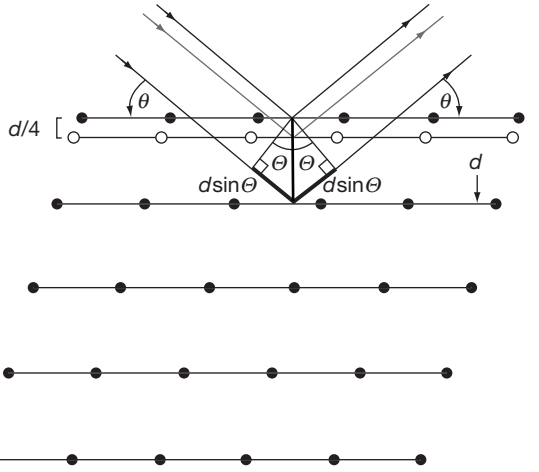


Figure 15 Basic principle of XRD. For constructive interference, that is, the observation of a Bragg-peak, the path length difference between the interfering X-rays has to be an integer multiple of the X-ray wavelength λ as shown for the lattice planes with the black atoms. If a second net plane (white atoms) is inserted, the diffracted intensities are influenced accordingly. See text for more details.

$$n\lambda = 2d\sin\Theta, \qquad [10]$$

where λ is the X-ray wavelength, and d the lattice spacing of the sample under investigation. In general, not only the fundamental wave ($n = 1$) but also higher harmonics ($n > 1$) are observed in a diffraction experiment. The interpretation of eqn [10] is illustrated in **Figure 15**, that is, constructive interference only occurs, if the optical path difference between the X-rays which are diffracted from neighboring lattice planes is equal to an integer multiple of the X-ray wavelength λ. The resulting diffraction pattern is representative of the crystal structure because each crystal structure is characterized by the distribution of its lattice atoms in the unit cell. If, for example, a second lattice plane is present in between those considered as first (open dots in **Figure 15**), this inserted plane may alter the diffraction pattern significantly as follows. If the distance between the two planes is equal to a quarter of the original *d*-spacing, then the path difference for the inserted net-plane equals half of the wavelength for $n = 2$, that is, destructive interference occurs. This is the case, for example, for the (222)-reflections of Si and Ge. In general, this behavior is mathematically described by the so-called structure factors F_{hkl}, which include the positions of all the atoms in the crystal structure under investigation as well as their occupation by different types of atoms according to eqn [11].

$$F_{bkl} = \sum_n f_n \exp(2\pi i \underline{Q} R_n) \quad [11]$$

$$R_n = n_1 a_1 + n_2 a_2 + n_3 a_3 \quad [12]$$

$$\underline{Q} = h_1 b_1^* + h_2 b_2^* + h_3 b_3^* \quad [13]$$

Here R_n is the lattice vector of the nth atom (eqn. [12]), with a_1, a_2, and a_3 as the basis of the unit cell unit and n_1–n_3 are integer numbers. $\underline{Q}$ is the reciprocal lattice vector which corresponds to the scattering processes at the (hkl)-lattice planes (eqn [13]).[25] Again, h_1–h_3 are integer numbers, and the reciprocal lattice is connected to the real space representation by $a_i b_j^* = 2\pi\delta_{ij}$ (for $i = 1, 2, 3$, $\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ for $i \neq j$).

If $\underline{Q}$ is given by the difference between the K-vector of the scattered X-rays and the impinging radiation, that is, $\underline{Q} = K_f - K_i$, as can be seen in **Figure 16**, where a schematic representation of different diffraction geometries is given, then we may expect a Bragg-peak for the corresponding angles. The summation in eqn [11] includes all the atoms in the unit cell, and f_n is the atom form factor, which represents the scattering power of an individual atom. Basically, f_n is determined by the number and the distribution of the electrons belonging to a single atom.[25] Values for f_n are tabulated elsewhere.[26] The diffracted intensity then is given by eqn [14]

$$I \propto |F_{bkl}|^2 \quad [14]$$

and thus not only the positions of the diffraction peaks are representative of the crystal structure, but also their intensities, and characteristic diffraction patterns result for each individual crystal structure.

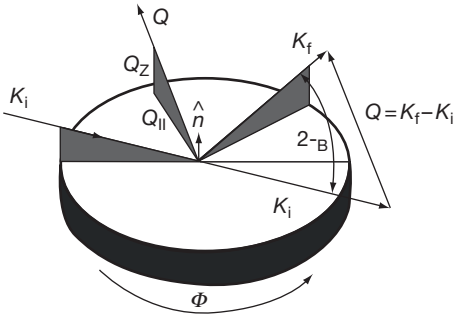


Figure 16 Schematic presentation of the scattering geometry in an XRD experiment (Laue-case). A Bragg peak is observed if the difference between the wave vector of the Bragg-reflected beam K_f and the impinging radiation K_i is identical to a reciprocal lattice vector Q. The Bragg angle Θ_B is identical to the angle between the diffracted wave and the (hkl)-lattice plane, which is not necessarily parallel to the surface of the crystal under investigation. Three different types of scans are indicated. (a) Out-of-plane scan (Q_z or $Q_\perp$-scan): the rotation (Φ) of the crystal is fixed, and data are measured by varying the component of Q normal to the surface (which is named Q_z or $Q_\perp$) while the component parallel to the surface ($Q_{||}$) is fixed. (b) In-plane radial scan ($Q_{||}$-scan): the component of Q along the surface is varied. The direction of Q is kept, but its magnitude is varied by changing the incident and diffracted angles (Θ and 2Θ) symmetrically. (c) Phi-scan: the crystal rotates about its surface normal with the positions of the incident and diffracted beams fixed. The magnitude of Q is fixed for this scan while the direction of Q rotates.

As an example from corrosion research, we present results determined from a diffraction experiment of passive layers on iron in borate buffer solution.[27,28] It should be mentioned here that a long and controversial discussion has been carried out and that contradicting results have been obtained.[6,24] Both amorphous as well as different crystalline structures have been proposed. Thus, an *in situ* diffraction experiment is very helpful for a structure determination. Due to the weak scattering intensities, focused SR as well as sophisticated detector equipment was used to obtain reliable data quality.[27,28] In **Figure 17(a)**, the structure factors obtained for different diffraction peaks for the passive film on Fe(110) are compared to those of several different model compounds. Based on the measured symmetry, lattice constants and the detected Bragg-peak intensities, the passive layer data are most consistent with oxides of the spinel type (e.g., Fe_3O_4 and γ-Fe_2O_3), while all other crystalline iron oxides, hydroxides or oxy-hydroxides are not compatible to the measured diffractograms. As can easily be seen, however, there are statistically significant deviations between the F_{bkl} of the measured data and those of the model compounds Fe_3O_4 and γ-Fe_2O_3. Furthermore, linear combinations of the Fe_3O_4 and γ-Fe_2O_3 model compounds are not able to fit the experimental data sufficiently, and thus, a refinement of the structure was performed, yielding the structure presented in **Figure 17(b)**, that is, the LAMM phase. Without going too much into the details, it should be mentioned here that the spinel unit cell which contains 32 oxygen anions, 16 octahedral, and eight tetrahedral cation sites is fully occupied in Fe_3O_4. For γ-Fe_2O_3, a quarter of the octahedral sites have only 33% occupancy, the remaining octahedral and tetrahedral sites being fully occupied. Both spinel reference compounds do not contain any interstitial ions. The authors found an octahedral site occupancy of (80 ± 10)% and a tetrahedral site occupancy of (66 ± 10)% for the passive film. While no evidence

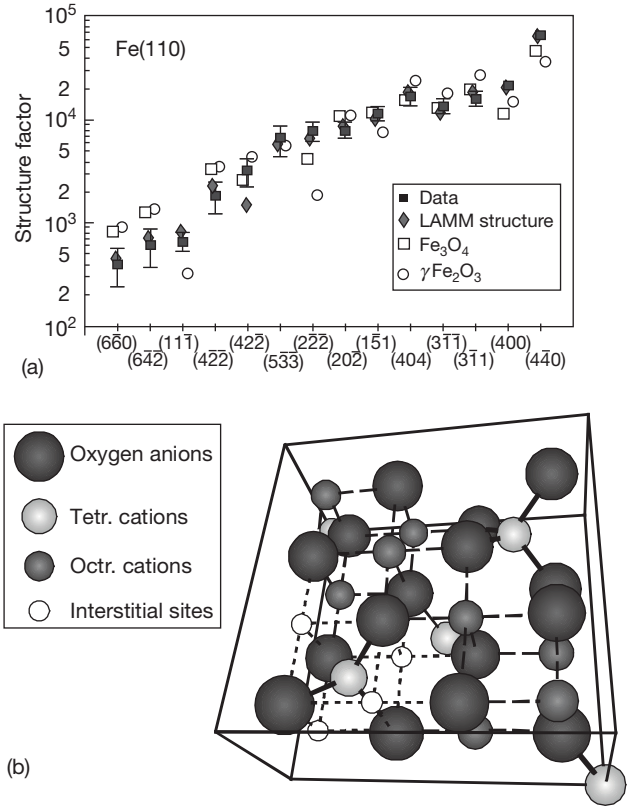


Figure 17 (a) Comparison of the experimental structure factors (filled squares) with those of selected Fe-oxide model compounds for the passive layer on Fe(110): open squares for Fe_3O_4, open circles for γ-Fe_2O_3, and filled diamonds for the 'LAMM'-phase. Reproduced from Toney, M. F.; Davenport, A. J.; Oblonsky, L. J.; Ryan, M. P.; Vitus, C. M. *Phys. Rev. Lett.* **1997**, *79*, 4282–4285. (b) Schematic illustration of the passive film structure (LAMM-structure). The solid black lines are the borders of the bottom half of the unit cell (big spheres: oxygen anions, fully occupied; smaller light grey spheres: tetrahedral cation sites, 66% occupancy; dark grey spheres: octahedral cation sites, 80% occupancy). Four of the eight octahedral interstitial sites are indicated by small white spheres, four additional sites are located in the upper right hand section which is not shown. Note that the interstitials are shown in well defined positions for clarity, while in the LAMM-structure used for the fitting of the diffractograms, they are distributed randomly in the structure. Bonds are also indicated for clarity: bold, dashed, and dotted lines for the tetrahedral, octahedral, and octahedral interstitial bonds, respectively. Adapted from Toney, M. F.; Davenport, A. J.; Oblonsky, L. J.; Ryan, M. P.; Vitus, C. M. *Phys. Rev. Lett.* **1997**, *79*, 4282–4285.

for tetrahedral interstitials was found, there are cations occupying $(12 \pm 4)\%$ of the available octahedral interstitial sites.[27,28] The large Debye–Waller factors (0.01 nm up to 0.028 nm) which are needed to fit the experimental data are the indication of the static disorder in the film, which results from the bond length variations induced by the interstitials and vacancies, and may explain why previous X-ray absorption spectroscopic investigations also found evidence for amorphous structures.

For real corroding systems, the interpretation of diffraction patterns in comparison to the data for passive layers on single crystal Fe-electrodes is even more complicated. A sample in practice consists of crystallites of different but eventually some preferential orientation, that is, a certain texture at the surface. As a consequence, the Bragg reflexes of all other orientations are suppressed in their intensity compared to those expected from the calculated structure factors. Such a behavior is well known for example, for vapor deposited thin films. Furthermore, a diffraction pattern is also modified by the presence of defects in the crystal lattice. For example, stacking faults in the sample cause systematic shifts and a characteristic broadening of the diffraction peaks.[29] Some of the peaks remain stable in position and half width, while others systematically move toward larger or smaller Bragg angles in conjunction with broadening. Even peak splitting is observed for special defect structures. Stress and strain also cause systematic Bragg peak shifts.[30] Therefore, the detailed analysis of peak positions shifts, their widths, and

intensities can be used for the identification of existing defects, as shown elsewhere.[27,31] For example, in the case of the passive film on Ni single crystal electrodes in sulfuric acid, a tilt of the oxide structure with respect to the underlying Ni-metal was found.[31]

When finite size effects are included, the width of measured Bragg peaks are found to be inversely related to the dimension of the diffracting region of the crystal. Thus, line broadening provides information about the particle size D of small crystallites which may be calculated by the application of the Scherrer formula of eqn [15]:

$$D = \frac{K\lambda}{\Delta\Theta\cos\Theta} \qquad [15]$$

where K is the Scherrer constant, which depends on the shape of the crystallites under investigation and in most cases is close to 0.9, and $\Delta\Theta$ is the line width at half maximum in radians after correction for instrumental broadening.[30] Although it is not trivial to separate the different contributions, information about the particle size distribution can be obtained from a more complete line shape Fourier-analysis.[30] In this regard, the use of SR may simplify the data interpretation. Compared to a laboratory experiment, the experimental line width is drastically reduced due to the extremely small divergence of SR-beams. In the case of the Fe passive layers presented above, the lateral crystallite size can be calculated to ~4.5 nm for films on Fe(110), whereas for Fe(001), a slightly larger size of ~6 nm was found.[27,28] Therefore the passive film is best described as a nanocrystalline material. In the case of passive films on Ni(111) in sulfuric acid, slightly larger crystallite sizes of ~8–10 nm were found.[31]

Even with the use of SR, XRD is limited to particles or domain sizes of at least 2–3 nm to yield a measurable diffraction pattern in the sense of the Bragg equation. Thus, particles or domains with smaller size will appear as X-ray amorphous in an XRD experiment and will not reveal any sharp Bragg reflexes. In conclusion, XRD is an excellent method for very detailed information on crystalline structures, although in practice, the interpretation of the measured diffraction patterns may be very complicated.

2.31.7.2 X-ray Reflectivity

Due to the large penetration depth of hard X-rays in matter, X-ray techniques are in general not surface sensitive; that is, they are not suited for the investigation of surface phenomena such as passive oxide film formation, adsorption, corrosion, etc. without any special precautions. However, the whole spectrum of X-ray analytical methods (e.g., XAS, XRD, X-ray fluorescence, X-ray topography) can be made surface sensitive using the grazing incidence geometry, which is schematically shown in the insert of **Figure 18**. The energy dependent index of refraction $n(E)$ in the

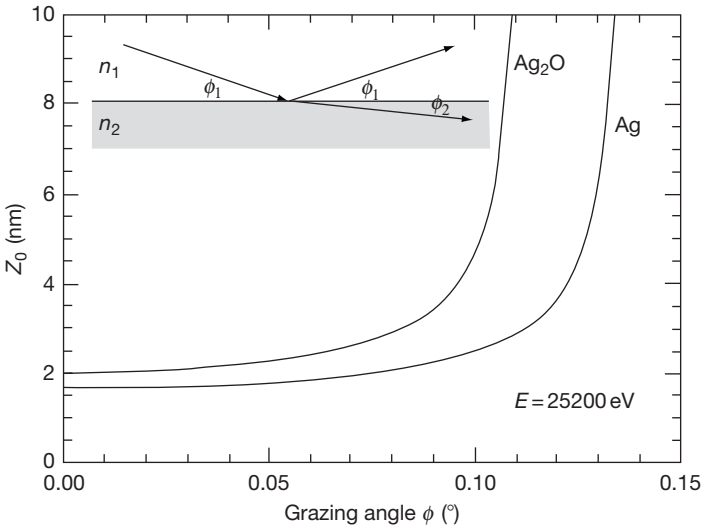


Figure 18 X-ray penetration depth z_0 for metallic silver and silver oxide Ag_2O at $h\nu = 25\,200$ eV, as a function of incidence angle. The critical angle is ≈0.13° for Ag, and ≈0.11° for Ag_2O. In the insert, a schematic presentation of the grazing incidence geometry is given: the refracted beam is diffracted toward the surface.

hard X-ray regime is given by eqn [16],

$$n(E) = 1 - \delta(E) - i\beta(E) \qquad [16]$$

where β and δ are small positive quantities. Depending on the material and the photon energy, typical values for β and δ range from $\sim 10^{-4}$ to 10^{-8}. The absorptive correction β is proportional to the linear absorption coefficient μ, through $\beta = \lambda\mu/4\pi$. Therefore, condensed matter is optically thinner than vacuum, and X-rays entering into the samples surface are refracted away from the surface's normal, as indicated in the insert of **Figure 18**. The application of Snell's law of refraction, which has the same form as in the visible spectral range (eqn [17]),

$$n_1 \cos\phi_1 = n_2 \cos\phi_2 \qquad [17]$$

suggests the existence of a glancing angle ϕ_c, the critical angle, for which $\phi_2 = 0$. Assuming $n_1 = 1$ for vacuum, ϕ_c can be calculated from $\cos\phi_c = n_2$. If the absorption term β can be neglected, a simple calculus leads to $\phi_c \approx (2\delta_2)^{1/2}$. Typical values for ϕ_c are 0.39° for Cu and 0.22° for Al for 8600 eV photon energy, so that total reflection can only be observed for grazing incidence angles below 0.5°.[32] In this case, there is no X-ray wave propagation inside the material, and the electric field amplitude within the sample is damped exponentially with depth. The decay constant, that is, the penetration depth z_0, amounts to only few nm below the critical angle. This relation is illustrated in **Figure 18** for Ag and Ag_2O, where a value of ~2–3 nm is obtained below the critical angle, while a steep increase is observed above ϕ_c. This behavior generally leads to a steep decrease of the XRR for angles larger than the critical angle, and a detailed evaluation of such angle dependent specular reflectivities may yield the density and the surface roughness of the material.[32] In the case of a thin film on a substrate, the X-rays that have penetrated the surface layer will be reflected at the inner interface and interfere with those beams that have been reflected at the outer side, that is, the air–vacuum surface, leading to a systematic variation of the measured reflectivity profiles. Their detailed analysis yields to the thickness of the film and the roughness of the inner interface, in addition to the densities of the film and the substrate and the surface roughness.[32] Without going into great detail, it should be mentioned here that multilayered film structures can be treated and analyzed in a straightforward manner, so that specular reflectivity measurements are ideal tools for the *in situ* observation of electrochemical growth of oxide and passive layers and their reduction.[33] In **Figure 19**, specular reflectivity profiles recorded in the course of several oxidation–reduction cycles of a thin film Cu electrode in borate buffer (pH 8.4) are shown for a photon energy of 7700 eV ($\lambda = 0.148$ nm). The incidence angle scale was transferred to the transferred momentum (Q_z) scale by $Q_z = 4\pi/\lambda \sin\phi$ (see also **Figure 16**). As can be seen in **Figure 19**, the reflectivity profiles change considerably upon the oxidation and reduction of the Cu electrode. For example, the thickness oscillations are much more pronounced for the

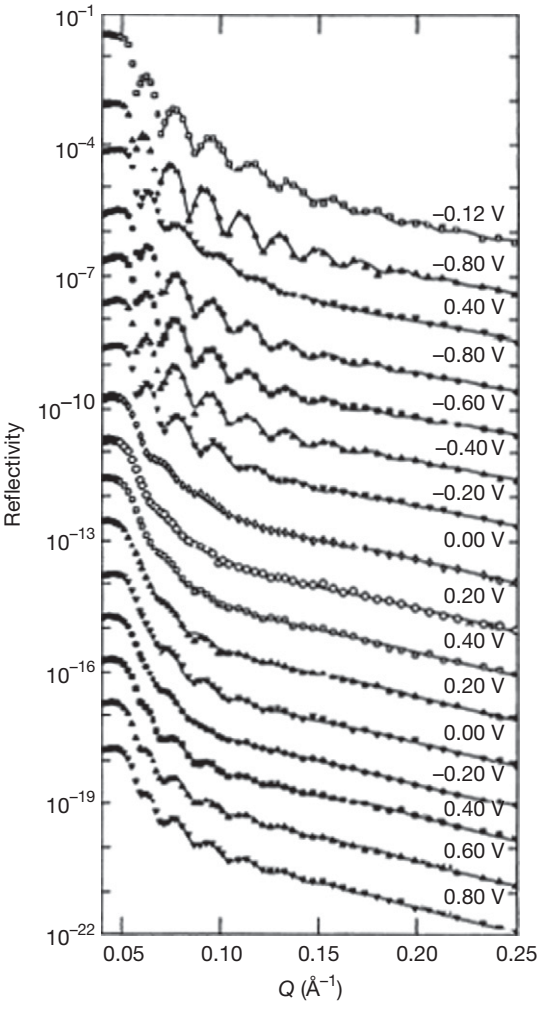


Figure 19 A series of specular X-ray reflectivity profiles measured *in situ* in the course of the anodic oxidation and reduction of a Cu thin film electrode in borate buffer solution (pH 8.4). The measurements were performed at the potentials indicated, and the curves are vertically offset for easier visualization. Reproduced from You, H.; Melendres, C. A.; Nagy, Z.; Maroni, V. A.; Yun, W.; Yonco, R. M. *Phys. Rev. B* **1992**, *45*, 11288–11298, with permission from American Physical Society.

reduced electrode, while only weak thickness fringes can be identified in the case of the oxidized sample, which are the first indication of an increased roughness at the interface to the electrolyte. Moreover, it can already be seen in these raw specular reflectivity data that the original state of the electrode is not recovered after completion of several oxidation/reduction cycles. The quantitative analysis of the presented data allows a microscopic access to the processes at the electrode/electrolyte interface. For example, the surface roughness of the electrode/electrolyte surface is ~1.1–1.2 nm in the reduced state prior to the oxidation, while it increases to ~2.5 nm in the fully oxidized state. After a completed oxidation–reduction cycle, the electrode remains with an irreversibly increased roughness of ~1.6 nm, which can be further increased by prolonged cycling between anodic oxidation and cathodic reduction conditions.[33] Such a behavior has been previously found for Cu and Ag electrodes in alkaline media using SEM and quartz microbalance techniques and is important with regard to surface enhanced Raman scattering (SERS). Furthermore, the increase of the oxide layer thickness with increasing potential at the expense of the thickness of the thin film metal electrode is evident. The thickness of the fully developed oxide layer on Cu amounts to ~3.0 nm, in agreement with previous *ex situ* XPS, UPS, and ISS experiments[5,10,16] and *in situ* Raman spectroscopy.[34] However, the reduction is fully reversible, as the initial thickness of the metal layer is recovered after a complete oxidation–reduction cycle. Such a behavior can be anticipated from the cyclic voltammogram, where balanced anodic and cathodic charges were found. In conclusion, we would like to mention that specular XRR measurements can be used for the *in situ* investigation of gaseous oxidation and corrosion processes in a similar way.

2.31.7.3 X-ray Absorption Spectroscopy

The basic process underpinning XAS is identical to that of the XPS-technique, that is, an incident photon with sufficient energy is absorbed by an atom and excites a core electron to unoccupied levels (bands) or to the continuum. This excited atom can relax by filling the empty core level with an electron from a higher occupied core level. The energy difference of the involved electronic levels can be released as a photon or may lead to emission of an Auger-electron. XAS is governed by the details of this photon absorption process. In a XAS experiment, the absorption of X-rays within the sample is measured as a function of the X-ray energy ($E = h\nu$). More specifically, the X-ray absorption coefficient, $\mu(E) = -\mathrm{d}\ \ln\ I/\mathrm{d}x$, is determined from the decay in the X-ray beam intensity I with distance (sample thickness) x. Monochromatic radiation is used, the energy of which is increased to the point at which core electrons can be excited to unoccupied states close to the continuum. In **Figure 20**, an absorption spectrum of a gold metal foil is shown for an extended photon energy range. The experimental data show three general features: (1) an overall decrease in X-ray absorption with increasing energy; (2) the presence of sharp rises at certain energies called edges, which roughly resemble step-function increases in the absorption (i.e., K-, L-, and M-edges in **Figure 20**); and (3) above the edges, a series of wiggles or an oscillatory structure occurs which modulates the absorption typically by a few percent of the overall absorption cross section as can be seen in the insert of **Figure 20**. It should be already stressed at this stage that the availability of intense new synchrotron X-ray sources permits the investigation of highly dilute systems such as impurities or trace elements within an alloy, or adsorbates and thin films on a surface of a foreign substrate. This capability of the XAS-technique is important and somehow unique with regard to the investigation of corrosion processes, where the

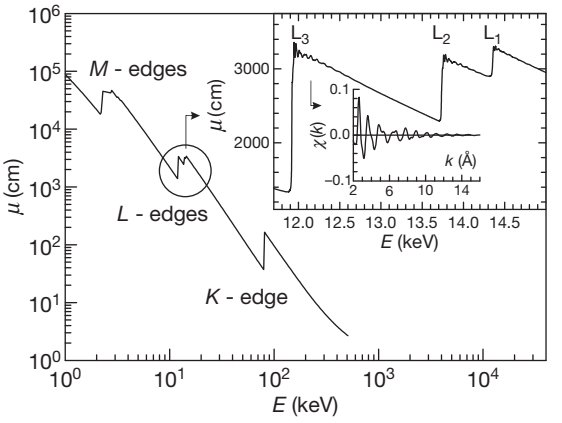


Figure 20 Schematic view of the absorption spectrum of a gold metal foil over a wide energy range. Steep increases of the absorption coefficient are related to the K-, L-, and M-edges, respectively. The inset shows the absorption at the Au-L edges on a magnified scale. The oscillatory behavior of the absorption coefficient $\mu(E)$ above the edge, known as EXAFS, is clearly resolved at each of the L-edges. The extracted fine structure oscillations $\chi(k)$ measured at the L_3-edge after the removal of the smooth background function $\mu_0(E)$ are also displayed as a function of the photoelectron wave vector.

alloyed elements often have a dramatic influence on the resulting corrosion protection properties. Various alloyed elements of a specimen can be investigated separately by XAS due to their different X-ray absorbing edges. However, for the investigation of flat electrode surfaces, a simple transmission mode experiment is not adequate. Fluorescence mode detection and/or grazing incidence techniques are required, which are well suited even for samples in an electrochemical environment.[22,24]

The general decrease of μ with E can be related to the well-understood quantum mechanical phenomenon of X-ray absorption by atoms, which can be described by Fermi's golden rule. As with the energy of fluorescence photons, the energy of the absorption edge is characteristic of the absorbing material and hence the measured edges are signatures of the atomic species present in a material. For example, for the transition metals Fe, Co, Ni and Cu, the K edges are located at 7112, 7709, 8333, and 8979 eV, that is, there is a separation of 600 eV between consecutive edges, while, for example, the Mo K edge is located at 20 000 eV. Each edge corresponds to a quantum-mechanical transition that excites a particular atomic core–orbital electron to free or unoccupied levels. The nomenclature for the X-ray absorption edges reflects the origin of the core electron; that is, K edges refer to transitions that excite the innermost 1s electron, while L edges are related to those of $2p_{3/2}$ (L_3), $2p_{1/2}$ (L_2), and 2s (L_1) initial states. These transitions always occur in unoccupied states, that is, in states with an excited photoelectron above the Fermi energy leaving behind a core hole. In a solid, such a photoelectron generally has enough kinetic energy to propagate freely through the material and independent from the structure and the state of matter of the sample under investigation. From the change of the absorption coefficient at the edge, the concentration of the respective element can easily be calculated (see, e.g., Ref. 20). For example, if a metal is corroded and dissolved in an electrolyte, the specific absorption of the element of interest within the electrolyte can be used for the analysis of soluble corrosion products in the active, passive or transpassive state of a metal or the selective dissolution of a single element in a multielement specimen such as an alloy.

Similar to an XPS experiment, the exact energy of the absorption edge is a sensitive function of the valency of the excited atom. In general, a shift of the absorption edge toward higher photon energies is observed as the chemical valency of the absorber atom is increased, i.e., with increasing oxidation state.[35,36] In many cases, a more or less linear shift of the edge with typically 1–3 eV per unit of the valency can be found in the literature for different elements. This effect provides an easy access to determine the valency of a selected element (see, e.g., Refs. 35–37) even for trace elements in complex systems. In addition, in some special cases, there are also features in the absorption spectrum below the edge. These so-called pre-edge peaks can be attributed to transitions from the excited photoelectron into unoccupied electronic levels of the sample, that is, an X-ray absorption spectrum probes the density of unoccupied states of the absorbing element.

The detailed structure of the absorption coefficient within ~50 eV above the edge is referred to as near edge X-ray absorption fine structure (NEXAFS) or X-ray absorption near edge structure (XANES). Besides the above mentioned transitions into unoccupied states, the near edge structure also involves transitions of the photoelectron into the continuum, that is, into states above the vacuum level. In the latter case, the photoelectron has a low kinetic energy and accordingly a large (inelastic) mean free path (see, e.g., Ref. 13), so that it can undergo multiple scattering events.[38] The shape of the edge is thus highly sensitive to both the valency and the coordination of the absorbing atoms, including also the angular arrangement of the neighboring atoms. Although the complexity of the involved processes makes a full theoretical treatment very difficult, it is currently possible to calculate near edge X-ray absorption spectra on an *ab initio* basis, that is, based on the assumption of a model structure.[38] This approach works quite well for molecules and complexes – which is an interesting feature for the investigation of corrosion products within the electrolyte as shown in Ref. 37. However, for crystalline materials it is still a quite challenging task to fit an experimental spectrum using *ab initio* calculations since the theory of XANES is not fully quantitative and requires several physical considerations.[38] Thus, for the practical analysis the near edge spectra are most commonly used as fingerprint techniques. Edges are collected for a number of standard compounds with known valence and crystal structure and these are compared to the spectra of the actual samples. Certain compounds such as Cr^{6+} or Mo^{6+} give rise for distinct pre-edge peaks, so that these species can be identified qualitatively and quantitatively. For the application of fingerprint techniques to electrochemical probes, see for example, 24, 37. As an example for the application of XANES spectroscopy in

corrosion research, we present a study dealing with the initial stages of high temperature corrosion of Y-implanted Fe–Cr steels.[39] This study intends to examine the so-called 'rare-earth effect.' It is well-known that the presence of a small percentage of active elements, especially rare-earth elements, may have a beneficial effect on the properties and the adherence of protective scales to the substrate. In the present experiment, the rare earth element was inserted into the Ni–Cr alloy by means of 200 keV Y^+-ion implanting, and the resulting samples were oxidized at 700 °C in a dilute oxygen atmosphere (1% O_2, 99% Ar) for varying times. In **Figure 21**, grazing incidence XANES spectra acquired from the as-implanted sample and samples oxidized for different times are compared with the reference spectra from metallic Y and Y_2O_3.[39] A progressive oxidation of the samples is observed as the oxidation time increases, as directly evidenced by the shift of the absorption edge toward increasing energy values. Furthermore, after ~2 min of oxidation, the obtained spectra closely resemble those of yttria Y_2O_3. A more detailed investigation of the extended X-ray absorption structure suggests that the formed crystallites are of nanosize, and parallel XRD experiments revealed

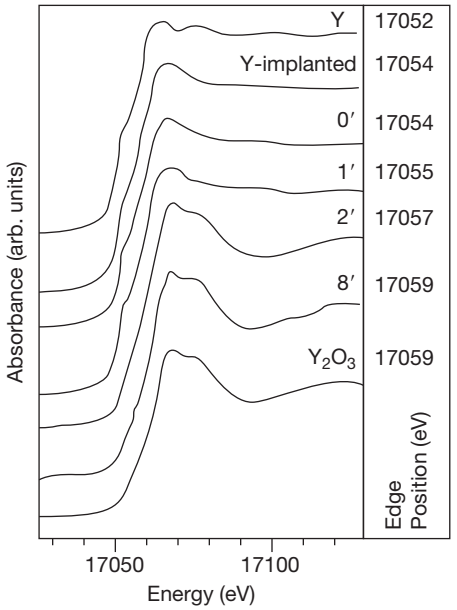


Figure 21 Near edge X-ray absorption spectra at the Y K-edge of a Y-implanted Ni–20Cr alloy in the as-prepared state as well as after oxidation in a gas atmosphere of 1% O_2 in Ar at 700 °C for 1, 2, and 8 min, respectively. (Metallic Y and Y_2O_3 reference spectra are also shown). Reproduced from Gibson, P. N.; Crabb, T. A. *Nucl. Instrum. Methods B*, **1995**, *97*, 495–498.

that the crystallization is *initiated* at the interface between the oxide scale and the bare alloy, proving in this case the positive effect of the dispersed rare earth element on the formation of a well ordered oxide.[39]

The third general feature in XAS spectra are the oscillations of the absorption coefficient which are visible up to ~1000 eV above the edge. This so-called extended X-ray absorption fine structure (EXAFS) originates from scattering of the emitted photoelectrons by neighboring atoms. The photoelectron, which is ejected from the absorbing atom, can be described as a propagating electron wave with a wave vector k and a de Broglie wavelength λ (eqn [18]):

$$k = \frac{\sqrt{2m(E - E_0)}}{\bar{h}}, \lambda = 2\pi/k \qquad [18]$$

E is the actual photon energy, E_0 the edge energy, m the electron mass, $\bar{h} = h/2\pi$ and h is Planck's constant. This wave can be scattered by neighboring atoms, so that the outgoing wave interferes with the backscattered wave. Depending on the distance, r_j, to the scattering atoms and the wavelength of the photoelectron, this interference can be constructive or destructive, resulting in an increase or a decrease of the absorption coefficient by some few percentage points. This phenomenon was first theoretically described by the short range order theory developed by Sayers, Stern, and Lytle.[40] In contrast to the near edge structures in the absorption coefficient, the theory of the EXAFS is today well understood with acceptable tolerances from the experimental results.[38] The EXAFS spectrum $\chi(E)$ above a given absorption edge is defined as the oscillatory part of the X-ray absorption:

$$\chi(E) = \frac{\mu(E) - \mu_0(E)}{\mu_0(E)} \qquad [19]$$

$\mu(E)$ is the energy dependent absorption coefficient, and $\mu_0(E)$ is the smoothly varying, atomic-like background absorption. $\mu_0(E)$ contains all the contributions from other edges and other elements and their absorption edges within the sample. According to the plane-wave concept and the single-scattering approximation, the EXAFS oscillations $\chi(\kappa)$ can be expressed by eqn [20]:

$$\chi(\kappa) = \sum_j S_0^2 N_j \frac{|f_j(k)|}{kr_j^2} \cdot \sin(2kr_j + \delta_j(k)) e^{-2r_j/\lambda(k)} e^{-\sigma_j^2 k^2} \qquad [20]$$

N_j is the coordination number of the central atom in the jth shell, $f_j(k)$ the backscattering amplitude, σ_j the mean relative displacement of the atoms in the jth

shell, and δ_j a phase shift associated with the interactions of the photoelectron wave with the potential of the absorbing and the backscattering atoms. Since different backscattering atoms have their own characteristic backscattering phases and amplitudes, EXAFS enables identification of the nature of the atoms in each coordination sphere around the absorbing atom. In contrast to XRD, EXAFS is very sensitive toward light backscattering atoms, which justifies its use for the investigation of, for example, oxides and passive layers. The disorder σ^2 is partly due to thermal effects, which cause all the atoms to oscillate around their equilibrium positions. Effects of structural disorder are similar, giving an additional contribution to σ^2. Both effects can be separated by means of temperature dependent measurements. As a consequence of the $1/r_j^2$ dependence of $\chi(k)$ and the limited mean free path of the photoelectrons characterized by $\lambda(k)$, EXAFS is useful in providing information about the local short-range order around the absorbing atom, that is, the fine structure contains precise information about the local atomic structure around the atom that absorbed the X-rays. The structural information is typically obtained by Fourier-filtering of the experimental $\chi(k)$ data into distance-space, giving a radial distribution function from which coordination numbers, interatomic distances, and the local disorders are determined.[40] Similar to the XANES, this turns out to be a unique signature of a given material, since it depends on both the detailed atomic structure and its vibrational properties. For this reason, EXAFS has become an important probe especially for disordered materials, where XRD can hardly be applied due to the absence of any long range order correlations. A complete data analysis comprises the separation of the individual shells contributing to the radial distribution function, and the fitting of these peaks with phases and amplitude functions.[41]

In this contribution, we will not show a detailed data evaluation, but only focus on the main features of EXAFS. As an example, in **Figure 22**, we show the *in situ* evolution of the EXAFS at the Ni K-edge (8333 eV) of a metallic Ni-layer on an Al_2O_3-substrate during the formation of a spinel-type oxide in an O_2-containing atmosphere at elevated temperatures.[42] First, the thin Ni film was oxidized to NiO at 700 °C, and the resulting X-ray absorption fine structure of this Ni-oxide thin film is similar to that of the crystalline NiO-reference material, as can be seen in **Figure 22**. In a second oxidation step at an elevated temperature of ~1000 °C, this NiO-layer interacts with the Al-oxide substrate, and a spinel-type

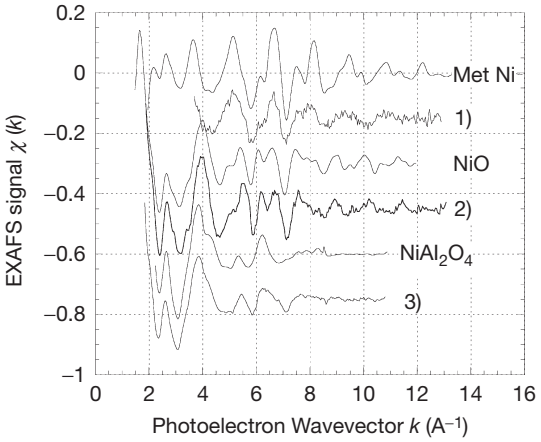


Figure 22 EXAFS of a Ni thin-film sample on a single crystal Al_2O_3 substrate after (a) initial deposition, (b) 20 min in 60 mbar O_2 at 973 K, and (c) additional 9 h oxidation at 1273 K. For comparison, spectra of metallic Ni and crystalline NiO and $NiAl_2O_4$ spinel are also presented. Reproduced from d'Acapito, F.; Davoli, I.; Ghigna, P.; Mobilio, S. *J. Synchrotron Rad.* **2003**, *10*, 260–264.

$NiAl_2O_4$-layer is formed as can be deduced by comparing this high temperature spectrum with that of the related spinel reference material.[42] In a similar way, the oxidizing effect of elevated temperatures on high-T_c superconductors has also been proven, showing clearly the further oxygenation of the superconductor as well as the appearance of new phases.[43] It is important to mention here that several *in situ* EXAFS studies have also dealt with electrochemically formed passive layers on various metals and alloys (see Refs. 24, 27, 28, 37, 44). Finally, due to the fact that XAS experiments are element specific and not restricted to investigations of crystalline materials (such as XRD studies), it is possible to investigate in detail the structure and the chemical state of corrosion products dissolved in the electrolyte in front of a corroding electrode.[20,37] The use of the state of the art detectors and stable monochromators also enables the investigation of highly *dilute* solutions, and time resolved investigations are easily possible, so that the change in concentration in solution can be followed in tandem with changes in the structure of the solid.

2.31.8 Glow Discharge Optical Emission Spectroscopy (GDOES)

If a sufficiently high voltage is applied, a GD can be generated between two physically separated

electrodes in a gas filled vacuum vessel with a pressure of typically 10–1000 Pa. Depending on the gas type and pressure, the distance of the electrodes and the applied voltage, electrical currents, and characteristic light emissions are observed. We will only focus on the most relevant points here; for more detailed information on the fundamentals of GDs, the reader can refer to the literature, for example, Ref. 45. As a result of the physical processes, GD plasma shows several luminous and dark regions. For analytical use of the GD, the optical emission in the luminous regions are studied and analyzed quantitatively. It is also essential that the cathode (sample under study) is exposed to a bombardment of ions and fast neutral species from the plasma, and therefore, the cathode is eroded and cathode material is released into the gas phase continuously – this effect is known as cathode sputtering. This sputtering of the cathode is the basis for analytical exploitation of GD. The sputtered species from the sample arrive in the plasma where they are subject to ionization and excitation. If the corresponding ions are measured in a mass spectrometer, they can be identified quantitatively. This method is called glow discharge mass spectrometry (GDMS) in literature. On the other hand, the excited atoms emit characteristic photons during their deexcitation which in principle allow their identification, if they are detected with an optical spectrometer. This technique is called GDOES. It should also be mentioned here that the sputtered atoms and ions can also be excited with an external light source for atomic absorption (GDAAS) or fluorescence (GDAFS) spectroscopy. Given that the sample is continuously eroded, a depth profiling of the sample is possible, if the erosion rate is calibrated properly. In contrast to other depth profiling techniques, matrix effects only play a minor role here because of the spatial separation of the sputtering and excitation of the sputtered atoms, as well as their dilution in the plasma. Using the Monte-Carlo simulation techniques, it is possible to calculate the profile of the sputtered crater on the sample, where reasonable agreement to the experiment was found for direct current GDs. These simulations also show that a homogeneous electrical field distribution in the active (sputtering) region is required to ensure that the sample surface is evenly eroded, resulting in a sputter crater with a nearly flat bottom. Only in this case, a well resolved depth profile of the sample can be obtained. It is noteworthy to mention here that for analytical GDs argon is most commonly used as the discharge gas, in order to avoid chemical reactions of the sputter gas with the sample and the system. In addition, a small but continuous flow is applied to ensure a clean atmosphere during the process of analysis.

A schematic representation of a GDOES analytical system is shown in **Figure 23**. The hollow anode configuration is composed of a grounded metal tube. An appropriate window confines the vacuum vessel on one side, while the flat sample closes the

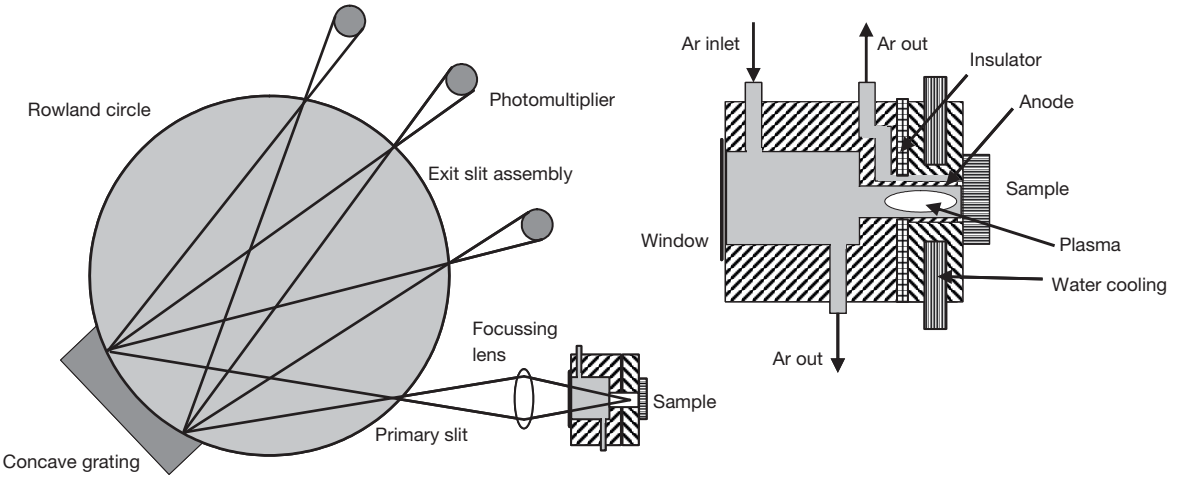


Figure 23 Schematic representation of a spectrometer for glow discharge optical emission spectroscopy (GDOES). The sample, cathode, is placed in a hollow, as shown, in an expanded view in the insert. Ions from the plasma are accelerated against the sample, and material is eroded and subsequently diffuses into the plasma. Here excitation of these atoms may occur, leading to the emission of characteristic photons during relaxation to the ground state. These optical emissions of the plasma are focused onto the primary slit of a grating spectrometer, where the spectrum is analyzed by means of a grating spectrometer, and the spectral contributions of different species are detected by an array of photomultipliers.

discharge cavity on the other side. This way, sample introduction or change is easy, and the electrical power that is needed for ignition and maintenance of the plasma is simply supplied to the sample. The sample holder is water cooled in order to limit the heating of the samples. Usually, a quartz window is used because of its transparency for photons of lower wavelength.

In the case of a conducting (metal) sample, a regulated DC-voltage of typically some hundred volts is applied. As for sputtering, etching, and coating processes, however, the GD can also be energized by alternating radiofrequency (RF) fields in order to investigate in particular poor/bad or nonconducting samples.[45] Assuming the presence of a nonconducting sample, a DC-discharge would cause a positive charging of the cathode because of the recombination of positive ions and electrons coming from the near surface region of the sample. As a consequence, the potential drop between the two electrodes would decrease and the discharge would extinguish, within some few milliseconds, with a positive charge remaining on the sample. If the direction of the voltage is reversed, the same processes will occur, resulting in a decrease of the positive charge on the sample, which in this case acts as an anode until a negative charge is accumulated so that the discharge is stopped again. If the switching of the voltage polarity is repeated continuously with a high frequency, that is, RF with a power of typically up to ~30 W, the discharge can be maintained continuously even for nonconducting samples. Using frequencies in the mega-Hertz range, the plasma with typical frequencies of the order of 1 kHz is not able to decay but will glow permanently. One may argue that the use of an RF-discharge could be disadvantageous for GDOES-investigations because the material of the hollow electrode will be sputtered as well and thus contribute to the measured optical emissions. However, due to the characteristic design of the two electrodes, the sputtered area on the sample is limited to some few square millimeter while on the other hand the area of the hollow tube exposed to electrons and ions is not restricted at all and can be extended as much as required by the size of the GD plasma, thus reducing the related contributions accordingly.

The spectrometer displayed in **Figure 23** employs a concave grating that is fixed in the Rowland geometry, and the light emissions from the GD are focused onto the entrance slit of the spectrometer using a lens of suitable focal length. Several photomultiplier tubes are mounted at well defined positions on the Rowland circle for light detection related to some few specific excitation lines, as it is not possible to measure a full spectrum in this configuration. However, due to the fast erosion of the samples, it is important to make use of an extremely fast and reliable detection method for the excited photons. State of the art systems employ up to 60 individual detectors in parallel. The use of solid state detectors, charged coupled devices or photo diode arrays have become a common alternative to photomultiplier tubes. These detectors allow the acquisition of the entire spectrum, or at least a large portion of it, but are usually slower than photomultiplier tubes and therefore not suitable in those cases where extremely short acquisition times are required, for example, for the analysis of adsorbates or very thin surface films. Light elements such as hydrogen, carbon, nitrogen, oxygen, or sulfur have important spectral lines in the vacuum ultraviolet (UV) region with wavelengths below ~200 nm. For example, oxygen species are in many cases identified using an oxygen emission line at 130.2 nm. Spectrometers in this spectral range are usually operated under vacuum, or at least a purged nitrogen atmosphere to obtain a sufficient transparency in this spectral range.

GDOES, as an analytical technique, was introduced by Grimm in 1968 with first applications being the determination of the bulk atomic compositions of metal alloys.[46] The capabilities of GDOES for surface analytical studies was proven a few years later with oxidation studies of steels being one of the first areas of research interest. In a depth profiling experiment, the intensities of the spectral lines of interest are measured as a function of the sputter time. For a quantitative analysis, these have to be converted to an analyte concentration and a sputter depth. Although the detection limits for GDOES are generally small with typical values in the range of 1–10 ppm, an exact determination of atomic concentration and film thicknesses is in many cases not straightforward and the use of more sophisticated calibration procedures for example, using certified materials in a wide range of elemental compositions, is recommended. Furthermore, spurious hydrogen in the plasma, either from a contamination of the plasma gas or the sample, may affect the measured concentrations of the analytes and have to be corrected accordingly. Nevertheless, in combination with the numerous efforts that have been made for quantitative modeling and understanding of GD-plasmas and sputtering for analytical purposes, and because of its easy use, GDOES has become an

interesting and important method for routine depth-profiling analysis in a wide range of applications, including corrosion phenomena.

Corrosion protection coatings – in particular on steel – are an important topic. The depth profiling capabilities of GDOES give a fast and direct access to a simple determination of the composition of the coating as well as its thickness. This is of special importance in those cases where the coating contains the same chemical elements as the underlying steel, such as for Fe/Zn-coating on iron based steels. For industrial applications, it is in many cases important to know the structure and the composition of a material in order to accomplish reproducible physicochemical properties such as, mechanical elasticity and plasticity, or corrosion resistance of metal sheets. These properties are often influenced by the pretreatment of the metal, such as milling or hot-rolling which alter the microstructure and the composition of the metal significantly. Depending on the details of the processing conditions, trace elements, impurities, or rolled-in oxide particles are located especially in the near surface region of the metal. In **Figure 24(a)**, GDOES depth profile of a hot-rolled steel surface is presented.[47] As can be seen, the oxide layer on the steel has a rather complex substructure, and even the alloyed elements Mn and Si are detectable, although their atomic concentrations are only in the 1% range. Besides the near surface region of some 10 nm, where Si is significantly enriched, the outer oxide scale has a lower Fe-content compared to the inner part of the duplex layer, which contains significantly more iron. The manganese concentration is smaller in the outer layer, and increases in the inner layer, but it is still depleted compared to the bulk concentration. While the entire oxide is almost free of Si and C, both elements are significantly enriched at the interface between the oxide and the bulk material.

In a similar manner, GDOES has fostered our understanding of high temperature corrosion processes and materials. Very recently, first studies on materials and coatings that are used in biological environments and biocorrosion have been published. The thickness and the composition of passive oxide layers were determined using GDOES for a variety of different materials and for various electrochemical environments – here the most important aspect of GDOES is its easy operation and data interpretation. This is also true for samples of archeological interest, where GDOES can easily be applied to conducting as well as to nonconducting samples such as, ancient bronzes, ceramics, pottery, etc., and GDOES may give information which is important for the conservation of such objects.

In conclusion, when considering GDOES for a practical application, one should always have in mind that the recorded depth profiles are destructive, which is sometimes critical for example, in the case of archeological objects. Furthermore, GDOES is always an *ex situ* method, that is, it relies on the assumption that the sample is not affected or altered by the transfer to the spectrometer and the bombardment with electrons and ions, which seems to be true for metals and alloys. Compared to UHV-techniques such as XPS or SIMS, GDOES is not very time-consuming, and the repetition of each experiment allows one to apply statistical techniques to the results, for example, to elemental concentrations or thickness *values*, which is again important for industrial applications. On the other hand, however,

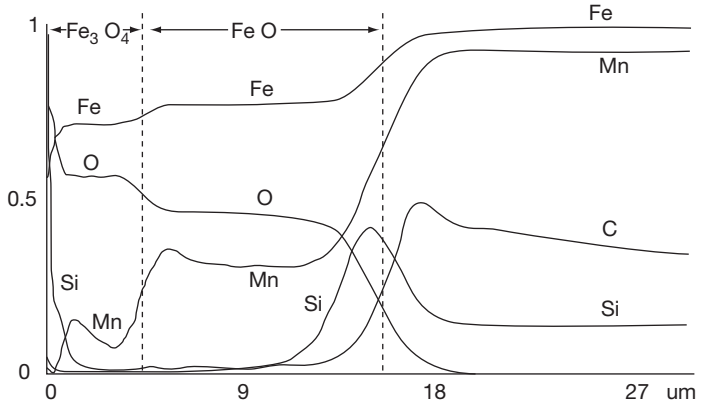


Figure 24 GDOES depth profile of an oxide layer on hot-rolled carbon steel. The concentration scale for each of the elements is given in relation to the following values: 100% for Fe, 50% for O, 0.5% for Mn, 0.2% for C, and 2% for Si. Reproduced from Weiss, Z.; Musil, J.; Vlcek, J., Fresenius J. *Anal. Chem.* **1996**, *354*, 188–192.

GDOES is not able to specify the binding state of an element, that is, GDOES is not able to distinguish between a metal cation and the reduced metal. Thus, it is useful to combine GDOES with other techniques in order to strengthen the conclusions drawn.

2.31.9 Infrared Spectroscopy

Infrared spectroscopy applies electromagnetic radiation of wavelength λ or wave number $\tilde{v}$ of the middle-IR range ($\lambda = 2.5$–$50\,\mu m$, $\tilde{v} = 4000$–$200\,cm^{-1}$) to a sample. Typically, the near-IR range ($\lambda = 0.8$–$2.5\,\mu m$, $\tilde{v} = 13\,000$–$4000\,cm^{-1}$) or the far-IR range ($\lambda = 5$–$1000\,\mu m$, $\tilde{v} = 200$–$10\,cm^{-1}$) are of minor importance for this spectroscopic method. The absorption of IR-radiation in the middle range by a sample leads to excitation of vibrational modes of its molecules as a whole (fingerprint region, $\tilde{v} = 600$–$1200\,cm^{-1}$) or of characteristic groups such as C=O and O–H (group frequency region, $\tilde{v} = 1200$–$3600\,cm^{-1}$). The excitation of vibrations is controlled by selection rules. For IR absorption, the dipole moment should change with the oscillation mode. Therefore, polar groups like C=O are IR-active. If, however, the changes in the dipole moment compensate internally within a molecule, the IR mode is inactive, for example, the symmetric stretch of O=C=O. The asymmetric stretching and the bending vibrations of O=C=O are IR-active. We note that the selection rules for Raman Spectroscopy depend on the polarizability of groups and molecules. Under the selection rules for Raman, the symmetric stretching of the O=C=O molecule, which has an inversion center, is active, whereas the asymmetric stretching and the bending mode are Raman-inactive. Thus, IR and Raman can give complimentary results.

An IR spectrometer usually has a strong light source like a Nernst Glower, which is an electrically heated rod of rare earth oxides, or a Globar, a heated rod of silicon carbide. A very intense IR source is SR. In this case, polarized light of high intensity is also available.

A more conventional IR spectrometer uses a monochromator, separating the incoming polychromatic radiation into its components and thus taking sequentially a wavelength dispersive spectrum. A Fourier transform IR (FTIR) spectrometer uses a Michelson Interferometer which creates an interferogram due to the different optical path length of the two interfering beams. Its Fourier transform leads to an IR spectrum. The advantages of this approach are a much better signal to noise ratio and the short time required to record a spectrum. Here the whole spectrum is obtained in parallel, rather than sequentially, with no loss of radiation at the entrance and exit slits of a monochromator. A spectrum may be obtained within a second, and thus it is possible to follow the kinetics of corrosion processes. An additional advantage of FTIR spectroscopy is the high spectral resolution of $0.1\,cm^{-1}$.

A variety of detectors are utilized in IR spectroscopy, based either on thermal or photonic response. Thermal detectors are IR absorbers that measure the produced heat via thermocouples, or bolometers, which measure the change of heat via the resistance change of a platinum wire. Pyroelectric detectors are also employed. They use chopped IR light, which changes the polarization of a condenser whose capacity is a measure of the absorbed radiation and the thereby produced heat. They have the advantage of suppressing background radiation efficiently. As regards photonic detectors, low band gap semiconductors such as $Cd_xHg_{1-x}Te$ (CMT) are typically used, where the incoming IR radiation excites electrons from the valence to the conduction band leading to photocurrents which are a measure of its intensity. CMT detectors are very sensitive and have a fast response so that they are widely used for FTIR spectrometers.

IR spectroscopy is routinely employed by synthetic chemists in the transmission mode, with specimens pressed as pellets often in KBr as an IR transparent medium or in solution with appropriate organic solvents. In aqueous corrosion and electrochemistry, the chemistry of a metal surface and its changes with the corrosion process are of decisive importance. Therefore, one needs to investigate the surface preferably in its natural environment. Two main problems arise. First, water strongly absorbs IR-radiation. Therefore, the transmission path in water should be avoided or made as short as possible. Furthermore, thin films or absorption layers are often of interest and not the substrate metal or the bulk electrolyte and the information should come from the electrode/electrolyte interface. Therefore two main approaches are used in electrochemistry and corrosion science, namely attenuated total reflection spectroscopy (ATR) and IR-Reflection-Absorption Spectroscopy (IRAS and IRRAS).

2.31.9.1 Attenuated Total Reflection

In ATR spectroscopy, the IR beam travels through a crystal such as Ge or ZnSe with total reflection at its

interfaces, that is, the angle of incidence, measured against the normal surface, is above its critical value of total reflection. One surface of the ATR crystal may be covered with a very thin metal layer which then is exposed to the electrolyte. Although the radiation is totally reflected, the electrical field vector penetrates the metal and samples the electrode–electrolyte interface with a depth of ~0.1–5 μm (evanescent wave). It thus picks up the IR spectrum of adsorbed molecules and thin layers and may follow changes at the interface even for well controlled electrochemical conditions with a potentiostat and counter and reference electrode.

2.31.9.2 IR Reflection Absorption Spectroscopy

In IRAS, the IR-beam strikes the specimen surface from the front side at grazing incidence, i.e., at a large angle relative to the surface normal. This geometry permits an intense interaction of the beam with the electrode–electrolyte interface. For IR radiation polarized parallel to the plane of incidence (containing the incoming beam and the surface normal), that is, p-type radiation, the electrical field vector changes in the direction of the axis of at least partially perpendicularly oriented molecules and thus gives its characteristic interaction and absorption lines. For perpendicularly oriented, s-type radiation one does not get a signal, for due to a 90° phase shift during reflection it cancels with the incoming radiation. These surface selection rules lead to an IRAS-signal for those molecules oriented on the surface with a component of a changing dipole moment in the direction of the normal surface. These characteristics permit one to distinguish between the randomly oriented molecules within the electrolyte and those adsorbed with a characteristic orientation at the metal surface. IR spectroscopy may also contribute to the monitoring and understanding of atmospheric corrosion.[48] In this case, the electrolyte is a thin water containing film and thus does not cause too many problems for the application of the method.

2.31.10 Summary

This chapter describes the most important surface analytical methods which provide information on the chemistry and structure of corroding surfaces and surface layers. Many methods work in vacuum, and one should be aware that the specimen has lost the contact with its environment and the electrode potential for the case of electrochemical corrosion systems. However, such broad and valuable information is urgently needed for an understanding of surface reactions and corrosion phenomena. It has been shown by numerous detailed studies that the results follow the systemically changed experimental parameters, for more extensive studies, when one examines the reaction, free of artifacts. *In situ* methods are free of this problem, but they cannot replace the rich information provided by the methods functioning only in vacuum. Generally, it is the combination of many methods which ensures understanding and a correct interpretation of the reactions and their mechanisms. XPS is a soft method and gives excellent data on the chemistry at the surface. AES allows a high lateral resolution, which is extremely important if localized corrosion phenomena occur. ISS may improve the depth profiles of thin films, and RBS will give a detailed insight into reactions on thick films, however, with a poor depth resolution of ~5 nm only. IR methods may give *in situ* information on the chemistry of the surface whereas XRD and XAS provide information on the structure, that is, the long range and the near range order of surfaces and surface films. In addition, *in situ* scanning tunneling microscopy and atomic force microscopy give exact information on their structure. Thus, all these methods supply details of the characteristics and the reactions occurring at a solid surface. Furthermore, some methods add spectroscopic information of the dissolved species in solution or in the gas phase like XAS in transmission mode. Table 1 summarizes the methods which have been discussed in this chapter and presents their advantages and disadvantages. A group might not have experts in all these methods, but laboratories may collaborate and thus the expertise and the equipment of the laboratories may be complementary. A collaboration will be useful when the different backgrounds of the scientists involved still allows them to communicate, which is a difficult task, as the methods range from corrosion and electrochemistry to surface physics and the application of SR facilities. It is often observed in science that different communities work parallel to each other, largely ignoring each other. However, the situation is improving. One requirement is crucial in all cases. The application of a sophisticated surface method to a system requires the full knowledge of its electrochemistry. Any study without a careful and reliable specimen preparation is useless. On the other hand, all these surface methods are needed to understand the surface reactions. Therefore, the surface physicist

should understand to some extent electrochemistry, corrosion, and materials science, and the corrosion engineer and electrochemist to some extent surface science and its related methods, which would be an excellent basis for a fruitful collaboration of experts.

References

1. Pourbaix, M. *Atlas of Electrochemical Equilibria in Aqueous Solutions*; Pergamon: Oxford, 1966.
2. Briggs, D.; Seah, M. P. *Practical Surface Analysis*; Wiley, 1983.
3. Marcus, P.; Mansfeld, F. Eds. *Analytical Methods in Corrosion Science and Engineering*; Taylor & Francis: Boca Raton FL, 2006.
4. Haupt, S.; Collisi, U.; Speckmann, H. D.; Strehblow, H.-H. *J. Electroanal. Chem.* **1985**, *194*, 179–190.
5. Haupt, S.; Calinski, C.; Collisi, U.; Speckmann, H. D.; Strehblow, H.-H. *Surf. Interface Anal.* **1986**, *9*, 357–365.
6. Strehblow, H.-H. In *Advances in Electrochemical Science and Engineering*; Alkire, R. C., Kolb, D. M., Eds.; Wiley-VCH: Weinheim, 2003; pp 271–378.
7. Hecht, D.; Strehblow, H.-H. *J. Electroanal. Chem.* **1997**, *436*, 109–118.
8. Strehblow, H.-H. In *Bergmann Schäfer, Lehrbuch der Experimentalphysik, Gase, Nanosysteme Flüssigkeiten* de Gryter: Berlin, New York, 2006; Vol. 5, pp 551–555.
9. Abels, J. M.; Strehblow, H.-H. *Corros. Sci.* **1997**, *39*, 115–132.
10. Strehblow, H.-H.; Marcus, P. In *Analytical Methods in Corrosion Science and Engineering*; Taylor & Francis: Boca Raton FL, 2006; pp 1–37.
11. Shirley, D. A. *Phys. Rev. B* **1972**, *5*, 4709–4714.
12. Scofield, H. *J. Electron Spectrosc.* **1976**, *8*, 129–137.
13. Seah, M. P.; Dench, W. A. *Surf. Interface Anal.* **1979**, *1*, 2–11.
14. Haupt, S.; Strehblow, H.-H. *Langmuir* **1987**, *3*, 873–885.
15. Vetter, K. J. *Electrochemical Kinetics*; Academic Press: New York, 1967; pp 780–783.
16. Speckmann, H. D.; Haupt, S.; Strehblow, H.-H. *Surf. Interface Anal.* **1988**, *11*, 148–155.
17. Strehblow, H.-H.; Melliar Smith, C. M.; Augustyniak, W. M. *J. Electrochem. Soc.* **1978**, *125*, 915–919.
18. Mcintyre, N. S.; Graham, M. J. In *Analytical Methods in Corrosion Science and Engineering*; Taylor & Francis: Boca Raton, FL, 2006; pp 65–102.
19. Margeritondo, G. *Introduction to Synchrotron Radiation Research*; Oxford University Press: Oxford, 1988.
20. Lützenkirchen-Hecht, D.; Frahm, R. *J. Phys. Chem. B* **2001**, *105*, 9988–9993.
21. Scherb, G.; Kazimirov, A.; Zegenhagen, J. *Rev. Sci. Instrum.* **1998**, *69*, 512–516.
22. Shi, Z.; Wu, S.; Lipkowski, J. *Electrochim. Acta* **1995**, *40*, 9–15.
23. Zegenhagen, J. *Surf. Sci. Rep.* **1993**, *18*, 200–271.
24. Davenport, A. J.; Sansone, M.; Bardwell, J. A.; Aldykiewicz, A. J., Jr.; Taube, M.; Vitus, C. M. *J. Electrochem. Soc.* **1994**, *141*, L6–8.
25. Kittel, C. *Introduction to Solid State Physics*, 7th ed.; Wiley: New York, 1996.
26. *International Tables for X-ray Crystallography: Mathematical Tables*; Kynoch Press: Birmingham, 1962; Vol. 3. (Published for the International Union of Crystallography.).
27. Davenport, A. J.; Oblonsky, L. J.; Ryan, M. P.; Toney, M. F. *J. Electrochem. Soc.* **2000**, *147*, 2162–2173.
28. Toney, M. F.; Davenport, A. J.; Oblonsky, L. J.; Ryan, M. P.; Vitus, C. M. *Phys. Rev. Lett.* **1997**, *79*, 4282–4285.
29. Paterson, M. S. *J. Appl. Phys.* **1952**, *23*, 805–811.
30. Klug, H. P.; Alexander, L. E. *X-ray Diffraction Procedures*; Wiley: New York, 1974.
31. Magnussen, O. M.; Scherer, J.; Ocko, B. M.; Behm, R. J. *J. Phys. Chem. B* **2000**, *104*, 1222–1226.
32. Parratt, L. G. *Phys. Rev.* **1954**, *95*, 359–369.
33. You, H.; Melendres, C. A.; Nagy, Z.; Maroni, V. A.; Yun, W.; Yonco, R. M. *Phys. Rev. B* **1992**, *45*, 11288–11298.
34. Schwartz, D. T.; Müller, R. H. *Surf. Sci.* **1991**, *248*, 349–358.
35. Chiu, N. S.; Bauer, S. H.; Johnson, M. F. L. *J. Catal.* **1984**, *89*, 226–243.
36. Ressler, T.; Wong, J.; Roos, J. *J. Synchrotron Rad.* **1999**, *6*, 656–658.
37. Lützenkirchen-Hecht, D.; Strehblow, H.-H. In *Analytical Methods in Corrosion Science and Engineering*; Marcus, P., Mansfeld, F., Eds.; Taylor & Francis: Boca Raton FL, 2006.
38. Rehr, J. J.; Albers, R. C. *Rev. Mod. Phys.* **2000**, *72*, 621–654.
39. Gibson, P. N.; Crabb, T. A. *Nucl. Instrum. Methods B* **1995**, *97*, 495–498.
40. Sayers, D. E.; Stern, E. A.; Lytle, F. W. *Phys. Rev. Lett.* **1971**, *27*, 1204–1207.
41. Koningsberger, D.; Prins, R. *X-ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS and XANES*; Wiley: New York, 1988.
42. d'Acapito, F.; Davoli, I.; Ghigna, P.; Mobilio, S. *J. Synchrotron Rad.* **2003**, *10*, 260–264.
43. Salluzo, M.; Natali, F.; Aruta, C.; Maglione, M. G.; Ricci, F.; Koller, E.; Fischer, O.; Saini, N. L. *Eur. Phys. J. B* **2002**, *27*, 467–472.
44. Kruger, J.; Long, G. G.; Zhang, Z.; Tanaka, D. K. *Corros. Sci.* **1990**, *31*, 111–120.
45. Bogaerts, A.; Neyts, R.; Gijbels, R.; van der Mullen, J. J. A. M. *Spectrochim. Acta B* **2002**, *57*, 609–658.
46. Grimm, W. *Spectrochim. Acta B* **1968**, *23*, 443–454.
47. Weiss, Z.; Musil, J.; Vlcek, J.; Fresenius, J. *Anal. Chem.* **1996**, *354*, 188–192.
48. Leygraf, C.; Johnson, M. In *Analytical Methods in Corrosion Science and Engineering*; Taylor & Francis: Boca Raton, FL, 2006; pp 237–268.

2.33 Scanning Probe Microscopies

P. Marcus and V. Maurice
Laboratoire de Physico-Chimie des Surfaces, CNRS (UMR 7045) – Université Pierre et Marie Curie, Ecole Nationale Supérieure de Chimie de Paris, 11 rue Pierre et Marie Curie, 75231 Paris Cedex 05, France

Glossary

Adsorption The first stage of interaction between a surface and its environment (gaseous or liquid), in which atoms or molecules are bound to the surface of a material.

Corrosion The degradation of a metallic material under the effect of the environment.

Dissolution The electrochemical reaction in which the metal is ionized and the cation thus formed goes into solution where it is hydrolyzed.

Localized corrosion Limitation of the corrosion of a metal surface to local areas.

Passivation The property of a metal to form a thin, continuous, compact, and adherent layer of oxide or oxyhydroxide (the passive film) at its surface that blocks the reaction of dissolution of metal atoms.

Passivity breakdown The local breakdown of the passive film.

Pitting The result of localized corrosion of passivated metal surfaces in local areas.

Reconstruction The change of atomic position when the surface of a material has a structure different from a bulk termination.

Superstructure The specific structure formed at the surface of a material.

Abbreviations

AFM Atomic force microscopy
CB Conduction band
DOS Density of states
ECAFM Electrochemical atomic force microscopy
ECSTM Electrochemical scanning tunneling microscopy
ECTS Electrochemical tunneling spectroscopy
SECM Scanning electrochemical microscope
SPM Scanning probe microscopy
STM Scanning tunneling microscopy
STS Scanning tunneling spectroscopy
UHV Ultrahigh vacuum
VB Valence band

Symbols

d Distance between tip and sample surface (m)
d_z Deflection along the axis Z (m)

e Charge of proton (1.6×10^{-19} C)
E_F Fermi level energy (eV)
h Planck's constant (6.63×10^{-34} J s)
I Tunneling current (A)
I_T Tip current (A)
k Spring constant (N m^{-1})
m_e Mass of the electron (9.11×10^{-31} kg)
P_X, P_Y, P_Z Components of the piezoelectric scanner controlling x, y, z position
T Tunneling transmission probability
U Intermolecular potential
U_S Sample electrochemical potential (V vs. reference electrode)
U_T Tip electrochemical potential (V vs. reference electrode)
V_{bias} Bias potential between sample and tip (V)
V_X, V_Y, V_Z Voltages applied to P_X, P_Y, P_Z (V)
X, Y, Z Orthogonal scan axes
ϕ Electron tunneling barrier height (eV)
κ Tunneling decay constant (m^{-1})
$\rho_S(E)$ Density of states of the sample at the energy E (m^{-2})
$\rho_t(E)$ Density of states of the tip at the energy E (m^{-2})

2.33.1 Introduction

Scanning probes microscopies (scanning tunneling microscopy (STM) and atomic force microscopy (AFM)) allow direct three-dimensional (3D) imaging of the surface structure and topography of materials at high resolution (atomic or molecular resolution). Their advent in the 1980s has opened up new prospects for the structural analysis of corroded surfaces, as these microscopes can be applied to both solid–gas interfaces under controlled environments (ultrahigh vacuum (UHV)) and to solid–liquid interfaces under electrochemical control. Moreover, time-resolved imaging allows the investigation of the dynamics of structure modifications produced by corrosion processes. The coupling of STM with spectroscopic measurements (scanning tunneling spectroscopy (STS)) provides information on a local scale on the electronic properties of surfaces.

This chapter describes the principles and operation of STM/STS and AFM and their electrochemistry implementation (ECSTM and ECAFM) for the analysis of corrosion processes at surfaces. Applications of ECSTM, ECTS, and ECAFM are illustrated by selected examples, emphasizing data obtained *in situ* at the solid–liquid interface on the anodic dissolution of metals, the growth mechanism and structure of passive oxide layers and their electronic properties at the (sub)nanometer scale, and on the initiation of localized corrosion at the (sub)micrometer scale.

2.33.2 Scanning Tunneling Microscopy

This technique is unique in that it provides 3D real space images and allows localized measurements of geometric and electronic structures at high resolution. The first atomically resolved images were published in 1983.[1]

2.33.2.1 Principle and Operation

STM is based on tunneling of electrons between two electrodes separated by a potential barrier and biased by a voltage V_{bias} (**Figure 1**). Electrons within eV_{bias} of the Fermi level tunnel from the occupied states of the negatively biased electrode to the unoccupied states of the positively biased electrode. Assuming constant density of these electronic states, the tunneling current through the planar barrier increases linearly with V_{bias} and decreases exponentially with the barrier width d according to eqn. [1]

$$I \propto V_{bias} \exp\left(-2\kappa d\right) \qquad [1]$$

where κ is the tunneling decay constant in the barrier, given by eqn. [2]

$$\kappa = \frac{2\pi}{h}\sqrt{2m_e\phi} \qquad [2]$$

with h being the Planck's constant, m_e the electron mass, and ϕ the tunnel barrier height. For electrons at the Fermi level, ϕ is the work function of the negatively biased electrode. Since most work functions are 4–5 eV, $\kappa \sim 1$ Å^{-1} we find that for a bias of 0.1 V, a tunneling current of ~1.5 nA can be observed for a separation between the two electrodes of ~0.9 nm. The tunneling current drops by nearly one order of magnitude for every increase of 0.1 nm of the barrier width. A direct consequence of this exponential decay is that if one electrode is a tip terminated by an atomic asperity protruding by ~0.1 nm at the apex (**Figure 1**), ~90% of the current is spatially limited to this atomic asperity.

STM operation is based on precise control of the position of the tip with respect to the surface to be analyzed. The tip is placed on a tripod formed by three independent piezoelectric elements, P_X, P_Y, P_Z, controlled by the voltages V_X, V_Y, V_Z respectively (**Figure 1**). At a given bias voltage, the tip is scanned

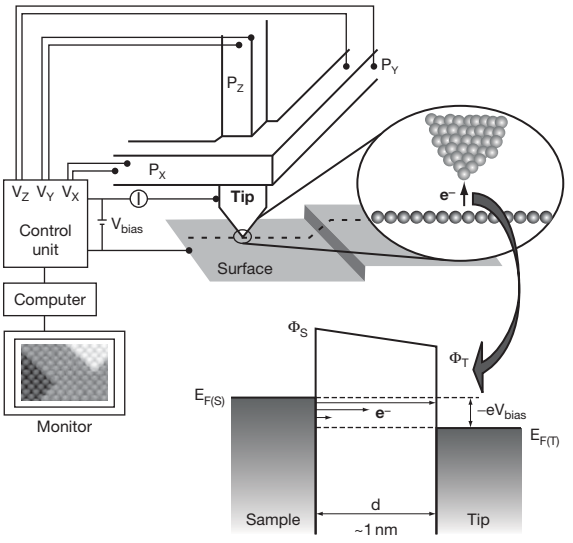


Figure 1 Schematic diagram of the scanning tunneling microscope and principle of electron tunneling through a biased potential barrier between sample and tip.

along the two lateral directions X and Y parallel to the surface, while a feedback loop constantly adjusts the tip vertical position along the Z-axis, so as to keep the current to a constant setpoint value chosen by the operator. This is the 'constant current' mode where κd is kept constant. If the local work function ϕ remains constant, a constant current corresponds to a constant distance d between tip and surface. The shape of the surface is then reproduced by the path of the tip. The values of $V_Z = f(V_X, V_Y)$ are used to produce a 3D image of the surface most often displayed on the monitor, using a color or gray scale. The microscope can also be operated in 'constant height' mode with the feedback slowed down or switched off, but at the risk of tip damage, as the tip height is not adjusted.

The setpoint current used can vary from a few tens of picoamperes for poorly conductive surfaces to a few nanoamperes for metal or doped semiconductor surfaces. Bias voltages can vary from a few volts to a few millivolts. High quality tips can be prepared from tungsten or Pt–Ir wires by electrochemical etching.[2] The lateral (spatial) resolution of the instrument depends on the quality of tip preparation; it is, at best, typically ~0.1 nm. The vertical (depth) resolution depends on the quality of the damping system used to isolate the instrument from external vibrations and on the precision of the control of the tip–sample distance; it is ~0.001 nm at best. Most commercial instruments adhere to specifications that allow imaging at atomic resolution on reference samples.

2.33.2.2 Scanning Tunneling Spectroscopy

STM images also contain information on the local electronic structures of the surface and tip. According to the theory of STM,[3] the tunneling current is expressed as (eqn. [3])

$$I = \int_0^{eV_{\text{bias}}} \rho_s(r, E)\rho_t(r, E - eV_{\text{bias}})\, T(E, eV_{\text{bias}})\mathrm{d}E \qquad [3]$$

where $\rho_s(r, E)$ and $\rho_t(r, E)$ are respectively the density of states (DOS) of the sample and tip at location r and energy E, measured with respect to their individual Fermi levels. $T(E,eV_{\text{bias}})$ is the tunneling transmission probability for electrons with energy E at applied bias voltage V_{bias}. For a fixed surface-to-tip distance and assuming to a first approximation, a constant $\rho_t(r, E)$, eqn. [4] can be deduced:

$$\frac{\mathrm{d}\ln I}{\mathrm{d}\ln V_{\text{bias}}} = \frac{\mathrm{d}I}{\mathrm{d}V_{\text{bias}}} \times \frac{V_{\text{bias}}}{I} \propto \rho_s(E) \qquad [4]$$

It follows that $\rho_s(E)$ can be extracted from measurements of the tunneling current as a function of the bias voltage applied to the tunneling barrier, that is, from tunneling spectroscopy I–V_{bias} measurements.[4]

In tunneling spectroscopy (TS), quantitative information on surface electronic states is obtained using either modulation techniques or numerical techniques to measure $\mathrm{d}I/\mathrm{d}V_{\text{bias}}$ at constant average tunneling distance as a function of V_{bias} at a single point of the surface. Such measurements can reveal structures in the surface density of states that can arise from critical points in the surface-projected bulk band structure or from true surface states associated with surface reconstructions, adsorbates, or ultrathin films formed on the surface. In scanning tunneling spectroscopy (STS), spectroscopic I–V_{bias} measurements are combined with topographic measurements to extract laterally resolved information and map $\rho_s(E)$ spatially. In the so-called current imaging tunneling spectroscopy (CITS) mode, I–V_{bias} curves are recorded for every pixel of the topographic image with the feedback loop switched off in order to operate at a fixed surface-to-tip distance. The interest of corrosion science in these techniques is the possibility of acquiring laterally resolved chemical information on surfaces.

2.33.2.3 Application to the Solid–Gas Interface and Limitations

STM was originally designed for application at the solid–gas interface under UHV. The environmental control of the interface inherent to UHV allows one to

locally characterize, at the atomic scale, the surface structure of semiconductors, superconductors, and metallic materials, along with the modifications induced by the adsorption of atoms and molecules or by the deposition of nanoparticles. The technique is also widely used to study nanostructuring processes resulting from surface modifications. Operation at low temperature (down to 4 K) permits manipulation of atoms, molecules, and nanoparticles. Operation at high temperature (up to ~900 K) limits space resolution due to acceleration of surface diffusion and thermal drift (uncontrolled variation of tip position with respect to the surface), and the time resolution due to activation of surface reactions.

For corrosion related studies, UHV-STM is mostly applied to study the initial stages of oxidation of metallic materials. The adsorption of oxygen on metallic materials has been the most studied topic. A review of data obtained on Cu, Ni, and Ag single-crystal surfaces has been published.[5] The technique allows one to determine the adsorption site of the adsorbate, the adsorbate superstructure including the reconstruction of the substrate possibly induced by adsorption, and the nature and role of defects.

The growth mechanism of the superstructure and surface diffusion can be studied by acquiring consecutive images from a selected area as a function of time (time-resolved imaging). However, the time resolution is of the order of several tens of seconds per image with most commercial microscopes and the application is restricted to relatively slow processes. Relatively fast physical processes can be slowed down by operating at low temperature, or alternatively, fast-scanning microscopes with a time resolution better than one second per image can be used. One such example of fast-scan STM applied to the solid–liquid interface is presented further on.

Time-resolved UHV-STM can also be applied to the solid–gas interface to study the mechanism of phase transformation from metals to oxides. The local nature of measurement is then most effective to characterize the sites of nucleation of the oxide phase and the role of surface defects, the growth mechanism of the nuclei until full coverage of the surface, and the structure of the oxide nuclei. STS can be combined with STM to locally characterize the oxide nuclei and the unmodified metallic areas.[6]

Because of its sensitivity to the topmost surface atomic layer only, STM is less effective in characterizing the structure of 3D films formed by oxidation at the surface of metals and alloys. Indeed, neither the inner structure of the film nor the buried metal–oxide interface can be accessed, thereby preventing a detailed investigation of the mechanism of 3D growth of the oxide film. Rare exceptions of imaging of the buried metal–oxide interface have been reported when an ultrathin oxide film transparent to the tunneling electrons is formed on the metallic substrate.[7] In most cases, only the surface of the 3D oxide films can be characterized, including its atomic structure and nanostructural defects. However, the technique is well suited to locally characterize the reactivity of 3D oxide films and their resulting surface structural modifications. It has been applied, for example, to the interaction of water vapor with thermal oxide films of chromia grown on Cr surfaces[8] and alumina on NiAl surfaces.[9]

2.33.3 Electrochemical STM

STM was first applied in solution in 1986[10] and to corrosion studies in 1988.[11] A major difference with UHV operation is the decrease of the tunneling barrier height in solution. Experimental values ranging from 0.1 to 0.3 eV have been reported,[12,13] which is much smaller than the values of ~4 eV for UHV operation.

2.33.3.1 Electrochemical Implementation and Limitations

In ECSTM, the tip is immersed in the electrolyte and acts as a fourth electrode added to a conventional three electrode cell with working reference and counter electrodes (**Figure 2**). Charging of the double layer and electrochemical reactions take place at the tip–electrolyte interface. They generate capacitive and faradic currents at the tip that superimpose onto the tunneling current and must be minimized as they affect STM operation based on the measurement of the total tip current. The most effective minimization is obtained by reduction of the tip surface in contact with the electrolyte. This is done by covering most of the immersed portion of the tip by an insulating coating. Glass, epoxy varnish, silicone polymer, Apiezon wax, nail polish, or electrophoretic paint can be used. Well-prepared coatings will leave only ~10 μm or less, uncoated at the very end of the tip. Further minimization of the electrochemical currents at the tip is achieved by polarizing the tip in the double layer charging region (PtIr) or at the corrosion potential (W). This combined minimization allows one to decrease the electrochemical current to values of 50 pA or less, extremely small relative to the usual set-point current of 0.5–10 nA used in ECSTM,

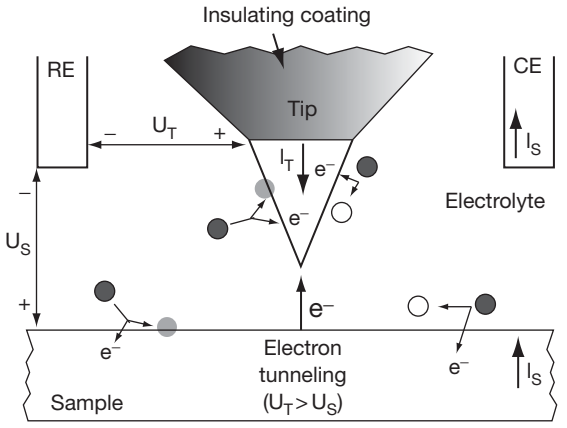


Figure 2 Schematic diagram of the four electrode cell used in ECSTM. RE is the reference electrode. CE is the counter electrode. U_S and U_T are electrochemical potentials applied to the sample and tip, respectively. I_S and I_T are the sample current and tip currents, respectively. For $U_T > U_S$, electrons tunnel from the occupied states of the sample to the empty states of the tip.

thus ensuring no interference with the functioning of the STM.

ECSTM, performed in a four-electrode cell, requires the use of a bipotentiostat that enables one to control independently the electrochemical potentials of the tip (U_T) and sample (U_S) relative to a common reference electrode. The bias voltage for STM operation is then given by the difference between U_T and U_S. The current resulting from the electrochemical reaction taking place at the sample–electrolyte interface is measured by the counter electrode, as in a conventional electrochemical cell.

There are limitations related to the local nature of the measurements. The high spatial resolution capability of the instrument, which provides the basis for the ability to image individual atoms or molecules, brings down the analyzed area to a tiny fraction of the total surface area (typically ($1 \times 1\,\mu m^2$) or less). A correlation between STM-derived and electrochemical information (integrated over areas of $1\,cm^2$ or less) is really meaningful only if the area probed by STM is representative of the whole electrode surface. Limitations can also arise from the possible interference of the tip with the electrochemical process at the working electrode surface. The close proximity of the tip can cause shielding effects for reaction at the sample–solution interface. An interference of sample and tip electric double layer is also possible.[12,14] Tip-induced corrosion processes can also occur.[14,15] Corrosion studies by ECSTM may also be limited by the small volume of the electrochemical STM cells. It is indeed not possible to investigate the substrate modifications resulting from extensive corrosion processes as the observation of the substrate surface would be masked by redeposition of corrosion products.

Despite these limitations, ECSTM investigations can provide invaluable information, in particular molecular scale information, on the structural modifications related to corrosion processes such as the active dissolution and passivation of metallic surfaces as illustrated below, as well as the adsorption of corrosion inhibitors[16] or the reactivity of anodic oxide films determining the initial stages of localized corrosion of passivated metallic materials.[17]

2.33.3.2 Application to Corrosion Analysis at the Solid–Liquid Interface

The high resolution provided by ECSTM allows structural characterization of corrosion processes at the molecular level. This requires appropriate control of electrochemical conditions in order to minimize the rate of corrosion processes to a point compatible with the STM time resolution. In contrast with UHV-STM, corrosion reactions cannot be slowed down by operating at low temperatures because of the liquid environment, but electrochemical potential can be used to control the reaction kinetics. High resolution ECSTM also requires high quality of surface preparation and control, that is, atomically flat surfaces of well-defined structure. This can be achieved with single-crystals of known orientation whose surface has been appropriately prepared to have a terrace and step topography.

2.33.3.2.1 Active dissolution of metals

ECSTM is very effective to study *in situ* the active dissolution of metals because the surface is free of passivating oxides and for pure metals, the electrode surface remains terminated by an atomically flat terrace and step topography during the corrosion process. This is a very favorable situation for atomic-scale imaging of dynamic processes at surfaces. Investigations can focus on the structural modifications of atomically flat surfaces resulting from adsorption of anions in the double layer potential region preceding the onset of dissolution and on the dissolution of metal atoms at specific sites. The formation of well-ordered superstructures, more corrugated than the metallic lattice, facilitates high resolution ECSTM imaging of the surface structure. For alloys,

dealloying can be studied. The dynamics of these processes can be studied by time-resolved imaging of a selected region. High-resolution ECSTM has been applied to the anodic dissolution of pure metals, Cu,[18–20] Ni,[21,22] Ag,[23] Co,[24] and Pd,[25] and alloys, CuAu.[26]

Figure 3 illustrates the process of dissolution, proceeding layer-by-layer *via* a step-flow mechanism at moderate potentials (i.e., slow etching rate). It was observed in this case at atomic resolution for Cu(001) in HCl_{aq} with a fast-scan ECSTM.[20] Preferential etching of the surface takes place at the defects

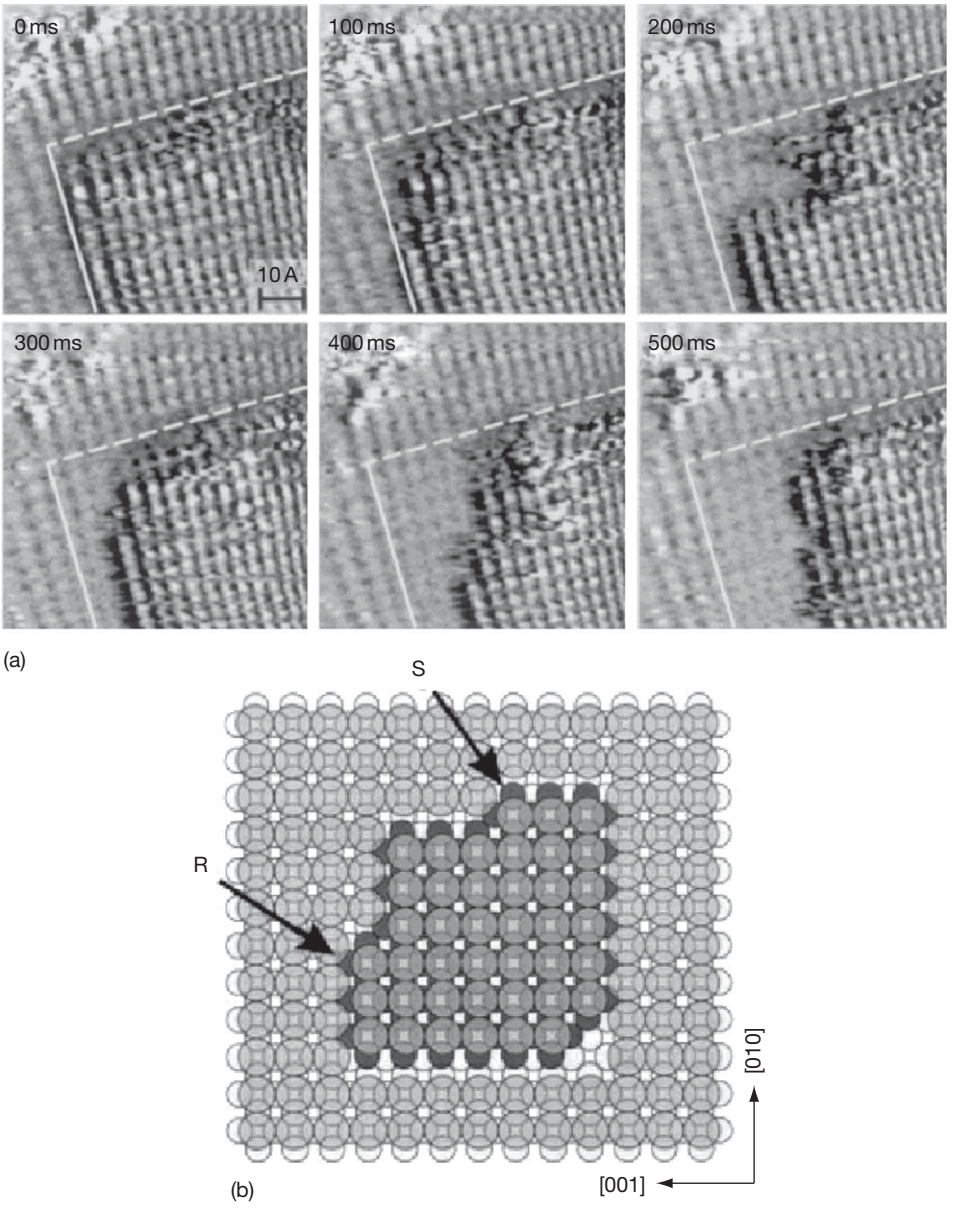


Figure 3 (a) Sequence of fast-scan ECSTM images (0.1 s per image) recorded on Cu(001) in 0.01 M HCl_{aq} at –0.17 V/SCE showing dissolution of a Cu terrace starting at the outer terrace corner. The observed lattice is formed by adsorbed chlorine. The markedly different dissolution behavior at the active and stable steps whose initial position is marked by solid and dashed lines, respectively, is clearly visible. (b) Model of the (001) surface of a face-centered cubic metal (e.g., Cu) covered by an ordered $c(2 \times 2)$ adlayer (e.g., Cl) showing different structure of steps and kinks along the [010] and [001] directions. R and S mark the reactive and stable Cu atoms at the outer kinks, respectively. Reproduced from Magnussen, O. M.; Zitzler, L.; Gleich, B.; Vogt, M. R.; Behm, R. J. *Electrochimica Acta* **2001,** *46,* 3725–3733.

corresponding to preexisting step edges. Selective etching of atoms at these sites results from their lower coordination to nearest neighbor atoms.

The superstructure formed by strongly adsorbed ions in the double layer region influences the anisotropy of the etching. In the absence of strongly adsorbed anions, the etching process, at moderate overpotentials, stabilizes the step edges oriented along the close-packed directions of the crystal where the nearest neighbor coordination of the atoms is maximum, e.g., along the atomically smooth <1–10> and <10–10> directions for (111)-oriented fcc metals and (0001)-oriented hcp metals respectively. On Cu(001), a highly ordered $c(2 \times 2)$ adlayer (observed in **Figure 3**) is formed in HCl_{aq} solutions. It stabilizes the step edges along the close-packed directions of the superstructure corresponding to the <100> substrate directions. Atomic-scale imaging shows that the dissolution proceeds at structurally well-defined kinks of the [100]- and [010]-oriented steps *via* removal of the primitive unit cells of the Cl adlayer. Another effect of the $c(2 \times 2)$ adlayer is an induced anisotropy of the dissolution process along the symmetrical <100> directions of the step edges. This is also visible in **Figure 3** where one orientation of the step edges dissolves more rapidly than the other. This results from the structural anisotropy of the <100> step edges induced by the presence of the $c(2 \times 2)$ adlayer, as illustrated by the model shown in **Figure 3**. This has been tentatively explained by the coordination of the outmost Cu atom forming the dissolving kink to the adjacent Cl adsorbates.[20]

In the double layer potential range, equilibrium fluctuations of these step edges are observed, which result from local removal/redeposition processes. These fluctuations are also illustrated by the sequence of fast-scan images shown in **Figure 3**. At more anodic potentials, the dissolution process prevails and a net removal is observed; but re-deposition still occurs.

2.33.3.2.2 Growth and structure of passive films

In the area of passivation, high-resolution ECSTM measurements have also been obtained with well-prepared single-crystal surfaces, bringing new insight into the mechanisms of self-protection of metal and alloys against corrosion in aqueous solutions. Data on this topic are available for Cu,[27–29] Ni,[21,22,30,31] Ag,[32] Co,[33] Fe,[34], Cr,[35] Al,[36] and ferritic[37] and austenitic[38] stainless steels.

The structural modifications of the electrode occurring at the active/passive transition as well as in the potential range preceding oxide formation are illustrated here for Cu(111) in aqueous sodium hydroxide solution.[27–29] **Figure 4** illustrates the growth of the adsorbed hydroxide layer on Cu(111) at potentials below that at which copper oxide is formed. The sequence of images shown in **Figure 4(a)** was obtained after stepping anodically the potential to $-0.6\,V_{SHE}$, below the active/passive transition at $\sim -0.2\,V_{SHE}$. The atomically smooth terraces of the surface (marked M) become progressively covered by darker-appearing islands (marked ad) that grow laterally and coalesce to cover the terraces completely. The darker appearance of these ad-islands is a typical effect of adsorbed O species also observed at the solid–gas interface.[5] The adsorbed layer preferentially grows at the step edges, confirming the preferential reactivity of these defect sites of the surface. The terraces grow laterally due to displacement of the step edges, and monoatomic ad-islands are formed at the end of the growth process. These two features are indicative of the reconstruction of the topmost Cu plane induced by the adsorption of the OH groups. The reconstruction causes ejection of Cu atoms from the original metallic surface layer. These ejected atoms diffuse on the surface and aggregate at step edges, which cause the observed lateral displacement of the step edges. In the final stages of the adsorption process, when most of the surface is already covered by the adlayer, the ejected atoms have reduced mobility on the OH-covered terraces and aggregate to form the observed monoatomic ad-islands.

Figure 4(b) shows an atomically-resolved image and model of the ordered OH_{ads}/Cu(111) surface that confirm the reconstruction of the Cu(111) topmost atomic plane. A hexagonal lattice with a parameter of 0.6 ± 0.02 nm is measured. The interatomic distance between the Cu atoms is $\sim$0.3 nm, which is larger than the interatomic spacing of 0.256 nm in Cu(111), and confirms the reconstruction of the topmost Cu plane into a plane of lower atomic density. A coverage of $\sim$0.2 OH per Cu(111) atom is deduced from the density of the OH_{ads}, in excellent agreement with the coverage of 0.19 obtained from the electrochemical charge transfer measurements. In addition, the STM data show that the OH_{ads} sit in the three-fold hollow sites of the reconstructed Cu plane on which they form a (2×2) structure.

Figure 5 illustrates the influence of the potential on nucleation, growth, crystallization, and structure of the Cu_2O oxide film formed in the potential range of Cu(I) oxidation. At low oversaturation (**Figure 5(a)**), monolayer thick islands poorly crystallized and partially

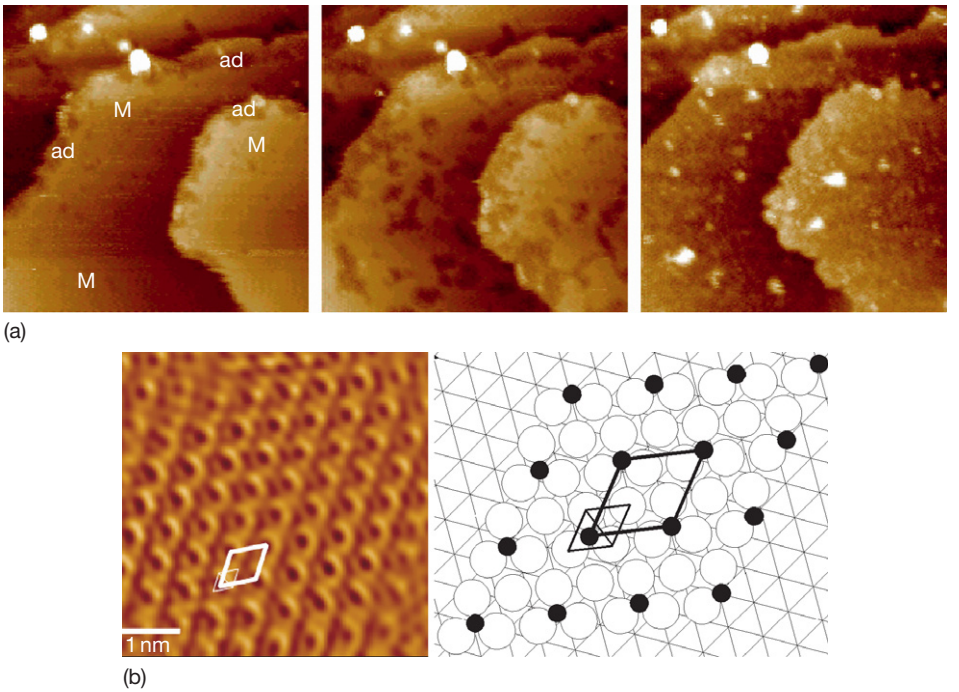


Figure 4 (a) Sequence of ECSTM images (38 s per image) showing growth of an adlayer of OH groups on Cu(111) at –0.6 V_{SHE} in 0.1 M $NaOH_{aq}$. Adapted from Maurice, V.; Strehblow, H.-H.; Marcus, P. *Surface Science* **2000,** *458,* 185–194. (b) ECSTM image and model of the ordered structure of adsorbed OH groups formed on Cu(111). The large and small cells mark the lattice of adsorbed OH and reconstructed copper, respectively. Adapted from Kunze, J.; Maurice, V.; Klein, L. H.; Strehblow, H.-H.; Marcus, P. *Electrochimica Acta* **2003,** *48,* 1157–1167.

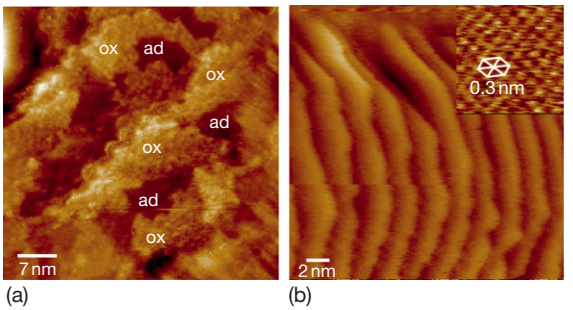


Figure 5 ECSTM images showing the effect of the potential on the growth and structure of the Cu(I) oxide formed on Cu(111) in 0.1 M $NaOH_{aq}$ at −0.25 V_{SHE} (a) and −0.20 V_{SHE} (b). In (a), noncrystalline oxide islands (ox.) separated by the adsorbed OH layer (ad) partially cover the substrate. In (b), a crystalline oxide layer fully covers the substrate. Its atomic lattice (shown in the inset) corresponds to $Cu_2O(111)$. Adapted from Kunze, J.; Maurice, V.; Klein, L. H.; Strehblow, H.-H.; Marcus, P. *Journal of Physical Chemistry B* **2001,** *105,* 4263–4269, with permission from American Chemical Society.

covering the substrate, are formed after preferential nucleation at step edges. They are separated by islands of the ordered hydroxide adlayer. At higher oversaturation (Figure 5(b)), a well crystallized multilayer thick film is formed, and the step edges are not preferential sites of nucleation. The equivalent thickness of the oxide layer can be deduced from subsequent measurements of the charge transfer during cathodic reduction scans. It was ∼0.5 and 7 equivalent monolayers (ML) of $Cu_2O(111)$ after growth at –0.25 and –0.2 V_{SHE}, respectively. The observed lattice of the ordered oxide layer is hexagonal with a parameter of ∼0.3 nm, consistent with the Cu sublattice in the (111)-oriented cuprite. The oxide grows in parallel (or antiparallel) epitaxy ($Cu_2O(111)$ $[1\bar{1}0]||Cu(111)$ $[1\bar{1}0]$ or $[\bar{1}10]$).

The crystalline Cu(I) oxide layer has a nanostructured surface consisting of facets. The surface faceting results from a tilt of a few degrees of the orientation of the oxide lattice with respect to the Cu lattice. A possible origin of the tilt is the relaxation of the epitaxial stress at the metal–oxide interface resulting from the large mismatch between the two lattices. The height of the surface steps of the oxide layer corresponds to 1 ML of cuprite, indicating an identical chemical termination of the Cu_2O (111) oxide terraces. It is thought that the surface of

the oxide layer is hydroxylated in the aqueous solution and that the measured lattice corresponds to OH and/or OH^- groups forming a (1 × 1) layer on the Cu^+ planes of the (111)-oriented cuprite layers.

2.33.3.3 Electrochemical Tunneling Spectroscopy

With the potential of the substrate fixed by the electrochemical process under study and that of the tip adjusted to minimize the electrochemical tip current and optimize the STM measurement, the tunnel voltage can no longer be adjusted in a wide range as in UHV conditions. Thus, spectroscopic I–V_{bias} measurements (TS) are most often restricted to a few hundreds millivolts range in an ECSTM cell.[12] However, recent developments allowing *in situ* single-point spectroscopic measurements (ECTS) over a larger range of tunnel voltage of nearly 3 V have been applied to Fe in a borate buffer solution.[13] The developments are based on the use of noble metals for tips (PtIr) combined with improved performance in tip isolation using electrophoretic coating. Ultrafast voltage ramps (up to $20\,V\,s^{-1}$) are applied to vary the tip potential (U_T) at a fixed sample potential (U_S) and at a tunneling distance defined by the initial set-point conditions. The tip current (I_T) is measured during the U_T ramp with the feedback loop switched off. I_T–U_T curves measured at distances beyond the tunneling gap (with the tip retracted) are subtracted from I_T–U_T curves measured at predefined tunneling distances. This procedure is used to eliminate capacitive and faradaic contributions to the tip current.

Tunneling spectra obtained by scanning U_T between −800 and +1500 mV/SCC at five different U_S values corresponding to various oxidation states of the Fe electrode versus U_T are plotted in **Figure 6**. The lower panel shows the tunneling conductance dI_T/dU_T, proportional to the Fe electrode DOS. It is obtained by numerical differentiation of the curves in the upper panel. The upper panel shows linear spectra for iron polarized at −800 and −650 mV. This results from a high and nearly constant density of surface states (or high surface conductance) of the metallic Fe(0) electrode. At −400 mV, an hydrated Fe(II) layer, a few nanometers thick and with unknown electronic properties, formed on the electrode surface in the specific conditions used in this study. A double exponential behavior is measured revealing a region of low current due to a decrease of the DOS. At 500 mV, the oxide layer is oxidized to

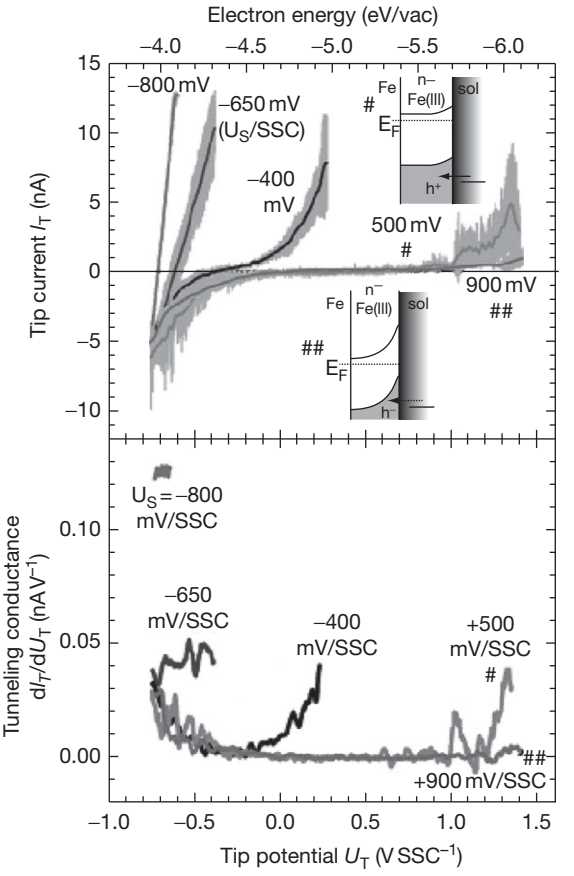


Figure 6 ECTS spectra of Fe polarized at U_S= −800, −650, −400, +500, and +900 mV in a borate buffer aqueous solution. The upper panel shows tip current vs. tip potential curves. Solid lines and gray bands show average curves and standard deviations, respectively. The lower panel shows the conductance curves. See text for details. Adapted from Díez-Pérez, I.; Güell, A. G.; Sanz, F.; Gorostiza, P. *Analytical Chemistry* **2006**, *78*, 7325–7329, with permission from American Chemical Society.

Fe(III) with well-characterized n-type semiconductive properties. A band gap in the DOS increases the separation between the exponential current branches. Its width is consistent with values reported for passivated iron (E_g = 1.6–1.9 eV). The positive current branch results from hole injection from the tip to the valence band (VB). At 500 mV, it is maintained by the slight depletion of the Fe(III) oxide film (upper inset marked #). At 900 mV, the positive current branch becomes extinguished due to stronger depletion conditions and demonstrates low density of charge carriers (upper inset marked ##). Both at 500 and 900 mV, the negative current branches result from electron injection from the tip to the conduction band (CB) of the oxide.

This spectroscopic data demonstrates the capability of ECTS to obtain *in situ* the electronic structure of the solid/electrolyte interface and measure the availability of charge carriers at the electrode surface. Series of tunneling spectra can be acquired stepwise over a large range of polarization potential and used to plot so-called conductograms (map of the conductance (i.e., the charge exchange properties) of the electrode as a function of potential), to study modifications resulting from a variation of the pH of the electrolyte or from the presence of chlorides. Up to now, such tunneling spectra have only been obtained in single-point mode, without combined imaging of the surface.

2.33.4 Atomic Force Microscopy

In contrast with STM, AFM, invented in 1986,[39] does not require conductive samples and tips. It can measure and image local forces between a surface and a tip, including van der Waals, Born repulsion, electrostatic and magnetic forces, friction, and adhesion.[40] The following presentation is restricted to topographic data derived from measurement of local forces on a length scale of 10^{-10} to 10^{-8} m.

2.33.4.1 Principle and Operation

AFM is based on the variation of the interactions between two atoms or molecules. At very short distances, the interaction is repulsive and dominated by the electrostatic repulsion felt by the electrons as the orbitals start to overlap. At larger distances, the interaction is attractive and dominated by the dipole–dipole interaction between molecules. These interactions vary as inverse power laws of the distance, $1/d^{12}$ and $1/d^{6}$ for repulsive and attractive interactions, respectively. Their sum defines the total intermolecular potential U.

In AFM, this is the force $((-\partial U/\partial d)$ that is measured. The tip is positioned at the end of a cantilever spring (**Figure 7**). The deflection, d_z, of the cantilever is measured by an optical setup consisting of a laser beam reflected from the back side of the cantilever and impinging on a partitioned photodiode. The force on the cantilever is simply $-k \cdot d_z$, k being the cantilever's spring constant. As in STM, a piezoelectric setup allows precise control of the tip position relative to the surface. A feedback loop is used to adjust the tip–surface distance in order to maintain the deflection (force) of the cantilever at a set-point

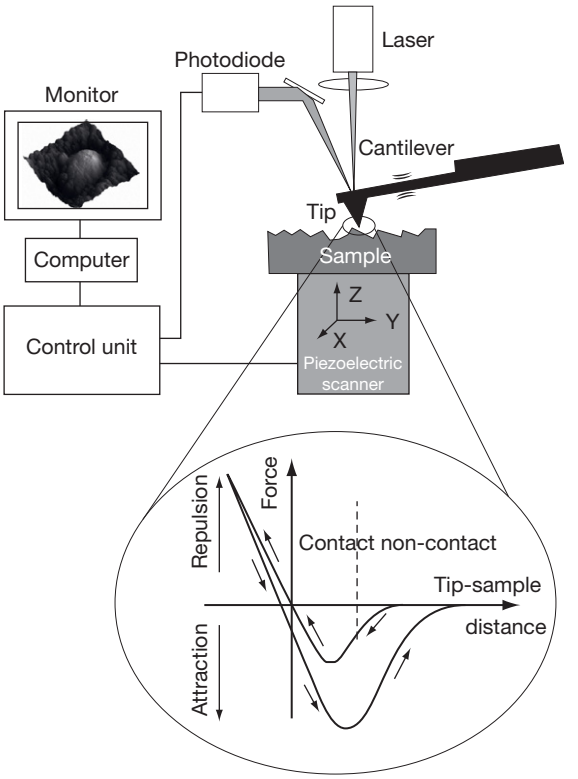


Figure 7 Schematic diagram of the atomic force microscope and typical force curve showing the force on the cantilever as the tip–sample distance is varied.

value chosen by the operator when scanning the X and Y axes. This is the 'constant force' mode of imaging in which the tip reproduces the topographic profile of the surface. The lateral spatial resolution of the instrument is ~0.1 nm when the interaction can be spatially limited to the atomic asperities of the tip. The vertical resolution is ~0.02 nm.

The AFM user can choose to operate the microscope at any point along the typical force curve obtained when varying tip–sample distance (**Figure 7**). Contact occurs at the inflection point of the approaching curve where the curvature (i.e., derivative) changes from attractive to repulsive. Contact increases after this point as the tip–sample distance decreases. Contact-mode imaging is at risk of a potentially destructive interaction that can damage tip and/or sample. In air, stray adhesive forces (10^{-7} to 10^{-8} N), resulting mostly from the presence of a condensed layer of water or other contaminants forming a capillary bridge between tip and surface, and at the origin of the hysteresis in the force curve, interfere to decrease the spatial resolution of the imaging. In addition, the dragging motion of the tip on the surface involves lateral shear force that can

distort measurement severely and induce lateral motion and/or tearing of surface features. The adhesive force is drastically reduced ($\leq 10^{-9}$ N) in liquid applications where the whole cantilever–tip/surface assembly is immersed in solution. However the energy dissipated (≤ 1 eV) by local interaction remains of the order of the energy of chemical bonds, and imaging in these conditions remains potentially destructive.

In the resonant (or Tapping®) mode, the cantilever–tip assembly oscillates at or near its resonance frequency with an amplitude of several nanometers. The oscillating tip is then moved toward the surface until contact. During scanning, the vertically oscillating tip alternately contacts the surface for force measurement and lifts off for implementation of the X,Y scan. This avoids dragging of the tip across the surface. The tip–sample separation is adjusted *via* the feedback loop of the microscope to maintain a constant oscillation amplitude and force on the sample. This resonant mode inherently prevents the tip from sticking to the surface and causing damage during scanning.

The noncontact mode is obtained for tip–surface distances larger than the one where the curvature of the force curve in **Figure 7** changes from negative to positive. Unfortunately, the attractive van der Waals forces are substantially weaker than forces used in contact mode and force resolution is poor. The consequences are a poor spatial resolution and a problematic thermal drift. The noncontact mode is more appropriate for UHV operation where the sample and tip can be cleaned from contaminants and the thermal drift is much less problematic. True atomic resolution can then be achieved when the tip–sample attractive force is such that the interaction is limited to a single atom at the tip apex.

2.33.4.2 Electrochemical Implementation and Limitations

In ECAFM (electrochemical atomic force microscopy) the cantilever-tip assembly does not constitute a fourth electrode through which the electrochemical currents could flow, interfering with the signal used to probe the surface. Immersing the cantilever tip assembly and surface in solution allows better control of the adhesive force in contact mode (see above). In resonant contact mode, the fluid medium tends to damp the normal resonance frequency of the cantilever, which complicates this mode of operation. The noncontact mode becomes impractical because the van der Waals forces are even smaller, which can be a substantial limitation (e.g., for biological applications).

The limitations of ECAFM regarding the correlation between local information and integrated macroscopic information, possible tip-shielding effects on the corrosion behavior, kinetics of mass transport and small volume of the electrochemical cell are the same as those described for ECSTM.

2.33.4.3 Application to Corrosion Analysis at the Solid–Liquid Interface

The application of ECAFM to corrosion analysis is much more widespread than that of ECSTM. ECAFM is often preferred because of its easier setup and because the obtained topographic information is independent of the conductivity of corrosion products. Moreover, ECAFM is most often used at the (sub) micrometer scale, i.e., at a lower level of spatial resolution that does not require preparation of atomically smooth surfaces as for ECSTM studies. ECAFM or AFM is often combined with a variety of other techniques to analyze corrosion and optimize corrosion protection properties. Corrosion protection by inhibitors[41] and coatings,[42] atmospheric corrosion,[43] passivation,[44] passivity breakdown,[45] localized corrosion,[46,47,48,49] and stress corrosion cracking[50] have been studied. We present examples of application to localized corrosion at inclusions.

The lower spatial resolution of AFM is appropriate to study pitting corrosion at microstructural defects (inclusions, grain boundaries, interphase boundaries) of metallic materials. **Figure 8** shows two images of localized corrosion observed *in situ* at the inclusion–matrix interface on an Al–Si(7.7 wt%) alloy under anodic polarization close to breakdown potential in 10 mM $NaCl_{aq}$ + 50 mM KI_{aq}. **Figure 8(a)** shows localized dissolution occuring at the boundary surrounding a round-shaped particle of ~1 μm, whereas **Figure 8(b)** shows localized dissolution occuring between two irregular shaped particles (1–2 μm size), while the boundary region to the matrix remains intact. The different nature of the particles present in the alloy (predominantly AlFeSi and eutectic Si phases) was advanced to explain the different corrosion behaviors of these particle/matrix boundaries.[49] A commonly advanced explanation for the formation of trenches around inclusions in Al alloys is that the inclusions act as cathodic sites for reduction of oxygen, causing the formation of an alkaline environment in the immediate vicinity of each inclusion, confirmed by local pH measurements.[46] This local increase of pH promotes local dissolution of the Al matrix.

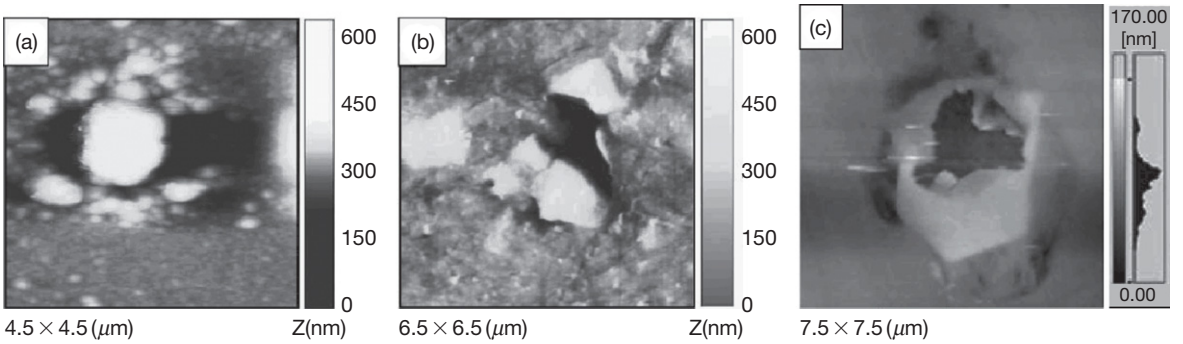


Figure 8 (a,b) ECAFM images of an Al–7.7 wt% Si alloy under anodic polarization close to breakdown potential in 10 mM $NaCl_{aq}$ + 50 mM KI_{aq}. Localized corrosion around (a) and in between (b) inclusions is observed. Reproduced from Davoodi, A.; Pan, J.; Leygraf, C.; Norgren, S. *Applied Surface Science* **2006,** *252,* 5499–5503. (c) ECAFM image of SUS304 stainless steel treated at a current density of 5 mA cm^{-2} for 1500 s in 0.6 M $NaCl_{aq}$. A crust of corrosion product is observed above a pit. Reproduced from Zhang, Q.; Wang, R.; Kato, M.; Nakasa, K. *Scripta Materialia* **2005,** *52,* 227–230.

On stainless steels, studies have focused on the initiation of pitting near or at sulfide and carbide inclusions. On 304 SS, *ex situ* AFM combined with scanning electrochemical microscopy (SECM) and EDX analysis has shown that certain inclusions (sulfides) concentrate chloride, by electromigration, under a sulfur crust as a result of their dissolution evidenced by AFM. It has been proposed that the high local current density (evidenced by SECM), the electromigration of chloride, and the sulfur crust (evidenced by EDX) generate an occluded extreme environment in which stainless steel depassivates.[47] Figure 8(c) shows a crust of corrosion product on a small pit observed *in situ* on SUS304 treated anodically in NaCl (3.5 wt%) to trigger localized corrosion.[48] A part of the crust was destroyed, possibly by the scanning probe, and the pit hole remained shallow indicating repassivation. This was explained by the decrease of the concentration of chloride and hydrogen ions in the pit due to the destruction of the crust and stirring of the solution in the initially occluded environment in the pit.

2.33.5 Conclusion

The principles, instrumentation, and electrochemical implementation of STM and AFM have been presented. Both methods can provide 3D imaging of surfaces in direct space during corrosion processes. ECSTM is better suited for high resolution studies at the (sub)nanometer scale but is restricted to conductive surfaces. Atomic scale studies require a high level of control to produce atomically smooth surfaces. ECSTM can be combined with ECTS to characterize the electronic properties of surface layers *in situ*. ECAFM is well-suited for (sub) micrometer scale studies of surfaces irrespective of their conductivity. Selected examples of applications to corrosion analysis show that both methods can greatly contribute to corrosion analysis and an improved understanding and control of corrosion of metals and alloys.

References

1. Binnig, G.; Rohrer, H.; Gerber, Ch.; Weibel, E. *Phys. Rev. Lett.* **1983**, *50*, 120–123.
2. Rohrer, G. In *Scanning Tunneling Microscopy and Spectroscopy. Theory, Techniques and Applications*; Bonnell, D. A., Ed.; Wiley-VCH: New York, 1993; pp 155–187.
3. Tersoff, J. In *Scanning Tunneling Microscopy and Spectroscopy. Theory, Techniques and Applications*; Bonnell, D. A., Ed.; Wiley-VCH: New York, 1993; pp 31–50.
4. Hamers, R. J. In *Scanning Tunneling Microscopy and Spectroscopy. Theory, Techniques and Applications*; Bonnell, D. A., Ed.; Wiley-VCH: New York, 1993; pp 51–103.
5. Besenbacher, F.; Norskov, J. K. *Prog. Surf. Sci.* **1993**, *44*, 5–66.
6. Wiame, F.; Maurice, V.; Marcus, P. *Surf. Sci.* **2007**, *601*, 1193–1204.
7. Maurice, V.; Despert, G.; Zanna, S.; Bacos, M.-P.; Marcus, P. *Nature Materials* **2004**, *3*, 687–691.
8. Maurice, V.; Cadot, S.; Marcus, P. *Surf. Sci.* **2001**, *471*, 43–58.
9. Maurice, V.; Cadot, S.; Marcus, P. *Surf. Sci.* **2005**, *581*, 88–104.
10. Sonnenfeld, R.; Hansma, P. K. *Science* **1986**, *232*, 211–213.
11. Lev, O.; Fan, F.-R.; Bard, A. J. *J. Electrochem. Soc.* **1988**, *135*, 783–784.
12. Halbritter, J.; Repphun, G.; Vinzelberg, S.; Staikov, G.; Lorenz, W. J. *Electrochim. Acta* **1995**, *40*, 1385–1394.
13. Diéz-Pérez, I.; Guell, A. G.; Sanz, F.; Gorostiza, P. *Anal. Chem.* **2006**, *78*, 7325–7329.

14. Xie, Z.-X.; Kolb, D. M. *J. Electroanal. Chem.* **2000**, *481*, 177–182.
15. Garcia, S. G.; Salinas, D. R.; Mayer, C. E.; Lorenz, W. J.; Staikov, G. *Electrochim. Acta* **2003**, *48*, 1279–1285.
16. Vogt, M. R.; Nichols, R. J.; Magnussen, O. M.; Behm, R. J. *J. Phys. Chem. B* **1998**, *102*, 5859–5865.
17. Maurice, V.; Klein, L. H.; Marcus, P. *Surf. Interf. Anal.* **2002**, *34*, 139–143.
18. Suggs, D. W.; Bard, A. J. *J. Am. Chem. Soc.* **1994**, *116*, 10725–10733.
19. Broekmann, P.; Anastasescu, M.; Spaenig, A.; Lisowski, W.; Wandelt, K. *J. Electroanal. Chem.* **2001**, *500*, 241–254.
20. Magnussen, O. M.; Zitzler, L.; Gleich, B.; Vogt, M. R.; Behm, R. J. *Electrochim. Acta* **2001**, *46*, 3725–3733.
21. Suzuki, T.; Yamada, T.; Itaya, K. *J. Phys. Chem.* **1996**, *100*, 8954–8961.
22. Seyeux, A.; Maurice, V.; Klein, L. H.; Marcus, P. *J. Electrochem. Soc.* **2006**, *153*, B453–B463.
23. Dietterle, M.; Will, T.; Kolb, D. M. *Surf. Sci.* **1995**, *327*, L495–L500.
24. Ando, S.; Suzuki, T.; Itaya, K. *J. Electroanal. Chem.* **1997**, *431*, 277–284.
25. Itaya, K. In *Interfacial Electrochemistry – Theory, Experiments and Applications*; Wieckowski, A., Ed.; Marcel Dekker, New York, 1999; pp 187–210.
26. Stratmann, M.; Rohwerder, M. *Nature* **2001**, *410*, 420–423.
27. Maurice, V.; Strehblow, H.-H.; Marcus, P. *Surf. Sci.* **2000**, *458*, 185–194.
28. Kunze, J.; Maurice, V.; Klein, L. H.; Strehblow, H.-H.; Marcus, P. *J. Phys. Chem. B* **2001**, *105*, 4263–4269.
29. Kunze, J.; Maurice, V.; Klein, L. H.; Strehblow, H.-H.; Marcus, P. *Electrochim. Acta* **2003**, *48*, 1157–1167.
30. Zuili, D.; Maurice, V.; Marcus, P. *J. Electrochem. Soc.* **2000**, *147*, 1393–1400.
31. Scherer, J.; Ocko, B. M.; Magnussen, O. M. *Electrochim. Acta* **2003**, *48*, 1169–1191.
32. Kunze, J.; Strehblow, H.-H.; Staikov, G. *Electrochem. Comm.* **2004**, *6*, 132–137.
33. Foelske, A.; Kunze, J.; Strehblow, H.-H. *Surf. Sci.* **2004**, *554*, 10–24.
34. Ryan, M. P.; Newman, R. C.; Thompson, G. E. *J. Electrochem. Soc.* **1995**, *142*, L177–L179.
35. Zuili, D.; Maurice, V.; Marcus, P. *J. Phys. Chem. B* **1999**, *103*, 7896–7905.
36. Bhardwaj, R. C.; Gonzalez-Martin, A.; Bockris, J. O. M. *J. Electrochem. Soc.* **1992**, *139*, 1050–1058.
37. Ryan, M. P.; Newman, R. C.; Thompson, G. E. *Philos. Mag. B* **1994**, *70*, 241–251.
38. Maurice, V.; Yang, W. P.; Marcus, P. *J. Electrochem. Soc.* **1998**, *145*, 909–920.
39. Binnig, G.; Quate, C. F.; Gerber, Ch. *Phys. Rev. Lett.* **1986**, *56*, 930–933.
40. Burnham, N. A.; Colton, R. J. In *Scanning Tunneling Microscopy and Spectroscopy. Theory, Techniques and Applications*; Bonnell, D. A., Ed.; Wiley-VCH: New York, 1993; pp 191–249.
41. Rocca, E.; Bertrand, G.; Rapin, C.; Labrune, J. C. *J. Electroanal. Chem.* **2001**, *503*, 103–140.
42. Donley, M. S.; Mantz, R. A.; Khramov, A. N.; Balbyshev, V. N.; Kasten, L. S.; Gaspar, D. J. *Prog. Org. Coat.* **2003**, *47*, 401–415.
43. Kleber, C. H.; Hilfrich, U.; Schreiner, M. *Appl. Surf. Sci.* **2007**, *253*, 3712–3731.
44. Ikemiya, N.; Kubo, T.; Hara, S. *Surf. Sci.* **1995**, *323*, 81–90.
45. Maurice, V.; Nakamura, T.; Klein, L. H.; Strehblow, H.-H.; Marcus, P. In *Local Probe Techniques for Corrosion Research*; Woohead Publishing and Maney Publishing, CRC Press, 2007; pp 71–83. EFC Publications Number 45.
46. Park, J. O.; Paik, C.-H.; Huang, Y. H.; Alkire, R. C. *J. Electrochem. Soc.* **1999**, *146*, 517–523.
47. Williams, D. E.; Mohiuddin, T. F.; Zhu, Y. Y. *J. Electrochem. Soc.* **1998**, *145*, 2664–2672.
48. Zhang, Q.; Wang, R.; Kato, M.; Nakasa, K. *Scr. Mater.* **2005**, *52*, 227–230.
49. Davoodi, A.; Pan, J.; Leygraf, C.; Norgren, S. *Appl. Surf. Sci.* **2006**, *252*, 5499–5503.
50. Prades, S.; Bonamy, D.; Dalmas, D.; Bouchaud, E.; Guillot, C. *Int. J. Solids Struct* **2005**, *42*, 637–645.

2.32 Electron and Photon Based Spatially Resolved Techniques

X. Zhou and G. E. Thompson
School of Materials, University of Manchester, PO Box 88, Manchester M60 1QD, UK

Abbreviations

BF Bright field
BSE Backscattered electron
CBED Convergent beam electron diffraction
CCD Charge-coupled device
DF Dark field
EBSD Electron backscatter diffraction
EDS Energy dispersive spectrometer
EDX Energy dispersive X-ray
EELS Electron energy loss spectroscopy
EPMA Electron probe microanalysis
ESEM Environmental scanning electron microscopy
FIB Focused ion beam
HRTEM High resolution transmission electron microscopy
SAD Selected area diffraction
SE Secondary electron
SEM Scanning electron microscopy
STEM Scanning transmission electron microscopy
TEM Transmission electron microscopy
WDS Wavelength dispersive spectrometer

2.32.1 Introduction

Understanding the corrosion and protection of materials often requires examination of the morphology,

composition and structure of the surface and near-surface regions of a material, as well as their relationship to the bulk material. The knowledge gained allows definition of features that may increase susceptibility to degradation, and hence, routes for their removal.

Electron- and photon-based, spatially resolved techniques are widely used in corrosion studies. The selection of a technique for corrosion studies requires consideration of various factors, including:

- sensitivity/detection limit,
- elemental range,
- spatial resolution,
- depth penetration,
- quantitative/qualitative information,
- sample preparation,
- destructive/nondestructive procedure, and
- cost.

Usually, a single technique will provide only a part of the required information. For example, common approaches for morphological characterization include

- visual observation,
- optical microscopy (resolution $\sim 1\,\mu m$),
- scanning electron microscopy (SEM) (resolution ~ 1–2 nm), and
- transmission electron microscopy (TEM) (resolution ~ 0.1 nm).

Typically, a combination of the techniques provides the comprehensive information required for the corrosion studies. **Figure 1** displays a set of optical, SEM and TEM images and an energy dispersive X-ray (EDX) spectrum obtained from a friction-stir-welded AA7108 aluminum alloy following corrosion testing for 72 h. The images reveal the propagation path of intergranular corrosion and the associated sensitized microstructure.

Figures 1(a) and **1(b)** show the optical micrographs of the weld surface after corrosion testing. Semicircular marks created by the tool shoulder on the weld surface, evident at the center of the weld, are revealed as a series of ridges in the cross-section of the weld. Severe localized attack is evident in the heat-affected zone (HAZ),

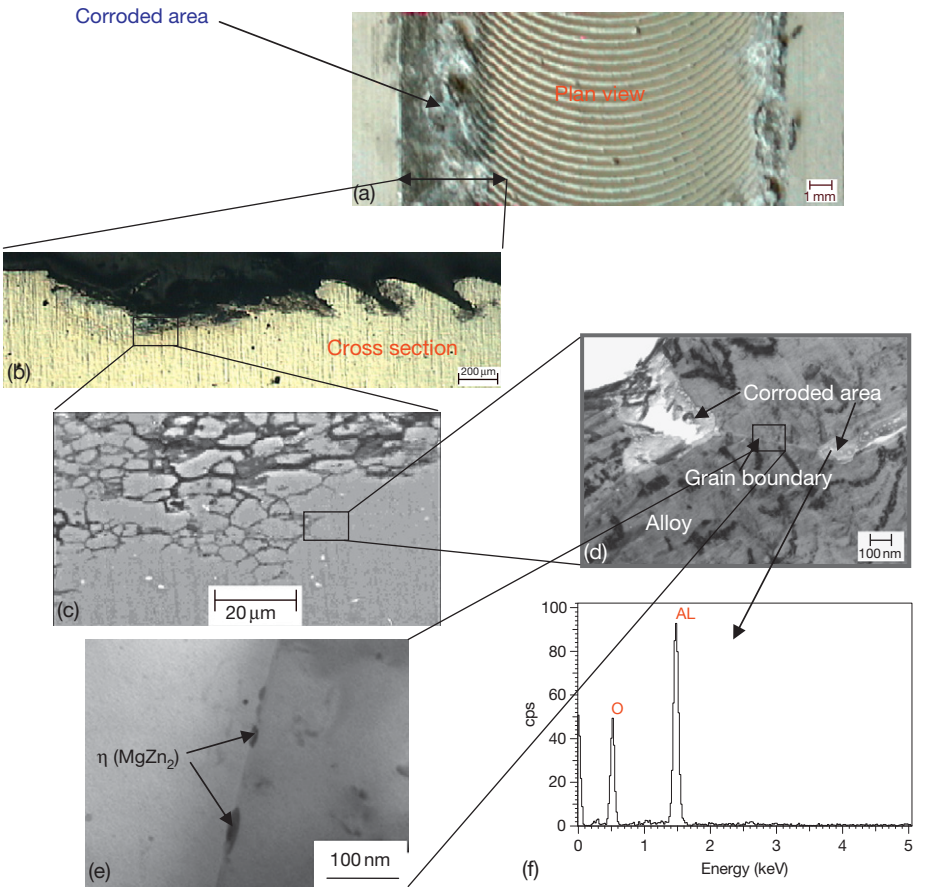


Figure 1 Corroded AA7108 aluminum alloy weld: (a) and (b) optical images; (c) scanning electron micrograph; (d) and (e) transmission electron micrograph; and (f) EDX spectrum taken from corrosion front.

where some ridges at the surface have been completely removed, with the attack extending through the HAZ to a depth of about 300 μm from the original surface. However, the corrosion path is not resolved in the optical images. A scanning electron micrograph of the framed region in **Figure 1(b)** displays intergranular corrosion within the HAZ (**Figure 1(c)**). The dark lines in this backscattered electron image represent areas associated with relatively light elements compared with those of the adjacent regions. In this case, the dark lines represent corroded areas as corrosion products (i.e., aluminum oxide/hydroxide) contain oxygen and hydrogen, which are lighter than aluminum.

Figure 1(d) shows a transmission electron micrograph of an ultramicrotomed section taken from the corrosion front located in the framed region in **Figure 1(c)**. Significant attack of the region between the grains is evident, with the alloy bulk now being penetrated by the resin used to embed the specimen. Further, corrosion of the individual grains has resulted in the development of hydrated alumina of thickness about 90 nm; EDX analysis of the corrosion products gives the yields of aluminum and oxygen as expected (**Figure 1(f)**).

Figure 1(e) displays a transmission electron micrograph of an electropolished foil taken from the HAZ. At the grain boundary, distinct $MgZn_2$ particles, about 100 nm length and 20 nm width, are evident, indicating sensitization of the HAZ that has led to the intergranular attack due to microgalvanic action.

2.32.2 Optical Microscopy

Optical microscopy is still a major tool in corrosion studies, being relatively cheap, readily available, and easy to use. Specimens can be observed in air. The images are of natural appearance, that is, color, with available resolution up to 1 μm (**Figures 1(a)** and **1(b)**). Optical microscopy is generally a reflection technique. A visible light beam is focused onto the area of interest. Reflected light beams pass through a single lens, or multiple lenses, to produce a magnified view of the examined area. The resultant image is detected directly by the eye, recorded on photographic film or captured by digital devices.

However, optical microscopes are limited in their ability to resolve fine detail by a phenomenon called diffraction. The resolution (δ) of an optical microscope is given by

$$\delta = \frac{0.5\lambda}{\mathrm{NA}} \quad \mathrm{NA} = \mu \sin \alpha \qquad [1]$$

where NA or A_N is the numerical aperture of the optical microscope. Thus, the resolution is controlled by three factors:

μ – the refractive index of the medium between the lens and the specimen;
λ – the wavelength of illuminating radiation;
α – the half angle subtended by the maximum cone of rays entering the objective lens.

The resolution can be improved by increasing μ; for example, μ can be increased from 1 to 2 by oil immersion. However, decreasing λ has the greatest potential to improve resolution. Visible light has wavelengths of 400–700 nm; with air as the medium, the highest practical NA is 0.95. Thus, the highest achievable resolution is ~0.2–0.4 μm.

In addition to the limitation on resolution, the depth of field is also relatively poor in optical microscopy. The depth of field is the portion of the object that can be seen acceptably sharply. The depth of field decreases as the magnification increases. Typically, the depth of field is 250 μm at a magnification of 15× compared with 0.08 μm at the magnification of 1200×.

In an optical micrograph taken at a relatively high magnification, only a section of the radiolarian is in sharp focus. Thus, flat specimens are required for optical microscopy. Therefore, metallography is employed to prepare a flat surface. Generally, metallic specimens that are to be examined in an optical microscope are mechanically polished to a 1-μm finish. Examination in the as-polished condition reveals structural features such as porosity, cracks, and inclusion of foreign matter. Etching with appropriate chemical agents may be used to reveal the microstructural features such as grain size and phase distribution as grain boundaries are preferentially attacked by the etchants. **Figure 2** displays an optical micrograph of an AZ91D magnesium alloy after etching in a mixed etching solution (5 ml acetic acid, 10 ml water, 6 g picric acid, and 100 ml ethanol). The typical α and β phases are clearly revealed.

A further limitation of conventional optical microscopy is that only dark or strongly reflecting objects can be imaged effectively. This limitation has, to some extent, been overcome by specific microscopy techniques using differences in the refractive indices of objects. To improve image contrast in optical microscopy, dark field or oblique illumination may be employed, but with no improvement in resolution. Other approaches include phase contrast microscopy, interference microscopy, and polarized light microscopy.

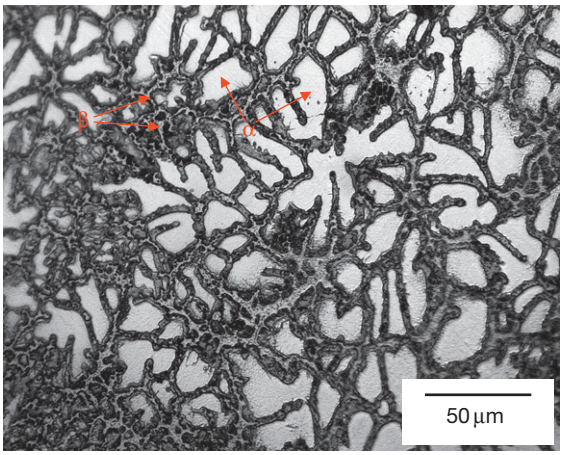


Figure 2 Optical micrograph of AZ91D magnesium alloy after etching in a mixed solution (5 ml acetic acid, 10 ml water, 6 g picric acid, and 100 ml ethanol).

2.32.3 Electron Microscopy

2.32.3.1 Introduction

The limitations of optical microscopy are, to an extent, overcome in electron microscopy. From eqn [1], it is evident that the resolution is determined by the wavelength (λ) of the radiation and the NA of the objective. To overcome the limitations set by the diffraction limit of visible light, electron microscopes that use electrons in place of light have been developed. This gives much increased resolution because the wavelength of the electrons is shorter, and hence, the diffraction limit is lower. At an accelerating voltage of 100 kV, λ is 0.0037 nm, giving the potential for significantly increased resolution in an electron microscope. Theoretically, a resolution of 0.02 nm is achievable in an electron microscope operated at 100 kV. Practically, however, such resolution cannot be obtained because of lens aberrations. The resolution limit presently is around 0.05 nm.

Further, the depth of field is also improved by using electrons for imaging. It is not unusual that the macroscopic specimen surface is in focus in a scanning electron microscope (SEM) because of the relevant large depth of field. However, as electrons can be readily scattered or absorbed in air, vacuum is required for electron microscopy, which makes the system more expensive with the need for effective maintenance.

Typically, an electron microscope consists of two components: the illumination system and the imaging system. The illumination system comprises the electron gun and the lenses that transfer the electrons from the gun to the specimen. Electrons are thermally

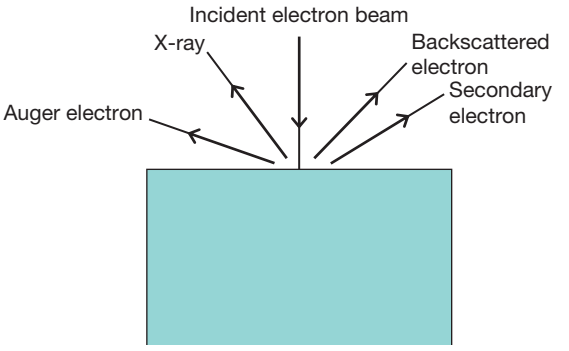


Figure 3 Schematic diagram illustrating the generation of electrons and X-rays by a bulk specimen.

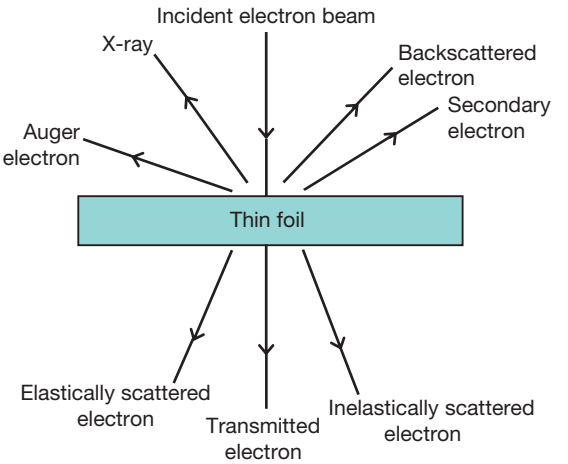


Figure 4 Schematic diagram illustrating the interaction between the incident beam and the thin foil.

generated from a tungsten or lanthanum hexaboride (LaB_6) cathode and are accelerated towards an anode. Alternatively, electrons can be emitted by field emission (FE). The latter produces an electron beam of fine size, increased brightness, and well-defined energies. The smaller energy spread of the FE gun, typically less than 0.5 eV compared with 1–2 eV for a thermionic source (e.g., tungsten or lanthanum hexaboride filament), is particularly important for electron energy loss spectroscopy (EELS).

In an electron microscope, incident electrons enter the specimen, interact with the material, and result in various secondary emissions, as illustrated in Figures 3–5. Each of the emissions may provide useful and different information about the material. Most of these emissions are used for imaging and analysis in the electron microscopy. Thus, electron microscopes generally require a source of incident electrons and a detector to collect the emitted

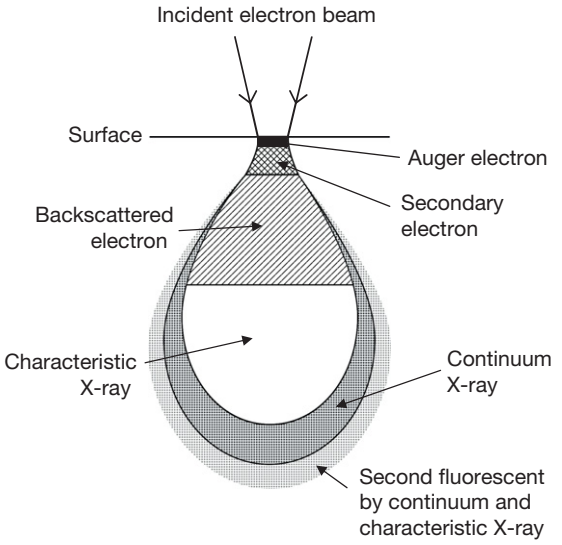


Figure 5 Schematic diagram illustrating the regions where electrons and X-rays are generated.

radiation; often several detectors are combined in one system using a common incident electron beam.

When an electron beam is projected onto a specimen, various interactions between the incident electrons and the atoms of the specimen occur, including (i) elastic scattering, (ii) inelastic scattering, and (iii) combinations of (i) and (ii). Elastic scattering, which is the main contributor to the electron diffraction patterns, is due to the Coulombic interaction between the incident electrons and the atoms of the specimen, resulting in a change of direction, but not the energy, of the incident electrons. Inelastic scattering leads to energy transfer from the incident electrons to the electrons or atoms of the specimen. A small proportion of the lost energy may escape from the specimen as secondary electrons or X-rays, which are very useful for imaging and compositional analysis. Further, the detection of energy losses of the incident electrons is the basis of EELS.

As illustrated in **Figures 3** and **4**, in scanning electron microscopes, the images are generally generated by secondary and backscattered electrons, and in transmission electron microscope, the images are produced by the transmitted electrons. Determined by the image generation mechanism, SEM is used for studying the surface and near-surface regions of bulk specimens, and TEM is employed to investigate the internal structure of thin foil specimens.

The electrons lose energy by inelastic scattering within a region that is termed the interaction volume, which has a typical onion shape and extends approximately up to a few micrometres into the examined material, as shown in **Figure 5**. The size of the interaction volume depends on the electron beam accelerating voltage, the atomic number of the elements in the specimen, and the material density.

2.32.3.2 Scanning Electron Microscopy

The application of SEM in corrosion studies is widespread, including (i) determination of structure–property relationships; (ii) location of corrosion; (iii) study of surface topography; (iv) examination of fracture sections and cracks; (v) and determination of the morphology of oxides, scales, corrosion products, etc.

In a typical SEM, a finely focused electron beam, of diameter down to 1 nm, is scanned across the surface of the specimen of interest. The scanning of the focused electron beam is achieved by controlling the current to the associated scan coils. The accelerating voltage in an SEM is typically in the range of a few hundred electron volts to 30 keV.

Figure 3 illustrates schematically the electrons and X-rays generated through the interaction of an electron beam with the examined material. Although all resultant radiations can be used to produce an image, for SEM the images of surfaces are formed by collecting the low-energy secondary electrons (SEs) and backscattered electrons (BSEs). SEs are electrons that are ejected from atoms which interacted with the incident electrons and to which a small amount of energy has been transferred by the incident electrons during inelastic scattering; consequently, SEs have relatively low energies, below 50 eV, and escape only from regions up to 50 nm depth below the surface. The yield of secondary electrons is heavily influenced by the nature of the surface. BSEs are incident electrons that are backscattered from the specimen interaction volume. BSEs carry relatively high energies and, therefore, emerge from deeper regions within specimens. BSEs are used for imaging and generation of diffraction patterns for structure and texture analyses. These electrons may be used to provide compositional information by detecting the contrast between areas with different chemical compositions, especially when the average atomic number of elements in various areas is different, since the yield of BSEs varies with the atomic number of the elements in the examined specimen. Further, as illustrated in **Figure 3**, the relatively high-energy BSEs, emerging from deeper regions in the specimen, have an increased sampling volume; thus, generally, the surface sensitivity and resolution

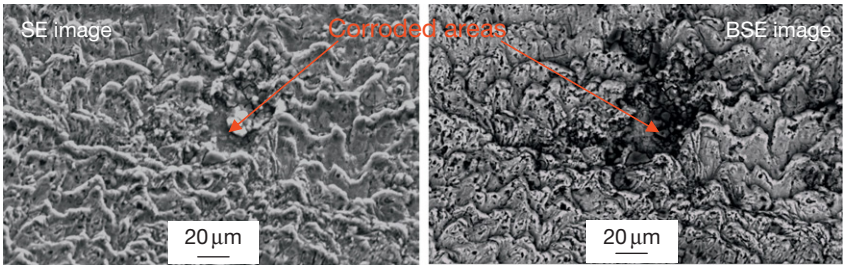


Figure 6 Corresponding secondary electron and backscattered electron images of a laser-surface-treated AA2024 aluminum alloy following accelerated corrosion testing.

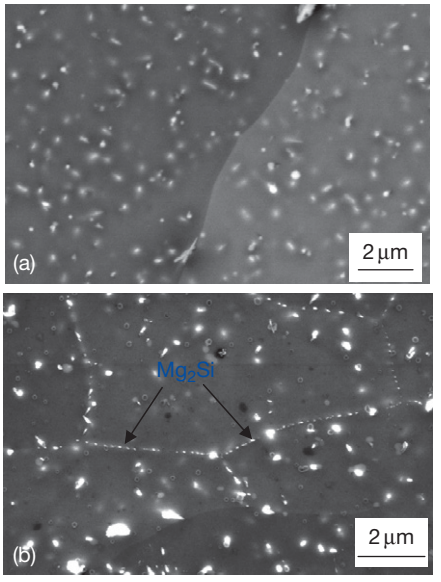


Figure 7 Backscattered electron images taken from (a) the parent alloy and (b) the HAZ of friction-stir-welded AA6082 aluminum alloy.

of the BSE approach are lower compared to that of the SE approach.

Figure 6 shows the corresponding SE and BSE electron images of the surface of a laser-treated AA2024 aluminum alloy following accelerated corrosion testing. The SE image reveals clearly the typical wavy appearance generated by laser treatment, but with little evidence of development of corrosion. The contrast in the BSE image readily shows localized corrosion. The relatively dark areas represent the areas where corrosion has occurred, since corrosion products (i.e., mainly aluminum oxides/hydroxides in this case) consist of relatively light elements of oxygen and hydrogen compared to metallic aluminum, thereby generating a relatively low intensity of BSEs.

Figures 7–9 display further BSE images taken from various specimens. Figure 7 shows the BSE images from the parent alloy and the HAZ of a friction-stir-welded AA6082 aluminum alloy, revealing the difference in the distribution, size and population density of intermetallics caused by thermal exposure during the welding process. Within the HAZ, a continuous network of β phase (Mg_2Si) is clearly evident along grain boundaries, indicating sensitization of the HAZ. Such information assists in the understanding of the reasons for the HAZ of the friction-stir-welded AA6082 aluminum alloy displaying a relatively high susceptibility to intergranular corrosion.

The backscattered electron image of Figure 8 displays the cross section of an AZ91D magnesium alloy after immersion in 3.5% NaCl for 4 h. As stated previously, the dark regions in this backscattered electron image represent areas of relatively light elements. In this case, the dark regions indicate corrosion products (i.e., magnesium oxides/hydroxides). It is clearly evident that corrosion occurred nonuniformly, with the areas of the α phase being preferentially attacked and the β phase remaining generally intact.

Figure 9 illustrates the backscattered electron images of a friction-stir-welded AA2024-T3 aluminum alloy following the immersion in an EXCO solution (ASTM G 34-01, a standard testing method for exfoliation susceptibility in 2xxx and 7xxx series aluminum alloys) for 8 h. Again, the dark regions in these BSE images represent corroded areas. At the surface, a severely corroded band is evident. On the right side of the corrosion band (Figure 9(a)), that is, in the parent alloy, little development of corrosion is observed. Conversely, on the left side of the corrosion band, that is, within the thermomechanically affected zone (TMAZ), a large number of randomly distributed pits, of dimension about 50–300 μm, is evident. Further, the weld cross-section revealed clearly that localized corrosion has penetrated into the alloy from both top and bottom surfaces.

A topographic image may be obtained using SEM. Figure 10(a) displays an SE image of an oxide film formed on a binary Al–Cu alloy, revealing the different topographies of the oxide film formed on three individual alloy grains of different crystallographic

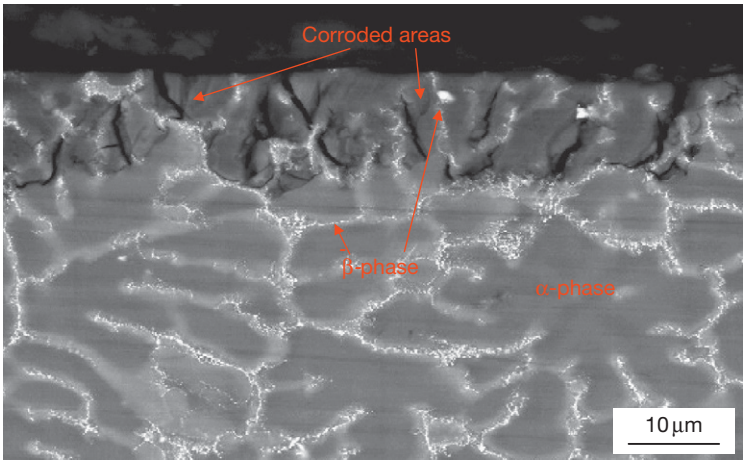


Figure 8 Backscattered electron image of the cross section of AZ91D magnesium alloy after immersion testing in 3.5% NaCl solution for 4 h.

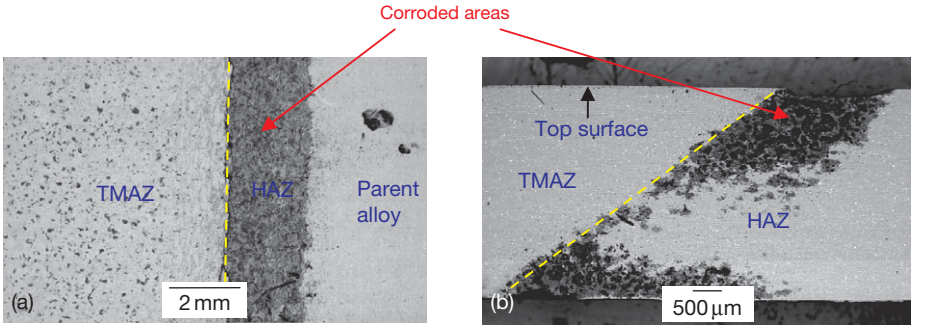


Figure 9 Backscattered electron images of friction-stir-welded AA2024-T3 aluminum alloy following corrosion testing in the EXCO solution for 8 h, revealing corrosion susceptible zones where the alloy was sensitized during welding: (a) plan view and (b) cross sectional view of (a).

orientations. Since the shape and size of the SE sampling volume is related to the angle between the surface and the incident electron beam, as the angle increases, the escape distance of one side of the beam will decrease and, consequently, more SEs will be emitted. Thus, steep surfaces and edges tend to be brighter than flat surfaces, resulting in images with a well-defined topographic contrast. It is essential to tilt the specimen to obtain the desirable topographic contrast, avoiding signals emitted from deeper regions within the specimen, and thereby enhancing the surface sensitivity in topographic studies. For example, **Figure 10(b)** was obtained from the framed area in **Figure 10(a)** without tilting the specimen, revealing little topographic contrast although features of subsurface regions of the specimen (i.e., voids in the oxide film in this case) are clearly displayed. This is because the surface topographic contrast is overwhelmed by internal features since a relatively large portion of signal is emitted from deeper regions within the specimen. **Figure 10(c)** illustrates schematically the cross section of the oxide film, revealing the distribution of the voids. It is clearly evident that the internal features of the oxide film overwhelmed the surface topographic features.

2.32.3.2.1 Electron backscatter diffraction

Electron backscatter diffraction (EBSD), also known as backscattered Kikuchi diffraction (BKD), allows crystallographic information to be obtained from any crystalline material using BSEs in the scanning electron microscope. EBSD can be used to determine grain sizes and shapes, as well as orientations and misorientations. Such information is very useful for corrosion studies. For example, the sensitization of an alloy may be correlated to misorientations between alloy grains. Because of the relatively weak channeling contrast used for EBSD, a carefully prepared surface is essential to obtain satisfactory information. Normally, a flat and strain-free surface is desirable.

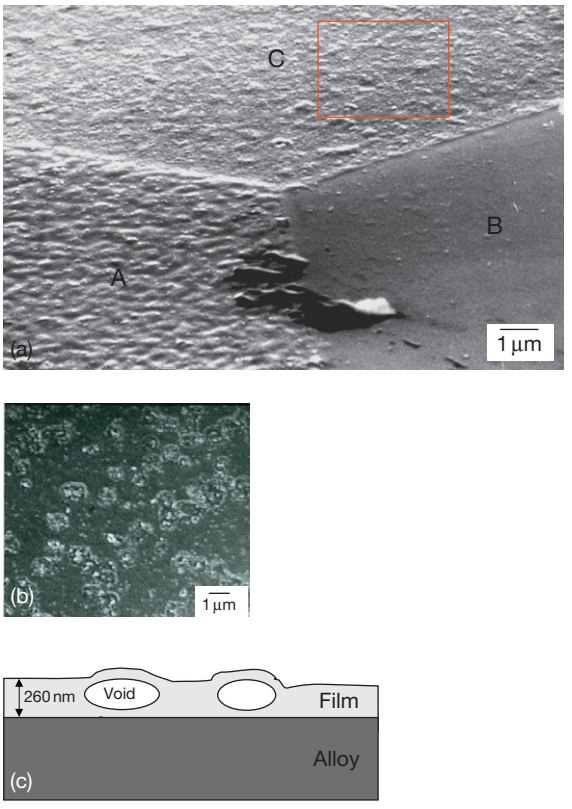


Figure 10 (a) and (b) Scanning electron micrographs of an oxide film formed on an Al–Cu binary alloy; and (c) schematic illustration of the cross section of the oxide film.

Such a surface may be produced by electropolishing. Mechanical polishing should be avoided in EBSD specimen preparation since a near-surface layer of deformed microstructure is normally introduced into the materials during the process. Figure 11 displays the EBSD maps recorded in various weld zones of a friction-stir-welded AA2024 aluminum alloy, revealing grain size and orientations. The parent alloy and the HAZ were shown to consist of a relatively coarse, elongated grain structure, ranging from approximately $20 \times 30\,\mu m$ to $50 \times 150\,\mu m$ in dimensions that were generated during fabrication. The nugget region, that is, the center of thermomechanically affected zone (TMAZ), displays relatively small, equiaxed grains of size about 1–10 μm. Outside the nugget, within the TMAZ, a heavily deformed parent microstructure is revealed.

2.32.3.2.2 Environmental scanning electron microscopy

Environmental scanning electron microscopy (ESEM), developed in the 1980s, allows operation in increased pressure gaseous environments, in the range of 1–10 torr, compared to a vacuum below 10^{-5} torr in a conventional SEM. This is achieved by the development of an SE detector that is capable of operating in the presence of gases and by the use of electron apertures and differential pumping systems to separate the high-vacuum regions around the gun and the column from the low-vacuum sample chamber. Therefore, in corrosion studies, ESEM allows imaging of specimens with a vapor or volatile component, or while *in situ* examinations involve gas or liquid phases. Additionally, ESEM allows examination of nonconductive specimens without coating since the excess electrons are dissipated by the collision between the electrons and gas molecules.

2.32.3.3 Transmission Electron Microscopy

The application of TEM in corrosion studies includes (i) structure–property relationships determination; (ii) sites of corrosion and surface roughness determination, (iii) morphological determination of oxides, scales, corrosion products, etc.

TEM forms images following the transmission of an electron beam through an electron-transparent specimen. An image is formed, magnified, and directed to appear either on a fluorescent screen or photographic film, or detected by a sensor such as a CCD (charge-coupled device) camera. In a typical TEM, electrons are generated from a tungsten or lanthanum hexaboride (LaB_6) filament or a FE gun which accelerates electrons through voltages in the range 40–300 kV. Over the years, high-voltage (up to 3 MV) TEMs have been developed to achieve improved resolution and increased penetration, but they have become rarer since the resolution of relatively low-voltage TEMs being improved through improved lens design, and thin specimen preparation techniques being developed. Further, FE guns are now becoming increasingly widespread because of the capability of producing electron beams of very fine diameter (0.2 nm), increased brightness, and well-defined energies. In a TEM, a parallel or convergent beam can be generated. The former is used for imaging and diffraction, and the latter is used for scanning transmission electron microscopy (STEM), microanalysis, and convergent beam electron diffraction (CBED).

As illustrated in Figure 4, when an electron beam is projected onto a thin foil specimen, various processes occur, including (i) the straight passage of electrons through the specimen, (ii) elastic scattering, (iii) inelastic scattering, and (iv) combinations of (ii) and (iii). The intensity of the transmitted beam is affected

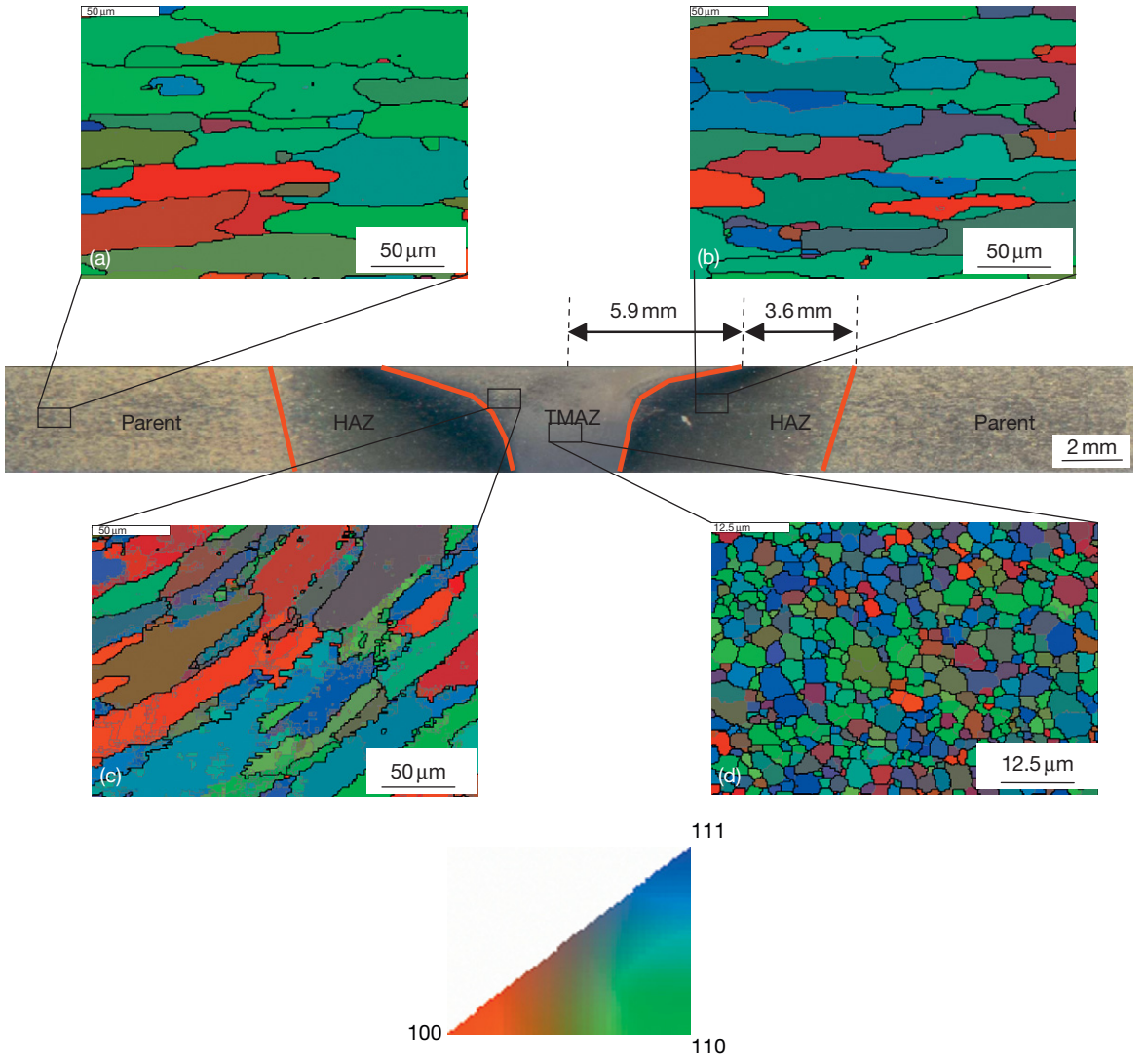


Figure 11 Electron backscatter diffraction (EBSD) maps recorded in various weld regions of friction-stir-welded AA2024 alloy: (a) the parent alloy; (b) the HAZ; (c) the TMAZ; and (d) the nugget region.

by the volume and density of the material through which the beam passes. The thick regions, or regions of increased density, scatter electrons strongly and appear dark in the image. This gives a mass-thickness contrast. However, a crystalline material can also interact with the electron beam by diffraction. The intensity of the diffraction depends on the orientation of the planes of atoms in a crystal relative to the incident electron beam. This is known as diffraction contrast.

In TEM, an objective aperture is placed at the focal plane of the objective lens to select electrons to form the image. The objective aperture can then be used to allow either the unscattered beam or the diffracted beam to form images. If the objective aperture is centered at the optical axis, the electrons that are deflected away from the optical axis are blocked, and only unscattered electrons pass through to form an image. This is known as a bright field (BF) image since in the absence of a specimen a bright background is seen on the screen. It is also possible to produce an image from electrons that are deflected by particular crystal planes. By either moving the objective aperture to the position to block the unscattered electrons and allow a particular diffracted beam to pass through, or tilting the incident electron beam so that a particular diffracted beam passes through the centered aperture, an image can be formed by the deflected electrons. This is known as a dark field (DF) image

since in the absence of a specimen the background appears dark. BF and DF images provide valuable complementary microstructural information.

Figure 12 shows a pair of BF and DF images of an ultramicrotomed section of an AA6111 aluminum alloy that was solution-heat-treated (SHT) at 813 K for 30 min followed by rapid quenching in ice water before mechanical grinding with 220 grit SiC paper and heat treatment at 180 °C for 30 min. The mechanical grinding introduced a deformed layer microstructure, which is characterized by very fine grains, approximately 50–150 nm diameter. Further, the DF image (**Figure 12(b)**), achieved by choosing the electrons deflected by a particular set of planes in the precipitate to form the image, reveals precipitate particles of approximately 20 nm diameter distributed along grain boundaries.

Using an annular detector to collect electrons that have been inelastically scattered through relatively large angles, compositional information of the specimen may be obtained by acquiring a *Z*-contrast image since the inelastic scattering is proportional to the square of the atomic number *Z*. Thus, areas with higher average atomic number elements appear brighter and areas with lower average atomic number elements appear darker. For example, etching or anodizing of copper-containing aluminum alloys normally results in the formation of a thin layer of alloy, about 2 nm thick, immediately beneath the residual or anodic alumina film, which is highly enriched in copper. Such enrichment arises from the less negative Gibbs free energy per equivalent for formation of copper oxide compared to alloys. **Figure 13(a)** shows a high angle annular dark field (HAADF) image of a cross-section of an AA6111 aluminum alloy after anodizing, revealing the *Z* contrast from the copper-rich band located at the alumina–aluminum interface. However, such a copper-rich band appears as a dark band in the BF image (**Figure 13(b)**) because of the mass contrast, that is, enhanced electron scattering by copper compared to aluminum.

A further example of the use of TEM in corrosion studies is given in **Figure 14**. A scanning electron micrograph of a cross-section taken from an AA6111 aluminum alloy is shown in **Figure 14(a)**, revealing localized corrosion which has proceeded into the alloy to a depth of ~4 μm. However, the SEM image provides little information about the corrosion initiation sites and propagation paths. **Figure 14(b)** displays the transmission electron micrograph of an ultramicrotomed section taken from the region indicated with the frame in **Figure 14(a)**. It is now clearly revealed that the corrosion is intergranular and develops preferentially along grain boundaries.

However, TEM requires consideration of the means of preparation of electron-transparent specimen. Thus, many materials require extensive specimen

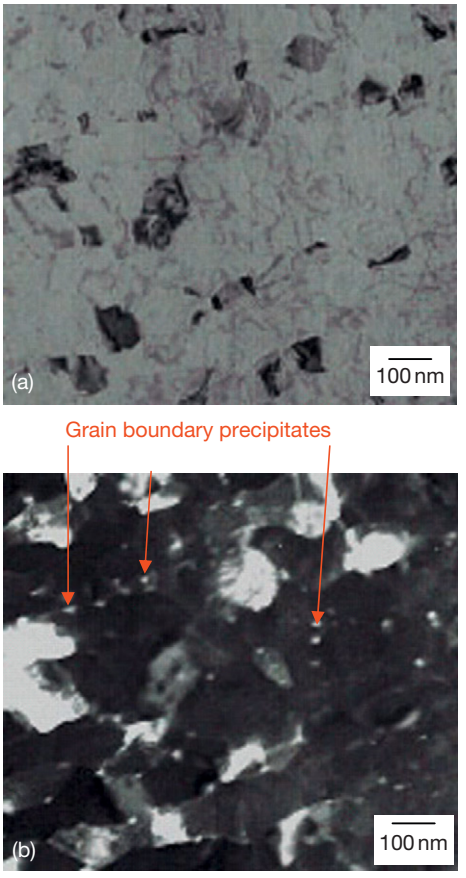


Figure 12 Transmission electron micrographs of ultramicrotomed section of the AA6111 alloy: (a) bright field image and (b) dark field image.

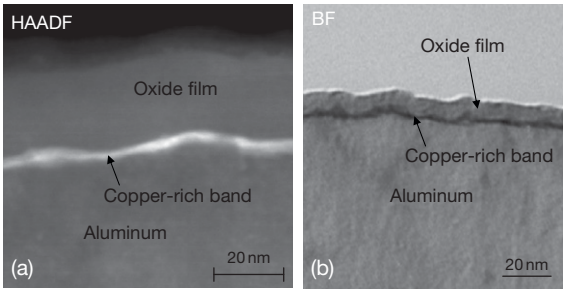


Figure 13 Transmission electron micrographs of cross-sections of AA6111 aluminum alloy after (a) anodizing and (b) caustic etching.

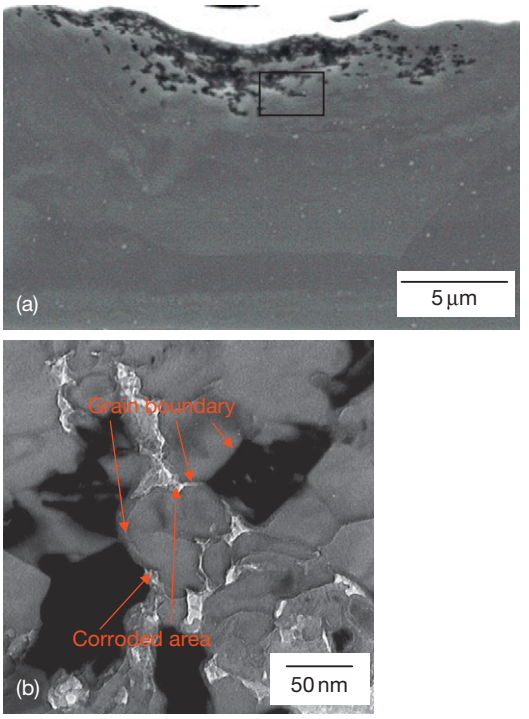


Figure 14 (a) Scanning electron micrograph of cross-section of corroded AA6111 aluminum alloy and (b) transmission electron micrograph of an ultramicrotomed section taken from corrosion front as indicated by the frame in (a).

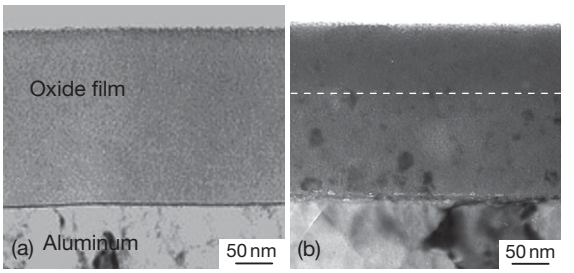


Figure 15 Transmission electron micrographs of an ultramicrotomed section of the anodic film formed to 200 V on the Al–2 wt% Cu alloy: (a) upon immediate examination in electron microscope, revealing a featureless amorphous film attached to the substrate and (b) after continued exposure to the electron beam in the microscope, the inner film region showing crystallization of alumina.

preparation to produce a specimen sufficiently thin to be transparent to electrons. Often, the specimen preparation is a time-consuming process with low throughput. Further, the composition and structure of the material may be changed during specimen preparation processes. Additionally, the sampling volume in TEM is relatively small, raising the possibility that the region examined may not be characteristic of the bulk material.

As stated previously, inelastic scattering generates useful signals. However, there is also the potential that the specimen may be damaged by the high-energy electron beam during inelastic scattering. Electron beam damage changes the structure and chemistry of the examined materials. Certain materials are more susceptible to damage than others. Inelastic scattering may break the chemical bonds of materials such as polymers. Further, the kinetic energy that is carried by the incident electrons may be transferred to heat in the specimen during the inelastic scattering processes. Electron-beam heating of thermally conductive materials such as metals is negligible under standard condition but can be quite substantial if thermal conduction is poor. For example, in Figure 15, a transmission electron micrograph of an ultramicrotomed section of the anodic film formed on the Al–2 wt% Cu alloy reveals a featureless amorphous film attached to the substrate. With continued exposure to the electron beam, the inner region of the anodic oxide film is crystallized to γ'/γ alumina, as shown in Figure 15(b)). However, such changes are not always problematic and advantage may be taken from such changes. From Figure 15(b)), it is evident that the thickness of the crystalline region adjacent to the alloy substrate is about 60% of the total film thickness; the outer region, representing about 40% of the total film thickness, showed delayed crystallization. This suggests that about 40% of the total film thickness is formed at the film–electrolyte interface by cation migration, and the remainder formed at the aluminum–film interface by anion migration since the outer region of the film contains incorporated, immobile electrolyte species that retard crystallization of the contaminated amorphous alumina under electron irradiation.

2.32.3.3.1 High-resolution TEM

High-resolution TEM (HRTEM) allows the imaging of the crystallographic structure of materials at an atomic scale. The image formation in HRTEM relies on phase contrast. As mentioned in Section 2.32.1, theoretically, a resolution of 0.02 nm is achievable in an electron microscope. However, such resolution cannot be obtained because of lens aberrations. At present, the highest resolution achieved is 0.05 nm on HRTEM with aberration correctors.

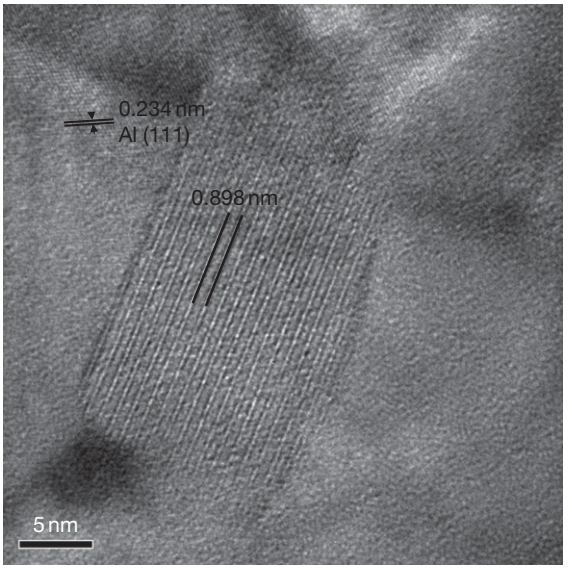


Figure 16 HRTEM image of a Q-phase particle in AA6111 aluminum alloy.

Figure 16 shows an HRTEM lattice image of a grain boundary precipitate in AA6111 aluminum alloy. The particle extends through the thickness of the section, with no consequent interference from the surrounding aluminum matrix. The lattice fringe spacing was determined accurately by using the (111) lattice plane spacing of the aluminum matrix as a reference. The measured fringe spacing of the precipitate is 0.898 nm, in good agreement with the d value of 0.899 nm for the $\{100\}_Q$ set of planes, thereby confirming that the precipitate is the quaternary Q phase ($Cu_2Mg_8Si_7Al_4$). The revelation of the Q phase at the grain boundary helps in explaining the occurrence of intergranular corrosion of the same alloy, as revealed in Figure 14. Where the Q phase particles are formed at grain boundaries, the microgalvanic coupling between the particles and the adjacent matrix to the particles provides the driving force for localized corrosion.

Further, although the high-angle annular dark field and BF image of Figure 13 revealed that a copper-enriched layer of 2 nm thickness formed at the aluminum–oxide interface by anodizing and caustic etching of the AA6111 aluminum alloy, it is impossible to tell how the copper species exist in the band from Figure 13. Such information can be readily obtained by using HRTEM. Figure 17 displays a lattice image of the copper-enriched layer, revealing clearly the lattice fringe spacing and confirming the presence of a $CuAl_2$ particle within the copper-enriched layer.

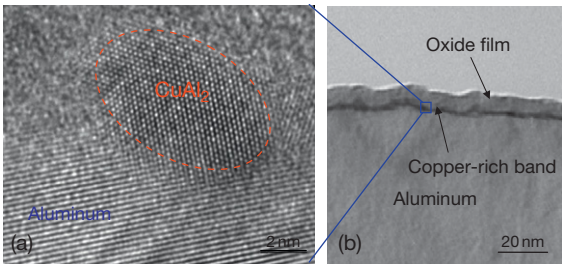


Figure 17 (a) Transmission electron micrograph of cross-section of AA6111 aluminum alloy after caustic etching, revealing a copper-rich band at the aluminum–oxide film interface; and (b) HRTEM lattice image of a $CuAl_2$ particle in the copper-rich band.

2.32.3.3.2 Scanning transmission electron microscopy

Scanning transmission electron microscopy (STEM) is a combination of SEM and TEM: that is, a transmission image is obtained using a scanning method. A TEM can be modified into an STEM by the addition of a system that scans a focused beam across the specimen to form the image. Further, by fitting a transmission stage and a detector, an SEM can also be operated in the STEM mode. Thus, a STEM combines some of the advantages of the SEM and the TEM. The image obtained in an STEM is similar to that obtained in a TEM, but the data are acquired serially.

In recent years, dedicated STEMs have also been built. In a dedicated STEM, since no objective or projector lenses are needed after the specimen, a relatively large space is available for various signal detectors to be positioned in. Thus, the dedicated STEM is normally used as a multisignal analytical microscope with high spatial resolution. Since the resolution of an STEM is determined by the diameter of the focused electron beam, the beam is focused to as small a size as possible. Nowadays, in a dedicated STEM, using a FE electron gun, an electron beam of diameter 0.2 nm can be generated. This is sufficient to image heavy single atoms. The employment of dedicated STEMs is limited since they are comparatively expensive and the improvement in the electron beam size in TEMs equipped with a FE gun has significantly reduced the difference in the resolution of the two types of microscope.

2.32.3.3.3 TEM tomography

Typically, TEM microscopy generates two-dimensional (2-D) images which locate the position of an internal feature in the examined specimen both vertically

and horizontally but with no information on the depth of that feature within the specimen being obtained. However, information concerning the depth of the feature can be obtained by tilting the specimen in small steps (small angle) and acquiring a series of 2-D images at each angle and subsequently reconstructing 3-D images by merging the individual 2-D images into a single 3-D block that can then locate any feature in the specimen in any direction. This technique is termed TEM tomography.

2.32.3.3.4 *Electron diffraction*

Electron diffraction is a technique that allows determination of the crystal structure of materials. When the electron beam is projected onto a specimen, its crystal lattice acts as a diffraction grating, scattering the electrons in a predictable manner, and resulting in a diffraction pattern. Electron diffraction patterns are mainly contributed by elastic scattering. In an electron microscope, the diffraction pattern appears either on a fluorescent screen or is recorded on a photographic film or detected by a sensor such as a CCD (charge-coupled device) camera. From the diffraction pattern, the arrangement of the atoms in a material can be determined. Electron diffraction can be acquired with either TEM or SEM (electron backscattered diffraction).

Selected area diffraction (SAD) is often employed to obtain a diffraction pattern from the area of interest, such as one grain in a polycrystalline specimen or a second phase embedded in the matrix. SAD is achieved by either inserting an aperture in the plane of the specimen itself or the first image produced by the objective lens, or illuminating only the area of interest. By using a convergent electron beam, diffraction patterns can also be obtained from fine regions. This technique is called convergent beam electron diffraction (CBED).

There are many reasons for obtaining diffraction patterns in corrosion studies. For example, by obtaining diffraction patterns, an unknown material may be identified, or the relationship between the grain orientations and the surface properties of a material may be established. Figure 18 displays diffraction patterns obtained from (a) a single crystal of tin along the $[001]_{Sn}$ zone axis, revealing a regular array of sharp spots since several sets of planes that are approximately parallel to the incident electron beam diffract electrons, thereby giving rise to a diffraction pattern; (b) a fine-grained polycrystalline AA 6111 aluminum alloy, exhibiting a ring-like diffraction pattern since the diffraction pattern is the sum of individual crystals of random orientations and (c) amorphous carbon, showing amorphous diffusion rings.

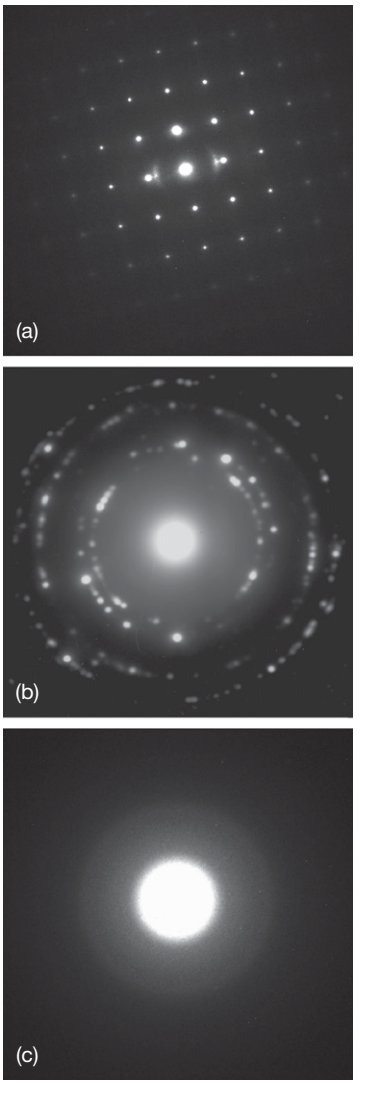


Figure 18 Electron diffraction patterns: (a) a single crystal of tin along the $[001]_{Sn}$ zone axis, (b) a fine-grained polycrystalline AA6111 aluminum alloy, and (c) amorphous carbon.

2.32.4 Chemical Analysis in the Electron Microscope

Chemical analysis in the electron microscope provides spatially resolved compositional information which is not available in conventional chemical or spectrographic methods.

2.32.4.1 X-ray Analysis

As stated previously, in an electron microscope, the energy lost by the incident electrons during inelastic scattering may escape from the specimen as X-rays. If an inner-shell electron has been knocked out of an atom during the inelastic scattering of an incident electron by that atom, the atom is in an excited high-energy state. The vacancy will be filled by an electron from an outer higher energy shell and the atom will relax. The excess energy of that electron is released in the form of an X-ray or an Auger electron. The X-rays are highly specific to individual elements, known as characteristic X-rays. For example, if a K shell electron has been removed from an iron atom and the vacancy is filled by an L shell electron, the energy difference, ΔE, is 6400 eV. The corresponding wavelength of the emitted X-ray can be calculated from $\lambda = hc/\Delta E$ to be 0.193 nm. The energy (wavelength) can be measured to determine the composition of the examined material since the energy (wavelength) is characteristic to a particular element. Thus, the emitted X-rays can be processed by two different approaches: using an energy dispersive spectrometer (EDS) or a wavelength dispersive spectrometer (WDS). Further, X-ray analysis is nondestructive and may be quantitative.

In an EDS, semiconducting silicon or germanium is used as the X-ray detector. The detector converts each absorbed X-ray photon into an electrical current pulse that is proportional to the energy of the X-ray photon. Each current pulse is amplified by a computerized analyzer. A histogram of energies of all X-rays arriving at the detector is collected and displayed as a smooth curve, which is the EDS spectrum. Normally, a window is necessary to separate the detector from the specimen chamber to prevent condensation on the detector since the detector works at very low temperatures. Unfortunately, such windows can absorb a significant portion of low-energy X-rays, thereby making light elements particularly difficult to be detected. However, windowless detectors or detectors with ultrathin windows are available to detect elements down to boron in the periodic table.

The main drawbacks of EDS systems are the relatively poor energy resolution, typically within the range of 100–200 eV, and a low signal-to-noise ratio, which make it difficult to resolve characteristic X-rays with sufficiently close energies and affect the detection, and therefore its use for quantitative analysis. The minimum detection limit for any element is about 0.1% using an EDS system and an order of magnitude lower using a WDS system. As an EDS detector can be placed very close to the specimen, thereby collecting X-rays very efficiently, a spectrum for qualitative analysis can be acquired within a couple of minutes.

In WDS, the arriving X-rays are filtered by a crystal spectrometer using diffraction to separate the X-rays on the basis of their wavelength. Then, X-rays of the energy resolution and signal-to-noise ratio are achieved, making it possible to resolve characteristic X-rays of close energies and to undertake more accurate quantitative analysis. Thus, the main drawbacks of EDS are overcome in a WDS system. However, since only one wavelength is detected at a time, it is very time consuming to acquire a spectrum.

Further, a characteristic X-ray is most efficiently generated when the primary electrons carry approximately 3 times the energy of the X-ray. Therefore, by using a SEM operating at 30 keV all elements are readily detectable since they have at least one strong characteristic X-ray line with energy less than 10 keV. In SEM, the surface sensitivity of X-ray analysis is related to the accelerating voltage since the penetration of the primary electrons is dependent on the beam energy: that is, accelerating voltage. With reducing accelerating voltage, electrons have less energy and hence penetrate to shallower depths; thus, X-rays are generated from regions close to the specimen surface.

Although secondary emission is generated throughout the entire interaction volume, which extends approximately a few micrometers into the examined materials, only that escaping from the specimen is detected. Most X-rays generated can escape from the specimen since X-rays are not readily absorbed. Thus, in the SEM, the X-rays are generated from a relatively large volume of material, of typical dimensions 2–3 μm. As illustrated in **Figure 3**, the sampling volume of X-rays is similar to the interaction volume, but the sampling volume of SEs and BSEs are significantly smaller. A large portion of X-ray signals are contributed by materials that are not seen in SE or BSE images. Thus, in SEM, an unseen phase below the surface may be detected during X-ray analysis, as illustrated in **Figure 19**.

Therefore, the spatial resolution of X-ray analysis in SEM or electron probe microanalysis is limited to the sampling volume of the X-rays, which is typically at the level of micrometers although the imaging resolution of SEMs can achieve 1 nm. It is difficult to reduce the sampling volume below $\sim 1\ \mu m^3$ without reducing the energy of the incident electron

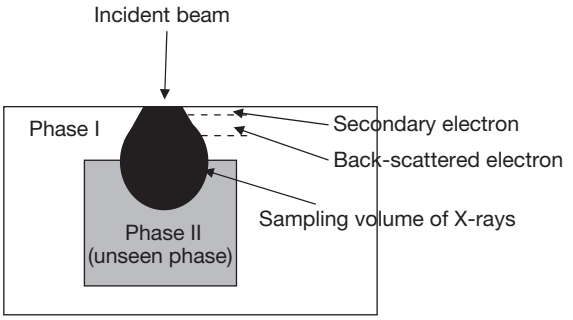


Figure 19 Schematic diagram illustrating the sampling volumes of imaging and X-ray analysis in SEM.

beam to the degree that no useful X-rays are generated. However, as illustrated in **Figure 5**, the lateral spread of the interaction volume and X-ray sampling volume increases with the depth of penetration. Thus, the problem can be overcome by using a sufficiently thin specimen in the TEM. With the use of such a specimen in the TEM, the X-ray sampling volume is approximately equivalent to the imaging region. Thus, the spatial resolution of X-ray analysis in the TEM is approximately the probe size, typically of diameter down to ~1 nm. However, the small amount of X-rays generated by the low current of such a small probe from a small sampling volume of the thin specimen makes meaningful analysis difficult. Practically, it is difficult to achieve meaningful analysis with a probe size of diameter much less than 10 nm for a typical specimen.

Further, in a TEM, the specimen is closely surrounded by the specimen holder and supporting grid, which can interfere with the analysis since X-rays can also be generated by BSEs from the specimen holder and the supporting grid. Such interference can be minimized by using a beryllium specimen holder which generates X-rays that are not detectable and by using grids made of a material that is irrelevant to the specimen. It is also critical to ensure that the grid is placed in a direction such that the specimen faces the incident beam, thereby minimizing interference.

If only one spot or area is analyzed, quantification of the relative proportions of elements under that spot or area is possible, which is known as point analysis (**Figures 1** and **20**). If a line of pixels is analyzed, changes in the proportions of elements with distance along the line can be determined, which is termed X-ray line scanning. Each pixel in an image can also be analyzed, illustrating the locations and amounts of a selected element across the image, which is known as X-ray mapping.

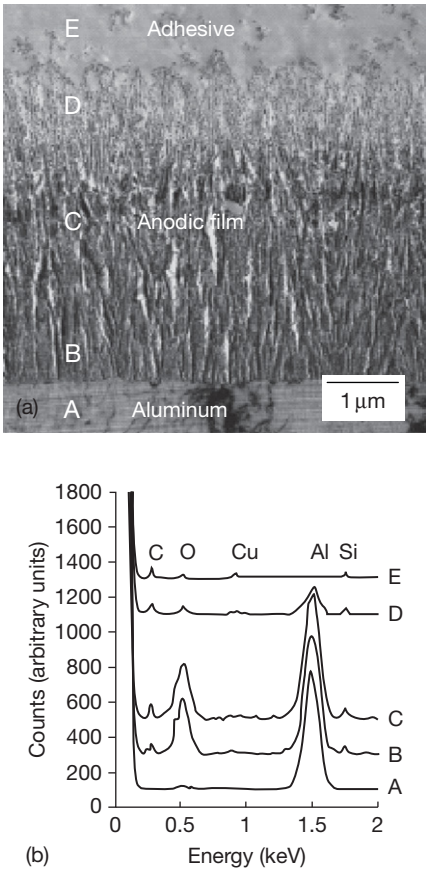


Figure 20 (a) Transmission electron micrograph of the cross section of the interface between the adhesive and the anodic porous alumina film formed on aluminum substrate and (b) EDX spectra taken from the areas indicated in (a).

Figure 20(a) illustrates the TEM image of the cross section taken from the interface region of adhesively bonded aluminum that was pretreated by chromic acid anodizing (CAA). The morphology of the porous anodic film formed by CAA is readily revealed from observation of the cross-section displayed in **Figure 20(a)**. The anodic film is well wetted by the adhesive, and the relatively open morphology of the outer region of the anodic film allows significant penetration of the adhesive. However, the extent of penetration of the adhesive into the pores of the anodic film cannot be determined precisely from the micrograph. Hence, energy dispersive analysis of X-rays was performed by placing an electron probe at various locations across the section of the anodic film/adhesive interface. **Figure 20(b)** shows the EDX spectra. Spectrum A, taken from the aluminum substrate, indicates that only aluminum is detected. Spectra B, C and D were taken within the

anodic film close to the interface with the aluminum substrate and in the middle and the outer region of section of the anodic film, respectively, indicated the presence of Al, O and, additionally, C and Si; C and Si arise from the adhesive (spectrum (E) taken from the bulk adhesive) that had penetrated into the film. This suggests that the adhesive penetrated significantly into the porous anodic film.

Figure 21(a) shows a transmission electron micrograph of an ultramicrotomed section of an AA5182 aluminum alloy after accelerated corrosion testing, revealing the corrosion front and the intact grain boundaries ahead of the corrosion front. Significant attack along a grain boundary is evident. Figure 21(b) shows the EDX spectra generated from a probe located at the grain boundary and within the bulk grains. For an electron probe of a nominal diameter 20 nm located on the boundary, an increased yield of magnesium was detected compared with that from the bulk grain, suggesting the presence of a magnesium-rich phase at the grain boundary since the relative peak height (area) is generally proportional to the concentration of individual elements in the specimen. The revelation of the magnesium-rich phase, probably β phase (Mg_2Al_3), at the grain boundary helps in explaining the occurrence of intergranular attack of the alloy. Since the magnesium-rich phase is anodic with respect to bulk grain, when such phases are formed at grain boundaries, the microgalvanic coupling between the magnesium-rich phase and the adjacent grain matrix provides the driving force for localized corrosion.

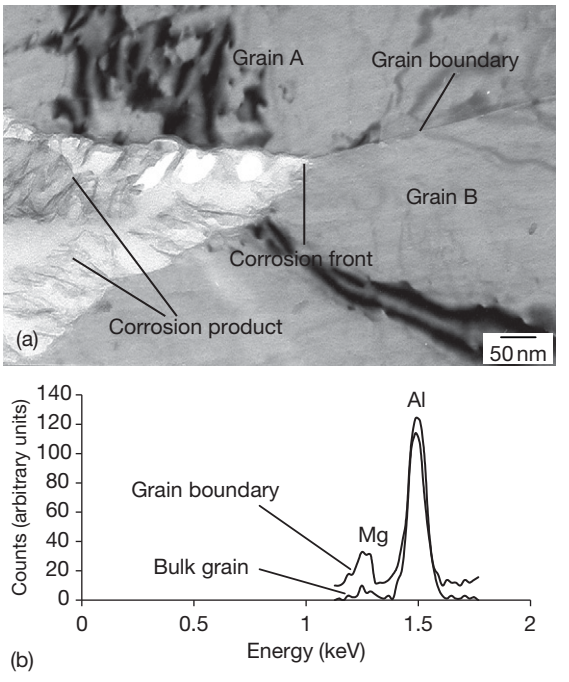


Figure 21 (a) Transmission electron micrograph of ultramicrotomed section of AA5182 aluminum alloy after accelerated corrosion testing, revealing the corrosion front and (b) EDX spectra taken from the intact grain boundary and bulk grain.

Figure 22 illustrates the line scan profiles across a precipitate revealed in the DF image of the AA6111 aluminum alloy shown in Figure 12. By scanning with an electron probe of nominal diameter ~2 nm across the precipitate and the adjacent grain matrix, changes in the proportions of elements with distance along the line are determined. Increased yields of copper, magnesium and silicon at the grain boundary particles are detected with respect to the grain matrix, suggesting the presence of the Q phase ($Cu_2Mg_8Si_7Al_4$). Notably, no solute-depleted zones are readily detected in the vicinity of the grain boundary precipitates or at the grain boundary itself.

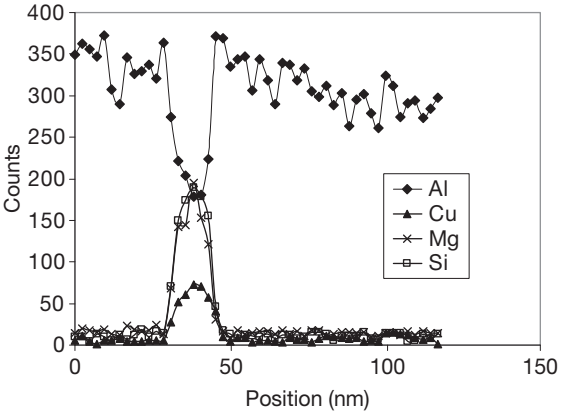


Figure 22 EDX line scanning profiles across a precipitate revealed in Figure 12 of the dark field image of AA6111 aluminum alloy.

2.32.4.2 Electron Probe Microanalysis

Electron probe microanalysis (EPMA) is a nondestructive technique that allows determination of the composition of small volumes of a solid material. In an electron probe microanalyzer, an electron beam is generated and directed to a specimen to be examined and a wavelength dispersive spectrometer is employed to detect the X-rays emitted from the specimen. By scanning the electron beam and fitting an electron detector, complementary images may also be obtained. In this mode, an EPMA is similar to the SEM. As stated

previously, WDS provides much improved energy resolution and signal-to-noise ratio compared to EDS. However, to achieve more accurate quantitative analysis and to cover the full X-ray spectrum, a range of crystals of different lattice spaces is required. Further, the mechanics of WDS place strict geometric restrictions on the positions of the specimen, crystal spectrometer, and the detector. These make it difficult, but not impossible, to fit such a system into an electron microscope. Thus, a purpose-built electron probe microanalyzer is necessary. Typically, four computer-controlled crystal spectrometers containing a range of crystals of different lattice spacing are installed in a dedicated EPMA system. Such equipment can achieve more accurate quantitative analysis than X-ray analyses in a standard electron microscope. EPMA can quantitatively analyze elements down to boron in the periodic table at levels as low as 100 ppm.

Mapping of elemental distributions by EPMA in the cross-section of a coating formed on an aluminum substrate anodized for 1800 s at $5\,A\,dm^{-2}$ in 0.05 M potassium hydroxide/0.15 M sodium metasilicate electrolyte is shown in **Figure 23**. Oxygen is revealed in all regions of the coating thickness, apart from cavities. Other elements are less uniformly distributed. Notably, aluminum is largely absent in some regions of size up to 50 µm, which are enriched in silicon. These regions were found at various depths within the coating, including regions close to the metal.

2.32.4.3 Electron Energy Loss Spectroscopy

As discussed previously, a primary electron that undergoes any inelastic interaction with the examined material loses energy. EELS determines the losses of electron energy. The most commonly used EELS is transmission EELS, in which the energy loss of incident electrons with a known, narrow range of kinetic energy passing through a thin specimen is determined by an electron spectrometer that is mounted after the specimen in the TEM.

Typically, an EELS spectrum consists of three regions: (i) zero loss peak, resulting from electrons that have suffered negligible inelastic scattering; (ii) low loss region, containing electrons that have suffered inelastic scatterings including phonon excitations, inter- and intraband transitions, plasmon excitations, and have lost up to ~50 eV energy and (iii) high loss region, containing electrons that have suffered interaction with tightly bound inner shell electrons and caused inner shell ionization, which is characteristic of the element. Although each region of the EELS spectrum provides complementary information, it is the characteristic edges in the high energy loss region that are most commonly used. In TEM, EELS is capable of obtaining spatially resolved information on composition, chemical bonding as well as valence and conduction band electronic structure since the spectra for an element in different compounds are different. This ability is a strong advantage of EELS compared with EDX.

A further advantage of EELS compared to EDX is its capability of detecting light elements that generate X-rays that are too soft to be detected, such as helium, lithium, and beryllium. The probability that a particular electron will be scattered by an atom is defined as the cross-section, which decreases for elements of high atomic number. Thus, light element edges are far more intensive than those of elements with higher atomic number. Thus, EELS is very powerful for the analysis of light elements.

In corrosion studies, spatially resolved EELS has been successfully applied, for example, to the study of local chemistry of anodic oxide films. A correlation between flaws in the oxide film and changes in the oxygen gas generation has been established via spatially resolved EELS. **Figure 24** shows a transmission electron micrograph of ultramicrotomed section of the anodic film formed at a constant current density of $50\,A\,m^{-2}$ on a binary Al–4 at.% Cr alloy in 0.01 M ammonium pentaborate solution to 32 V, revealing an anodic film of ~40 nm thickness. Variation of film morphology is clearly evident. Light regions of different sizes are revealed within the film section. **Figure 25** displays the EELS spectra at the oxygen K-shell electron ionization energy region acquired from locations across the film from the film surface to the alloy/film interface; the separation of the spectrum line is 0.8 nm. The spatially resolved oxygen K-edge fine structure encompassing a region of light film materials reveals a sharp pre-edge peak at 529 eV before the OK edge; such pre-edge peaks are characteristic of oxygen gas. Thus, it is concluded that the light regions in the film are bubbles filled with oxygen gas.

Further, the valence state of chromium species that have been incorporated into the film from the substrate alloy has also been determined by EELS. Using the data generated from anodic alumina, Cr_2O_3 and $K_2Cr_2O_7$, spectra have been simulated for anodic alumina with associated Cr^{3+} or Cr^{6+} species. From the simulation and measurement of the experimental spectrum acquired at a location in the anodic film, as shown in **Figure 26**, it is evident

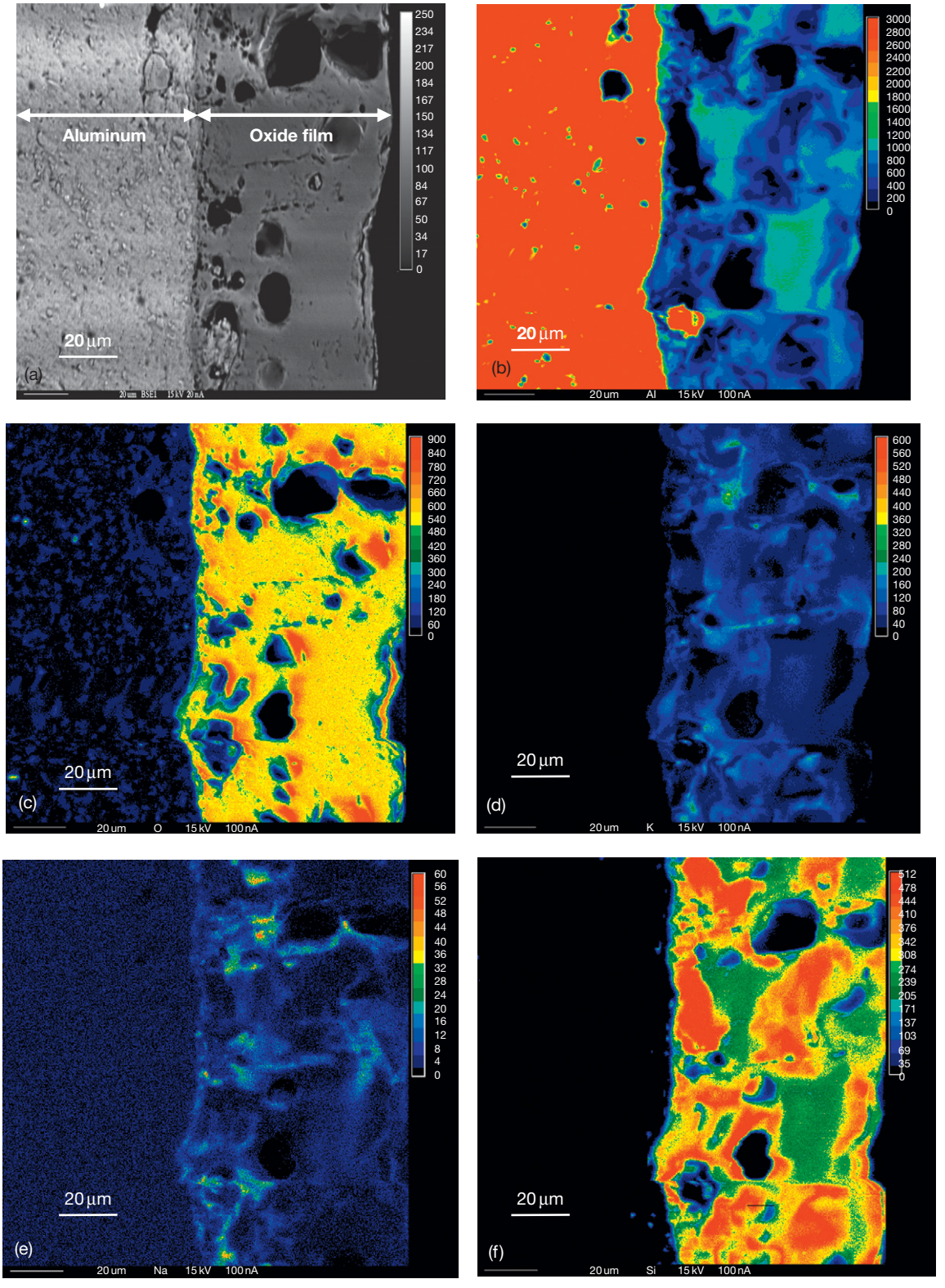


Figure 23 EPMA of the cross section of aluminum anodized for 1800 s at 5 A dm^{-2} in 0.05 M potassium hydroxide/0.15 M sodium metasilicate electrolyte at 293 K: (a) scanning electron micrograph, (b) aluminum map, (c) oxygen map, (d) potassium map, (e) sodium map, and (f) silicon map.

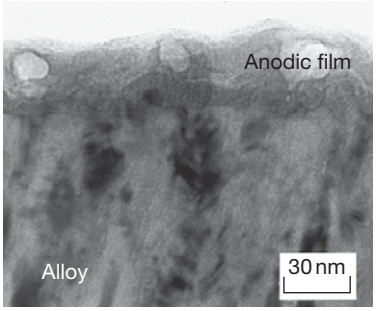


Figure 24 Transmission electron micrograph of an ultramicrotomed section of the anodic film formed at a constant current density of 50 A m^{-2} on a binary Al–4 at.% Cr alloy in 0.01 M ammonium pentaborate solution to 32 V.

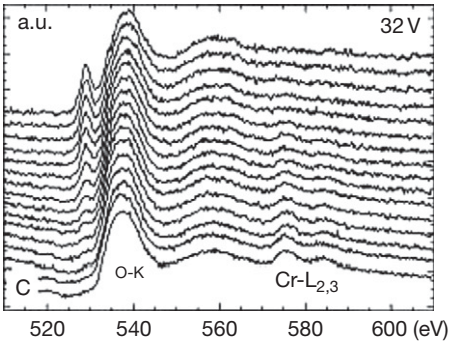


Figure 25 EELS spectra acquired at a series of locations from close to the film surface to the alloy–film interface for the film displayed in **Figure 24**.

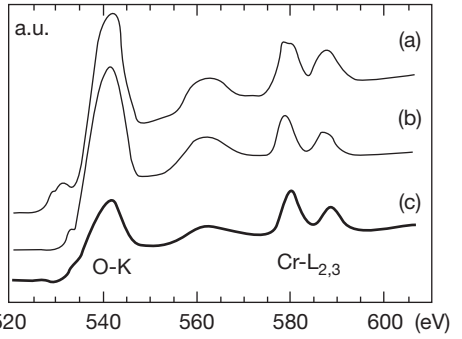


Figure 26 Simulated and experimental EELS spectra: (a) simulated spectrum for anodic alumina containing Cr^{3+} species, (b) simulated spectrum for anodic alumina containing Cr^{6+} species, (c) experimental spectrum acquired at a location in the anodic film close to the alloy–film interface.

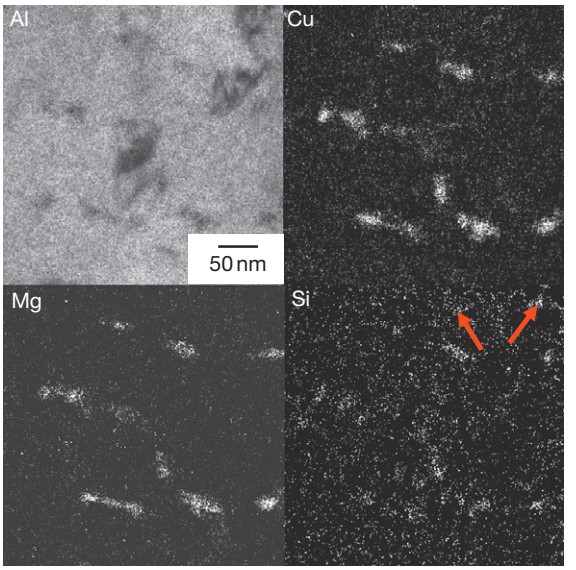


Figure 27 EELS maps showing the distribution of precipitates in AA6111 aluminum alloy. The bright areas represent relatively high intensity of signals.

that Cr^{3+} species was incorporated into the anodic film during the anodization of Al–4 at.% Cr alloy.

If EELS spectra are acquired from each pixel in an image, the distribution of a chosen element at certain oxidation state across the image can be analyzed, which is known as EELS mapping. For example, the DF image of **Figure 12(b)** reveals precipitate particles of ~20 nm diameter distributed along grain boundaries in the examined AA6111 aluminum alloy. However, no compositional information is revealed in **Figure 12(b)**. Such information can be readily obtained by EELS mapping. **Figure 27** shows EELS maps of precipitates in the AA6111 aluminum alloy revealed in **Figure 12(b)**. The bright areas represent relatively high intensities of signals. Mapping images of 508 × 508 pixels were acquired, giving a pixel size of 0.92 × 0.92 nm for the selected magnification. The images were selected at the Al K, Cu $L_{2,3}$, Mg K and Si K edges, with onsets at 1560, 931, 1305 and 1839 eV, respectively, and energy resolution of about 0.6 eV. The EELS maps of **Figure 27** confirm that the fine particles contain increased levels of copper, magnesium and silicon compared to the matrix. Examination of the images also reveals regions in which only silicon was detected, as shown by arrows, indicating the presence of two different types of precipitates in the alloy.

Comparing **Figure 27** with the X-ray maps (EPMA or EDX, **Figure 23**), it is clearly evident that EELS mapping has significantly improved the spatial resolution. This is because the spatial resolution of X-ray analysis in SEM or electron probe

microanalysis is limited to the sampling volume of the X-rays, which is typically at the level of micrometers although the imaging resolution of the SEMs is as high as 1 nm.

2.32.5 Specimen Preparation Techniques

Successful electron microscopy is essentially dependent on the employment of appropriate specimen preparation techniques that can produce revealing specimens and avoid, or minimize, change to the materials.

2.32.5.1 TEM Specimen Preparation

Since TEM forms images following transmission of an electron beam through the specimen, thin foil specimen of thickness from a few nanometers to a few hundred nanometers must be prepared. For example, at an accelerating voltage of 100 kV, the maximum thickness of a transparent specimen is about 100 nm depending on the density of the material. The task is made even harder by the need to precisely prepare thin foil specimens from interesting regions within a large sample. Thus, preparation of TEM specimen can be a very challenging and time-consuming operation. The commonly employed TEM specimen preparation techniques include

- electropolishing,
- chemical polishing,
- ion beam thinning,
- deposition,
- film stripping,
- ultramicrotomy, and
- replication.

Depending on the type of materials being studied, appropriate preparation techniques should be chosen to generate an electron-transparent specimen.

TEM specimens from materials that have dimensions sufficiently small to be transparent to an electron, such as powders or nanotubes, can be quickly produced by the deposition of a suspension containing the material onto support grids. The suspension is normally made by diluting the sample in a volatile solvent such as ethanol. The solvent rapidly evaporates after depositing the suspension onto support grids, thereby producing a specimen that is suitable for TEM. Dry powder may also be directly dusted onto support grids.

The surface of a solid sample cannot be observed in TEM directly. The examination of a surface of a solid sample is achieved by replication. The replication process typically involves molding the surface of interest in a thin replica material which is transparent to electrons and subsequently removing the replica from the surface. Then the replica, which has copied the features on the surface, is ready for examination in the TEM.

Thin oxide films can be stripped from metal substrates by a chemical stripping process and transferred to supporting grids for TEM examination. Electropolishing is the most commonly employed technique for preparing thin foil specimen from electrically conductive materials such as metals. Chemical polishing is frequently used for generating TEM specimens from nonconductive materials such as semiconductors and ceramics. The main drawback of both electropolishing and chemical polishing is that a certain component of the material may be preferentially leached to change the specimen, thereby leading to inaccuracies in the subsequent TEM characterization.

Ion beam thinning is widely used to prepare an electron-transparent foil from hard materials including metals and ceramics. Focused ion beam (FIB) thinning is a relatively new technique that can be employed to prepare thin foils from precisely selected regions in a relatively large sample of any material. However, the highly energetic ion beam may introduce artifacts such as amorphization of the materials or contamination to the resultant thin foil specimen. Such artifacts introduce noticeable effects when lattice images are examined in HRTEM or EELS analysis is carried out.

In corrosion studies, ultramicrotomy is a very powerful tool. With experience, very revealing specimens can be produced. The principal advantages of ultramicrotomy are that it can create uniform thin foils and maintain the integrity of a multiphase sample (**Figure 20(a)**), for example, enabling thin foil specimens to be generated displaying the interface between the metal and its protective coatings (**Figure 13**) or exhibiting corrosion sites with corrosion products being kept in its original location (**Figures 1, 14,** and **21**). However, ultramicrotomy may introduce mechanical damage, mainly compression, to the thin foil specimen; also reaction may occur between the materials and water in the bath behind the diamond knife employed in ultramicrotomy. The latter can be used to advantage in corrosion studies in the case of employing thin foil specimen to investigate the relationship between the microstructure and corrosion initiation.

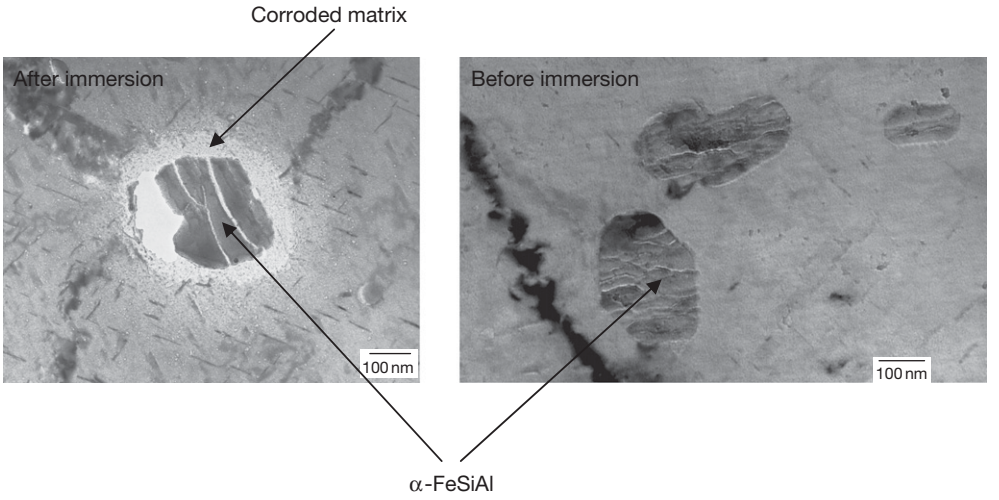


Figure 28 Transmission electron micrographs of an ultramicromomed section of an aluminum alloy before/after immersion in water.

Figure 28 shows transmission electron micrographs of the ultramicromomed section of an aluminum alloy before and after the immersion in water; α-FeSiAl-particle-induced localized corrosion of the alloy matrix, due to the cathodic nature of the particle with respect to surrounding matrix, is revealed.

2.32.5.2 SEM Specimen Preparation

Compared to TEM, SEM allows much larger specimens to be examined, and specimen preparation is much simpler. Examination of bulk specimens makes it possible to image a comparatively large area. In corrosion studies, it is often needed to examine the cross-section of a specimen to determine corrosion propagation paths, the extent of corrosion beneath the surface, the thickness of a protective coating and its morphology in inner regions, its composition profile and adhesion to substrate, etc. In addition to conventional methods, such as mechanical polishing, for the preparation of the cross-section, new methods such as ultramicrotomy are increasingly employed to produce cross-section specimens with significantly reduced roughness, mechanical damage and contamination, thereby producing revealing specimens for the new generation of SEMs that provides very high surface sensitivity by avoiding interference between signals from different depths of specimens.

Further, with the typical operation voltages (20–30 kV), the number of electrons that escape from the specimen is smaller than that of electrons that hit the specimen. Consequently, there is a buildup of electrons at the specimen surface if the excess electrons are not conducted away from the specimen, leading to the specimen surface becoming negatively charged. As a result, the incident electrons are repelled, and the image is distorted. Thus, for a nonconductive specimen surface, such as a metal with corrosion products or oxides films on the surface, the surface must be coated with a thin conducting layer, such as carbon or gold–palladium, which does not readily form an insulating oxide layer in air during its transport from the coating unit to the microscope. Thus, the build-up of electrons at the specimen surface is avoided by conducting the excess electrons away from the specimen by the conducting coating. However, there are drawbacks for the employment of such coatings. The coatings may mask fine surface features. The coating may also absorb some of the X-rays emitted from the specimen during X-ray analysis. The absorption is more pronounced by coating with elements of higher atomic number such as gold/palladium. An alternative approach to the surface charging problem is to operate the SEM at a relatively low voltage (1–5 kV) so that excessive electron build-up at the specimen surface is minimized.

2.32.6 Other Techniques

2.32.6.1 X-ray Microscopy

2.32.6.1.1 Introduction

X-ray microscopy is a nondestructive technique that generates an image of the internal features of the

examined materials, with the image contrast being determined by the difference in absorption of X-rays by the different components in the materials. The first X-ray microscope was developed in the late 1940s. In the X-ray microscope, images are recorded on photographic film or detected by a sensor such as a CCD camera. Similar to visible light, X-rays are a form of electromagnetic energy, but with much shorter wavelength and much higher energy. The resolution of X-ray microscopy lies between that of light microscopy and electron microscopy since an X-ray microscope uses soft X-rays with a wavelength in the range of 10 to 0.01 nm. Resolutions of nanometer level are achievable using the Fresnel zone plate lens which forms the image using the soft X-rays emitted from a synchrotron or from a laser-produced plasma. Further, because of their high energy, X-ray microscopes have the ability to penetrate relatively deep into materials, allowing images to be generated from the internal features. Compared to TEM, X-ray microscopy can generate images from relatively thick samples. A further advantage over electron microscopy is that X-ray microscopy does not require particular specimen preparation or a high-vacuum sample chamber.

2.32.6.1.2 X-ray tomography

Standard X-ray microscopy can generate 2-D images which locate the position of an internal feature in the examined specimen both vertically and horizontally but with no information on the depth of that feature within the specimen being obtained. However, the information concerning the depth of the feature can be obtained by X-ray tomography. X-ray tomography is a 3-D scanning technique that allows nondestructive, high-definition scans of a specimen to be made. 3-D tomography requires the collection of a series of 2-D X-ray images on all sides of the specimen. Subsequently, highly detailed 3-D images are reconstructed by merging the individual 2-D X-ray images into a single 3-D block which can then locate any feature in the specimen in any direction. Systems with zone plate lenses can provide spatial resolution better than 50 nm.

The ability of nondestructive imaging of the internal features of materials in three dimensions gives X-ray tomography critical advantages in applications such as failure analysis as well as fundamental research. For example, internal cracks can be observed in three dimensions and correlated to the microstructure of a material; intergranular corrosion may be examined and correlated to grain boundary compositions; the 3-D approach can also lead to an improved understanding of interfacial reactions that occur at the intermetallics–matrix interface during pitting. This may lead to significant insight and understanding of the relationship between the microstructure and performance, that is, susceptibility to corrosion. X-ray tomography may also be extended to the observation of real-time processes of crack growth and corrosion propagation. Figure 29 shows an image of isotropic nuclear graphite obtained on the Swiss Light Source TOMCAT synchrotron, revealing the porosity introduced by the fabrication process.

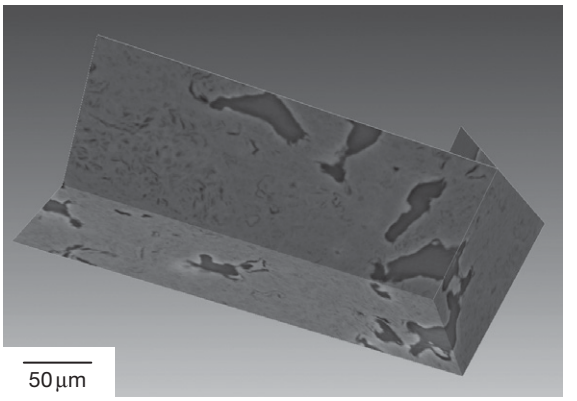


Figure 29 X-ray tomography image of isotropic nuclear graphite, revealing porosity introduced by the fabrication process. Courtesy D. James, A. Jone, T.J. Marrow, and N. Stevens, Materials Performance Centre, School of Materials, The University of Manchester.

2.32.6.2 Infrared Microscopy

Infrared spectroscopy is a technique based on the specific resonant frequencies of molecules, that is, the frequencies at which they rotate or vibrate. The resonant frequencies can be related to the strength of specific bonds and the mass of the atoms forming the bonds. Thus, the frequency of the vibrations can be associated with a particular bond type. When an infrared beam is directed to the material to be examined, the rotation or vibrations can lead to infrared absorptions at characteristic frequencies that are related to specific chemical groups in the material. The infrared spectrum is obtained by measuring the absorption at each wavelength. Thus, infrared spectroscopy can provide direct molecular identification of the examined materials.

In an infrared microscope, spatially resolved infrared spectra are obtained to form images, with the contrast being determined by the response of

individual specimen regions to particular wavelengths, generating a molecular fingerprint from the specimen. For infrared microscopy, as well as Raman microscopy, which is discussed in the next section, there are two approaches to acquire the spectrum. These approaches are known as parallel acquisition and serial acquisition. For parallel acquisition, the area of interest is fully illuminated by the incident beam; then, a micrograph is generated at a chosen wavelength which is characteristic of a molecular compound. For the serial mode, a finely focused incident beam is rastered across the surface of the materials to be examined, generating spatially resolved spectra. Subsequently, the micrograph is reconstructed to display the locations and amount of molecular compounds within the area of interest. Infrared microscopy employs focal plane array detection for infrared chemical imaging. The spatial resolution of infrared microscope is limited to about 10 μm depending on the wavelength of light and the aperture. A typical infrared microscope operates over the range 650–4000 cm^{-1}.

In corrosion studies, infrared microscopy can provide valuable information on interfacial reactions such as inhibition of metals and chemical bonding between a protective coating and the substrate. Further, by measuring at a specific frequency over time, changes in the character or quantity of a particular bond can be measured. This is especially useful in measuring the degree of polymerization in an organic protective coating. The infrared microscope operates in three different modes, transmission, reflection, and absorbance, depending on the transparency of the material to be examined. Materials that are transparent to infrared may be investigated in the transmission mode, which usually requires that the specimen is polished on both sides. For the reflectance mode, little specimen preparation is required.

2.32.6.3 Raman Microscopy

Like infrared spectroscopy, Raman spectroscopy is also based on the vibration or rotation of molecules within the material of interest. Thus, similar information is generated by Raman spectroscopy. However, infrared spectroscopy and Raman spectroscopy yield complementary information since they are based on different physical principles. Raman spectroscopy is based on inelastic scattering, or Raman scattering, of the incident monochromatic radiation by the material. Normally, a monochromatic source of light is used as the incident radiation. Highly monochromatic and intense lasers in the visible, near-infrared or near-ultraviolet regions are usually employed as incident radiation in a modern Raman spectroscope. The interaction between the material and the laser beam results in phonons or other excitations, leading to the energy of the laser photons being shifted. Thereafter, the laser beam is collected with a lens and transferred through a monochromator. Wavelengths close to that of the incident radiation, due to elastic Rayleigh scattering, are filtered out, while the rest of the collected light is dispersed onto a detector to generate the spectrum. Thus, Raman spectroscopy allows the molecular identification of species present in the examined materials since the energy shift of the laser beam is associated with specific molecules in the materials.

A typical Raman microscope consists of a standard optical microscope, a laser, a monochromator, and a detector. A CCD or photomultiplier tube (PMT) is normally used as the detector. Relatively long acquisition times are required when a photomultiplier tube is used. The acquisition time is significantly reduced with CCD detectors. In Raman microscopy, the micrograph is obtained by generating spatially resolved spectra in the parallel or serial acquisition mode. The spectra are then analyzed and used to form a corresponding chemical image that shows the location and amount of different components within the examined area. Raman microscopy has relatively high spatial resolution. The lateral and depth resolutions are approximately 0.3 and 2 μm, respectively. Raman microscopy cannot be performed on materials that exhibit fluorescence, since fluorescence yields a much more intense signal than Raman scattering and, therefore, masks any Raman bands that might be present. For incident radiation of wavelengths less than 785 nm, fluorescence interference can be very common. Moving to long-wavelength lasers can mitigate much of the fluorescence interference. However, the efficiency of Raman scattering decreases as the fourth power of wavelength, making long-wavelength Raman microscopes much less sensitive than systems based on visible wavelengths. Since Raman spectroscopy is a scattering technique, specimens do not need to be specially prepared.

Recently, Raman spectroscopy has been successfully combined with SEM and scanning probe microscopy (SPM), allowing morphological, compositional, structural and chemical analyses without moving the specimen between instruments. The features of interest are precisely located using the SEM and SPM and are chemically identified using Raman spectroscopy.

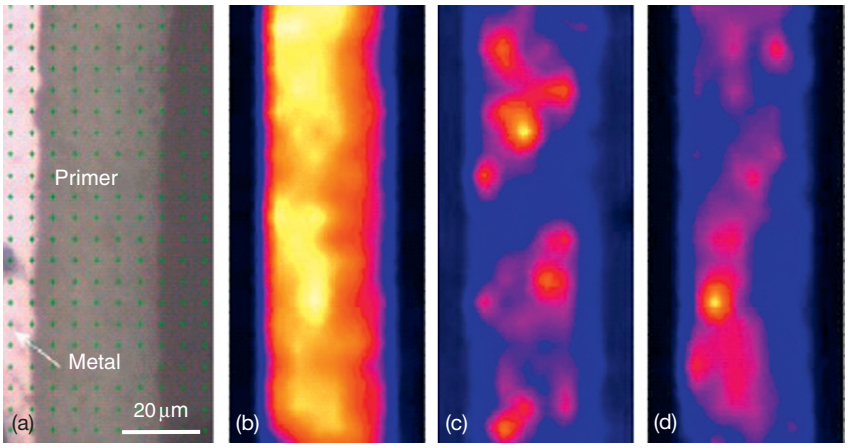


Figure 30 Raman micrograph of chromate-inhibited primer: (a) optical image, (b) TiO_2, (c) $BaSO_4$, and (d) $SrCrO_4$. Courtesy Dr A.E. Hughes, CSIRO Manufactory and Infrastructure Technology, Australia.

Figure 30 displays the Raman micrographs of a cross-section of the interface between an AA2024 aluminum alloy and chromate-inhibited epoxy polyamide primer, revealing the distribution of various components, including TiO_2, $BaSO_4$ and $SrCrO_4$ in the primer. Raman micrographs were generated using peaks at 866, 987, and 638 cm^{-1} for $SrCrO_4$, $BaSO_4$ and TiO_2, respectively. It is clearly evident that TiO_2 is distributed relatively uniformly throughout the primer, whereas $BaSO_4$ and $SrCrO_4$ exhibit locally high concentration corresponding to relatively large particles of these two phases in the primer.

2.32.7 Concluding Remarks

This chapter is written mainly on the basis of experiences in the Corrosion and Protection Centre at The University of Manchester. The authors acknowledge the many colleagues who have also contributed significantly to the application of electron- and photon-based spatially resolved techniques in corrosion studies.

Leading scientists around the world continue to raise expectation for electron microscopy, and the manufacturers of electron microscopes continue to meet the expectation with improved design of the electron microscope. The introduction of the FEI Titan TEM in 2005 is a good example. In the Titan microscope, aberration correction breaks through the fundamental optical limitations of magnetic lenses. The Titan's extra wide column was specifically designed to give the mechanical stability required by the added height of probe and image correctors. Its proprietary constant power lenses and advanced power supplies provide the necessary thermal and electronic stability. Titan enables sub-Angstrom microscopy in both TEM and STEM modes. The Titan can also achieve energy resolution down to 0.1 eV for EELS, obtaining information of the electronic properties of materials, such as bonding states or band gaps, with unprecedented spatial resolution. Titan ETEM extends the capabilities of conventional TEM by permitting gas in the specimen chamber (up to 4 kPa) along with specimen heating (up to 1000 °C), thereby allowing investigation of interactions between the material and the local gaseous environment.

Although this chapter is entitled 'electron- and photon-based spatially resolved techniques,' it would not be complete without mentioning the increasing sophistication in microscopy, for example, the recently developed Carl Zeiss ORION helium ion microscope. The scanning ion microscope uses a beam of helium ions as the incident radiation. Like a conventional SEM, in a scanning helium ion microscope the ion beam is rastered across the surface of the material to be examined. Since ion beam can be focused to a finer size, with less surface interaction than electrons, the scanning ion microscope can generate images with increased resolution and better contrast, more surface sensitivity, and better depth of field. Further, as in the case of Rutherford backscattering spectroscopy, backscattered helium ions can be collected in the helium ion microscope, with the total backscattered ion yield being directly proportional to the mass of the sample atoms. Thus, it is possible to easily differentiate between elements.

For those who wish to pursue the subject further, there are many textbooks that provide much more detailed information on electron- and photon-based spatially resolved techniques.

Further Reading

Clarke, A. R.; Eberhardt, C. N. *Microscopy Techniques for Materials Science*; CRC Press: Boca Raton and Cambridge, 2002.
Egerton, R. F. *Physical Principles of Electron Microscopy: An Introduction to TEM, SEM, and AEM*; Springer: Boston, 2005.
Goodhew, P. J.; Humphreys, J.; Beanland, R. *Electron Microscopy and Analysis*, 3rd ed.; Taylor & Francis: London, 2001.
Lawes, G. In *Scanning Electron Microscopy and X-ray Microanalysis*; James, A. M., Ed.; Wiley: Chichester, 1987.
Reimer, L. *Transmission Electron Microscopy: Physics of Image Formation and Microanalysis*, 4th ed.; Springer: Berlin and London, 1997.
Reimer, L. *Scanning Electron Microscopy: Physics of Image Formation and Microanalysis*, 2nd ed.; Springer: Berlin and London, 1998.
Stuart, B. H. *Infrared Spectroscopy: Fundamentals and Applications*; Wiley: Chichester, 2004.
Turrell, G.; Corset, J. *Raman Microscopy: Developments and Applications*; Academic Press: London, 1996.
Williams, D. B.; Carter, C. B. *Transmission Electron Microscopy*; Plenum Press: New York and London, 1996.

2.34 Corrosion Testing and Determination of Corrosion Rates

P. J. McIntyre
NPL, Hampton Road, Teddington, Middlesex TW11 0LW, UK
A. D. Mercer
Formerly NPL, Hampton Road, Teddington, Middlesex TW11 0LW, UK

This article is a revision of the Third Edition article 19.1 by P. McIntyre and A. D. Mercer, volume 2, pp 19:3–19:118,

Glossary

σ-phase A brittle intermetallic phase that forms in stainless steels in some heat treatment conditions.

Abbreviations

ASTM American Society for Testing and Materials
BS British Standard
CCT Critical crevice temperature
CPT Critical pitting temperature
EC Electrolytic corrosion
EPR Electrochemical potentiokinetic reactivation
gma g per m^2 per annum
IEC International Electrotechnical Commission
ISO International Organization for Standardization
LME Liquid–metal embrittlement
mdd mg per dm^2 per day
mpy mils (0.001 inch) per year
NACE National Association of Corrosion Engineers
PTFE Polytetrafluoroethylene
SCE Saturated calomel electrode
SHE Standard hydrogen electrode

Symbols

b_a, b_c Tafel constants for anodic and cathodic processes
E Potential
E_b Pitting breakdown potential
E_{corr} Corrosion potential
E_p Pit repassivation potential
i Current density
i_a, i_c Net anodic and cathodic current densities
i_{corr} Corrosion current density
i_L Limiting current density
i_o Exchange current density
I Total current
Q Electrical charge
R Electrical resistance
R_p Polarization resistance
β_a, β_c Slopes of anodic and cathodic polarization curves
$\Delta\psi$ Potential drop across the double layer
ΔK Stress intensity factor range
ρ Density

2.34.1 Introduction

Corrosion tests provide the basis for the practical control of corrosion and therefore deserve a more exhaustive discussion than limitations of space will permit. A detailed description of all the procedures and devices that have been employed in corrosion studies in many countries will not be attempted. Instead, attention will be directed principally to underlying principles and to comments on the significance and limitations of the results of the test methods that are considered. Further details may be obtained from the references and from the comprehensive works by Champion[1] and Ailor.[2] (See also: Baboian, R.; Dean, S. W. *Corrosion Testing and Evaluation: Silver Anniversary Volume*, ASTM STP1000-EB; ASTM: Philadelphia, 1990; 436 p. and Cramer, S. D.; Covino, B. S., Jr., Eds. *ASM Handbook Volume 13A: Corrosion: Fundamentals, Testing, and Protection*; ASM International, Ohio, 2003; 1135 p.)

Tests may be classified conveniently under three headings.

1. Laboratory tests, in which conditions can be precisely defined and controlled.
2. Field tests (tests in real environments), in which replicate test samples of metals or alloys – referred to as test coupons or specimens – are exposed to the actual environmental conditions expected in service, for example, the atmosphere, the ground, the sea, etc.
3. Service tests, in which the test specimens – which may often take the form of manufactured components – are exposed to the particular conditions in which they are to be used, for example, in process streams of chemical plant.

Laboratory tests, although often necessarily conducted under conditions that are not met in service, have a number of advantages over the other types of tests. Since conditions can be controlled at will, it is possible to identify the separate effects of a number of factors on the corrosion behavior. These factors include the type and condition of the metal surface, the environmental composition, temperature and pressure, movement of the specimen relative to the environment, time of exposure, and so on. Laboratory tests, at least in principle, also enable comparisons to be made under identical conditions of the relative corrosion behavior of different metals and alloys and different protective schemes, for example, coatings, environmental treatments, etc.

In many cases, attempts will be made to accelerate the test to produce results in a shorter time than might otherwise be possible in field or service tests. Such acceleration is usually achieved by intensifying one or more of the controlling factors. Tests might be conducted at a higher temperature, with more corrosive media, with activation of the corrosion process by electrochemical methods, etc., with the object of enhancing the aggressivity of the test conditions. While accelerated test procedures are often used, the results should always be treated with careful consideration. It is not unknown for a protective system to fail to meet the requirements of an accelerated test, although showing satisfactory performance in normal conditions of use. Nevertheless, a number of such tests, particularly for atmospheric corrosion, in which rates of corrosion in real conditions are often low, are accepted and correlations have been established with real conditions. (Such tests are typically designed and standardized for very specific conditions, such as the performance of nickel–chromium-plated components for automotive applications, and care should be taken when applying them for other situations. A key requirement for any accelerated test is that it should produce the same type of corrosion (uniform, pitting, etc.) as in the service exposure.)

Field tests do not have the uncertainties attached to accelerated laboratory tests, since there is no attempt to adjust the controlling environmental conditions. The chief problem is obtaining reproducible conditions from one test to another. This is particularly the case with tests in the atmosphere. While broad classes of terrestrial atmospheres have long been recognized, for example, tropical, rural, urban, marine, etc., difficulties remain that are associated with variations within these classes. In the 1980s, steps were taken within ISO to rationalize the situation by producing a standard on the classification of atmospheres (ISO 9223). Other standards that provide guidance on the mounting and disposition of specimens for field tests (ISO 8565) and for the statistical treatment of results where large numbers of specimens are used (see Ailor[2] and ASTM G 16) are available.

Service tests will be used: (1) where the operating conditions cannot be successfully reproduced in laboratory tests, (2) where the environment does not occur naturally, (3) where real components, as opposed to test specimens, need appraisal, and (4) to confirm laboratory and/or field tests.

Often, all the three types of test will be used sequentially. An example might be the development of a coating to protect suspension cables for use on a bridge in a coastal region. The test program could involve salt spray testing of candidate treatments in the laboratory, followed by field trials of the most successful materials at a site similar in aggressivity to the location of the final product and eventual testing at the site with loadings and positioning matching those of the end use. It would be expected that the number of candidate materials would decrease through this sequence of tests.

Irrespective of the method of test or the purpose for which it is made, there are certain practical features that require attention and will be necessary to achieve good reproducibility (by one operator) and repeatability (by different operators).

2.34.2 Test Procedures

2.34.2.1 Preparation of Surface

When the test is to predict the performance of a material in a particular service, the ideal procedure would be to have the surface of the test-pieces duplicate the surface of the material as it would be used. Here, however, a complication is presented by the fact that materials in service are commonly used in several forms with different conditions of surface. Where the number of materials to be compared is large, it will usually be impractical to test all the conditions of surface treatment of possible interest. The best practical procedure, then, is to choose some condition of surface more or less arbitrarily to allow the materials to perform near the upper limits of their ability. If all the materials to be tested are treated in this way, and preferably with uniform surface treatment, the results of the test will indicate the relative abilities of the different materials to resist the test environment when in a satisfactory condition of surface treatment. Then, if it should be considered prudent or desirable to do so, the most promising materials can be subjected to

further tests in a variety of surface conditions so that any surface sensitivity can be detected.

These remarks apply as well to the treatment of the surfaces of specimens to be used in tests in corrosion research projects, except that here, selection of a particular method of surface preparation is required to achieve reproducibility of results from test to test and among different investigators. Methods of preparing specimens are described in ASTM G 1 and ISO 7539-1.

The final step in surface preparation should ordinarily be a cleaning and degreasing treatment to remove any dirt, oil, or grease that might interfere with the inception or distribution of corrosion. The simplest test of a satisfactory surface condition in this respect is for the specimens to be free from 'water break' when rinsed with water after cleaning. As a final treatment for specimens to be weighed prior to exposure, a dip in a mixture of water and acetone or of alcohol and ether will facilitate quick drying and avoid water-deposited films. Specimens to be stored prior to weighing should be placed in a desiccator, which, in best practice, should be sealed without grease.[3]

In addition to the preparation of the principal surfaces of the specimen, it is essential to machine or grind any cut or sheared edges, since these could become sites of preferential attack. As a general rule, edge effects should be kept to a minimum by using specimens in which the ratio of surface area to edge area is large. With flat specimens, a disc is best from this point of view, but other shapes may be more convenient and acceptable in many practical instances. When mass loss is to be used as a measure of corrosion, precision will be improved by providing a large ratio of exposed area to mass, and thin flat specimens or fine wires have obvious advantages. (Wires also have the disadvantage of a metallurgical structure that is usually not typical of bulk material.)

For accuracy of weighing, it is usually necessary to restrict the dimensions of specimens to what can be accommodated on the common analytical balances. It must be borne in mind that where attack occurs in the form of a very few pits or in crevices under supports, the extent of this localized attack may be determined by the total area of the test-piece as it establishes the area of passive metal acting as a cathode to the few anodic areas. Thus, larger specimens, or the much larger surfaces that will often be involved in field or service tests, may give rise to much more severe localized attack under nominally the same conditions of exposure.

In certain tests, it is sometimes desirable to eliminate any effects of a mechanically achieved surface condition by chemical treatment or pickling of the surface prior to test. This may be done in a pickling solution; alternatively, the test itself may be interrupted after sufficient corrosion has occurred to remove the original surface, the specimen then being cleaned and reweighed and the test started over again. Wesley[4] found it to be desirable to pickle off about 0.008 mm from the surface of specimens in acid to improve the reproducibility of the tests.

With materials the stainless steels that may be either active or passive in a test environment, it is common practice to produce a particular initial level of passivity or activity by some special chemical treatment prior to exposure. With stainless steels this objective may be subsidiary to eliminating surface contamination, such as iron from processing tools, by treatment in a nitric acid solution, which might also be expected to achieve substantial passivity incidental to the cleaning action (ASTM A 380).

In studies of the behavior of materials that may be either active or passive in the test environment, there would seem to be a real advantage in starting with specimens in an activated state to see whether they will become passive, and to ascertain how fast they are corroded if they remain active. If passivity should be achieved after such an activated start, the material can be considered to be more reliable in the test environment than would be the case if by chance it managed to retain an originally induced passivity for all, or most of, the test period. It may also be valuable to know how fast the metal will be corroded by the test medium if activity should persist.

A procedure for testing previously activated specimens applied in studies of titanium was described by Bayer and Kachik.[5] Renshaw and Ferree[6] also employed prior activation in their studies of the passivation characteristics of stainless steels.

In many cases, there is a need to test metal-coated specimens, for example, galvanized steel, tin-plated copper, nickel-plated zinc, etc. It will then be necessary to test specimens in the completely coated condition and also with the coating damaged so that the basis metal is exposed. The latter condition will provide the conditions for galvanic action between the coating and the basis metal. With sheet specimens, this condition is most readily achieved by leaving cut edges exposed to the test environment.

There may also be a need to consider the performance of precorroded test specimens. Apart from the fact that these conditions frequently arise in service, it is also important for other reasons:

- The presence of corrosion products or other surface layers may affect the access of constituents of the environment to the underlying metal surface where the corrosion process occurs.
- In the case of alloys, some precorrosion may lead to compositional changes in the surface.

(Additionally, in the case of inhibitor tests, preexisting corrosion product often absorbs a significant proportion of the inhibitor, leading to a lower inhibition efficiency than for a clean surface.) These factors should be taken into account in the application of any test method.

2.34.2.2 Marking Specimens for Identification

The simplest way to identify a specimen is to mark it with letters or numbers applied by stamping with a stencil or number punch. (Both of these methods have potential problems – in the case of stenciled identification the paint used may act as a crevice-former, while stamping leads to an area of higher plastic strain, which may modify its corrosion behavior.) There is, of course, always the danger that the identification marks will be obliterated by corrosion. To guard against this, the several specimens in a test should be identified further by a record of their positions relative to each other or to their supporting device. Before specimens are taken from test, their identity should be established in this manner unless inspection has already shown that the identification marks have been preserved.

Other means of identification can be used on specimens exposed to atmospheric corrosion. For example, where stamped letters cannot be expected to persist, identification may be provided by holes drilled in particular positions or by notching the edges of specimens in particular places, both in accordance with a template. Where severe corrosion is encountered, the identification by drilled holes is more permanent than that achieved by notching edges.

Other means of identification sometimes used satisfactorily involve chemical etching of the surface (not to be generally recommended), or the formation of letters or numbers by means of a vibrating stylus. The former is advantageous in studies of stress-corrosion cracking, in which stamped symbols could lead to regions of stress concentration.

2.34.2.3 Number of Replicate Specimens

Practical considerations usually limit the number of replicate specimens of each kind that can be exposed for each period of test. At least two are recommended for obvious reasons, and if a larger number can be accommodated in the program more valuable results can be secured – especially when it is desired to establish the reality of small differences in performance. For statistical analysis, five replicates are desirable. Accounts of statistical planning and analysis are given by F. H. Haynie in Ailor[2] and in ASTM G 16.

In providing replicates for tests to be subjected to statistical analysis, it is necessary, in the original sampling of the materials to be tested, to ensure that normal variations in those qualities of the metals that might affect the results are represented in each set of samples.

In order to secure information as to changes in corrosion rates with time, as in atmospheric exposure tests, it is necessary to expose sufficient specimens to allow sets to be taken from test after at least three time intervals.

For preliminary tests where the number of test specimens that can be accommodated is limited, yet numerous materials are of possible interest, it is in order to expose single specimens. This may be more advantageous than limiting the compositions that can be investigated by exposing half the number of materials in duplicate. Probably the greatest advantage in exposing two specimens of a material instead of only one is in detecting gross errors, as in weighing, etc. rather than in any considerable improvement in the precision of the observations that may be made as to the relative behaviors of the metals tested.

2.34.2.4 Test of Fusion Welds

In view of the widespread use of welded joints in equipment and structures exposed to corrosion, it is necessary to know whether such welded joints will demonstrate satisfactory resistance to attack. It is not necessary to include welded specimens of all materials in a preliminary study to discover which of them have satisfactory resistance to a particular environment. Weld tests can be postponed until the preliminary selection has been made, or alternatively, those materials expected in advance to be most likely to be resistant can be exposed in the welded condition to expedite the final answer.

There are several reasons for testing welded specimens. The first is to discover whether the weld itself will resist corrosion satisfactorily. A second purpose is to discover whether the heat effects associated with welding operations have been in any way detrimental to the corrosion resistance of the parent metal near the

weld as in the case of the so-called 'weld decay' of stainless steels. Since weld deposits may themselves be subject to a weld thermal cycle, it is necessary to include cross welds in the design of welded specimens for such corrosion tests. Further, a weld will generally constitute a stress concentration, and unless postweld heat treated, will contain residual contraction stress. Thus, testing may be necessary to generate appropriate corrosion fatigue or stress-corrosion cracking data. The latter may follow ASTM standard G 58, 'Standard Practice for the Preparation of Stress-Corrosion Test Specimens for Weldments,' for example.

A weld bead included in a test-piece is, to some extent, peculiar to itself and may not necessarily be representative of nominally similar welds to be made by other welders under other circumstances. To this extent, results of tests on welds must be subject to some qualification in interpretation, having in mind that what will be disclosed principally will be the overall ability of the composition of the weld metal to resist the corrosive environment. In some cases, entrapped flux, craters, fissures, folds, surface oxides, etc. may introduce localized corrosion that may or may not occur with all welds of the type studied.

The heat effects of welding are to an even greater extent peculiar to the particular test specimens used. They will be influenced by the welding process, by the skill of the welder, by the thickness of the metal welded, by the type of joint made, and by the geometry and mass of the surrounding structure insofar as they affect heating and cooling rates and areas over which these effects apply. Consequently, what happens to a particular welded test-piece has a questionable general significance, especially when the result shows no apparent damage to a material known to be susceptible to welding heat effects in corrosive environments. It should not be assumed that high heat input during welding will represent the worst case. For example, with ferritic steels that are sensitive to hydrogen embrittlement stemming from environmental action, low welding heat input can be most detrimental because of the formation of hardened structures in the weld area. Moreover, the possible effects of multipass welding with attendant repeated thermal cycles must be recognized in the design of a suitable test-piece.

With some materials, there are specific heat treatments that are known to reproduce the worst effects of the heat of welding. It is recommended, therefore, that in tests made to qualify a material for a particular service environment, in addition to the exposure of welded test specimens in order to observe effects of welding heat, specimens that have been given a controlled abusive or sensitizing heat treatment should be included. As an illustration, austenitic stainless steels may be held at 650–700 °C for 0.5–1 h, followed by testing for susceptibility to intercrystalline attack as in ISO 3651-1 or-2.

If such sensitized specimens remain as free from accelerated corrosion as the welded specimens do, then it can be concluded that no detrimental effects of the heat of welding need to be anticipated in the environment covered by the test. However, if the sensitized specimens are corroded while the welded specimens are not, there will remain the possibility that, under some conditions of welding, difficulties due to the effects of the welding heat may be encountered, and appropriate action or the substitution of more reliable compositions will be required. Having in mind the effect of time in damage of this sort, it will be necessary to make a careful examination of the corroded specimens to detect the first signs of attack before it can be concluded that none has occurred. In assessing the significance of attack observed on drastically sensitized specimens, it is necessary to keep in mind that no similar sensitization may result from good welding practice. Likewise, it should not be concluded that attack in a specific test environment will occur to a similar extent or will not occur at all, in some quite different environment.

The evaluation of heat treatments or the effectiveness of stabilization by limiting carbon content of these stainless steels can be determined by subjecting specimens to the ASTM standardized acid copper sulfate test or boiling nitric acid test (ASTM A 262).

2.34.2.5 Duration of Exposure

The duration of a particular test is likely to be determined by practical factors such as the need for some information within a particular limit of time, or the nature of the operation or process with which the test is concerned. Tests are rarely run too long; however, this can happen, particularly in laboratory tests where the nature of the corrosive environment may be changed drastically by the exhaustion of some important constituent initially present in small concentration, or by the accumulation of reaction products that may either stifle or accelerate further attack. In either case, the corrosivity of the environment may be altered considerably. Gross errors may result from the assumption that the results apply to the original conditions of the test rather than to

some uncertain and continually changing conditions that may exist during the course of too extended a test period.

Rates of corrosion rarely remain constant with time. More often than not, rates of attack tend to diminish as a result of the formation of adherent insoluble corrosion products or other protective films originating in the environment (**Figure 1**). Therefore, extrapolation of results of tests that are too short is more likely to indicate a lower resistance to attack than will actually be observed over a prolonged period of exposure. To this extent, such extrapolation may be considered as conservative. At the worst, it may lead to the use of a more resistant material or a heavier section than is actually needed, or to the exclusion from consideration of some materials that might be much better than the short-time test results would indicate.

Tests should be of sufficiently long duration to permit demonstration of the possible protective nature of films. Lengthy tests would not normally be required for materials that experience severe corrosion, although there are cases where this is not so. For example, lead exposed to sulfuric acid corrodes at an extremely high rate initially, while building up a protective film, then the rate decreases considerably and further corrosion is negligible. Short tests on such materials would indicate a high corrosion rate and would be completely misleading. Short-term tests can also give misleading results on alloys such as stainless steels that form passive films. With borderline conditions, a prolonged test may be needed to permit breakdown of the passive film and the subsequent more rapid attack. Consequently, tests run for long periods are considerably more realistic than those conducted for short periods.

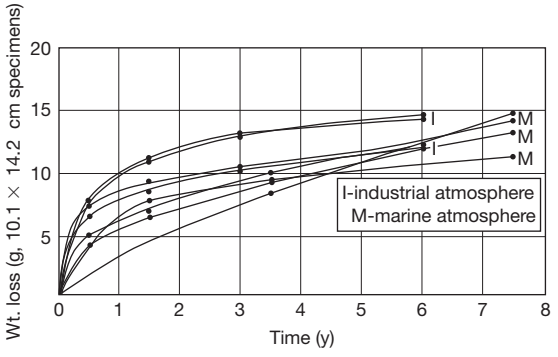


Figure 1 Corrosion rate vs. time curve showing diminishing rate of attack. Reproduced from *Proc. ASTM* **1951**, *51*, 500.

Where anticipated corrosion rates are moderate or low, ASTM G 31 suggests that the following equation be used to estimate a suitable test duration:

$$\text{Test duration(h)} = 50/(\text{Corrosion rate, mm year}^{-1})$$

For example, where the corrosion rate is 0.25 mm year^{-1}, the test should run for at least 200 h.

Due to the relatively slow rate of the atmospheric corrosion process, it is recommended in ISO 8565 that test exposures be on a schedule such as 1, 2, 5, 10, and 20 years, depending on the corrosion resistance of the metal or coating being tested. In some cases, a total exposure of less than 2 years may be suitable. It should be noted that, especially for short-term testing, the results may depend on the season of initiation of exposure. Therefore, it is recommended that exposures are commenced in the period of highest corrosivity (usually autumn).

Tests in waters and soils should ordinarily be allowed to run for extended periods in excess of 3 years, with removals of specimens in groups after different time intervals. A desirable schedule for any extended test in a natural environment is one in which the interval between successive removals is doubled each time. For example, the first removal would be after 1 year, the second after 3 years, and the third after 7 years, and so on. On the other hand, test periods should not be significantly longer than the process or exposure time of the end-use requirement. The testing of inhibitors for use in pickling or cleaning treatments should be of a period commensurate with the practical requirement, which may be for only a few minutes.

In any event, the actual duration of a test must be reported along with the results, so that those who may wish to make predictions based on them will have an accurate idea of the extent to which they may undertake any extrapolation or interpolation.

2.34.2.6 Heat Treatment

Many alloys are subject to drastic changes in their response to the effects of corrosive media when they have undergone certain heat treatments. The principal effect of interest is a loss of corrosion resistance to some degree. This commonly takes the form of concentration of corrosion in particular regions or along certain paths – as in the vicinity of grain boundaries – where phases formed by heat treatment are most likely to be concentrated. In other instances, and particularly in castings, homogenizing heat treatments may improve corrosion resistance by

eliminating 'coring' or major differences in composition from point to point in the original dendritic cast structure. Heat treatments that eliminate internal stresses are obviously helpful in connection with stress corrosion, but may induce structural changes that can affect corrosion in other forms.

Heat treatments involving heating to a temperature high enough to take harmful phases into solution, followed by cooling (e.g., by quenching) at a rate high enough to hold such phases in solution, may also be helpful in improving resistance to corrosion by avoiding attack that would otherwise be associated with a precipitated phase or compound.

Obviously, some knowledge of the possible effects of such heat treatments is essential for a complete understanding of the corrosion behavior of an alloy. Studies along this line should follow the initial selection of a material considered to be possibly useful for a particular service. Thus, it should be tested in the condition most likely to resist corrosion. Sometimes the obtaining of this condition may require annealing at a temperature sufficiently high to take any possibly harmful phases or compounds into solution followed by quenching to prevent them from precipitating. Following this preliminary selection, it would be prudent to carry out additional corrosion tests on specimens that have been deliberately subjected to any possibly detrimental heat treatments to which the material may be subjected during processing, fabrication, or use.

Heat treatment may also affect the extent and distribution of internal stresses. These may be eliminated by appropriate annealing treatments that can remove susceptibility to stress-corrosion cracking. This must be explored in any study of the performance of materials in environments where stress-corrosion cracking is a hazard. In particular cases, stress-relief annealing treatments may result in the appearance of new phases that, while eliminating the stress-corrosion effects, will induce another type of path of attack. This possibility must be kept in mind in assessing the overall benefits of heat treatments applied primarily for stress relief.

In other instances, heat treatments involving quenching, tempering, or holding at some temperature to precipitate an age-hardening compound are employed to secure some desired level of hardness or other mechanical properties. It is obviously necessary to explore what effects such heat treatments may have on the corrosion resistance of the material in the condition, or conditions, of heat treatment in which it is to be used.

2.34.2.7 Stress Effects

Techniques for studying the effects of stress on corrosion are covered in some detail elsewhere in this work. So far as attention to stress effects in a general materials-selection program is concerned, it is suggested that this should be a supplement to the initial selection of processing materials by exposing specimens to what approaches their best condition to resist corrosion, that is, free from stresses. Materials found to be worthy of further consideration in this way can be subjected to tests for stress effects. Where it is desired to discover whether severe internal stresses can be satisfactorily accepted, it will suffice to expose specimens in such a condition of stress. For example, a crucial test can be made by using a specimen in the form of a heavily cold-drawn tube in the as-drawn condition flattened on one end to introduce some additional multiaxial stresses. If such a severely cold-worked specimen suffers no stress-corrosion cracking in a test, then the danger of this occurring on any structure of that metal in the environment represented by the test is extremely remote.

2.34.2.8 Appraisal of Damage

There are many ways of determining the extent or progress of corrosion. The choice may be determined either by convenience or on the basis of some special interest in a particular result of corrosion or in a particular stage of a corrosion process.

Probably the most frequently made observation is the change in mass of a test-piece. This may take the form of a mass gain or a mass loss.

Mass-gain determinations are most common in studies of the extent and rate of oxidation or scaling at elevated temperatures. Very precise studies of this sort can be made by continuous observation of mass changes, as in the use of microbalances, such as used and described by Gulbransen.[7] Such data have quantitative significance only when the exact composition (metal content) of the scale is known or can be determined and when there has been no loss of loose scale during or after the test. Fundamental studies of the initial stages of corrosion when films of a few monolayers are formed have made use of an ellipsometer to follow the increase of thickness of corrosion products without disturbing the specimen.[8]

In most other cases, data on gains in mass due to the accumulation of corrosion products have little quantitative significance, since there is usually a question as to how much of the corroded metal is

represented in the corrosion products that remain attached to the specimen at a particular time. There are also uncertainties as to the chemical composition of corrosion products, which may consist of mixtures of several compounds with varying amounts of combined or uncombined water, depending on the humidity of the atmosphere at the time.

For these reasons, it is much better to determine the amount of metal removed by corrosion by weighing what is left after the removal of all adherent corrosion products by some method that will not cause further attack in the process, or by making proper correction for losses in the cleaning process. (Removal of corrosion products is dealt with in detail in Appendix A.) Subtracting this final mass from the original mass will give the loss in mass during the test. Since the extent of this loss in mass will be influenced by the area exposed, as well as by the duration of exposure, it is desirable, to facilitate comparisons between different tests and different specimens, to report the loss in mass in a unit that includes both area and time. A most commonly used unit of this sort is milligrams weight loss per square decimeter of exposed surface per day (24 h) (mdd).

The unit gram per square meter per annum (gma) is sometimes used in atmospheric corrosion tests (see ISO 9226) where 'a' represents 'year.' It must be recognized that these units embody two assumptions that may not, in fact, be true. The first is that corrosion has occurred at a constant rate throughout the test period. This is rarely the case, since most rates of attack tend to diminish with time, but if the duration of the test and the actual loss in mass are also reported, the user of the data can take this into account. The second probable error in a mass loss/unit area unit is that it implies that corrosion has proceeded uniformly over the whole surface. These units, therefore, will give the wrong impression as far as the probable depth of attack is concerned, if corrosion has occurred at only a few spots on the surface of the specimen. Obviously, the mdd and gma units have limited significance when corrosion has taken the form of scattered pits or has been confined to the crevices where the specimen was supported. This should be covered by appended notes describing the nature and location of the corrosion represented and should be supplemented by data on the actual depths of the pitting or crevice attack. Here, again, the report should include data on the actual mass losses and duration of exposure.

Expression of mass loss in terms of a percentage of the original mass of a test-piece is usually meaningless except for comparing specimens of the same size and shape, since it does not take into account the important relationship between surface and mass.

As indicated, it is necessary to measure and report the depths of any pitting or other localized corrosion, such as in crevices, that may have occurred. It is also useful to provide information on the frequency of occurrence, distribution, and shape of pits, since these features are likely to have practical significance. Champion[1] has produced charts in which the number of pits/unit area, the size of pits, the depth of pitting, cracking, and general attack can each be rated by the numbers 1–7. Where the number of pits is very large, it is obviously impracticable to measure the depths of all of them. Consequently, the practice of choosing 10 of the deepest pits and reporting their average depth and that of the deepest of them has developed. All surfaces of the specimen should be examined in selecting the 10 deepest pits.

There are several ways of measuring pit depths, but in all cases these measurements are facilitated if corrosion products are first removed (see Appendix A). If the pits are large enough, their depths may be measured directly with a pointed micrometer or with an indicating needle-point depth gauge. Otherwise, they may be measured optically with a microscope by focusing in turn on the surface of the specimen and, on the bottom of the pit using a calibrated wheel on the fine-focus adjustment rack for this focusing operation. In some instances, the small dimensions or shapes of pits may require metallographic examination of a cross-section for a precise measurement of depth. Such metallographic examination may also be useful in detecting an association of pitting with a structural feature of the metal.

Since it is often difficult to visualize the extent of attack in terms of depth from such mass-loss units as mdd, it is common practice to convert these mdd figures into others to indicate depth of penetration, for example, millimeter per year or mils per year (mpy), where 1 mil = 0.001 in. Such calculations suffer from the same defects as the mdd figures in that they take into account neither changes in corrosion rates with time nor nonuniform distribution of corrosion. However, since such conversions are often made, it is desirable for the initial reporter of the test results to make the calculations accurately and to report corrosion rates in both mdd and mm $year^{-1}$ or similar units.

The basic formula for making such calculation is:

$$\text{mdd} \times 0.0365/\rho = \text{mm year}^{-1}$$

where ρ is the density of the metal ($g\,cm^{-3}$). Some values showing the relationship between mdd, and mpy and mm $year^{-1}$ are given in Table 1.

Losses in mass will also not disclose the extent of deterioration that may result from the distribution of a very small amount of attack concentrated along grain boundaries or in transgranular paths (as in some cases of stress-corrosion cracking). In such instances, an apparently trivial or even undetectable loss in mass may be associated with a practically complete loss of the strength or ductility of the corroded metal. Where this may be suspected, or in any doubtful cases, the mass-loss determinations must be supplemented by other means of detecting this sort of damage, including simple bend tests followed by visual or metallographic examination to disclose surface cracking, quantitative tension tests, and direct metallographic examination of cross-sections. Changes in electrical resistance have been used as a measure of intergranular attack.[9] Because of the nature of such resistance determinations,[10,11] they have been more useful for comparing specimens of a particular kind and size than as a basis for quantitative expression of rates of attack.

Table 1 Relationship between corrosion rate in mg dm^{-2} d^{-1} (mdd) and penetration in mpy and mm y^{-1}

Material	*Density ($g\,cm^{-3}$)*	*Penetration equivalent to a corrosion rate of 1 mdd*	
		mpy	*mm $year^{-1}$ $\times 10^2$*
Aluminum 2S	2.72	0.528	1.342
Ambrac (Cu-6.5Si)	8.86	0.162	0.412
Brass (admiralty)	8.54	0.168	0.427
Brass (red)	8.75	0.164	0.416
Brass (yellow)	8.47	0.170	0.432
Bronze, phosphor (5% Sn)	8.86	0.162	0.412
Bronze (silicon)	8.54	0.168	0.427
Bronze, cast (85–5–5–5)	8.70	0.165	0.419
Cast iron	7.20	0.200	0.508
Copper	8.92	0.161	0.409
Cu-30Ni	8.95	0.161	0.409
Hastelloy *A*	8.80	0.163	0.414
Hastelloy *B*	9.24	0.155	0.394
Hastelloy *C*	8.94	0.161	0.409
Inconel 600	8.42	0.171	0.434
Iron–silicon alloy	7.0	0.205	0.521
Lead (chemical)	11.35	0.127	0.323
Monel	8.84	0.163	0.414
Nickel	8.89	0.162	0.412
Nickel silver (18% Ni)	8.75	0.164	0.417
Ni-resist	7.48	0.192	0.488
Silver	10.50	0.137	0.348
Stainless steel Type 304	7.92	0.181	0.462
Stainless steel Type 430	7.61	0.189	0.480
Steel (mild)	7.86	0.183	0.465
Tin	7.29	0.197	0.500
Zinc	7.15	0.201	0.510

The characteristic mode of corrosion of some alloys may be the formation, as a corrosion product, of a redeposited layer of one of the alloy constituents, as in the case of the brasses that dezincify, or of a residue of one of the components, as in the case of the graphitic corrosion of cast iron. Particularly in the case of the dezincified brass, the adherent copper is not likely to be removed with the other corrosion products, and therefore, the mass-loss determination will not disclose the total amount of brass that has been corroded. This is especially important because the copper layer has very little strength and ductility and the extent of weakening of the alloy will not be indicated by the mass loss. In these cases, also, the mass-loss determinations must be supplemented by, or replaced by, mechanical tests or metallographic examination, or both, to reveal the true extent of damage by corrosion. Difficulties in obtaining accurate mass losses of heavily graphitized specimens have been reported.[12]

Whenever changes in mechanical properties, such as performance in tension tests, fatigue tests, and impact tests, are to be used as a measure of corrosion damage, it is obviously necessary to provide test data on the relevant properties of the uncorroded metal. When tests extend over long periods during which the alloys being tested may be subject to changes in mechanical properties due to ageing effects, entirely aside from corrosion, it will be necessary to provide sets of specimens that may be subjected to similar ageing in a noncorrosive environment so that by direct comparison with corroded specimens of the same age the changes due to corrosion can be separated from those due to ageing. Preferably the control specimens should be stored so that they will be subjected to the same thermal experience as the specimens undergoing corrosion. This is usually very difficult to accomplish while maintaining the control specimens completely protected from corrosion.

In calculating the strength properties of the corroded specimens and comparing them with those of the uncorroded control specimens after appropriate

mechanical tests, it will be necessary to take into account the actual area of the cross-section of the corroded metal and report results on this basis instead of, or as well as, on the basis of the original cross-section before exposure, such as would be represented by the uncorroded control specimens.

In view of possible or probable variations in mechanical properties among different specimens of the same metal cut from different sheets or other pieces, or even from different sections of the same sheet or piece, it is necessary to pay careful attention to the initial sampling of stock to be used for control, as well as exposure, specimens. An interesting case in which several of these considerations were involved was provided by the long-time atmospheric exposure tests of nonferrous metals carried out by Subcommittee VI of ASTM Committee B-3 on Corrosion of Non-Ferrous Metals and Alloys[13] in which changes in tensile properties were used as one of the means of measuring the extent of corrosion.

Tests carried out for particular purposes may make use of other special means to measure the progress of corrosion. For example, changes in the reflectivity of polished surfaces[14,15] have been used as a sensitive means of following changes in the very early stages of corrosion in laboratory studies. A similar technique has been applied on a practical scale in connection with the direct evaluation of the relative merits of different alloys as used for mirrors in searchlights exposed to corrosive natural atmospheres.

Kruger,[16] at the then US National Bureau of Standards, used an ellipsometer to follow the growth of very thin corrosion-product films (oxides) during the initial stages of corrosion. This requires a knowledge of the composition of the oxide and its refractive index.

In some cases, the principal interest is in the possibility of undesired contamination or other alteration of an environment rather than in the rate of destruction of the metals being tested. Here, in addition to paying attention to the usual factors that influence rates of corrosion, it is also necessary to consider the ratio of the area of the test specimen to the volume or mass of test solution, and the time of contact. All of these factors may be quite different in a test from what would be obtained in a practical case, and any distortions of the test in these ways must be taken into account in planning the test and in interpreting the results.

In cases such as this, the possible contamination of the solution by corrosion products may be estimated from the loss in mass of the test specimen. This, however, does not make any distinction between soluble and insoluble corrosion products, which may have different effects and which can be studied best by chemical analysis of the test solution and the materials filtered from it. Similarly, chemical analysis may be required to detect any other changes in the composition of the test solution that may be of interest.

Particularly in theoretical studies of corrosion processes, it has been useful to measure the progress of corrosion in terms of the rate or extent of consumption of oxygen in the corrosion reactions. This technique has been very useful in following the progress of wet corrosion or of oxidation in its initial stages.[17]

Somewhat along the same lines is the measurement of the volume of hydrogen generated as corrosion proceeds.[18,19] This technique has been used not only in theoretical studies, but also as a means of comparing some corrosion-resisting characteristics of different lots of steel that seem to affect their behavior when used as a base metal for tin cans.[20–22]

The polarograph has been found to be a very useful tool for following the progress of corrosion, especially in its early stages, by measuring minute changes in the composition of the solution, as in the consumption of some constituent, such as oxygen, or by the accumulation of metal salts or other reaction products, such as hydrogen peroxide.[23]

An electrical resistance method that directly measures loss of metal from a probe installed in the corrosive system under study is described as follows. It is reported that corrosion equivalent to a thickness loss of as little as 2.5×10^{-7} cm can be detected.[24,25] This technique is most useful as a means of monitoring steps taken to reduce corrosion, for example, by inhibitors, or to detect changes in the corrosivity of process streams. Electrical methods of determining corrosion rates are considered subsequently.

Temperature effects may also be used in test methods, notably for assessing the effects of inhibitors in acid solutions. The technique is based on the one first proposed by Mylius,[26] which records the temperature–time behavior associated with the exothermic reaction resulting from the initial contact of a metal with a corrosive acid solution. The effectiveness of inhibitors may then be determined from their effects on the temperature–time behavior.[27]

2.34.2.9 Removal of Corrosion Products

An ideal method for removing corrosion products would be one that would remove them completely

without causing any further corrosion or other deterioration of a test specimen in the process. Procedures that achieve this ideal or approach it very closely have been developed for many of the common alloys.

There are numerous satisfactory methods of cleaning corroded specimens, but whatever the method, its effect in removing base metal should be determined for each material[28–30] taking into account possible differences between the behavior of 'as-new' and corroded base metal (see Appendix A). The various methods may be classified as follows:

1. Mechanical treatment:
   a. scrubbing with bristle brush,
   b. scraping,
   c. wire brushing,
   d. grit, shot, sand blasting;
2. Chemical treatments:
   a. organic solvents,
   b. chemical reagents;
3. Electrolytic treatments as cathode in solutions of the following:
   a. sulfuric acid, usually inhibited,
   b. citric acid,
   c. potassium cyanide,
   d. sodium hydroxide.

Further details of removing corrosion products are given in Appendix A.

2.34.3 Laboratory Corrosion Tests

2.34.3.1 Total-Immersion Tests

The total-immersion corrosion test is most adaptable to rigorous control of the important factors that influence results. This control may be achieved in different ways and it is unnecessary and undesirable to seek a standardized method or apparatus for universal use. All that is required is a recognition of what is essential, as covered, for example, by the ASTM procedure G 31. This represents a code of minimum requirements without insisting on the use of any particular kind of apparatus or specifying the exact conditions of aeration, temperature or velocity to be used. Since different metals respond differently to effects of aeration, temperature and velocity, the setting up of standard test conditions in terms of these factors would be inappropriate. Depending on the environment, such standardized testing conditions would favor maximum corrosion of some materials and minimum corrosion of others and thus lead to gross errors in indicating any general order of merit applicable under conditions differing from those of a standardized test.

In some instances, it may be possible, although it is usually very difficult, to undertake laboratory corrosion tests under conditions that will be the same as those encountered in some practical application and thus to secure some directly applicable data. More often, the conditions of service are so variable or so difficult to appraise accurately and duplicate in the laboratory that it is impractical and probably unwise to attempt to do so. A better procedure is to examine the individual effects of the several controlling factors by varying them one at a time so as to provide a picture of their influence on the behavior of the materials of interest in the corrosive medium being investigated. This information will be helpful in deciding whether the conditions of a particular use are favorable or unfavorable to the materials being considered. It will also serve as a guide to account for behavior in service and to suggest changes in the operating conditions that may be expected to reduce corrosion of a material being used.

In many cases, and particularly in aqueous solution, the most important controlling factors will be solution composition, temperature, aeration, and velocity.

2.34.3.1.1 Solution composition

When designing tests to determine the effects of solution composition on corrosion, it is important to understand the nature of the controlling process. In the case of many metals and alloys, the rate-determining step will be the supply of cathodic reactant to the metal surface. This is particularly true in neutral solutions where corrosion will often be under oxygen diffusion control. Thus, tests in stagnant (unstirred, quiescent) conditions may be inappropriate, since the effects of solution composition will be insignificant compared with the oxygen diffusion effect. In stagnant conditions, corrosion rates of mild steel in, for example, sodium chloride, sodium sulfate and other salt solutions will be effectively the same over a range of concentrations. The effects of anion type and concentration begin to be shown only with movement of the solution, that is, when oxygen access to the metal surface is facilitated to the point where it may be no longer rate controlling. Specific effects of anions in stagnant solutions will, however, be found when the anion has oxidizing properties, as in the case of nitrate and in systems where the dominant cathodic reaction is hydrogen evolution.

Care must therefore be taken in designing tests to study the effects of solution composition since

different results will be obtained depending on the degree of aeration and/or movement of the solution.

Variations in solution composition throughout a test should be monitored and, if appropriate, corrected. Variations may occur as a result of reactions of one or more of the constituents of the solution with the test specimen, the atmosphere or the test vessel. Thus, it is important that the composition of the testing solution is what it is supposed to be. Carefully made-up solutions of pure chemicals may not act in the same way as nominally similar solutions encountered in practice, which may, and usually do, contain other compounds or impurities that may have major effects on corrosion. This applies particularly to 'artificial' seawater, which is usually less corrosive than natural seawater. This subject is discussed in detail in a Special Technical Publication of ASTM,[31] and tests with natural, transported, and artificial seawater have been described.[32] Suspected impurities may be added to the pure solutions in appropriate concentrations or, better still, the testing solutions may be taken directly from plant processes whenever this is practical.

It should also be pointed out that in exploring the effects of the concentration of a particular acid or other chemical on its corrosivity, it is necessary to cover the full possible variation of concentrations thoroughly, since it frequently happens that particular ranges of concentration are especially corrosive to some metals. This extends to the highest degrees of concentration where sometimes the complete elimination of water may increase corrosion a great deal – as in the case of aluminum in acetic acid. On the other hand, the presence of a trace of water may make other chemicals much more corrosive – as in the case of bromine and other halogens.

It should be noted, also, that exposing a specimen to a solution of some chemical while it is being concentrated by evaporation practically to dryness will not suffice to explore the effects of the complete range of concentration, simply because the period in which any particular concentration range exists is not likely to be long enough to permit any especially corrosive effects to be detected in the overall result.

2.34.3.1.2 Temperature control

Of the factors mentioned, temperature is probably the easiest to control; this can be accomplished by means of a thermostat or by operating at the boiling point of the testing solution with an appropriate reflux condenser to maintain the solution at a constant concentration. Control to $\pm 1\,^{\circ}C$ is not hard to accomplish.

The need for temperature cycling should be taken into account when designing or conducting tests. The nature of the test vessel should be considered for tests in aqueous solutions at temperatures above about $60\,^{\circ}C$ since soluble constituents of the test vessel material can inhibit or accelerate the corrosion process. An inhibiting effect of soluble species from glass, notably silica, on the behavior of steel in hot water has been shown.[33] Pure quartz or polymeric materials are often more appropriate for test vessel construction. (It should also be recognized that the control of gas concentrations becomes more difficult as temperature approaches the boiling point of the solution.)

2.34.3.1.3 Aeration

Control of aeration is more difficult. Aeration here means the amount of oxygen supplied either as such or, more commonly, in air. In some situations, it may not require a large amount of air bubbled through a solution to accommodate even a modest rate of corrosion of a small test-piece. **Figure 2** shows the relationship between the rate of supply of air used for aeration and the rate of corrosion of Monel alloy in 5% sulfuric acid.

To facilitate rapid solution of oxygen from air bubbles it is desirable to make these as small as possible, for example, by having the air enter through a porous thimble or sintered glass disc. Much less satisfactory results are secured by simply letting air escape into the solution from a tube drawn to a fine tip.

It is also undesirable to permit air bubbles to impinge directly on the test-pieces. This can be avoided by placing the aerator inside a chimney.

When it is desired to study effects of various degrees of aeration, it is better to do this by varying the oxygen content of the saturating gas (e.g., by using controlled mixtures of oxygen and nitrogen)

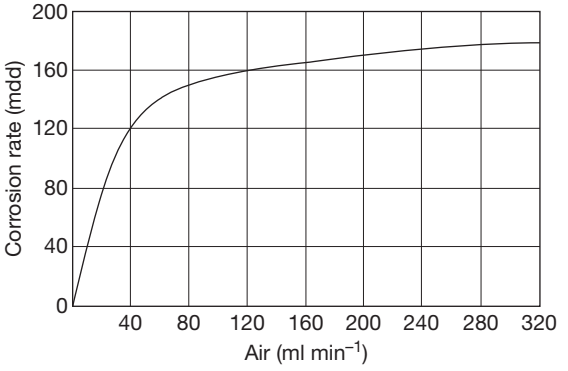


Figure 2 Effect of rate of supply of air used for aeration on corrosion of Monel alloy in 5% sulfuric acid.

introduced at a constant and adequate rate than by attempting to vary the rate of admission of a gas (such as air) of constant composition. This extends as well to zero aeration, which can best be accomplished by saturating the test solution with deoxygenated nitrogen or other inert gas. It is unwise to assume that, because no air is purposely added, oxygen has been excluded from a test solution in a vessel open to air. Such a practice provides a low oxygen availability that is not sufficiently under control to ensure reproducible results.[3]

2.34.3.1.4 Velocity

Controlled flow conditions can be achieved either by having the test-piece move through a notionally stationary liquid or by having a moving liquid come into contact with a stationary test-piece. Occasionally tests may involve both types of exposure. Details of test procedures are presented elsewhere.

The achievement of zero velocity in a test set-up is about as difficult as the accurate control of some high velocity. It is a common mistake to assume that by not making any attempt to move either the specimen or the testing liquid, the relative velocity between them will be zero. This neglects such effects as convection currents and the agitation due to the effects of corrosion products streaming under the influence of gravity. The most common difficulty arising from this situation is that these uncontrolled effects in tests made under presumably quiet or stagnant conditions make it very difficult to secure reproducible results from test to test. Therefore, even when there is no practical interest in the effects of any appreciable velocity, it is desirable to provide for some controlled movement of either the specimens or the solution at some velocity such as 7.5 cm s^{-1}, readily achieved with a vertical circular-path machine.

Equipment of this type in which the specimens are moved in a vertical circular path with all portions of the surface of a specimen moving at the same speed has been used in instances where such moderate test velocities are required. Statistical analysis of data from tests with an apparatus by Wesley[4] has demonstrated satisfactory reproducibility of results not only among specimens in a particular test, but also from test to test undertaken at different times.

Where effects of much higher velocities are to be studied, various devices have been used to move test-pieces through the testing solution at high velocity.

One procedure is to use test specimens in the form of discs, which can be rotated at the desired speed while either wholly or partly immersed in the testing solution, and Freeman and Tracy described a device of this sort in a contribution to the ASTM Symposium on Corrosion Testing Procedures.[28] With their apparatus the specimen discs were mounted on horizontal shafts and were partially immersed in the testing solution.

A similar method of test was used at the International Nickel Company's Corrosion Laboratory at North Carolina. The specimen discs are mounted on insulated vertical spindles and submerged in seawater, which is supplied continuously to the tank in which the specimens are immersed. The maximum peripheral speed of the spinning disc is about 760 cm s^{-1}, and the characteristic pattern of attack is shown in **Figure 3(a)**. Studies of variation of depth of attack with velocity indicate that at low velocities (up to about 450 cm s^{-1}) alloys such as Admiralty brass, Cu–10Ni and cupronickel alloys containing iron maintain their protective film with a consequent small and similar depth of attack for the different alloys. At higher velocities the rate increases due to breakdown of the film.

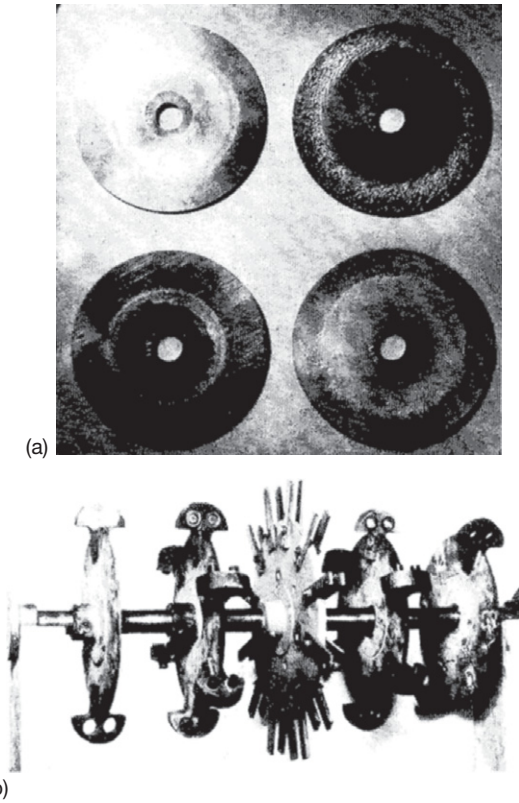


Figure 3 (a) Distribution of corrosion on surfaces of rotating disc specimens and (b) assembly of specimens attached to rotating discs.

Tests of this sort indicate a sort of critical velocity for each material that marks the boundary between the maintenance and loss of protective films. These apparently 'critical' velocities must be considered as relative and only applicable to the conditions of test in which they are measured. Because of the complex effects associated with the differences in velocity from point to point on such rotating specimens, the apparent 'critical velocity' obtained in a given test may be quite different from what might be indicated by another test in which the same velocity is achieved in some other way – as by moving the liquid past a stationary specimen at a uniform velocity from point to point. The apparent 'critical velocity' indicated by this latter method of test will likely be higher for many materials than that shown by the spinning disc test. Thus, the establishment of critical velocities by a particular method of test will afford only qualitative data regarding the relative abilities of a number of materials to resist the destructive effects of high velocity. Furthermore, the critical velocity at which severe attack commences has been found to depend on the diameter of the disc so that no quantitative significance can be attached to it. This restriction extends as well to tests with iron discs, where attack is concentrated at the centre of the disc rather than at the periphery, irrespective of its diameter. (Small variations in solution composition may also affect the value of any critical velocity. In laboratory tests using recirculating artificial seawater the presence of dissolved copper from copper alloy test-pieces has been shown to affect the critical velocity for such materials.)[34]

Somewhat similar tests may be made by attaching specimens to discs that can be rotated at some desired velocity in the testing medium. A machine of this sort that is used extensively in studying corrosion of metals by seawater at high velocity was developed by the staff of the US Naval Engineering Experiment Station at Annapolis, Maryland.[35] A typical assembly of discs and specimens is shown in **Figure 3(b)**.

The action of the rotating discs with their attached specimens causes violent agitation of the liquid in the tank. Depending on the height of liquid above the specimens, as determined by the location of the overflow pipe, there may be considerable whipping of air bubbles into the liquid or none at all, as desired. The heat of agitation causes the temperature to rise. This may be controlled readily by adjusting the amount of fresh cold liquid, for example seawater, allowed to pass into the tank and out through the overflow pipe. It is not difficult to hold the temperature within 1–2 °C of the desired value.

The use of rotating discs to carry test specimens has been extended to studies of protective coatings in what are considered to be 'accelerated' tests of such coatings for service underwater.[36]

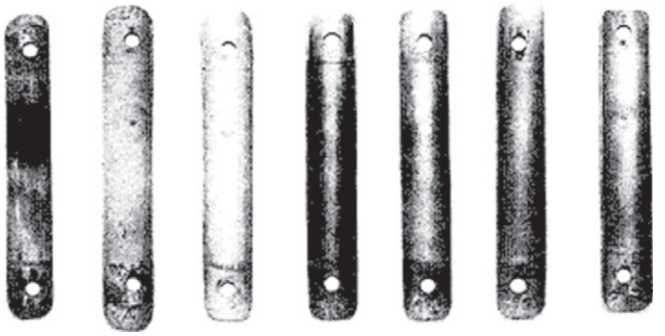
Velocity effects involving a high differential in velocity between adjacent areas are achieved simply by exposing a test specimen to the action of a submerged jet. This sort of test has been very popular and very useful in studying impingement attack or erosion of condenser tube alloys. It was introduced originally by Bengough and May[37] and later modifications were described subsequently by May and Stacpoole.[38] The appearances of typical specimens from this test are shown in **Figure 4**. In this test the dimensions of test specimens should be standardized, since the depth of attack has been found to be influenced by the extent of the immersed area of the specimen that is outside the impingement zone.

Along the same general lines is an apparatus employed by Brownsdon and Bannister[39] in which a stream of air at high velocity is directed against the surface of a submerged test specimen.

A straightforward way to study velocity effects is to force the testing liquid through tubular specimens, which may be arranged to form model piping systems for studying the peculiar corrosion that may result from severe turbulence effects downstream of valves, reducers, branch connections, elbows, and other fittings. In such systems the rates of flow can be measured by suitable orifice meters and regulated by control valves. A somewhat similar technique applied to condenser-tube alloys is to test them as installed in model tube-bundle assemblies.[40] Butler and Ison have described a laboratory test rig for studying the effects of flowing water on steel pipework.[41]

Other methods involve holding specimens in suitable fixtures so that they form the walls of channels through which the test solution can be passed at controlled rates of flow. Such devices have been used at the Harbor Island Test Station in North Carolina primarily for studying the electrode potential and polarization characteristics of metals and alloys, but they are also suitable for observing effects of velocity on corrosion. This is illustrated in **Figure 5** in which the specimen and Pt electrode are of the same size and are placed parallel to one another in the holder. When required potentials are measured by inserting a capillary through the hole in the Pt, it is removed to avoid shielding effect.

Effects of velocity are sometimes aggravated by the presence of abrasive solids in suspension, which increases deterioration by straight mechanical abrasion as well as accelerating corrosion by continually



Sample	*Alloy*	*Mass loss after 1000 h (g)*	*Penetration of impingement pit (% of cross-section)*	*Appearance*
A	Arsenical 70/30 brass	0.262	40	Deep pit opposite jet, surrounding area darkly stained
B	Arsenical admiralty brass	0.422	44	Deep pit opposite jet
C	Arsenical aluminum brass	0.010	1	Very slightly attacked opposite jet
D	Cu 10Ni	0.026	<1	No measurable attack opposite jet
E	Cu–30Ni–1Fe 1Mn	0.019	<1	No measurable attack opposite jet
F	Cu–30Ni–2Fe 2Mn	0.014	<1	Very slight pitting attack beneath brown tubercles of corrosion products. Also pitting in areas next to sample holder. No measurable attack opposite jet
G	Cu–30Ni	0.061	9	Pitted opposite jet

Materials conform to BS 2871 and ASTM B111

Figure 4 (Top) Pattern of corrosion of jet-impingement test specimens and (bottom) sample test data and results.

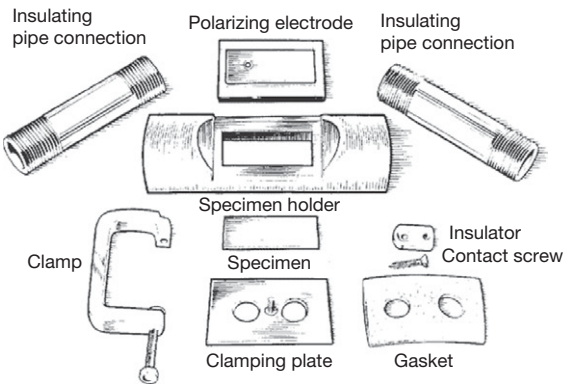


Figure 5 Components of apparatus in which specimens form walls of channel for test solution.

exposing fresh surfaces to attack. Such attack is especially serious with pumps, agitators, and piping systems. Special apparatus has been designed to measure the performance of materials under such conditions, as described, for example, by Fontana.[42]

Special devices have been used to study erosion-corrosion by boiler water moving at high velocity, and an example is the method used by Wagner *et al.*[43]

Where high rates of flow are desired with a small volume of testing liquid, a specimen may be mounted in the form of a tube inside a large glass tube and a small mass of liquid may be forced to flow through the restricted annular space between the two tubes. Such a method was used successfully in

studying corrosion by milk where the volume of milk was small, the required movement being achieved simply by the use of an air lift to return the milk to an overhead reservoir from which it flowed by gravity through the test set up.[44] Velocities as high as 0.6 m s^{-1} were studied in this way.

2.34.3.1.5 Volume of testing solution

If exhaustion of corrosive constituents that may be present in minute concentrations and the accumulation of reaction products which may either accelerate or stifle further attack are to be avoided, the volume or mass of testing solution must be sufficiently large to avoid effects caused by these factors. In laboratory tests, however, practical considerations limit the volume of testing solution that can be provided for. A minimum of 250 ml of testing solution for each 6.3 cm^2 of specimen area is suggested in NACE TM0169.

2.34.3.1.6 Support of specimens

Since crevices set up where specimens are in contact with their supports and may become the seats of accelerated corrosion by concentration cell effects, special attention should be given to this detail in setting up tests. The area screened by the supporting members should be kept to a minimum, for example, by making contact at a point or along a line rather than over any appreciable area. In some instances it may be desirable to apply some protective coating to the areas that are in contact with the supporting members. In any event, any corrosion that has occurred in the area of the supports should be taken into account in appraising and reporting the results of a test.

Somewhat along the same lines are techniques that have been employed to avoid edge effects by having the specimen come into contact only with a pool of testing solution, which does not cover its complete surface – as described, for example, by Brennert.[45] A more elaborate technique in which the pool of testing solution was circulated by thermal currents was described by Smith.[46]

2.34.3.2 Alternating-Immersion Tests

One means of ensuring aeration of a testing solution in contact with a specimen is provided by an alternating-immersion corrosion test in which the specimen is alternately immersed in a solution and withdrawn from it in some predetermined cycle. This procedure also has the effect of allowing the test solution that clings to the specimen to become concentrated by evaporation while the specimen is out of the liquid, and in addition it permits corrosion products to remain and reach greater concentrations and undergo more chemical changes in immediate contact with the metal than can occur in continuous-immersion tests. In these ways, an alternating-immersion test may simulate certain circumstances of practical corrosion better than a continuous-immersion test and may, therefore, be preferred.

Since the conditions of this test can be standardized fairly readily, it has also been used as a routine test in comparing different alloys of the same general kind in the course of studies of effects of composition on properties, as, for example, in the researches by Hanawalt *et al.* on the corrosion of magnesium.[47]

The ASTM have established a recommended procedure for alternating-immersion stress corrosion tests in 3.5% NaCl solution (ASTM G 44).

The alternating immersion may be accomplished either by moving specimens held in a suitable suspension rack into and out of containers holding the test solution, by leaving the specimens fixed and raising and lowering the solution containers around them so as to immerse them or leave them suspended above the solution, or by using a combination of pump and siphon to move the solution between two reservoirs.

To favor reproducibility of results, the cycles of immersion and withdrawal must be kept the same from test to test. It is necessary to control the temperature and humidity of the atmosphere surrounding the test setup as these affect the rate of evaporation of the solution and of drying of the specimens when they are out of the solution. It is also necessary to provide for replenishing losses of water from the test solution resulting from evaporation.

2.34.3.3 Water-Line Tests

Materials may be subject to intense localized attack at the liquid level when they are partially immersed in a solution under conditions where the water line remains at a fixed position for long periods. This attack may be the result of concentration cell effects complicated by differences in the nature and adherence of corrosion-product films as they form in the water-line region as compared with those that form above or below this region.

The testing technique is very simple since it involves no more than providing means of supporting a specimen or specimens in a fixed position of partial immersion, and of maintaining the liquid level constant by the continuous addition of distilled water to

make up for evaporation losses. For maximum reproducibility of results, the dimensions of the specimens, and especially the ratio of areas above and below the liquid level, should be held constant, as should be the depth of immersion.

2.34.3.4 Heat-Flux Effects

When heat flows into or out of a fluid through a containing wall, the wall surface reaches a temperature that differs from that of the bulk of the fluid. The wall's corrosion resistance at this temperature may be significantly different from its resistance at the bulk-fluid temperature. Tubes or tank walls heated by steam or direct flame have failed in service in which similar materials, not so heated, performed acceptably.

The name 'hot wall effect' was given to this phenomenon by Benedicks[48] who observed separation of dissolved gas from aerated water in boiler tubes. The metal wall was insulated from the cooler boiler water by the gas, its temperature rose substantially, and the more severe corrosion took the form of pitting. It was recognized subsequently that even without such gas formation, a hot-wall effect resulted when heat flowed through a wall into a fluid. The temperature difference that had to exist in order for heat to flow was increased by the insulating effect of the thin film of almost stagnant fluid at the wall surface. This film is thinned by rapid flow of a fluid through a tube, but is not eliminated at any finite velocity. Boiling of a liquid, by either bubble nucleation or coverage of the heating surface by a vapor film, increases the skin temperature further; the second mechanism provides much more severe insulation and greater temperature rise.[49]

High rates of heat flow through heat-transfer surfaces in atomic energy installations studied by Groves made him develop an appropriate corrosion test method.[50] In this, a small sheet specimen in contact with a hot liquid, usually boiling, is heated externally so that its surface reaches the desired test temperature. The surface or skin temperature is close to the temperature at mid-thickness, which is measured by a thermocouple inserted in a drilled hole. The heating source is an electric soldering iron in which the tip is replaced by a flat-surfaced metal block. The voltage to the heat source is controlled by a variable transformer. Supplementary heating may be provided to the liquid in the test vessel by winding the liquid container with resistance wire, or by an immersion heater (**Figure 6**).

A similar unit, modified in details such as location of condenser, use of an agitator, and shape of the vessel, was used by Fisher and Whitney.[51] Further substantial modifications to permit interface location of specimens, cooling of specimens and operation under applied pressure, have been described by Fisher.[52] Earlier laboratory test methods tried by Fisher and Whitney[51] included exposure of specimens heated by their own electrical resistance and of tubular specimens containing a pencil-type resistance-wire heater in a quartz tube.

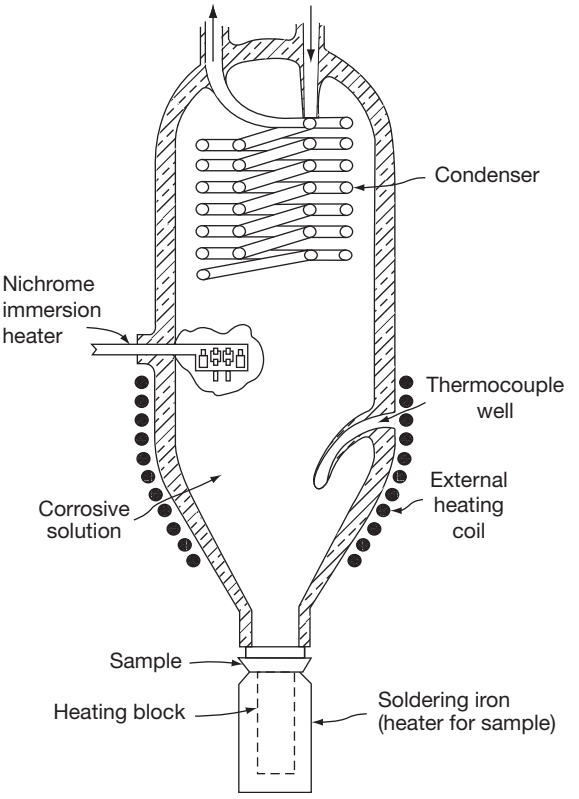


Figure 6 Device for studying hot wall effects.

To investigate corrosion in heated crevices filled with wetted paste a sandwich test assembly was designed by Gleekman and Swandby[347] simulating a slotted cylindrical steam-heated drier. Two plates are bolted together, the lower being heated by an electric hot plate.

It has been concluded from data reported in these studies that the skin temperature is the major controlling factor in corrosion, not the rate of heat flow through the metal.[51] It has also been concluded, however, that corrosion rates at a given mid-specimen temperature do depend on the presence or absence of thermal flux.[53] The difference between temperatures at skin and mid-specimen positions may account for this discrepancy.

Heat-flux corrosion rates can also be determined in plant tests using steam-heated tubular specimens which are weighed or callipered.

In addition to the direct effect of film temperature on corrosion rate, an indirect effect has been observed in the heating of some foods and chemicals, in which insulating solid corrosion films on different metals. By raising the metal surface temperature, these films may, when pervious, lead to further corrosion.

Apparatus and procedures for testing the corrosion resistance of alloys in brines at temperatures up to 120–150 °C are described by Hart.[54]

Testing procedures for corrosion inhibitors in heat flux conditions are discussed later.

2.34.3.5 Composition of Testing Solution

Changes in corrosivity with time may be observed by exposing fresh specimens to a solution that has already been used for testing. Where such changes are known to occur, or are suspected, it will be necessary to arrange for replacement of the testing solution after appropriate intervals or replenishment of constituents that may be consumed in the corrosion processes.

2.34.4 Electrochemical Measurements (General)

In view of the electrochemical nature of corrosion it is not surprising that measurements of the electrical properties of the interface metal–solution (electric double layer) are used extensively in fundamental studies of the mechanism of corrosion, in corrosion testing and in monitoring the control in service. In the context of this section, electrical measurements in the laboratory are used to assess the corrosion behavior of metals and alloys in service and to avoid the more tedious and prolonged field testing. Determinations of the corrosion rate, susceptibility of a metal to bimetallic corrosion, pitting, intergranular attack, stress-corrosion cracking, etc. are examples of corrosion phenomena that are studied in the laboratory by means of electrochemical methods in order to anticipate behavior in service.

Progress in this field has been made possible with increase in knowledge of the detailed mechanism of corrosion and by the developments that have taken place in instrumentation. The widespread use of potentiostatic control and the availability of a range of commercial potentiostats have given a tremendous impetus to electrochemical testing, and have perhaps led to the unfortunate belief that corrosion testing in the laboratory and in the field can be replaced completely by electrochemical measurements in the laboratory under conditions of controlled potential. Indeed, La Que[55] in 1969 was prompted to express concern about the proliferation of publications describing electrochemical techniques for corrosion testing and to advise caution regarding the extrapolation of results obtained in the laboratory with a potentiostat for the performance of metals in service.

Although important contributions in the use of electrical measurements in testing have been made by numerous workers, it is appropriate here to refer to the work of Stern and his coworkers[56,57] who have developed the important concept of linear polarization, which led to a rapid electrochemical method for determining corrosion rates, both in the laboratory and in plant. Pourbaix[58] and his coworkers on the basis of a purely thermodynamic approach to corrosion constructed potential–pH diagrams for the majority of metal–H_2O systems, and by means of a combined thermodynamic and kinetic approach developed a method of predicting the conditions under which a metal will (a) corrode uniformly, (b) pit, (c) passivate, or (d) remain immune. Laboratory tests for crevice corrosion and pitting, in which electrochemical measurements are used, are discussed later.

2.34.4.1 Techniques

Electrochemical methods of testing involve the determination of specific properties of the electrical double layer formed when a metal is placed in contact with a solution, and these can be summarized as follows.

1. The potential difference across the electric double layer $\Delta\psi$. This cannot be determined in absolute terms but must be defined with reference to another charged interface, that is, a reference electrode. In the case of a corroding metal, the potential is the corrosion potential which arises from the mutual polarization of the anodic and cathodic reactions constituting the overall corrosion reaction.
2. The reaction rate per unit area, expressed as the current density i. For a corroding metal the partial anodic and cathodic current densities cannot be determined directly by means of an ammeter unless the anodic and cathodic areas can be separated physically, for example, as in a bimetallic couple. If the metal is polarized a net current i_c

for cathodic polarization, and i_a for anodic polarization, will be obtained and can be measured by means of an ammeter.
3. The capacitance. The electrical double layer may be regarded as a resistance and capacitance in parallel, and measurements of the electrical impedance by the imposition of an alternating potential of known frequency can provide information on the nature of a surface. Electrochemical impedance spectroscopy is now well established as a powerful technique for investigating electrochemical and corrosion systems.

The most commonly used measurements are as follows:

1. determination of the steady-state corrosion potential E_{corr};
2. determination of the variation of E_{corr} with time;
3. determination of the $E-i$ relationships during polarization at constant current density (galvanostatic), the potential being the variable;
4. determination of the $E-i$ relationships during polarization at constant potential (potentiostatic), the current being the variable;
5. determination of the electrochemical impedance under alternating potential conditions as a function of frequency;
6. determination of electrochemical noise: fluctuations of the free potential or fluctuations of the current when a constant potential is maintained.

2.34.4.2 Instruments

The techniques and instruments used may be classified as follows:

1. *Potential measurements:* a reference electrode and a potentiometer or electrometer that require only a small current to give a measurement of e.m.f. and thus minimize polarization of the electrodes.
2. *Current measurements:* milliammeters or the measurements of the *IR* drop across a conductor of known resistance.
3. *Galvanostatic polarization:* constant direct current power units, or banks of accumulators or dry cells used in conjunction with a variable resistance.
4. *Potentiostatic polarization:* potentiostats with varying output currents.
5. *Determination of impedance:* dedicated instruments for the measurement of impedance as a function of frequency.

Impedance measurements, originally used in fundamental studies of anodic oxidation, have great power for deriving corrosion rates and information about reaction mechanisms for processes that occur at a corroding surface using readily available instrumentation. Armstrong *et al.*[59] have used impedance measurements for studying the active–passive transition of chromium and Epelboin *et al.*[60] describe its use for determining the instantaneous corrosion rate of a metal. Sathyaharayana[61] has described a method using Faradaic rectification to determine the instantaneous corrosion rate, in which no reference electrode is required; the electrodes consist of the metal under study and a counter electrode of a large area of the same metal. Macdonald[62] in a review of electrochemical impedance spectroscopy has highlighted the enormous advantage it offers over other electrochemical methods of being able to evaluate the validity of the data using the Kramers–Kronig transforms. While the use of impedance techniques is usually restricted to homogenous corrosion situations, applications to localized corrosion have been reported. These included pitting,[63–65] abrasion,[66] other types of corrosion[67–69] and stress-corrosion cracking.[70] The potentiostatic technique has a number of variations and the potential may be increased or decreased incrementally, changed continuously at a predetermined rate (potential sweep) or applied as pulses of very short duration.

The application of a stochastic approach to the study of pitting was pioneered by Sato.[71] Shibata and Takeyama[72] then developed stochastic theory to study the statistical variation of the pitting potential determined by the potential sweep method. Subsequently, Williams and his coworkers[73–75] described how electrochemical noise (current fluctuations in systems under potentiostatic control or potential fluctuations at the corrosion potential) can be analyzed to provide information on the initiation and propagation of pitting corrosion on stainless steels. In a review Gabrielli *et al.*[76] discuss published results for both stable pitting, where statistical counting is the major technique of investigation, and unstable pitting during the prepitting stage, where the techniques that can be applied include statistical counting, ensemble averages, and spectral analysis.

2.34.4.3 Electrochemical Cells

Since the single potential of a metal cannot be measured, it is necessary to use a suitable reference

electrode such as the $Hg/Hg_2Cl_2/KCl$ electrode or the Ag/AgCl/KCl electrode, and although potentials are commonly expressed with reference to the standard hydrogen electrode (SHE), the use of this electrode in practice is confined to fundamental studies rather than testing. Details of the preparation of reference electrodes, salt bridges, capillaries, etc. are given in the book by Ives and Janz[77] and elsewhere.[78–81]

Measurements of the corrosion potential of a single metal corroding uniformly do not involve an *IR* drop, but similar considerations do not apply when the metal is polarized by an external e.m.f., and under these circumstances the *IR* drop must be minimized by using a Luggin capillary placed close to the surface of the electrode. Even so, the *IR* drop is not completely eliminated by this method, and a further error is introduced by the capillary shielding the surface from the current flow with a consequent decrease in current density. At high current density this error due to the *IR* drop can conceal the Tafel region by distorting the measured overpotentials, a difficulty that can be overcome by determining the resistance of the solution at the capillary tip and by making an appropriate correction for each value of the current density. Alternatively, electronic feedback circuits may be employed for automatic compensation of the *IR* drop, and this method is attractive if rapid variations in overpotential are being studied.[82–85] Other methods include using a Piontelli capillary[86,87] or a rearside capillary[88] (**Figure 7**). However, in testing in electrolyte solutions of low resistivity these errors are normally small, and the conventional Luggin capillary is used in conjunction with a salt bridge and reference electrode.

For polarization studies, the cell must make provision for the metal electrode under study, an auxiliary or counter electrode, and a Luggin capillary. Provision must also be made for introducing a gas such as oxygen-free nitrogen or argon, which serves to remove dissolved oxygen and to prevent its introduction during the test (or to introduce it if required at predetermined partial pressures) and to agitate the solution; additional agitation if required can be obtained by means of a stirrer (electric or magnetic).

Figure 8 shows the design of an all-glass cell, which has been listed as the standard polarization cell in the ASTM Recommended Practice G 5, which makes provision for the essential requirements listed earlier; this cell is typical of those used for fundamental studies and for testing, although details of design may vary.

The metal electrode to be studied must be carefully prepared, attached to an electrical lead and mounted so that a known surface area of one face is presented to the solution. Several procedures such as mounting in a cold curing epoxy or polyester resin or inserting into a close-fitting holder of PTFE are used. In the case of metal–solution systems that have a

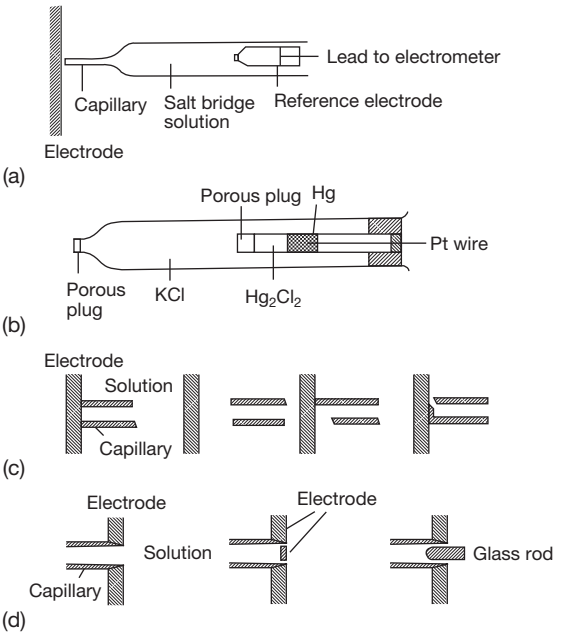


Figure 7 Reference electrodes and capillaries. (a) Reference electrode, salt bridge and Luggin capillary, (b) calomel electrode, (c) frontal types of capillaries and positions, and (d) rearside capillaries. Reproduced from von Fraunhofer J. A.; Banks, C. A. *Potentiostat and Its Applications*; Butterworths: London, 1972.

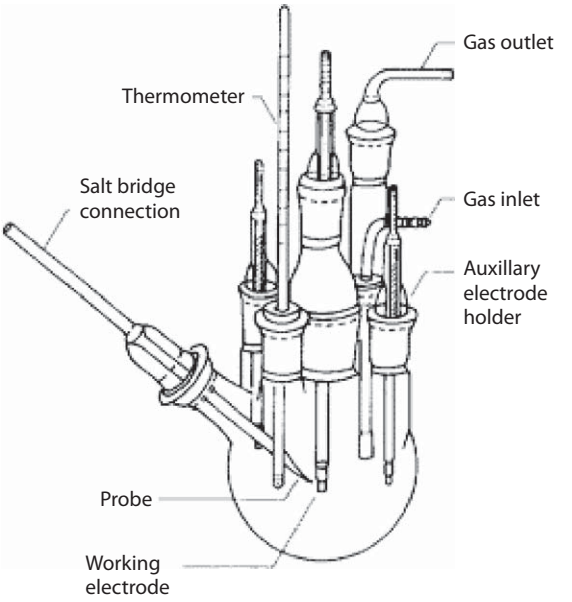


Figure 8 All-glass cell for studies of polarization of metal electrodes. Reproduced from ASTM G 5–94 (2004).

propensity for pitting, care must be taken to avoid a crevice at the interface between metal specimen and the mounting material (discussion of approaches to the solution of this problem is provided elsewhere). In view of the widespread use of the cell shown in Figure 8 suitable electrode holders, based on the use of a compression gasket of PTFE or a similar inert polymer, have been designed for bulk metal specimens,[89,90] wires,[91] tubes,[92] sheets and foil,[93] and for high temperature high-pressure assemblies.[94] Examples of methods of mounting specimens to give a defined area of surface are shown in Figure 9.[88]

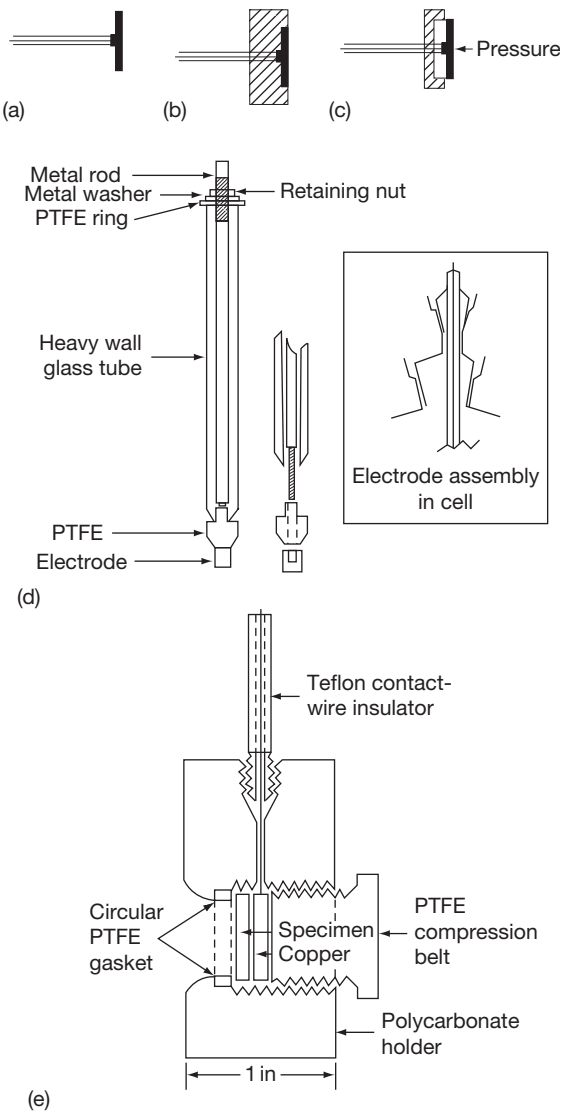


Figure 9 Methods of mounting specimens. (a) Wire soldered to metal specimen, wire being enclosed in glass tube, (b) specimen completely encapsulated in cold-setting resin that is ground down to expose one face; (c) specimen clipped into machined PTFE holder; (d) Stern Makrides pressure gasket for cylindrical specimen; and (e) pressure gasket for sheet or foil. (d) Reproduced from Stern, M.; Makrides, A. C. *J. Electrochem. Soc.* **1960**, *107*, 782. (e) Reproduced from France, W. D., Jr.; *J. Electrochem. Soc.* **1967**, *114*, 818.

Various types of reference electrodes are available of which the saturated calomel electrode (SCE), the silver–silver chloride electrode and the copper–copper sulphate electrode are the most widely used in corrosion testing and monitoring. It is appropriate, however, to point out here that the saturated calomel electrode (SCE), the silver–silver chloride electrode and the copper–copper sulfate electrode are the most widely used in corrosion testing and monitoring.

To avoid contamination of the solution under study, and to minimize the liquid-junction potential, it is usual to use a salt bridge, but in many cases this can be dispensed with; thus, if corrosion in a chloride-containing solution is being studied a Ag/AgCl electrode immersed directly in the solution could be used; similarly a Pb/PbO_2 electrode could be used for studies of corrosion in H_2SO_4.

2.34.4.4 Measurements of the Corrosion Potential

Pourbaix[58] has provided a survey of potential measurements in relation to the thermodynamics and kinetics of corrosion, and an example of how they can be used to assess the pitting propensity of copper in water.

The determination of the corrosion potential of the two metals constituting a bimetallic couple will provide information on which one of the two will be predominantly anodic and will suffer enhanced corrosion when they are coupled. Similarly the effect of microscopic heterogeneities in alloys on corrosion has been investigated by measuring the corrosion potentials of selected small areas of the surface of the metal. Smith and Pingel[95] coated the surface of the metal with a lacquer of ethyl cellulose (deposited from a volatile solvent), allowed it to dry and then perforated selected areas with a microhardness tester using a steel stylus ground to a truncated cone, and Budd and Booth[96] using a similar technique were able to produce punctures in the film down to 35 μm diameter. Microelectrodes for potential measurements were described by Cleary[97] who used a Ag/AgCl/Cl-electrode with capillary tips down to 15 μm; Cleary also described microglass electrodes of approximately 30 μm diameter for the determination of pH. Doig and Edington[98] used microelectrodes with an internal diameter at the tip of about 0.2 μm to measure localized corrosion potentials in the grain boundary regions of aged Al–Mg and Al–Cu alloys,

and Davis[99] used microelectrodes with a tip diameter of 1 μm to measure pH during stress corrosion of aluminum alloys.

Potential–time relationships have been widely used for studying film formation and film breakdown, as indicated by an increase or decrease in the corrosion potential, respectively. May[100] studied the corrosion of 70/30 brass and aluminum brass in seawater and showed how scratching the surface resulted in a sudden fall in potential to a more negative value followed by a rapid rise due to reformation of the film; conversely, the pitting of stainless steel in chemical plant may be detected by a sudden decrease in potential.[101]

Hoar and his coworkers[102,103] used potential changes to study film breakdown and repair during the stress-corrosion cracking of austenitic stainless steels in boiling saturated $MgCl_2$ solution. Horst, *et al.*[104] have used potential measurements as a test to predict, the stress-corrosion susceptibility of 2219 aluminum alloy products (alloys containing approximately 6% Cu and tempered to give maximum strength and resistance to stress-corrosion cracking). The test solution used was methanol plus carbon tetrachloride and, it was shown that susceptible alloys were 200–500 mV more positive than nonsusceptible alloys, and that this difference in potential was revealed in less than 1 h. They claim that the test is more sensitive than that in which sodium chloride plus hydrogen peroxide is used as the test solution, since the latter gives potential differences of only 20 mV.

ASTM G 5 outlines standard methods for making potentiostatic and potentiodynamic anodic polarization measurements and ASTM G 3 gives conventions applicable to electrochemical measurements in corrosion testing. Further information and references are provided elsewhere.

2.34.5 Polarization Resistance

It is evident from previous considerations that the corrosion potential E_{corr} provides no information on the corrosion rate, and it is also evident that in the case of a corroding metal in which the anodic and cathodic sites are inseparable (cf. bimetallic corrosion) it is not possible to determine i_{corr} by means of an ammeter. The conventional method of determining corrosion rates by mass-loss determinations is tedious and over the years attention has been directed to the possibility of using instantaneous electrochemical methods. Thus, based on the Pearson derivation,[105] Schwerdtfeger, *et al.*[106,107] have examined the logarithmic polarization curves for 'potential breaks' that can be used to evaluate the corrosion rate; however, the method has not found general acceptance.

Skold and Larson[108] in studies of the corrosion of steel and cast iron in natural water found that a linear relationship existed between potential and the applied anodic and cathodic current densities, provided the values of the latter were low. However, the recognition of the importance of these observations is due to Stern and his coworkers[56,57] who used the term 'linear polarization' to describe the linearity of the E versus i curve in the region of E_{corr}, the corrosion potential. The slope of this linear curve, ΔE versus ΔI or ΔE versus Δi, is termed the polarization resistance, R_p, since it has dimensions of ohms, and this term is synonymous with 'linear polarization' in describing the 'Stern–Geary' technique for evaluating corrosion rates.

Stern and Geary[56,57] on the basis of a detailed analysis of the polarization curves of the anodic and cathodic reactions involved in the corrosion of a metal, and on the assumption that both reactions were charge-transfer controlled (transport overpotential negligible) and that the IR drop involved in determining the potential was negligible, derived the expression:

$$\frac{1}{R_p} = \left[\frac{\Delta i}{\Delta E}\right]_{E_{corr}} = 2.3\left[\frac{b_a + |b_c|}{b_a|b_c|}\right] i_{corr} \qquad [1]$$

where R_p is the polarization resistance determined at potentials close to E_{corr}, and b_a, b_c are the Tafel constants; note that in the case of b_c the negative sign is disregarded. This equation shows that the corrosion rate is inversely proportional to R_p (or directly proportional to the reciprocal slope of the ΔE versus Δi curve) at potentials close to E_{corr} (typically within 10 mV), and that i_{corr} can be evaluated provided the Tafel constants are known. For a process that is controlled by diffusion of the cathode reactant (transport control) and in which the anodic process is under activation control a similar linear relationship applies:

$$\frac{1}{R_p} = \left[\frac{\Delta i}{\Delta E}\right]_{E_{corr}} = \frac{2.3 i_L}{b_a} = \frac{2.3 i_{corr}}{b_a} \qquad [2]$$

where i_L is the limiting current density of the cathodic reaction and it is assumed that $i_L = i_{corr}$

Stern and Weisert[109] by taking arbitrary values of the Tafel constants showed that corrosion rates

determined by the polarization resistance techniques are in good agreement with corrosion rates obtained by mass loss methods.

The importance of the method in corrosion testing and research has stimulated other work, and since the appearance of Stern's papers there have been a number of publications many of which question the validity of the concept of linear polarization. The derivation of linearity polarization is based on an approximation involving the difference of two exponential terms, and a number of papers that have attempted to define the range of validity of polarization resistance measurements have appeared. Barnartt[110,111] derived an analytical expression for the deviations from linearity and concluded that it varied widely between different systems. Leroy,[112] using mathematical and graphical methods, concluded that linearity was sufficient for the technique to be valid in many practical corrosion systems. Most authors emphasize the importance of making polarization resistance measurements at both positive and negative overpotentials.

Oldham and Mansfeld[113] approached the problem of linearity in a different way and their derivation avoids the approximation used by Stern and Geary. They conclude that although linearity is frequently achieved, this is due to three possible causes: (a) ohmic control due to the *IR* drop rather than control according to linear polarization, (b) the similarity of the values of b_a and b_c, and (c) a predisposition by the experimenter to assume that the ΔE versus Δi curves near E_{corr} must be linear. In a later paper Oldham and Mansfeld[114] showed that linearity of the ΔE versus Δi curve is not essential and that i_{corr} can be evaluated from the slopes of the tangents of the nonlinear curve determined at potentials of about 20–30 mV more positive and negative than E_{corr}.

Hickling,[115] in attempting to study the corrosion of steels under thin film conditions that simulate atmospheric exposure, took into account the time-dependence of polarization measurements, and developed a technique using galvanostatic transients.

2.34.5.1 Tafel Constants

It is evident from eqns [1] and [2] that the evaluation of i_{corr} from R_p determinations requires a knowledge of the Tafel constants b_a and b_c, which may not be available for the system under study and which may change in value during the progress of the determination. The determination of the Tafel constants from complete E versus i curves for each system studied is time consuming, and may not be particularly accurate owing to resistance and mass transfer effects. Hoar[116] has criticized the method on these grounds and has pointed out that the complete Tafel equations for the anodic and cathodic reactions, which have to be determined to evaluate the Tafel slopes, can be used to calculate i_{corr} without resorting to the polarization resistance technique. Mansfeld[117–120] suggests that polarization curves obtained in the R_p region can be fitted to various theoretical curves, preferably by computer analysis, to give the separate value of both b_a and b_c, which since they are determined simultaneously with the R_p values avoids the criticism that they may change substantially during the corrosion test. (An alternative approach uses the distortion of a small amplitude sine wave as a means of estimating the Tafel constants.)

The controversy that arises owing to the uncertainty of the exact values of b_a and b_c and their variation with environmental conditions, partial control of the anodic reaction by transport, etc. may be avoided by substituting an empirical constant for $(b_a + |b_c|/b_a|b_c|)$ in eqn [1], which is evaluated by the conventional mass-loss method. This approach has been used by Makrides[121] who monitors the polarization resistance continuously, and then uses a single mass-loss determination at the end of the test to obtain the constant. Once the constant has been determined it can be used throughout the tests, providing that there is no significant change in the nature of the solution that would lead to markedly different values of the Tafel constants.

2.34.5.2 Applications

The method, in spite of its limitations, has a number of significant advantages and provides a method of rapidly monitoring the instantaneous corrosion rates; furthermore, it has the advantage of the small changes in potential required in the determination not disturbing the system significantly. It is capable of measuring both high and low corrosion rates with accuracy, and may be used as a laboratory tool for testing or research, or for monitoring corrosion rates of plant.

Stern[57] pointed out that the polarization resistance method could be of value for determining the effect of the changes of environment (composition, temperature, velocity) and alloy composition on the corrosion rate and for evaluating inhibitors, and since his original publications the method has been widely used for a variety of studies. Thus, Legault and Walker[122] used the method for studying the

inhibition of the corrosion of steel in chloride solutions by $NaNO_2$, and Walker and France[123] extended this approach to a study of the in situ corrosion of the various metals in automotive-engine cooling systems. Jones and Greene[124] developed the theory of transient linear polarization to study very low corrosion rates, such as those that occur with surgical implant materials, and have shown how polarization resistance data can be used to monitor the onset of pitting or other forms of localized corrosion.

Wilde[125] has applied the Jones d.c.-bridge technique[126] to compensate for errors due to the *IR* drop, and has obtained meaningful corrosion rates from polarization resistance data in high temperature high-purity water in nuclear reactors.

Bureau[127] and others[128] have tried to apply the technique for evaluating the corrosion rate of painted metals, and although the results are controversial, the method has also been used successfully in the study of canning materials and lacquered surfaces.[129–131]

Rowlands and Bentley[132] have provided an account of the possibilities for continuously monitoring corrosion rates by polarization resistance measurements, and they also describe the development of a commercial instrument, which uses low-frequency square-wave current to polarize the test specimens.

2.34.5.3 Derivation of Linear Polarization Method for Determining Corrosion Rates

It is assumed that

1. The corrosion current i_{corr} (it is also assumed that the area of the metal is 1 cm^2 so that $I_{corr} = i_{corr}$) occurs at a value within the Tafel region for the anodic and cathodic reaction, that is, transport overpotential is negligible.
2. E_{corr} is remote from the reversible potentials of the anodic and cathodic reactions.
3. The *IR* drop in measuring the polarized potential is negligible.

Following Oldham and Mansfeld,[113] but using the symbols that have been adopted in the present work, it is required to show that

$$\left(\frac{\mathrm{d}i}{\mathrm{d}E}\right)_{E_{corr}} = Z = i_{corr}\left[\frac{1}{\beta_a} + \frac{1}{\beta_c}\right] \qquad [3]$$

where β_a, β_c are the Tafel slopes of anodic and cathodic reactions constituting the overall corrosion reaction, that is, i_{corr} is linearly related to the polarization resistance ($\mathrm{d}E/\mathrm{d}i$) at potentials close to E_{corr}. (Note that β is used to represent the Tafel slope expressed in terms of natural logarithms, so $b = 2.3\beta$.)

At any potential E the net current is given by

$$i = \overleftarrow{i_1} - \left|\overrightarrow{i_1}\right| + \overleftarrow{i_2} - \left|\overrightarrow{i_2}\right| \qquad [4]$$

where $\overleftarrow{i_1}$ is the anodic current for metal dissolution and $\overrightarrow{i_1}$ is the reverse cathodic current, and $\overrightarrow{i_2}$ is the cathodic current for the reduction of the cathode reactant (dissolved O_2, H^+, H_2O, etc.) and $\overleftarrow{i_2}$ the reverse current.

The rate of the anodic reaction at a potential E is given by

$$\overleftarrow{i_1} = i_{0,1}\exp\left(\frac{E - E_{r,1}}{\beta_a}\right) \qquad [5]$$

where $E_{r,1}$ is the reversible potential for the anodic dissolution reaction, β_a is the Tafel slope and $i_{0,1}$ is the exchange current density. Similarly for the cathodic reaction

$$\overrightarrow{i_2} = i_{0,2}\exp\left(\frac{E_{r,2} - E}{\beta_c}\right) \qquad [6]$$

Similar expressions may be written for the partial reverse rates $\overrightarrow{i_1}$ and $\overleftarrow{i_2}$, but under the conditions assumed here they may be neglected. Hence, substituting eqns [5] and [6] in eqn [4]:

$$i = i_{0,1}\exp\left(\frac{E - E_{r,1}}{\beta_a}\right) - i_{0,2}\exp\left(\frac{E_{r,2} - E}{\beta_c}\right) \qquad [7]$$

At the corrosion potential E_{corr} the net current i becomes zero, since $\overleftarrow{i_1} = |\overrightarrow{i_2}|$. Thus, the two terms on the right-hand side of eqn [7] become equal to one another and equal to i_{corr}, the corrosion current density. Thus, replacing E in eqn [7] by E_{corr} gives

$$i = i_{0,1}\exp\left(\frac{E_{corr} - E_{r,1}}{\beta_a}\right) - i_{0,2}\exp\left(\frac{E_{r,2} - E_{corr}}{\beta_c}\right) \qquad [8]$$

Differentiating eqn [7] with respect to E gives

$$\frac{\mathrm{d}i}{\mathrm{d}E} = \frac{i_{0,1}}{\beta_a}\exp\left(\frac{E - E_{r,1}}{\beta_a}\right) + \frac{i_{0,2}}{\beta_c}\exp\left(\frac{E_{r,2} - E}{\beta_c}\right) \qquad [9]$$

which for $E = E_{corr}$ becomes

$$\left(\frac{\mathrm{d}i}{\mathrm{d}E}\right)_{E_{corr}} = \frac{i_{0,1}}{\beta_a}\exp\left(\frac{E_{corr} - E_{r,1}}{\beta_a}\right) + \frac{i_{0,2}}{\beta_c}\exp\left(\frac{E_{r,2} - E_{corr}}{\beta_c}\right) \qquad [10]$$

Combining eqns [8] and [10] gives eqn [3] the Stern–Geary equation by simple algebra. However, Oldham and Mansfield point out that further differentiation of eqn [9] gives

$$\frac{d^2i}{dE^2} = \frac{i_{0,1}}{\beta_a^2}\exp\left(\frac{E - E_{r,1}}{\beta_a}\right) - \frac{i_{0,2}}{\beta_c^2}\exp\left(\frac{E_{r,2} - E}{\beta_c}\right) \quad [11]$$

an equation that demonstrates that there is only one point (a point of inflection, corresponding to a minimum slope) at which the i–E curve has no curvature and is linear. It follows that

$$\left(\frac{d^2i}{dE^2}\right)_{E_{corr}} = \frac{i_{0,1}}{\beta_a^2}\exp\left(\frac{E_{corr} - E_{r,1}}{\beta_a}\right) - \frac{i_{0,2}}{\beta_c^2}\exp\left(\frac{E_{r,2} - E_{corr}}{\beta_c}\right) \quad [12]$$

and combining this equation with eqn [8] gives

$$\left(\frac{d^2i}{dE^2}\right)_{E_{corr}} = i_{corr}\left(\frac{1}{\beta_a^2} - \frac{1}{\beta_c^2}\right) \quad [13]$$

For the E versus i plot to be linear at $E \approx E_{corr}\left(d^2i/dE^2\right)_{E_{corr}}$ must be zero, but eqn [13] shows that this will be true only if $\beta_a = \beta_c$. (Note that this does not invalidate the basic Stern–Geary equation, it merely shows that the d^2E/di^2 is only zero at E_{corr} if $\beta_a = \beta_c$, so that it could be argued that the use of the word 'linear' is not strictly accurate.)

2.34.5.4 Simultaneous Determination of Tafel Slopes and Corrosion Rates from R_p Determinations

Mansfeld[120] points out that a major limitation of the polarization resistance is that the factor $b_a b_c/2.3(b_a + b_c)$ must be determined in order to evaluate I_{corr} and has devised a procedure in which this can be achieved by a graphical method.

The Stern–Geary equation can be written in the form

$$I_{corr} = \frac{b_a b_c}{2.3(b_a + b_c)} \times \frac{1}{R_p} = \frac{B}{R_p} \quad [14]$$

where $B = b_a b_c/2.3(b_a + b_c)$ and $R_p = (dE/dI)$ at E_{corr}. Equation [14] is valid only if the relationship between I and E can be expressed as

$$I = I_{corr}\left\{\exp\left(\frac{2.3(E - E_{corr})}{b_a}\right) - \exp\left(\frac{-2.3(E - E_{corr})}{b_c}\right)\right\} \quad [15]$$

Combining eqns [14] and [15] and rearranging gives

$$2.3R_pI = \frac{b_a b_c}{b_a + b_c}\left\{\exp\left(\frac{2.3\Delta E}{b_a}\right) - \exp\left(-\frac{2.3\Delta E}{b_c}\right)\right\} \quad [16]$$

where $\Delta E = E - E_{corr}$ Since the right-hand side of eqn [16] depends only upon the Tafel slopes it should be possible to evaluate b_a and b_c from plots of R_pI versus ΔE.

Figure 10(a) shows a theoretical plot of the right-hand side of eqn [16] versus ΔE in which the cathodic

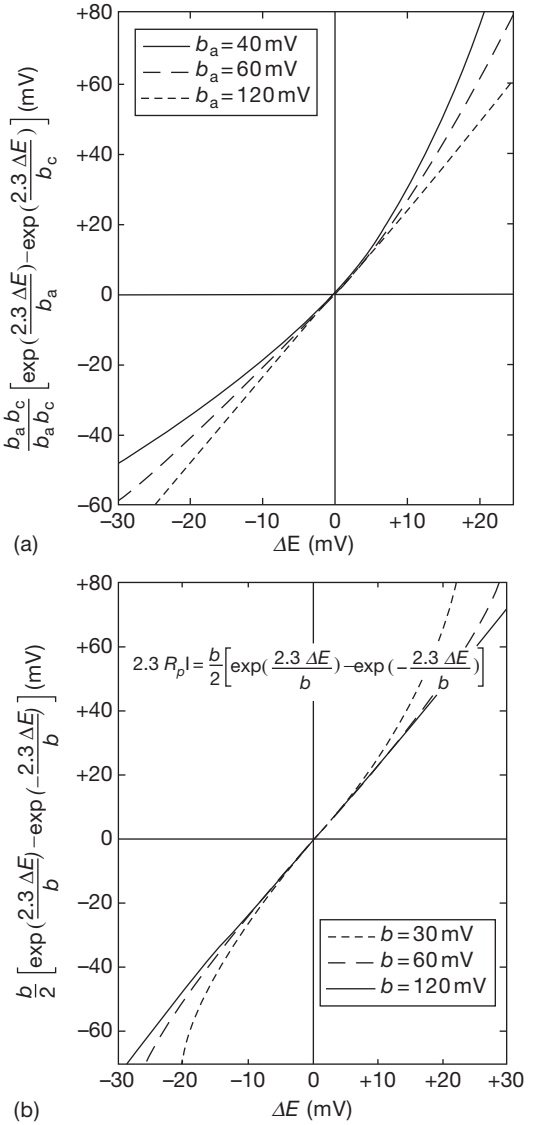


Figure 10 Plots of right-hand side of equation 16 versus $\Delta E = E - E_{corr}$ for various combinations of Tafel slopes. (a) b_c constant at 120 mV, b_a varied and (b) $b_a = b_c = b$. Reproduced from Mansfeld, F. *J. Electrochem. Soc.* **1973**, *120*, 515.

Tafel slope has been assumed to be constant at 120 mV and the anodic Tafel slope to have the arbitrary slopes of 40, 60, and 120 mV. It can be seen that linearity over a range of positive and negative potentials ΔE is achieved only when $b_a = b_c$ and that linearity is confined to $\Delta E \approx 0$ when b_a and b_c differ.

In **Figure 10(b)** it has been assumed that the Tafel slopes are equal, that is, $b_a = b_c = b$ and the modified expression for the right-hand side of eqn [16] has been plotted against ΔE for different values of b (e.g., 30, 60, and 120 mV). Comparison of **Figure 10(a)** and **10(b)** shows how the curvature of the plots differs at cathodic potentials, that is, $\Delta E < 0$. Thus, the kinetic behavior of a corroding metal, as expressed by different combinations of Tafel slopes, can be organized by this method of plotting curves. This theoretical approach has been confirmed experimentally by Mansfeld for the system Fe/H_2SO_4.

Mansfeld points out that I_{corr} can be calculated from the measured polarization curve by the following four steps, which are based on eqns [14] and [15]. (With the advent of low cost computing power, and computerized systems for the acquisition of polarization data, this process can be automated by fitting eqn [8] (or rather a simplified form of this equation, taking account of the fact that $i_{0,1}$ and $E_{r,1}$ (and $i_{0,2}$ and $E_{r,2}$) can be replaced by any current–potential pair that lies on the Tafel line, one such pair being i_{corr} and E_{corr} for both anodic and cathodic reactions), or a more sophisticated model of the E–i relationship, to the measured data by adjusting the variable parameters (for a recent example see Guinea, D. M.; Moreno, B.; Chinarro, E.; Guinea, D.; Jurado, J. R. *Int. J. Hydrogen Energy* **2008**, *33*(11), 2774–2782). Alternatively the harmonic analysis and electrochemical frequency modulation methods provide other approaches to obtain this information.)

1. Determine R_p from:

$$\left[\frac{\partial I}{\partial E}\right]_{E_{corr}} = \frac{1}{R_p}$$

   by drawing a tangent at $\Delta E = 0$ that is, at E_{corr}.
2. Multiply the current I measured at a certain value of ΔE by 2.3 R_p and plot 2.3 R_p I versus ΔE for various values.
3. Determine from this plot the Tafel slopes b_a and b_c by curve fitting using the theoretical curves calculated for various values of b_a and b_c.
4. Calculate I_{corr} from eqn [14] using the R_p, value evaluated in Step 1 and the Tafel slopes determined in Step 3.

2.34.6 Tests for Bimetallic Corrosion

The extent of galvanic effects will be influenced by, in addition to the usual factors that affect corrosion of a single metal, the potential relationships of the metals involved, their polarization characteristics, the relative areas of anode and cathode, and the internal and external resistances in the galvanic circuit.

The results of a galvanic corrosion test on a small scale are as a general rule no more than semiquantitative. A principal reason for this is that the magnitude of the galvanic effect is a function of galvanic current density which is usually determined by the relative areas of the metals forming the couple. There may also be major differences in circuit resistances in tests as compared with practice – especially if current-measuring shunts of substantial resistance are made part of the circuit in the test. The geometric relationship between the metals in the test will also influence the result through effects on electrolyte resistance and the distribution of the galvanic currents.

The simplest procedure in studying galvanic corrosion is a measurement of the open-circuit potential difference between the metals in a couple in the environment under consideration. This will at least indicate the probable direction of any galvanic effect, although no information is provided on the rate. A better procedure is to make similar open-circuit potential measurements between the individual metals and some appropriate reference electrode, which will yield the same information and will also permit observations of any changes in the potential of the individual metals with time that will affect the overall potential difference in the couple. For most practical laboratory testing, the saturated calomel half cell is most convenient. The precision of the determinations is adequate and it is easy to maintain a constant concentration of potassium chloride.

The preferred potential-measuring instruments are high-input impedance voltmeters that permit measurements to be made without flow of sufficient current to polarize the electrodes during the determinations.

Open-circuit potential measurements do not indicate the all important effects of continued current flow, and much more information is derived from frequent or continuous determinations of the magnitude of the galvanic current. In making these measurements it is necessary to avoid the use of instruments that will introduce sufficient resistance to exert a controlling effect on the magnitude of the galvanic current being measured. Instruments (zero-resistance ammeters) that permit current measurements to be made with

zero resistance in the measuring circuit (see Ailor)[2] are available.

In less critical cases it will suffice to include in the circuit a low value resistance over which *IR*-drop potential measurements can be made for calculation of the current. This technique permits measurements to be made as required without opening the circuit even momentarily for the introduction of current-measuring devices.

An obvious method for studying galvanic corrosion either with or without supplementary electrical measurements is to compare the extent of corrosion of coupled and uncoupled specimens exposed under identical conditions. Such measurements may use the same techniques for estimating corrosion damage, such as mass-loss determinations, as have been described in connection with ordinary corrosion tests.

A convenient method of carrying out such a galvanic test in the laboratory has been described by Wesley[133] in which the vertical circular-path machine is used. Each assembly includes two pairs of dissimilar metals – one pair coupled galvanically while the other pair is left uncoupled to determine the normal corrosion rates under the same environmental conditions. The type of motion provided (specimens moving in a vertical circular path) enables electrical connections to be made without mercury cup or commutator and the leads can be connected to a calibrated resistance for current measurements attached to the specimen carrier.

It is often of interest also to measure both the external and internal resistances of the galvanic circuit by the use of appropriate resistance-measurement bridges or by even more elaborate techniques such as those described by Pearson.[105]

It is often desirable to know something about the probable distribution of galvanic effects in a galvanic couple. This will be determined, of course, by the size and shape of the different metals and how they are placed relative to each other – whether more or less parallel in the electrolyte, close together or far apart, or joined along some line of contact. The distribution from such a line of contact may be observed directly if the test couples are exposed in this way and for long enough for sufficient corrosion to occur for it to be observed and measured. Alternatively, the distribution of the galvanic currents in terms of the current density on different portions of both the anode and cathode surfaces may be estimated from data derived from surveys of the potential field in the electrolyte around the couple. Such a potential survey may be made using a fixed and a movable reference electrode so that equipotential lines in several planes may be measured and plotted as was done by Copson[134] using a technique originally proposed by Hoar. By analysing the data from the potential surveys, it is possible to calculate the current distribution over different areas near and remote from the contact of the dissimilar metals. This technique has been used by Rowe[135] to study the corrosion behavior of coated and uncoated couples.

Guidance on conducting and evaluating galvanic corrosion tests in electrolytes is given in ASTM G 71.

2.34.7 Soil Tests

Soil corrosion does not lend itself readily to direct study in the laboratory. However, indirect methods involving the action of differential aeration cells have yielded valuable information in comparing the probable corrosivities of different soils towards steel. The details of this technique were described by Denison,[136] Ewing,[137] Schwerdtfeger,[138–141] and by Logan *et al.*[142]

The Schwerdtfeger[138–141] 'polarization break' and the polarization resistance methods have been studied by Jones and Lowe[143] in relation to their effectiveness in evaluating corrosion rates of buried metals. A Holler bridge circuit was used to remove *IR* contributions during the measurement of the polarized potential. Jones and Lowe, on the basis of their studies of buried steel and aluminum specimens, concluded that the polarization resistance was the most useful, and that the polarization break had the serious limitation difficult to making it identify the breaks in the curve.

2.34.8 Accelerated Tests – Electrolytic Tests

In view of the electrochemical nature of corrosion, it has seemed reasonable to many investigators to assume that suitable accelerated corrosion tests could be made by observing the response to electrolytic stimulation of the corrosion processes, or by attaching particular significance to the results of quickly made electrode potential and current measurements.

Acceleration of corrosion by electrolytic stimulation has sometimes been found to distort normal corrosion reactions to such an extent that the results bear no consistent relationship to ordinary corrosion and are, therefore, quite inconsistent and unreliable. This was shown, for example, by a series of tests sponsored by ASTM Committee B-3.[144,145] Nevertheless, considerable success has been achieved with the development

of the electrochemical potentiokinetic reactivation (EPR) test as an accelerated method for the detection of sensitization in austenitic stainless steels. This is discussed subsequently in the section concerned with intergranular attack of Cr–Ni–Fe alloys.

Measurements of open-circuit potentials relative to a reference electrode have been assumed on occasion to provide a means of rating metals as to their relative resistance to corrosion on the basis that the more negative the measured potential, the higher will be the rate of corrosion, but this assumption is obviously invalid, since it disregards polarization of the anodic and cathodic areas. (More importantly, it ignores the possibility of passivation, which leads to a more positive potential, but a reduced corrosion rate compared to active corrosion.)

Some examples of the use of potential measurements in corrosion tests have been given and it is of interest to outline here certain test procedures that are used industrially to supplement or replace the more tedious and prolonged laboratory and field tests. These tests frequently rely on changing the potential as a means of accelerating the test, and although, as emphasized earlier, this is capable of distorting the mechanism, it is less likely to do so than a change in the nature of the environment, an increase of temperature, etc. The majority of these tests are used for evaluating electrodeposits, anodized coatings, and paint films.

2.34.8.1 Electrolytic Oxalic Acid Etching Test

This test has been developed and used by Streicher[146–149] as a screening test to be used in conjunction with the tedious boiling nitric acid test for assessing the susceptibility of stainless steels to intergranular attack as specified in ASTM A 262, and will be considered subsequently in the section concerned with intergranular attack of Cr–Ni–Fe alloys.

2.34.8.2 The Electrolytic Corrosion (EC) Test

The EC test was developed by Saur and Basco[150–153] for decorative Cu (optional) + Ni + Cr electrodeposits. After an appropriate area is masked off and cleaned with a slurry of MgO, the specimen is immersed in test solution A or B (**Table 2**). It is held by means of a potentiostat at +0.3 V (versus SCE) and taken through cycles of 1 min anodically polarized, 2 min unpolarized.

The extent of pitting is estimated by a special microscopic technique, or by the attack on the substrate using an appropriate indicator. Thus, in the case of steel 1,10-phenanthroline hydrochloride is added to the electrolyte (solution B) to detect the formation of Fe^{2+} ions. Alternatively, the specimens can be removed from the corrosion test solution and placed in an indicator solution, that is, solution C for zinc-base die castings and solution D for steels.

The test is much faster than the CASS test and is probably more reproducible; more important is the fact that it has been correlated with service exposure.

Table 2 Electrolytic corrosion test: composition of test solutions (A, B) and indicators (C, D)

	Concentration			
	A	*B*	*C*	*D*
$NaNO_3$ (g l^{-1})	10	10		
NaCl (g l^{-1})	1.3	1.0		
HNO_3 (conc.; g l^{-1})	5	5		
1,10-Phenanthroline hydrochloride (g l^{-1})		1.0		
KCNS			3	3
Acetic acid (glacial, ml l^{-1})			2	2
Quinoline (ml l^{-1})			8	
H_2O_2(30%) (ml l^{-1})				3

Source: Saur, R. L.; Basco, R. P. *Plating* **1966**, *53*, 33.
Saur, R. L.; Basco, R. P. *Plating* **1966**, *53*, 320.
Saur, R. L.; Basco, R. P. *Plating* **1966**, *53*, 981.
Saur, R. L. *Plating* **1966**, *54*, 393.

2.34.8.3 Impedance (Aztac) Test

The impedance test[154] for anodized aluminum (ASTM B 457) employs a 1 V r.m.s. 1 kHz source applied to a test cell in an impedance bridge; the electrolyte solution is 3.5% NaCl. The results are expressed as kiloohms, and while a bare Al specimen will give a value of about 1 kΩ, a well scaled anodized coating will give a value of 100 kΩ. The admittance test (BS BS EN 12373-5 and BS 3987) is essentially the same as the impedance test but uses 3.5% K_2SO_4 rather than 3.5% NaCl. An admittance of $< 500/t$ μS (where t is the thickness of the film in micrometers) denotes good sealing. It should be noted that thickness of the sealed coating should be specified in both tests.

2.34.9 Accelerated Tests – Simulated Environments

2.34.9.1 Spray Tests

The most common of the spray tests is the salt-spray or salt-fog test, which was developed originally by

Capp[155] in 1914 for studying the protective values of metallic coatings on steel under conditions that he hoped would simulate exposure to a sea-coast atmosphere. Since then, the test has been used for a number of purposes, for many of which it is not well suited.[156–158]

Although there is no standard size or shape of salt-spray box, certain other features of the test have been standardized in ASTM B 117. Various factors affect the rate of attack and **Figure 11** shows the effect of the angle of exposure of the specimen to the salt droplets, which fall vertically from the spray nozzle, based on early work by May and Alexander.[159] It can be seen that maximum corrosion occurs at angles between about 30° and 80° to the horizontal; vertical exposures were found to give erratic results.

There may be variations from box to box, depending on differences in fog characteristics as influenced by the design of the spray nozzles, the pressure at which the brine enters the nozzle, and the constancy of this pressure.[160]

The results will also be influenced by the concentration of NaCl solution sprayed, some metals are affected more by one concentration than another, for example, zinc is corroded most by a concentrated brine (20%), while iron is corroded most by a dilute brine (3%); synthetic seawater is less corrosive to these metals than either brine. In view of the many other ways by which the conditions within a salt-spray box differ from those of exposure to a natural sea-coast environment, there seems to be no great advantage in making-up complicated synthetic seawaters for use in salt-spray testing. However, tablets for this purpose are commercially available.

Some modified brines have been used in salt-spray boxes for particular purposes, such as the acetic acid-modified brine developed by Nixon[161] to reproduce the type of blistering frequently encountered on chromium-plated zinc-base die castings. An acetic acid-salt-spray test has been adopted by ASTM (ASTM G 85) and other modifications of the spray test covered by the same standard practice include the cyclic acidified salt spray test, the acidified synthetic seawater spray test and the SO_2 salt spray test. Corresponding tests in the BS series are covered by BS EN 12540. In addition, the original acetic acid-salt-spray test has been modified by including cupric chloride in the brine. This is called the CASS (Copper Accelerated Acetic Acid Salt-Spray) Test; ASTM B 368 test. It is used extensively for testing nickel–chromium coatings on steel and zinc. The original acetic acid-salt-spray test, is modified in an important aspect by the addition of $0.25\,g\,l^{-1}$ of $CuCl_2{\cdot}2H_2O$, to the 5% NaCl test solution, which substantially increases the corrosivity of the solution, especially to nickel. The addition of $FeCl_3$ to the acetic acid-salt-spray solution, such as is used in the Corrodkote test, was earlier noted to be troublesome in that it tended to precipitate. For this reason, ferric iron is not included in the CASS test solution. Essential details[162,163] include control of cabinet temperature at 49 °C, control of saturation temperature at 57 °C, control of collection rate at $1.5 \pm 0.5\,ml\,h^{-1}$ per $80\,cm^2$ of specimen surface, control of pH by addition of acetic acid to 3.2 ± 0.1, and an operating air pressure of $103.4 \pm 6.9\,kN\,m^{-2}$. Higher pressure may be required to achieve the specified collection rate in 'walk-in' cabinets.

It has been shown that chromium is virtually unattacked by the CASS test solution.[164] Nickel, on the other hand, is corroded at a substantial rate (about $0.072\,mm\,year^{-1}$), the presence of the copper ions tending to maintain the nickel in an active state.[164,165] Thus, in the CASS test (and in the Corrodkote test as well) accelerated galvanic corrosion of the nickel occurs at any discontinuity in the chromium layer. Good correlation between the results of the CASS test and the performance of plated parts in service has been reported.[164]

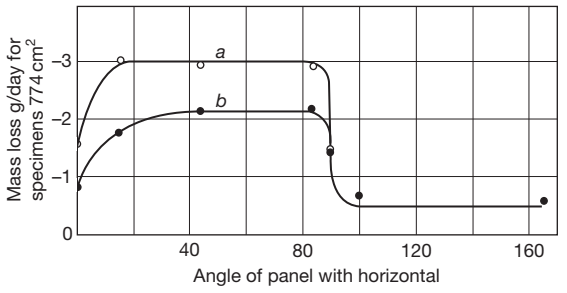


Figure 11 Effect of cold-rolled steel specimen position on corrosion in salt-spray test; a is 20% NaCl and b is synthetic seawater.

2.34.9.2 Corrodkote Test[162,166]

This is a refinement of an earlier test in which melted street slush together with its contained dirt, salts, etc. was splashed upon plated parts by means of a rotating paddle wheel. Parts soiled in this manner were then exposed to a warm, humid atmosphere. The results were striking and significant in that they closely paralleled service experience. The 'paddle-wheel test' was

intended to simulate the conditions to which plated parts on automobiles are subjected when cars are garaged, unwashed, after being driven over salted slush-covered city streets on typical winter days. Platings of inadequate thickness and quality have frequently been observed to show signs of failure after only a few weeks' or even a few days' use under such circumstances. Despite certain limitations, good correlation has been reported between the results of the Corrodkote test and service performance of plated components,[133] and it is now included in ASTM B 380 and BS EN 12540 and ISO 4541. Reagents for use in this test are prepared as follows.

1. The cupric nitrate reagent contains 2.5 g $CuNO_3 \cdot 3H_2O$, dissolved in 500 ml distilled water.
2. The ferric chloride reagent contains 2.5 g $FeCl_3 \cdot 6H_2O$, dissolved in 500 ml distilled water. This reagent should not be kept longer than 2 weeks.
3. The ammonium chloride reagent contains 50 g NH_4Cl, dissolved in 500 ml distilled water.
4. The Corrodkote slurry is prepared by mixing 7 ml of the cupric nitrate reagent, 33 ml of the ferric chloride reagent, and 10 ml of the ammonium chloride reagent with 30 g of kaolin to form a homogeneous slurry, which is sufficient to cover about 2.79 m^2 of plated surface. A fresh batch of slurry should be made up each day.
5. The surfaces to be tested should be coated with the slurry by brushing in circular motion, finishing with brush strokes in one direction. The coating should then be allowed to dry for 1 h, after which the coated specimen should be put into a noncondensing humidity cabinet at 38 °C and from 90 to 95% RH. After 20 h in the cabinet the specimen should be removed for inspection. Zinc die castings should be cleaned with running water and dried before inspection. Steel specimens should be examined before cleaning and the number of rust spots counted. Since most of the rust will come off with the Corrodkote coating, it may be difficult to distinguish between surface pits and pin holes reaching the basis metal after cleaning. Steel parts may be returned to a condensing humidity cabinet for 24 h or to a salt-spray cabinet for 4 h. Either supplementary exposure will bring out rust spots again.

One cycle of the Corrodkote test will reliably reveal coatings that will not endure one winter's normal use in a typical city that uses salt to deice its streets. In contrast, several cycles of the Corrodkote test are generally required to 'fail' coatings which will withstand one or more such winter's use.

In this connection, there is some indication that while the Corrodkote test can be depended upon to reveal coatings of unsatisfactory durability, there has been some question of its ability to distinguish between, or to predict the relative protective value or length of useful life of different coating systems in the very good or excellent durability range.

Also of questionable significance is the practice of shortening the Corrodkote cycle, to say 4 h, for the purpose of evaluating the durability of relatively thin coatings intended for use under comparatively mild conditions such as indoors or the interiors of automobiles, since by far the greatest amount of corrosion (of the nickel) appears to occur during the early part of the Corrodkote humidity cycle. Good correlation between the Corrodkote test and service performance has been obtained by Bigge.[166]

2.34.9.3 Sulfur Dioxide Tests

Historically, two tests in which sulfur dioxide is the principal corroding agent have been used, that is, the BNFMRA sulfur dioxide test[167,168] and the Kesternich test.[169] These tests were investigated by the American Electroplaters' Society Research Project 15 Committee early in its search for an acceptable accelerated corrosion test. They were soon abandoned, however, largely because the types of corrosion failures developed did not resemble those that occurred in actual service. Furthermore, the extreme corrosivity of the test environment to nickel (some 8.38 mm $year^{-1}$) appeared to place an undue premium on the integrity of the overlying chromium deposit, which is virtually unattacked in the test. Thus, coatings that were substandard in respect of nickel or copper–nickel thicknesses might easily pass the test provided the chromium topcoat was completely continuous and remained so for the duration of the test. Conversely, coatings of proven merit on the basis of service experience, such as 0.039 mm of semibright plus bright nickel (duplex) with 0.00025 mm of conventional chromium, could be expected to fail in these SO_2 tests relatively quickly at any discontinuities in the chromium. In this connection, it is well to keep in mind that, even though the chromium may be nonporous initially, it can hardly be expected to remain so in service on an automobile for example, where it is subject to impact from sand, gravel, etc.

The BNFMRA test was used in Europe for testing Ni + Cr coatings, but since 1970 it has been omitted from revisions of BS 1224. The test was also used to ensure the quality of anodic coatings on aluminum, but in the current British Standard (BS EN 12373) the

acetic acid salt spray test or the copper-accelerated salt spray test are specified for corrosion testing (BS EN ISO 9227). In the United Kingdom, the Kesternich test, which has obtained the status of an ISO Standard (ISO 6988), is only specified for checking the extent of porosity in tin (BS 1872), tin/lead (BS 6137) and tin/nickel (BS 3597) coatings. In the United States, another version with a much more aggressive atmosphere is used to detect porosity in gold coatings on copper, nickel or silver (ASTM B 735) but the equivalent British Standard (BS EN ISO 27874) specifies electrographic tests for this purpose.

Sulfur dioxide is not included as a corrosion test medium in the current ISO Standard for electrodeposited coatings of nickel + chromium and of copper + chromium (ISO 1456). However, one important use of sulfur dioxide atmospheres as a controlled accelerated test has been for electrical (and, particularly, electronic) contacting surfaces. In this case, the concentration of SO_2 is much less than for the Kesternich test and the time of testing is much longer (IEC 60068-2-49).

Because moist air containing sulfur dioxide quickly produces easily visible corrosion on many metals in a form resembling those that occur in industrial environments, ASTM have issued a Standard Practice for conducting tests in moist SO_2 (G 87). This suggests that such tests are well suited to detect pores or other sources of weakness in protective coatings and deficiencies in corrosion resistance associated with unsuitable alloy composition or treatments. However, it is stressed that the results obtained should only be taken as a general guide to the relative corrosion resistance of these materials in moist SO_2 service.

2.34.9.4 General Considerations of Spray and SO_2 Tests

The salt-spray test has seemed to yield the most consistent results when used to establish the relative merits of different aluminum alloys in resisting attack by marine atmospheres. The best results have been secured when the spray has been interrupted for many hours each day.[170]

Salt-spray boxes are also used for studying the deterioration and protective value of organic coatings, although this test is of doubtful value for such purposes, since it fails to include many factors, for example, sunlight, which affects the life of such coatings. Methods of testing organic coatings are discussed elsewhere. The variable responses of different metals and coatings to the conditions that can be set up in salt-spray boxes, as well as to the conditions that exist in natural atmospheres, make it impossible to determine the equivalent in some natural atmosphere of say an hour in a salt-spray box.

For additional information on some of the features of the salt-spray test and its limitations in respect of certain purposes for which it may be used, reference should be made to the book Corrosion Testing for Metal Finishing[171] prepared by the Institute of Metal Finishing.

There have been several attempts to develop rather elaborate testing machines in which specimens may be subjected to various sprays of fogs with cycles of condensation, heating, and drying. The object has been to reproduce the conditions encountered by metals exposed in polluted industrial atmospheres. Such devices have been experimented with in the United Kingdom[172] and the United States.[173] While it is sometimes possible for such tests to rate steels in a rough order of resistance to atmospheric corrosion, it should be appreciated that the nature of rust formed may differ from that obtained during actual exposure. It is only in rare cases that the resistance to attack by the sprays is analogous to their resistance in the natural atmosphere, which the tests seek to simulate. Such parallelism is not common enough to make these tests very reliable.[174]

Dennis and Such[175] point out that the BNFMRA SO_2 test was really a means of detecting discontinuities in the chromium layer of a Cr + Ni coating system and it therefore gave unfavorable results when used for testing microcracked or microporous Cr, since the Cr was rapidly undermined, with consequent flaking. Conversely, the test exaggerates the beneficial effect of crack-free Cr. The test also fails to indicate the improved corrosion resistance of duplex Ni when compared with bright Ni. A critical account of laboratory corrosion testing methods for Ni–Cr coatings is given by Dennis and Such.[176]

2.34.9.5 Accelerated Tests for Weathering Steels

Interest in weathering steels has stimulated work on accelerated laboratory tests that can be used to investigate the effect of alloy composition on performance. It is well established that a wetting and drying cycle should be an integral part of any laboratory test in which the characteristic properties of weathering steels are revealed,[177] and Bromley, Kilcullen, and Stanners[178] have designed a test rig (**Figure 12**) that provides results that can be correlated with actual atmospheric exposure data. The rig has been designed to investigate

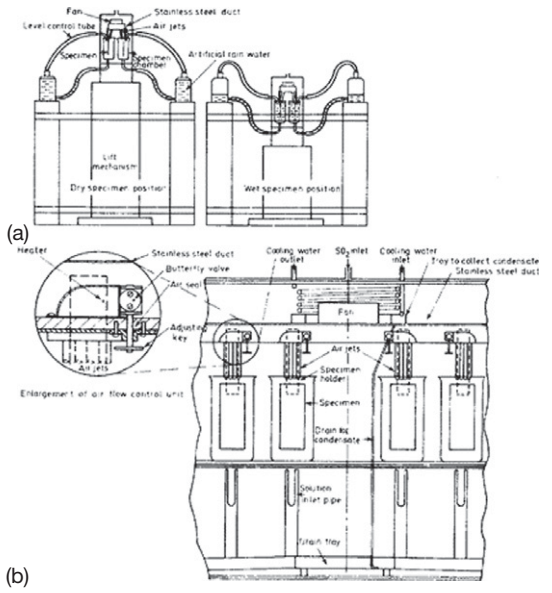


Figure 12 Rig used for laboratory evaluation of weathering steels. (a) General layout showing unimmersed and immersed positions and (b) detailed view of central position of cabinet. Reproduced from Bromley, A. F.; Kilcullen, M. B.; Stanners, J. F. In 5th European Congress of Corrosion, Paris, September 1973.

a wide range of alloying elements in a development program on slow-weathering steels for which it was essential to have a rapid, reliable and reproducible test that incorporated the specific atmospheric factors responsible for rust formation.

The results obtained with this equipment show that the corrosion rate in the rig is about four times that encountered in an industrial UK atmosphere. This acceleration, however, is not achieved by accentuating any of the environmental factors, but rather by holding them near the worst natural conditions for as long as possible. The procedure used ensures that the rust film is completely dried for short periods, thus simulating the conditions that bring out the beneficial effects of protective rust films on the steels under study.

The use of electrochemical tests for rapid assessment of the performance of these steels has attracted interest, and Pourbaix[179] has devised an apparatus in which potential measurements are used to evaluate the protective nature of corrosion products formed on low-alloy steels, such as the weathering steels, during periodic wetting and drying. The apparatus (**Figure 13**) consists of a glass tank containing an appropriate electrolyte, such as a natural or artificial water. Two specimens of the metal or alloy under study are attached to a spindle that rotates slowly

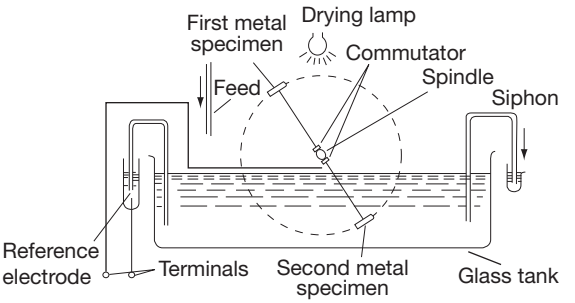


Figure 13 Equipment for studying patina formation on low alloy steels. Reproduced from Pourbaix, M. CEBELCOR RT; **1969**, 160.

(about 1 rev h^{-1}) so that the specimen is immersed in the solution for approximately half the time and exposed to the atmosphere for the remaining time. An electric lamp is placed above the tank so that the specimens remain wet for a time after withdrawal from the solution, but are completely dried during the cycle. Measurement of the potentials of the specimens at the beginning and end of the immersion period is effected by means of the commutators, which are attached to the spindle but electrically insulated from it, and a reference electrode. The e.m.f. taken from the terminals can be fed to a multipoint recorder so that a recording of the *E*–time relationship may be obtained for each specimen. The solution can be made to circulate slowly by allowing it to drip from a feed and overflow via a siphon. In a variation of the apparatus, a Luggin capillary is attached to the sample so that the potentials can be measured during the period when the specimen has emerged from the solution but is still wet.

Figure 14 shows results obtained from the apparatus for different steels some of which (nos. 1, 2, 3, and 7) form a protective patina for corrosion products, while others (nos. 4, 5, and 6) form patinas that are nonprotective; the criterion adopted is that the more positive the potential the more protective is the rust patina.[179]

Legault *et al.*[180,181] have used open-circuit potential versus time measurements and cathodic reduction of rust patinas for the rapid laboratory evaluation of the performance of low-alloy weathering steels. The steel specimens are first exposed for 48 h to the vapor of an 0.001 mol dm^{-3} sodium bisulfite solution maintained at 54 °C (humid SO_2-containing atmosphere) to stimulate corrosion under atmospheric conditions. They are then subjected to two types of test: (a) open-circuit potential–time tests for periods up to 3000 s in either distilled

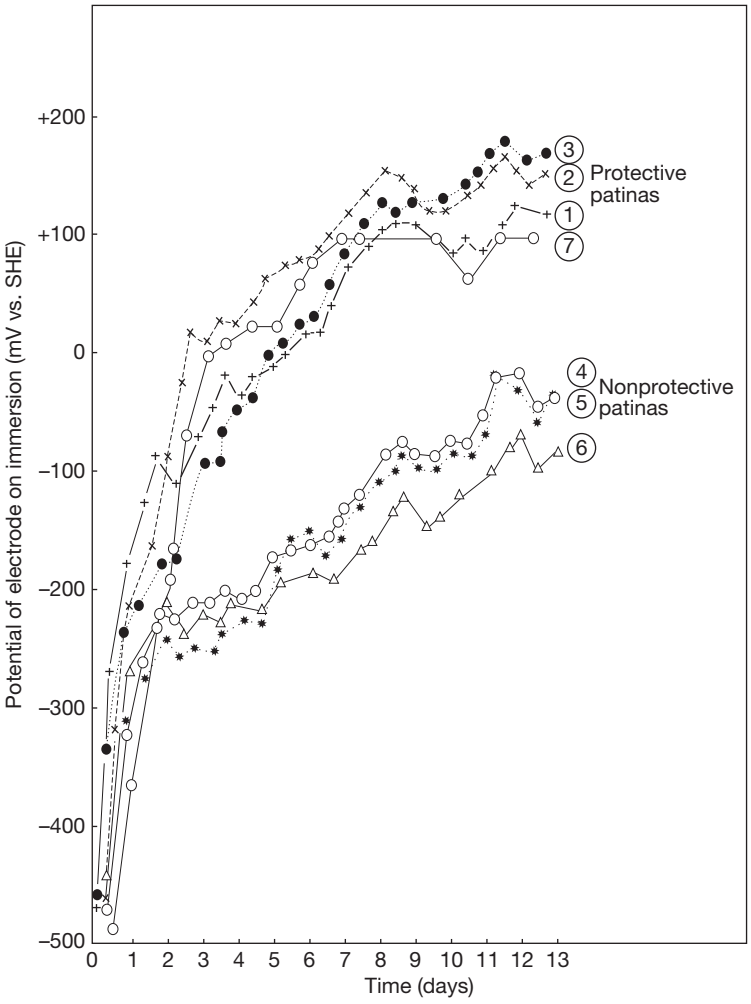


Figure 14 Relationship between potential (determined during initial immersion using the equipment in **Figure 13**) and time for steels that form protective (1, 2, 3, 7) or nonprotective (4, 5, 6) patinas. Reproduced from Pourbaix, M. CEBELCOR RT; **1969**, 160.

water or 0.1 mol dm^{-3} Na_2SO_4 and (b) cathodic reduction in 0.1 mol dm^{-3} Na_2SO_4 at 1 mA cm^{-2} c.d. In the cathodic reduction experiments, which provide a means of evaluating the degree of rusting,[182] both the potential and time are recorded, the onset of hydrogen evolution at constant potential being taken as the end point and giving the oxide-reduction time.

In order to evaluate the tests, determinations were carried out on the steels that had been exposed to the atmosphere for 1, 2, 3, 4, and 6-month periods. It was established that the initial open-circuit potential and the decrease in potential (more negative) with time varied with the nature of the steel and the time of exposure to the atmosphere, and the maximum negative potential was taken as a measure of corrosion resistance; the more negative the potential the lower the resistance of the alloy. In the case of three alloy steels that differed only in copper content it was found that the open-circuit potential was related to the corrosion rate as assessed by conventional weight loss.

A relationship was also established between the oxide-reduction time and time of exposure, and the results for a mild steel and a 1Cu–3Ni weathering steel were similar to those obtained by mass loss. The authors give various expressions that relate oxide-reduction time (min) with corrosion rate (mm $year^{-1}$), and claim that a short exposure to a laboratory SO_2 atmosphere followed by determining the E versus time and oxide-reduction time provides a rapid method of evaluating weathering steels.

2.34.10 Intergranular Attack of Cr–Ni–Fe Alloys

Early in the history of stainless steels it was recognized that they were highly susceptible to intergranular attack resulting from the precipitation of Cr–Fe carbides with the consequent depletion in the chromium content at grain boundaries when the alloy was heated in a specific range of temperature. It was necessary, therefore, to develop methods of testing that would detect susceptibility to intergranular attack as influenced by variations in processing and/or composition. As will be seen, most reagents used for these tests are highly aggressive, and it is important to note that an alloy found to be susceptible during testing will not necessarily be attacked intergranularly under the milder environmental conditions that may prevail in service.

Brown[183] has pointed out that Du Pont used evaluation tests for (a) as-received unstabilized alloys containing more than 0.03% C to check the effectiveness of the final heat treatment and (b) stabilized or special low-carbon grades after a sensitizing treatment (1 h at 677 °C) to determine whether susceptibility might develop during a subsequent welding operation.

Intergranular corrosion of Fe–Ni–Cr alloys has been the subject of a comprehensive review by Cowan and Tedmon[184] who summarized the various tests used for determining susceptibility (**Table 3**).

Of these tests, nos. 1–5, which are regarded as reliable test procedures by the ASTM, have been incorporated into ASTM A 262 'Standard Practices for Detecting Susceptibility to Intergranular Attack in Austenitic Stainless Steels' as follows:

Practice A – 10% oxalic acid, electrolytic etching at ambient temperatures;
Practice B – Boiling 50% H_2SO_4 + 25 g l^{-1} $Fe_2(SO_4)_3$;
Practice C – Boiling HNO_3;
Practice D – 10% HNO_3 + 3% HF at 70 °C;
Practice E – Boiling 16 wt% H_2SO_4 + 5.7% $CuSO_4$ + metallic copper.

It should be noted that although ASTM A 262 provides details of test procedures, no information is given on typical corrosion rates or acceptable limits for various heat-treated alloys, which are regarded as outside the province of a specification that describes test procedures. **Table 4**, taken from a paper by Brown,[183] shows the maximum acceptable evaluation test rates specified by the Du Pont Company for various alloys tested by the acid ferric sulfate test and by the Huey test. It should be noted that evaluation tests are specified by Du Pont when it is known or suspected that the environmental conditions in service are conducive to intergranular attack of susceptible material.

All the reagents used in Practice B to E have a high redox potential and Cowan and Tedmon[184] have presented schematic E–log i curves (**Figure 15**) showing the range of potentials of the various tests and the relative rates of attack on the matrix (Fe–18Cr–10Ni) and the chromium-depleted alloy at the grains boundaries, which has been assumed for this purpose to have a composition Fe–10Cr–10Ni, in a hot reducing acid. Although this diagram cannot show the effect of alloy composition, nature of test solution, conditions of test, etc. on intergranular attack, it serves to illustrate the electrochemical principles involved in the test procedures, all of which are based on reagents that attack the intergranular sensitized areas at a higher rate than the matrix; this may lead to the dislodgement of whole grains with a consequent high mass loss.

2.34.10.1 Boiling HNO_3 Test

This test, which is commonly referred to as the Huey test, was first described and used by Huey[185] in 1930, and since that time it has had wide application, particularly in the United States. The test consists of exposing the specimens (20–30 cm^2) to fresh boiling 65% HNO_3 (constant boiling mixture) for five successive periods of 48 h each under a reflux condenser. The specimens are cleaned and weighed after each period, and the corrosion rate (as a rate of penetration) is calculated for each period of test and for the average over the five periods; corrosion rates are expressed as millimeter per year. The reason for this procedure is the fact that Cr(VI) ions, produced from the oxidation of Cr^{2+} and Cr^{3+} by the HNO_3, if allowed to accumulate in the HNO_3 markedly increase its aggressiveness so that severe intergranular attack with grain dislodgement can occur even with solution-annealed steel free from precipitated carbides; hence, the necessity for the periodic changing of the solution and for a minimum ratio of solution volume to area of specimen (at least 20 ml HNO_3/cm^2 of stainless steel). Brown[183] points out that during a normal test the Cr(VI) content will not reach a level where an acceleration in rate occurs unless the specimen is in the sensitized condition, and under these circumstances the presence of Cr (VI) is an advantage in discriminating between sensitized and unsensitized material.

Table 3 Summary of chemical tests used for the determination of susceptibility to intergranular corrosion of iron–nickel–chromium alloys (Data from Cowan and Tedmon)

Test name	*Usual solution composition*	*Test procedure*	*Quantitative measure*	*Potential range (V vs. SHE)*	*Species selectivity attacked*
1. Nitric acid test[a]	65 wt% HNO_3	Five 48-h exposures to boiling solution; refreshed after period	Average mass loss per unit area of five testing periods	+0.99 to +1.20	1. Chromium-depleted areas 2. σ-phase 3. Chromium carbide
2. Acid ferric sulfate (Streichet) test[b,c]	50 wt% H_2SO_4 + 25 g l^{-1} ferric sulfate	120 h exposure to boiling solution	Mass loss per unit area	+0.7 to +0.9	1. Chromium-depleted areas 2. σ-phase in some alloys
3. Acid copper sulfate test[b,d]	16 wt% H_2SO_4 + 100 g l^{-1} $CuSO_4$ (+ metallic copper)	72 h exposure to boiling solution	1. Appearance of sample upon bending 2. Electrical resistivity change 3. Change in tensile properties	+0.30 to +0.58	Chromium-depleted area
4. Oxalic acid etch[b]	100 g $H_2C_2O_4{\cdot}2H_2O$ + 900 ml H_2O	Anodically etched at 1 A/cm^2 for 1.5 min	1. Geometry of attack on polished surface at × 250 or × 500	+1.70 to +2.00 or greater	Various carbides
5. Nitric-hydrofluoric[e] acid test	10% HNO_3 + 3% HF	4 h exposure to 70 °C solution	Comparison of ratio of mass loss of laboratory annealed and as-received samples of same material	Corrosion potential of 304 steel = +0.14 to + 0.54	1. Chromium-depleted areas 2. Not for σ-phase 3. Used only for Mo-bearing steels
6. Hydrochloric acid test[e]	10% HCl	24 h in boiling solution	1. Appearance of sample after bending around mandril 2. Mass loss per unit area	(a) Redox potential = + 0.32 (b) Corrosion potential = − 0.2 ± 0.1	1. Alloy-depleted area 2. Not for σ-phase
7. Nitric acid Cr^{6+} test[f]	5N H_2SO_4 + 0.5 N KCr_2O_2	Boiling with solution renewed every 2–4 h for up to 100 h	1. Mass loss per unit area 2. Electrical resistivity 3. Metaflographic examination	(a) Redox potential = +1.37 (b) Corrosion potential of 304 steel = +1.21	Solute segregation to grain boundaries

[a]Data after Cowan and Tedman.[184]
[b]A262-02a (2008) and Practice G 28-02 (2008).
[c]M. A. Streicher, ASTM Bulletin No. 229, 77–86 (1958) G28-1985.
[d]A262-02a (2008).
[e]D. Warren, ASTM Bulletin No. 230, 45–56 (1958).
[f]J. S. Armiju, *Corrosion*, **24** (1968).

Maximum corrosion rates used by Du Pont for various alloys are given in **Table 4**, and most users of the test consider average corrosion rates of 0.46–0.61 and 0.76 mm year^{-1} to represent the upper limits for satisfactory resistance for wrought austenitic alloys and cast austenitic alloys, respectively. Streicher[148] considers that if the corrosion rate for each period increases over that for the previous period the alloy is susceptible.

The mechanism of the corrosion reaction is not clear, particularly in view of the changes in composition of the HNO_3 that take place during the 48 h period of the test. Streicher[148] reports that the corrosion potential of the steel ranges from 1.00 to 1.20 V (versus SHE) during the test owing to the accumulation of Cr(VI), and it can be seen from **Figure 15** that the sensitized areas will have a higher corrosion rate than the matrix throughout this potential range, although they will become similar at the higher potentials. The high corrosion rates obtained in the test are due partly to intergranular attack and partly to the undermining and dislodgement of grains.

Stainless steels and Ni-base alloys containing Mo, such as Type 316L (0.03% C max.) and Hastelloy C, are found to give very high corrosion rates in the HNO_3 test even when they are immune to intergranular attack when subjected to other tests that reveal sensitization due to chromium-depleted zones; furthermore, such alloys even after being subjected to a sensitizing heat-treatment do not give rise to intergranular attack in most conditions of service. This high corrosion rate is considered to be due to the formation of a submicroscopic σ-phase, and although positive proof is not available, its presence is substantiated by the fact that the phase becomes identifiable after longer periods of sensitizing temperatures, although in this form it has little effect on the corrosion rate. It would appear that the σ-phase dissolves rapidly during the HNO_3 test, and since it has a high chromium content, the solution becomes enriched in Cr(VI) with a consequent increase in the corrosion rate of the alloy. It follows that the test is unsuitable for evaluating the behavior of stainless steels that may precipitate σ-phase, unless the alloy is to be used in service for nitric acid plant.

Henthorne,[186] in considering the corrosion testing of weldments, points out that the test will also give high rates due to (a) end-grain attack, which is particularly prevalent in resulfurized or heavily cold-worked material and (b) dissolution of Ti(C, N) such as occurs in Type 321 weldments and leads to knife-line attack. Since most service conditions do not cause attack on the alloy in these conditions the test can be misleading.

Thus, under the circumstances already outlined, the test can be misleadingly severe, but it is

Table 4 Maximum acceptable evaluation test rates specified by Du Pont for services where susceptible material would be intergranularly attacked

Type	*Condition*	*Max. corrosion rate (mm year^{-1})*
120 h acid $Fe_2(SO_4)_3$ test (ASTM A.262, Practice *B*)		
304	As received	1.22
304L	20 min at 677 °C	1.22
316	As received	1.22
316L	20 min at 677 °C	1.22
317L	20 min at 677 °C	1.22
CF-8	As received	1.22
CF-8M	As received	1.22
240 h HNO_3 test (ASTM A-262, Practice *C*)		
304	As received	0.457
304L	20 min at 677 °C	0.305
304L	1 h at 677 °C	0.610
309S	As received	0.305
316	As received	0.457
347	1 h at 677 °C	0.610
CF-8	As received	0.610
CF-8M	As received	0.762

Source: Brown, M. H. *Corrosion* **1974**, *30*, 1.

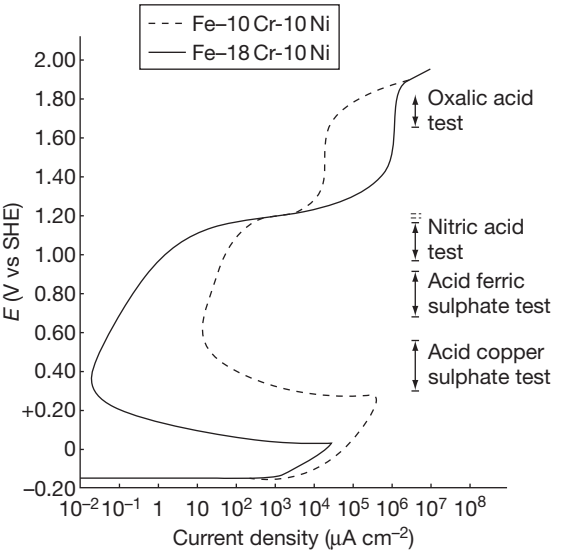


Figure 15 Schematic representation of ranges of corrosion potential expected from various chemical tests for sensitization in relation to the anodic dissolution kinetics of parent (Fe–18Cr–10Ni) and grain boundary (assumed Fe–10Cr–10Ni) material in a hot, reducing acid. Reproduced from Cowan, R. L.; Tedmon, C. S., Jr. In Fontana, M. G., Staehle, R. W., Eds.; *Advances in Corrosion Science and Technology*; Plenum Press: New York, 1973; Vol. 3.

particularly valuable for evaluating alloys for use in HNO_3 or in other strongly oxidizing acid solutions to ensure that they have received the correct heat treatment and have an appropriate composition, that is, a low carbon content or the correct ratio of (Ti or Nb)/C.

2.34.10.2 Boiling $H_2SO_4 + CuSO_4$ Tests

The use of boiling $H_2SO_4 + CuSO_4$ for detecting intergranular sensitivity was first described by Strauss, *et al.*[187] in 1930, and is frequently referred to as the Strauss test, although the conditions of the test have been modified; whereas the Huey test is most widely used in the United States the Strauss test has been the preferred test in Europe. The test is mild compared to the Huey test and intergranular attack takes place with little grain dislodgement.

The use of metallic copper chips placed in contact with the steel to speed up the test and thus decrease the time of testing was first described by Rocha,[188] and subsequent work by Streicher[148] showed that its presence significantly increased the rate of intergranular attack even when it was not in contact with the steel. Approximate mass losses for a sensitized Type 316 stainless steel during a 240 h testing in boiling $H_2SO_4 + CuSO_4$ are as follows:

No metallic copper present 0.1 g dm^{-2};
Metallic copper present but not in contact with the steel 1.0 g dm^{-2};
Metallic copper in contact with steel 4.0 g dm^{-2}.

As used in Germany the composition of the solution is 110 g $CuSO_4 \cdot 5H_2O$, 100 ml H_2SO_4 (spec. grav. 1.84) and 1 l of water, the test being conducted for 168 h in the boiling solution. The ASTM Tentative Procedure A393–63T specified a similar composition containing 100 g $CuSO_4 \cdot 5H_2O$, 100 ml H_2SO_4 (spec. grav. 1.84) with water added to make a total volume of 1 l. The test time was 72 h, and with the high carbon contents of the earlier steels this was adequate for detecting susceptibility. However, with the decrease in the carbon contents of stainless steels a more prolonged boiling time was found to be necessary, and Scharfstein and Eisenbrown[189] showed that a Type 304 stainless steel containing 0.068% C would pass the 72-h Strauss test even after a sensitizing treatment of up to 4 h at 677 °C. For this reason, A393–63T has been discontinued and in ASTM A 262 Practice E, the specimens are placed in contact with metallic copper chips to increase the rate of intergranular attack.[148,183] This test is of comparable sensitivity to the other tests, and is far more discriminating than the older tentative standard; furthermore, it is more severe so that the testing time is decreased from 72 to 24 h. This test has been incorporated in the international standard ISO 3651-2.

Figure 15 shows that the corrosion potential of stainless steel in the $H_2SO_4 + CuSO_4$ test lies in the range 0.30–0.58 V, and that while the corrosion rate of the unsensitized alloy is approximately 10^{-1} μA cm^{-2}, that of the sensitized material is 10 μA cm^{-2}; for heavily sensitized material the ratio of rates[184] of sensitized:unsensitized alloy may be as high as 10^5:1. This large difference in rates leads to rapid attack, which is confined to the depleted zone having a thickness of the order of 1 μm, and under these circumstances there will be little grain dislodgement. Thus, the mass change will be so small that it cannot be used as a criterion of susceptibility. For this reason assessment of intergranular attack is normally carried out (ASTM 262 Practice E) by bending the specimen around a mandrel through 180° and inspecting the bend surface for cracks. Measurements of changes in electrical resistivity[146] and in ultimate tensile strength[190] are used as quantitative methods of assessment, but according to Ebling and Scheil[191] they are not as discriminating as the qualitative bend test.

The H_2SO_4–$CuSO_4$ test, unlike the Huey test, is specific for susceptibility due to chromium depletion and is unaffected by the presence of submicroscopic σ-phase in stainless steels containing molybdenum or carbide stabilizers. It can be used, therefore, with confidence to test susceptibility in austenitic (300 series) and ferritic (400 series) stainless steels and in duplex austenoferritic stainless steels such as Types 329 and 326.

The mechanism of the action of metallic copper was investigated by Streicher[148] who determined the potential of a Type 314 stainless steel, the redox potential of the solution (as indicated by a platinized-Pt electrode) and the potential of the copper. The actual measurements were made with a saturated calomel electrode, but the results reported below are with reference to SHE. In the absence of copper the corrosion potential of the stainless steel was 0.58 V, whereas the potential of the Pt electrode was approximately 0.77 V. When metallic copper was introduced into the solution (not in contact with the steel) both the corrosion potential of the steel and that of the Pt electrode attained the same high negative potential of 0.37 V, the copper attaining a steady value of 0.30 V. Finally, when the stainless steel was placed in contact with the copper it took up

a more negative potential of 0.30 V, the potential of the copper being unaffected. These potentials have been interpreted by Streicher and have been expressed in *E–I* diagrams (**Figure 16**) showing how the corrosion potential and the corrosion rate varies with conditions of the test. Introduction of metallic copper into the solution results in the disproportionation reaction:

$$Cu + Cu^{2+} \leftrightarrows 2Cu^{+}$$

and the accumulation of the Cu^{+} ions in the solution produces a decrease in the polarization of the local anodes on the stainless steel, which are polarized to the redox potential of the solution (**Figure 16**). Contact of copper with the stainless steel results in a further decrease in the corrosion potential of the stainless steel to that of copper, indicating that cathodic polarization of the steel has occurred since the steel is the cathode of the stainless steel/copper bimetallic couple. Simultaneously, there is a reduction in the anodic polarization of the susceptible grain boundaries, and a consequent increase in the corrosion rate (**Figure 16**). Thus, contact of the steel with the copper results in intergranular attack of the steel at constant potential, the copper acting in the same way as a potentiostat.

The Huey test is widely used in the United States while the H_2SO_4–$CuSO_4$ test is preferred in the United Kingdom, with an increasing tendency to use the metallic copper variant. The H_2SO_4–$CuSO_4$–Cu test procedure is given in BS EN ISO 3651-2.

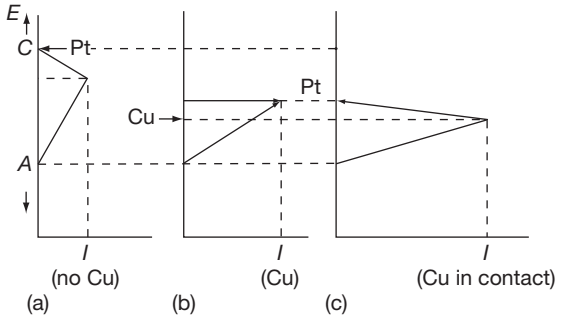


Figure 16 Schematic *E–I* diagrams of local cell action on stainless steel in $CuSO_4$ + H_2SO_4 solutions showing the effects of metallic copper on corrosion rate. *C* and *A* are the open-circuit potentials of the local cathodic and anodic areas and *I* is the corrosion current. The potentials of platinized platinum and copper in the same solution are indicated by arrows. (a) Corrosion of stainless steel in isolation. (b) The rate when copper is introduced into the acid without contact with the steel. (c) The rate when copper contacts the steel. Reproduced from Streicher, M. A. *J. Electrochem. Soc.* **1959**, *106*, 161.

2.34.10.3 HNO_3–HF Test

This test was first described by Warren[192] in 1958, and consists of two 2-h periods in 10% HNO_3 + 3% HF solution at 70 °C using fresh solution for each period. The test is therefore more rapid than the others, and it is specific for chromium depletion by carbide precipitation since it is unaffected by the submicroscopic σ-phase formed in molybdenum-bearing steels; as described in A262, its use is confined to Types 316, 316L, 317, and 317L stainless steels.[183] Since the corrosion rates of stainless steels in the acid are high and vary greatly from test to test, it is necessary to run two tests and to compare the corrosion rates of the specimen to be evaluated ('as received' for Types 316 and 317 and in the sensitized condition for Types 316L and 317L) and another laboratory-annealed specimen of the same alloy shown to be free from precipitated carbides by the step structure produced after electrolytic etching in oxalic acid. Intergranular attack is assessed by the rate of penetration evaluated from the mass loss, and if the mass loss of the specimen to be evaluated is greater than 1.5 times that of the standard, the former is considered to be susceptible.

The solution has a low redox potential and the corrosion potentials for austenitic stainless steels will be in the range 0.14–0.54 V, according to composition. Thus, it can be seen from **Figure 15** that all but the highest chromium steels will be in the active region, so that the test relies on vigorous corrosion of the grain boundary zones while the matrix remains somewhat passive and corrodes at a slower rate.[184] Although the test gives constant and reliable results, it has not been used widely for routine evaluations for the following reasons: (a) the need to use a ratio of two test rates, (b) inconvenience of handling solution containing HF, and (c) the availability of the $H_2SO_4 + Fe_2(SO_4)_3$ test.

2.34.10.4 $H_2SO_4 + Fe_2(SO_4)_3$ Test (Streicher Test)

This was described in 1959 by Streicher,[148] and consists of one period of exposure to a boiling solution of 50 mass% $H_2SO_4 + 25\ g\ l^{-1}$ $Fe_2(SO_4)_3$ for 120 h, assessment being based on mass loss (see **Table 4**). Streicher, however, usually reports a ratio of mass

loss of sample to be assessed/weight loss of annealed sample, and as for the HNO_3–HF test considers that a ratio >1.5–2.0 indicates susceptibility; for Type 304 Streicher considers a rate $>0.76\,mm\,year^{-1}$ to indicate susceptibility, but Brown considers a higher figure to be acceptable (see **Table 4**).

Accumulation of corrosion products does not stimulate attack and so several specimens may be tested in the same solution, but additional $Fe_2(SO_4)_3$ may have to be added (or the solution changed) if there is considerable attack on severely sensitized specimens, as indicated by a color change of the solution from brown to dark green.

The redox potential of the solution is that of the Fe^{3+}/Fe^{2+} equilibrium and lies within the range 0.80–0.85 V (versus SHE). The severe weight loss of susceptible alloys is due to undermining and grain dislodgement at the sensitized zones, which occurs at about twice the rate of that in the Huey test. Another difference is that while in the Huey test corrosion products [Cr(VI)] increase the rate by raising the potential of the alloy into the transpassive region, the converse applies in the acid ($Fe_2(SO_4)_3$) test, since reduction of Fe^{3+} to Fe^{2+} during the test will result in a decrease in the redox potential and the whole sample will corrode with hydrogen evolution.

According to Cowan and Tedmon[184] the test can selectively attack some types of σ-phase. Those of Types 321 and 347 are readily attacked, whereas the molybdenum-bearing σ-phase of Type 316 is unattacked. The test will also show Hastelloys and Inconels to be susceptible to intergranular attack when there are either chromium- (or molybdenum-) depleted grain boundaries or grain-boundary σ-phase present. Ferritic (200 series) and austenoferritic stainless steels can also be tested for chromium-depletion sensitization in this reagent, but whether σ-phases formed in these alloys affect the test has not been established.

In conclusion it must be emphasized again that all the tests used are accelerated tests and only provide information on susceptibility to intergranular attack under the prevailing precise test conditions. They are quality control tests that may be used to demonstrate either that heat treatment has been carried out adequately or that a steel will withstand the test for a certain sensitizing heat treatment.

2.34.10.5 Electrolytic Oxalic Acid Etching Test

This test, which was developed by Streicher,[148] is used as a preliminary screening test to be used in conjunction with the more tedious testing procedures such as the boiling HNO_3 test. The specimens are polished (3/0 grit paper) and then anodically polarized for 1.5 min at $1\,A\,cm^{-2}$ at room temperature in a solution prepared by dissolving 100 g of $H_2C_2O_4{\cdot}2H_2O$ in 900 ml of distilled water. The surface is then examined at about ×500 magnification and the structure is classified as 'step,' 'ditch,' or 'dual' (both 'step' and 'ditch'). If the surface shows a 'step' structure it is immune to intergranular attack and no further testing is necessary; if the structure is 'ditch,' further testing by the Huey test or some other chemical test is necessary; if 'dual' further testing may be necessary. Thus, the test, by identifying structures that are immune to intergranular attack, eliminates unnecessary testing, although where a 'ditch' (or possibly a 'dual' structure) is obtained, final confirmation by the Huey test is essential. **Figure 17** shows the 'ditch' and 'step' structures diagrammatically, and **Figure 18** shows photomicrographs of these structures and a 'dual' structure.[193]

The test operates at a potential above 2.00 V (versus SHE), and the 'ditch' structure obtained with sensitized alloys must be due, therefore, to the high rate of dissolution of the sensitized areas as compared with the matrix. The 'step' structure is due to the different rates of dissolution of different crystal planes, and the 'dual' structure is obtained when chromium carbides are present at grain boundaries, but not as a continuous network.

2.34.10.6 Electrochemical Tests

The difficulties associated with the ASTM 'Standard Practices for Detecting Susceptibility to Intergranular Attack in Austenitic Stainless Steels' (A 262) are that the methods are destructive and qualitative in nature. Early attempts to develop quantitative, nondestructive electro-chemical techniques to detect

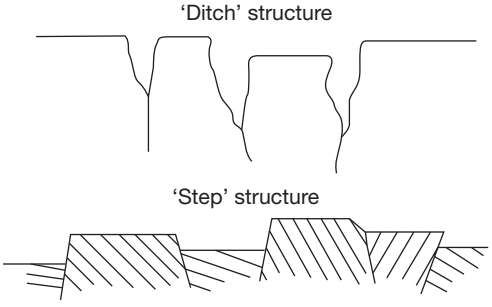


Figure 17 'Ditch' and 'step' structures. Reproduced from Streicher, M. A. *J. Electrochem. Soc.* **1959**, *106*, 161.

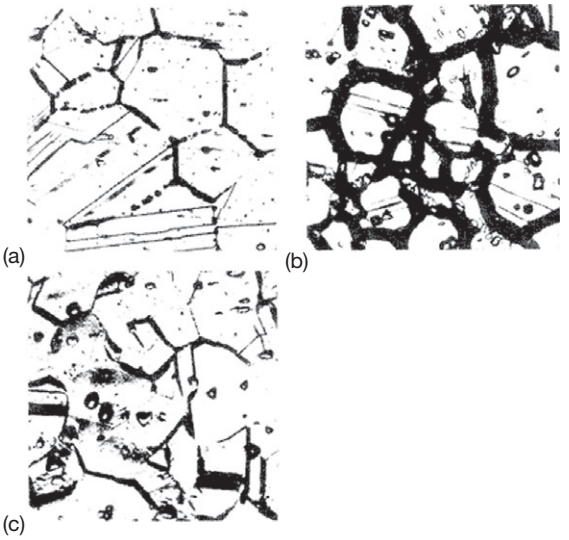


Figure 18 Micrographs of (a) step, (b) ditch, and (c) dual structures. Reproduced from Streicher, M. A. *J. Electrochem. Soc.* **1959**, *106*, 161; Cowan, R. L.; Tedmon, C. S., Jr. In Fontana, M. G., Staehle, R. W., Eds.; *Advances in Corrosion Science and Technology*; Plenum Press: New York, 1973; Vol. 3.

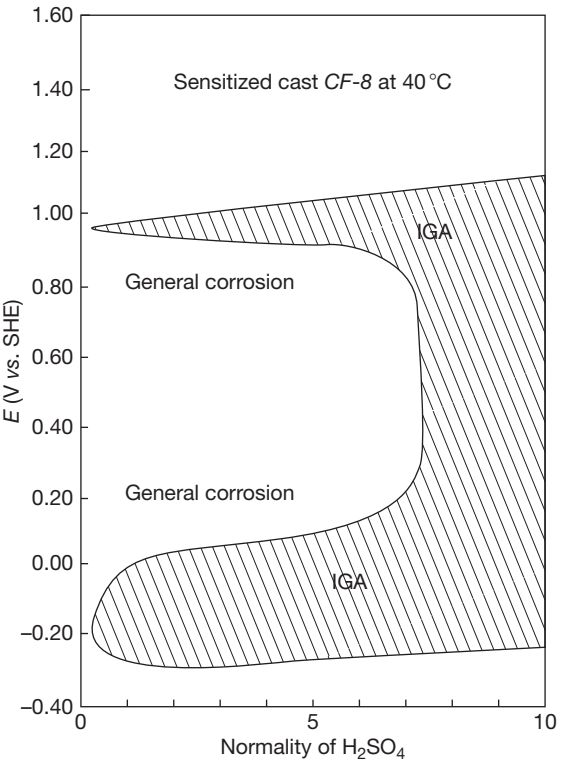


Figure 19 Intergranular corrosion plot for a sensitized cast CF-8 stainless steel (0.08%maxC 8–11%Ni 18–21%Cr) in H_2SO_4 at 40°C as a function of potential and acid concentration. Reproduced from France, W. D.; Greene, N. D. *Corros. Sci.* **1968**, *8*, 9.

sensitization by Clerbois *et al.*[194] employed potentiostatic techniques and it was observed that sensitized 18–8 stainless steel when anodically polarized potentiostatically in 1.0 mol dm^{-3} H_2SO_4 gave rise to a secondary active peak in the range 0.14–0.24 V (versus SHE) that was not present in the curve for the annealed alloys. This observation was criticized by France and Greene,[195] who consider that the active peak is due to the dissolution of Ni that had accumulated at the surface during active dissolution at lower potentials. Clerbois, *et al.*[194] also noted that if a sensitized sample is held at 0.14 V in 1.0 mol dm^{-3} H_2SO_4 for 24 h and then bent around a mandrel, it fissures and cracks, and it can be seen from **Figure 15** that at this potential the chromium-depleted grain boundary will corrode actively, whereas the matrix will be passive. The potentiostatic test using cracking to detect susceptibility is thus analogous to the acid-copper sulfate test.

France and Greene[195] proposed that it should be possible to predict service performance by potentiostatic studies of steels in the environments encountered in practice coupled with metallographic examination of the surfaces. They argued that many environments do not selectively attack the grain boundaries of sensitized stainless steels so that the use of costly preventative measures is unnecessary. Since the intergranular attack of austenitic stainless steels occurs only in limited potential regions it should be possible to predict service performance provided these regions are precisely characterized.

In their studies, specimens of different sensitized steels were held at various constant potentials in different concentrations of the acid under study at various temperatures and the surfaces were then examined metallographically for intergranular attack. Data obtained in this way enabled E-concentration of acid diagrams to be produced showing the zones of general corrosion, fine intergranular corrosion and coarse intergranular corrosion for a given sensitized stainless steel in a given acid at various constant temperatures (**Figure 19**).

Streicher,[196] however, considered this approach to be unsound and pointed out that the short duration of the potentiostatic studies carried out by France and Greene cannot be used to predict long-term behavior in service. The prolonged dialogue between these workers[197,198] was well summarized in the review article by Cowan and Tedmon[184] who concluded that these particular potentiostatic tests cannot be

regarded as accelerated tests for service environments and that predicting future industrial service for periods longer than the test is not advisable.

2.34.10.7 EPR Test

The electrochemical potentiokinetic reactivation (EPR) test was proposed by Cihal *et al.*[199] and developed by Novak and others[200–202] as a fast, quantitative and nondestructive technique for establishing the degree of sensitization of austenitic stainless steels.

The test is accomplished by a potentiodynamic sweep from the passive to the active regions of electrochemical potential (a process referred to as reactivation) for a given alloy in a specific electrolyte, during which the amount of current resulting from the corrosion of the chromium-depleted regions surrounding the precipitated chromium carbide particles is measured. In a sensitized microstructure, the bulk of these particles are located at the grain boundaries and are particularly susceptible to corrosion in oxidizing acids. Proposed national and international standards on EPR testing specify 0.5 M H_2SO_4 + 0.01 M KSCN at 30°C as the EPR test environment for sensitized austenitic stainless steels.

Three different forms of EPR test can be employed, designated as the single loop, double loop and reactivation ratio methods in Figure 20.

2.34.10.7.1 Single loop EPR test

The single loop method requires the sample to be polished to a 1 μm finish and then passivated at +200 mV (SCE) for 2 min following which the potential is decreased at 1.67 mV s^{-1} until the corrosion potential of approximately −400 mV (SCE) is reached. The reactivation process results in the preferential breakdown of the passive film in the chromium-depleted grain boundaries of sensitized material and an increase in the current through the cell. The area under the E versus log I curve (Figure 20a) is proportional to the electric charge, Q, measured during the reactivation process. On nonsensitized materials, the current density during the reactivation step is very low because the passive film remains essentially intact. A measure of the degree of sensitization is obtained by calculating the normalized charge, P_a, where:

$$P_a(\mathrm{C\,m^{-2}}) = Q/A$$

where Q = integrated charge during the reactivation scan, and A = grain boundary area (5.1×10^{-3} exp 0.35 G, where G is the ASTM grain size at 100 × magnification).

Pitting caused by the dissolution of nonmetallic inclusions can increase the P_a value. Consequently, the microstructures of specimens with a high P_a value must be examined to identify the source of the elevated value. In general, P_a values below 0.10 are characteristic of unsensitized microstructures, while sensitization is indicated if P_a exceeds 0.4. Single loop tests are sensitive to mild degrees of sensitization but do not readily distinguish between medium and severely sensitized materials.

2.34.10.7.2 Double loop EPR test

Details of this procedure are given in Japanese Industrial Standard JIS G 0580. The sample is ground to a 100 grit finish then placed in the test solution for about 2 min to establish the rest potential (about −400 mV (SCE) for AISI Types 304 and 304L stainless steel).

The sequence of polarization steps is shown in Figure 20b. The surface is first polarized anodically from the corrosion potential to +300 mV (SCE) at a rate of 1.67 mV s^{-1}. As soon as this potential is reached, the scanning direction is reversed and the potential is decreased at the same rate to the corrosion potential. The ratio of the maximum current in the reactivation loop, I_r, to that in the larger anodic loop, I_a, is used as a measure of the degree of sensitization.

2.34.10.7.3 Reactivation ratio EPR test (Figure 20c)

This is a simpler and more rapid method than the single or double loop tests, and depends on the fact that the value of I_a determined during the anodic scan of a double loop test (which produces general dissolution without intergranular attack on sensitized material) is essentially the same for all AISI Type 304 and 304L steels.

The specimen is ground to a 100 grit finish then, after 2 min at the corrosion potential (about −400 mV (SCE)), it is conditioned by a 2-min treatment at −230 mV (SCE) in order to eliminate the need for polishing prior to the reactivation procedure. Passivation is then accomplished at +200 mV (SCE) for 2 min after which the specimen is reactivated by scanning back to the corrosion potential at 1.67 mV s^{-1}. During this reactivation scan, the maximum current, I_r, is measured and is divided by the surface area as an indication of the degree of sensitization.

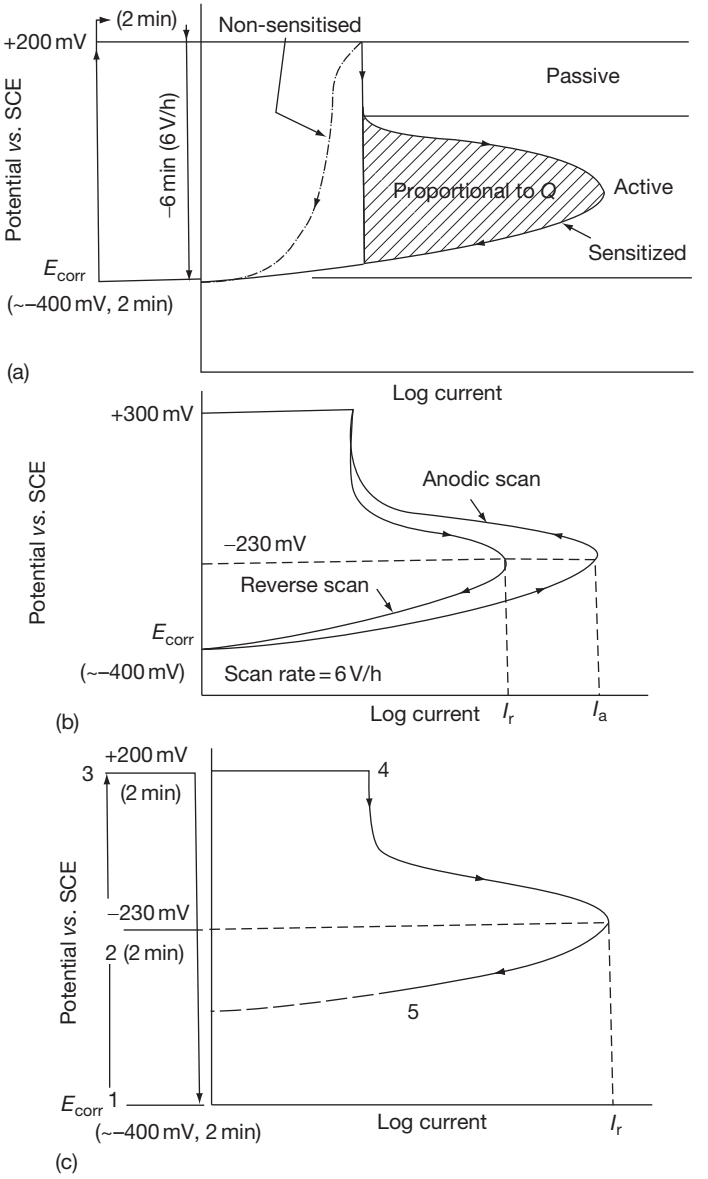


Figure 20 Schematics of reactivation polarisation curves. (a) Single loop EPR test method, and (b) double loop EPR test method, and (c) reactivation ratio EPR test method.

2.34.10.8 EPR Tests for Ferritic Stainless Steels

Lee[203] has demonstrated that in slightly modified forms the single loop EPR test can be used to quantify the degree of sensitization in ferritic stainless steels. For AISI Types 430, 430Ti, 430Nb, and 446 stainless steels, the test consists of passivating the specimen in deaerated 3 N H_2SO_4 solution at 30 °C at +400 mV (SCE) for 10 min, followed by a reactivation at a scan rate of 250 mV min^{-1}. The EPR test for AISI Type 434 stainless steel requires a reactivating scan rate of 150 mV min^{-1} (the other test conditions remaining unchanged). For AISI Type 444 stainless steel, the test is conducted in deaerated 5 N H_2SO_4 solution at 30 °C and involves passivation at +400 mV (SCE) for 2 min followed by a reactivation at a scan rate of 100 mV min^{-1}.

2.34.11 Crevice Corrosion and Pitting

Crevice corrosion and pitting are dealt with in some detail elsewhere, and it is not appropriate here to

discuss the nature of the phenomena nor the methods that have been used to determine the mechanisms of these forms of localized attack. However, it should be noted that many of the methods of testing follow directly from the concepts such as E_b (the critical pitting potential), E_p (the protection potential) have been investigated by a number of workers as possible criteria for the resistance of metals and alloys to pitting and crevice corrosion in service. It should also be noted that since crevice corrosion and pitting have similar mechanisms and since the presence of a crevice is conducive to pitting of alloys that have a propensity, to this form of attack, it is appropriate to consider them under the same heading. (Editor's note: A group of tests that have gained favor in recent years are based on the observation that, providing the solution is sufficiently oxidizing (or, equivalently, the potential is held at a suitably positive value), there is a critical temperature below which pitting will not occur, and a (more negative) critical temperature below which crevice corrosion will not occur. These are known as the critical pitting temperature (CPT) and critical crevice temperature (CCT), respectively, and the higher the temperature, the greater the resistance to pitting or crevice corrosion. There are various methods of measuring these temperatures (see ISO 17864, ASTM G 48, and G 150).)

In general, the tests may be classified as follows:

1. Laboratory tests, in which the specimen is immersed in a solution conducive to pitting, such as an acidified $FeCl_3$ solution (redox potential above the critical pitting potential E_b).
2. Laboratory tests, in which the specimen is anodically polarized in a chloride-containing solution to evaluate E_b and E_p.
3. Field tests, in which the specimen (with or without a crevice) is exposed to the environment that it will encounter in service.

As far as tests for crevice corrosion are concerned, all that is required is a geometrical configuration that simulates a crevice, which may be achieved in a variety of ways using either the metal itself or the metal and a nonmetallic material. Streicher[204] studied the crevice corrosion of Cr–Ni–Fe alloys, in which two plastic cylinders are held on the two opposite faces of a sheet metal specimen by two rubber bands, thus providing three different types of crevice in duplicate. A simple method of testing for crevice corrosion produced by contact with different materials is to use a horizontal strip of the metal under study and place on its upper surface at intervals small piles of sand, small piles of sludge, pieces of gasket material, rubber, etc. More precise crevices can be produced by bolting together two discs of the metal, which are machined on the facing surfaces so that there is a flat central portion followed by a taper to the periphery of the disc, the flat central portion providing a very fine crevice and the tapered portion a coarser one.[205]

Figure 21 shows the types of crevices used by Wilde[206] for studying crevice corrosion and pitting of Cr–Ni–Fe alloys in the laboratory and in the field. Types 1 and 5 were used for anodic polarization studies in nitrogen-saturated 1 mol dm^{-3} NaCl and in aerated 3.5 wt% NaCl, respectively, and it can be seen that attachment to the conducting lead is by means of a Stern-Makrides pressure gasket; Types 3 and 4 were used for field tests in seawater for periods up to 4½ years; Type 2 was used for laboratory studies in which the specimens were immersed in acidified $FeCl_3$ (108 g l^{-1} $FeCl_3{\cdot}6H_2O$ with the pH adjusted to 0.9 with HCl).

The value of electrochemical evaluation of the critical pitting potential as a rapid method of determining pitting propensity is controversial. France and Greene[207] studied the pitting of a ferritic steel (Type 430) using a controlled potential test in 1 M NaCl and a conventional immersion test in oxygen-saturated 1 M NaCl but found that at the same potential (−0.17 to 0.09 V versus SCE) the corrosion rates were 390 and 5.2 mm $year^{-1}$, respectively. Similar studies were carried out on Zr using 0.5 M H_2SO_4 + 1 M NaCl for the controlled potential test and 0.5 M $H_2SO_4 + FeCl_3{\cdot}6H_2O$ for the immersion test, and again the former gave a much higher corrosion rate than the latter. France and Greene conclude that these two types of test give rise to significantly different results under identical test conditions. To explain the results obtained with the ferritic stainless steel, they pointed out that, during the controlled potential test, the anodic reaction occurs at the metal's surface whereas the interdependent cathodic reaction takes place at the counter-electrode. Under these circumstances, the metal ions produced anodically result in increased migration of Cl^- to maintain electroneutrality, and this in turn results in a higher concentration of Cl^- at the metal–solution interface with consequent increase in the rate of pitting. A similar situation does not arise during the immersion test where the anodic and cathodic sites are in close proximity, and charge balance is maintained without the migration of Cl^- from the bulk solution.

Potentiostatic tests[208–210] have been used, and Wilde and Williams[208] in potentiokinetic studies of

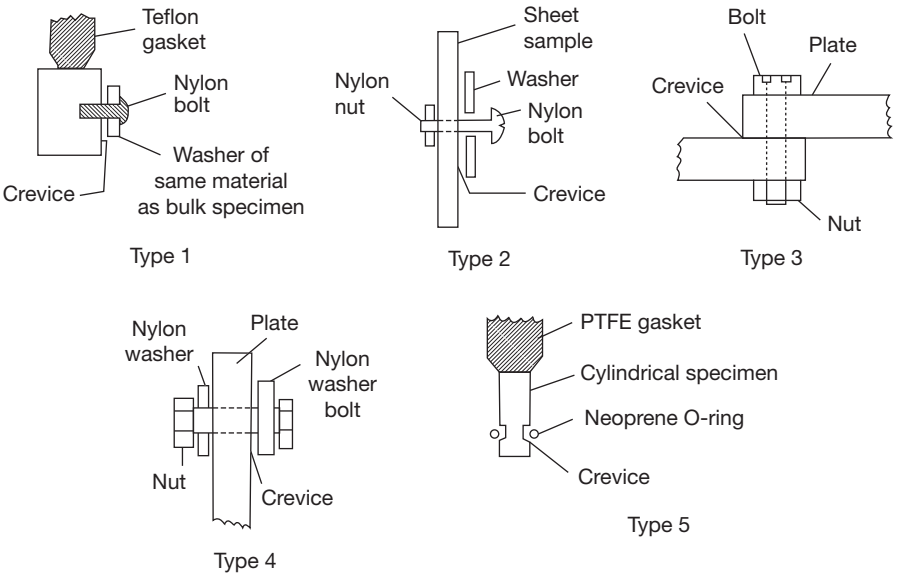


Figure 21 Various types of crevices used for investigating crevice corrosion of stainless steels. Reproduced from Wilde, B. E. *Corrosion* **1972**, *28*, 283.

the critical breakdown potential of stainless steels (Types 430 and 304) in 1.0 M NaCl showed that the nature of the gas used to purge the solution has a pronounced effect on the value of E_b (**Table 5**). In particular, they have established that the presence of dissolved O_2 enhances passivity thus causing E_b to become more positive, and consider that this explains the failure of France and Greene to obtain accord between controlled potential tests in hydrogen-saturated chloride solutions and immersion tests in oxygenated chloride solutions at the same potentials.

Wilde and Williams[208] have used the redox system 0.1 M $Fe(CN)_3^{6-}/Fe(CN)_4^{6-}$ for their immersion tests, which for Type 403 stainless steel gives a corrosion potential of −0.100 V (versus SCE); selection of this system was based on the premise that being large anions they would be less likely than dissolved O_2 to be involved in the adsorption processes that stabilize the passive state. Pitting occurred within 60 s, and equivalent tests on the same alloy conducted potentiostatically at −0.1 V (versus SCE) in hydrogen-saturated 1.0 M NaCl gave similar results. They conclude that these two tests give comparable results, but that extreme caution must be used in utilizing E_b as an index of pitting, since its value is dependent upon environmental variables and in particular the nature of the dissolved gas in the corrodent. Wilde and Williams[211] have also shown that the critical pitting potential can be used to predict the behavior of alloys exposed for long periods to seawater or to industrial chemical environments.

Table 5 Variation in *Eb* (V) for stainless steels in 1.0 mol dm^{-3} NaCl at 25 °C with nature of dissolved gas (*Eb* vs. SCE)

Gas	*Type 430 stainless steel*	*Type 304 stainless steel*
Hydrogen	−0.185	−0.050
Nitrogen	−0.130	−0.020
Argon	−0.100	+0.050
Oxygen	−0.035	+0.065

Data after Wilde and Williams[208]

In a subsequent paper, Wilde[206] pointed out that although E_b is qualitatively related to resistance of a material to breakdown of passivity and pit initiation, it is of questionable value in predicting performance when crevices are present. Wilde found that although the Fe–30Cr–3Mo alloy appeared to indicate total immunity to breakdown when tested anodically in 1 M NaCl and in the freely corroding condition in 10% $FeCl_3$, it pitted within the crevice when an artificial crevice was present. Exposure in seawater for a 16-month period showed that AISI Types 304 and 316 stainless steels and the Fe–30Cr–3Mo alloy all pitted to the same extent when a crevice was present, although the former two alloys are considered to be less resistant to pitting than the Fe–30Cr–3Mo alloy. Pourbaix *et al.* have defined the protection

potential E_p as the potential below which no pits can initiate and preexisting pits cannot propagate, since they are passive at that potential. However, Wilde, using cyclic potentiodynamic sweeps at varying sweep rates, has established that E_p is not a unique parameter and that it varies in a semilogarithmic manner with the extent of localized attack produced during the anodic polarization, that is, E_p-log(extent of pit propagation) is linear. Thus, at a sweep rate of $10\,V\,h^{-1}$, E_p was found to be $-0.290\,V$ (versus SCE) while it fell to a more negative value of $-0.410\,V$ at the slower sweep rate of $1\,V\,h^{-1}$. This was explained by Wilde as being due to the chemical changes that occur in the growing pit by hydrolysis of corrosion products and by the increased migration of Cl^- ions. Since E_p is a variable that depends upon experimental procedures, it cannot be used on its own as a criterion for protection against the propagation of preexisting pits or crevices in an engineering structure. Wilde considers that a more useful parameter appears to be the 'difference potential' ($E_b - E_p$), which is used as a rough measure of the hysteresis loop area produced during the cyclic determination of E_b and E_p. (Editor's note: The parameter $E_b - E_p$ has recently formed the basis of a predictive model of localized corrosion of Alloy 22 for the storage of nuclear waste over periods of up to a million years.) The area of the hysteresis loop obtained in a potentiodynamic sweep using a specimen with an artificial crevice provides a measure of the resistance to crevice corrosion in service, that is, the greater the area the lower the resistance. Figure 22 shows the linear relationship between the 'difference potential' and the mass losses of various stainless alloys containing an artificial crevice that have been exposed to seawater for 4½ years.

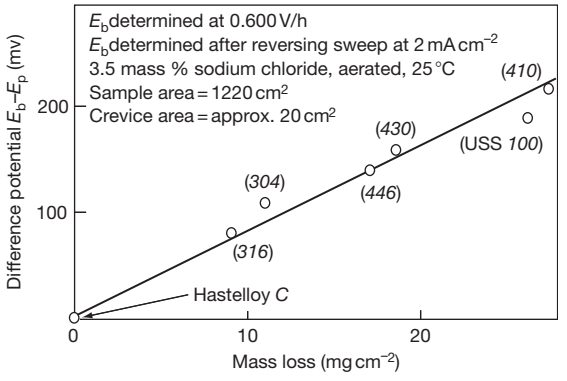


Figure 22 Linear correlation between difference potential and mass loss of various stainless steels containing crevices exposed to seawater for 4.25 years. Reproduced from Wilde, B. E. *Corrosion* **1972**, *28*, 283.

These considerations show that although considerable advances have been made in developing laboratory controlled potential tests for evaluating crevice corrosion and pitting, the results must be interpreted with caution.

Guidance on crevice corrosion testing of iron-base and nickel-base stainless alloys in seawater and other chloride-containing aqueous environments is given in ASTM G 78, while ASTM G 61 provides a standard test method for conducting cyclic potentiodynamic polarization measurements for localized corrosion susceptibility (i.e., pitting and crevice corrosion) of iron-, nickel-, and cobalt-based alloys. Guidance on the selection of procedures for the identification and examination of pitting corrosion to determine the extent of its effect is available in ASTM G 46.

2.34.12 Impingement Tests/Erosion Corrosion

The method most commonly used for testing condenser materials is the BNFMRA May jet impingement test[212] in which small sections of tube, abraded to a standard finish, are immersed in seawater and subjected to an underwater jet of seawater containing air bubbles. However, at high velocities, cavitation can occur in the water box in this test. An alternative design has been described to overcome this.[213] Resistance to impingement attack is also assessed by the Brownsdon and Bannister test[214] in which a stream of air bubbles is directed onto the surface of the test specimens immersed in seawater or sodium chloride solution. Special tests for resistance to corrosion under localized heat transfer conditions (hot-spot corrosion) have been described by Breckon and Gilbert[215] and by Bem and Campbell,[216] but temperature effects are usually ignored when comparing condenser tube materials.

Campbell[217] points out that in evaluating condenser tube materials a test apparatus is required that will include all the principal hazards likely to be encountered in service and should thus cater for the following conditions: impingement, slow moving water, heat transfer, and shielded areas. Furthermore, the internal surfaces should not be abraded, as in the jet impingement test, but should be tested in the 'as-manufactured' condition, particularly in view of the deleterious effect of carbon films produced during manufacture. LaQue has pointed out the importance of specimen area in impingement tests.[218]

The general arrangement of the apparatus is shown in Figure 23. It accommodates 10 vertical

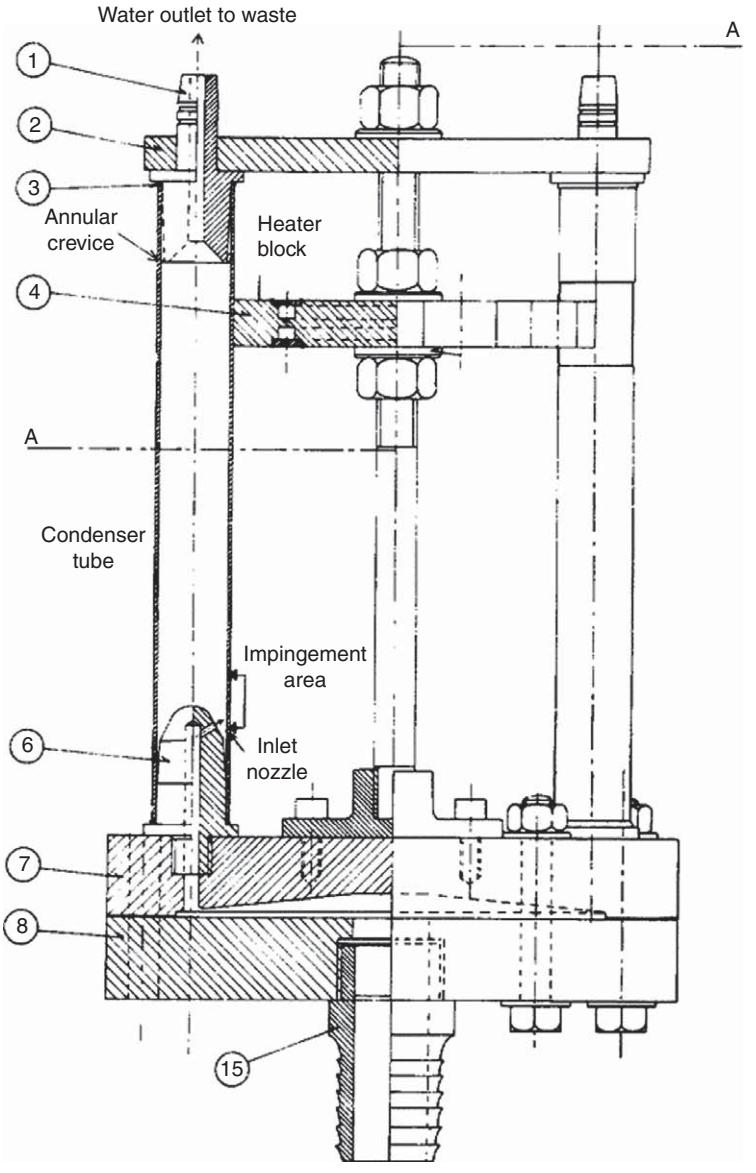


Figure 23 Campbell test apparatus for determining the various forms of attack suffered by condenser tubes in service. Reproduced from Campbell, H. S. MP577, BNFMRA, 1973.

200 mm lengths of condenser tube spaced equally around a 125-mm diameter circle. Water enters the bottom of each tube through an inlet nozzle (Part No. 6 in Figure 23) which fits inside the tube and also locates it. The nozzle has a 5-mm diameter blind hole up the center connecting with a 2.4-mm diameter hole, set at 45° to the vertical, through which the water emerges at a velocity of $10\,m\,s^{-1}$ to impinge on the wall of the tube. The water then rises up through the tube at a mean velocity of $0.1\,m\,s^{-1}$ (in a 22–24-mm diameter condenser tube) and leaves through an outlet nozzle (Part No. 1) fitted into the top end of the tube. Half the length of each outlet nozzle has a 2° taper on the outside to provide a reproducible annular crevice between it and the inside of the condenser tube. Neoprene 'O'-rings (Part No. 3) provide seals between the tube and the top and bottom nozzles, and the tubes are held in place by a common clamping plate (Part No. 2) at, the top. The 10 inlet nozzles are fed with water through a distributor (Part Nos. 7, 8, and 15) of the design used in the May jet impingement apparatus, which ensures equal

distribution of water between them. The distributor and nozzles are all of nonmetallic materials. The part of each test piece between 40 and 65 mm from the top is fine-machined externally to fit a semicircular notch in a 15-mm thick brass heater block (Part No. 4), the tubes being held in contact with the block by a circumferential clip to ensure efficient and equal heat transfer between the block and each tube. The diameter of the inlet and outlet nozzles and that of the semicircular notches in the heater block are made to suit the size of condenser tube to be tested.

The common heater block shown in Figure 23 can itself be subject to corrosion leading to different heat transfer conditions for different tubes, and in some later versions of the apparatus, individual short heating jackets are used for each tube, which are heated with oil from either a steam-heated or electrically heated heat exchanger. This modification not only avoids corrosion problems but also obviates the necessity to machine a length of the outside of each tube to fit the semicircular notches in the single heater block. The oil flow is adjusted to give an oil temperature of 95 °C at each outlet.

The test usually lasts 8 weeks, after which the tubes are sectioned longitudinally and their interiors inspected for accumulated deposits. Loose deposits are then removed by washing in water and the internal surfaces are examined for impingement attack, pitting and blistering or flaking of the corrosion-product film, using a low-power binocular microscope. After cleaning the section in 10% H_2SO_4, the depth of impingement attack, pitting, or other localized corrosion, is determined. Observations and measurements are recorded for each of the following five areas of the section: (a) impingement area opposite the inlet nozzle where water velocity and turbulence are greatest, (b) the slow-moving cold water area from the impingement area upward to the heated area, (c) the heated area including the heat-transfer area itself and the warm-water area above, and (d) the two annular crevices formed between the tapered portions of the cold-water inlet and warm-water outlet nozzles and the tube wall.

The Campbell apparatus is cheap to construct and easy to use and can be installed on site to assist the selection of condenser or heat-exchanger tube materials, or to monitor changes in the corrosivity of the cooling water. The information that it provides on the various forms of attack is more comprehensive than that of any other existing apparatus for corrosion testing condenser tubes, and it is therefore particularly suitable also for assessing new materials or the effect of surface conditions arising from changes in manufacture.

Impingement and erosion–corrosion forms of attack will usually be intensified by the presence of solid particles in the fluid. Variations of the jet test have been proposed to take this effect into account.[219]

Test equipment for the study of erosion-corrosion by liquids with sand content, as met in formation waters in oil and gas production, has been described by Kohley and Heitz.[220]

2.34.13 Corrosion Fatigue

The simultaneous action of alternating stresses and corrosion usually has a greater effect than when either is operating separately, and in this respect corrosion fatigue is analogous to stress-corrosion cracking. The important factors in corrosion fatigue include the following:

1. environmental conditions;
2. magnitude of the alternating stress;
3. magnitude of mean stress;
4. frequency of reversal of the stress;
5. load-versus-time waveform;
6. characteristics of the metal.

Depending on the intended purpose, corrosion fatigue tests can be conducted on smooth, notched or precracked specimens as well as on components and parts joined by welding. Because of the time-dependent nature of corrosion processes, it is essential that the mechanical variables employed during corrosion fatigue testing, including cyclic frequency and load-versus-time waveform, as well as the chemical and electrochemical conditions, are relevant to the intended application. For example, it is unlikely that data generated in a laboratory test at a frequency of 10 Hz would be applicable for predicting corrosion fatigue behavior in a structure that is cycled at 0.1 Hz.

Laboratory corrosion fatigue tests can be classified as either cycles to failure (crack initiation) or crack propagation tests.[221] Cycles to failure tests employ plain or notched specimens to provide data on the intrinsic corrosion fatigue crack initiation behavior of a metal or alloy. Crack propagation tests use precracked specimens to provide information on the threshold conditions for the propagation of preexisting defects by corrosion fatigue and on the rates of corrosion fatigue crack growth.

It is often difficult to conduct laboratory tests in which both the environmental and stressing conditions approximate to those encountered in service.

This applies particularly to the corrosive conditions, since it is necessary to find a means of applying cyclic stresses that will also permit maintenance around the stressed areas of a corrosive environment in which the factors that influence the initiation and growth of corrosion fatigue cracks may be controlled. Among these factors are electrolyte species and concentration, temperature, pressure, pH, flow rate, dissolved oxygen content, and potential (free corrosion potential or applied).

For tests on plain or notched specimens, a simple approach can be to use a conventional Wöhler rotating cantilever beam modified so as to permit the specimen to be brought into contact with the corrodent. This may be achieved by surrounding the specimen with a cell through which the corrosive solution is circulated or by applying it by a pad,[222] wick,[223] or drip feed.[224] Four-point loading or push–pull machines can be used in a similar way and have the advantage over the Wöhler machine when testing plain specimens that the length of the test-piece between the two points of loading is subjected to an approximately uniform stress.

Rawdon[225] used flat specimens that were subjected to repeated flexure while they were being immersed periodically in the corrosive solution. Kenyon[226] used a rotating wire specimen in the form of a loop, the upper part of which was attached to the motor while the lower part of the loop passed through the corrodent, and a somewhat similar device was developed by Haigh-Robertson and used in several studies.[227,228] Gough and Sopwith[229,230] used this machine in their studies, the corrodent being applied as a spray.

Figure 24 shows a slow fatigue machine[231] that has been developed to study the performance of welded butt and fillet joints for steels used in the construction of North Sea oil drilling rigs; the bending stress and frequency have been selected to simulate the forces produced by the wave motion. The specimens, 1500 × 100 × 12.5 mm with the weld 25 cm from the base, are clamped at the lower end, and the stress is applied as a variable bending moment at the upper end by rams. The rams, which are attached to a sliding frame, are activated by a pneumatic cylinder that can be automatically programmed for stroke and frequency and the stress level is monitored by strain gauges. The stress range is up to 300 MPa, and the frequency can be varied from 0.5 to 0.05 Hz. The corrodent is artificial seawater, and provision is made for studying the effect of cathodic protection by means of Zn anodes.

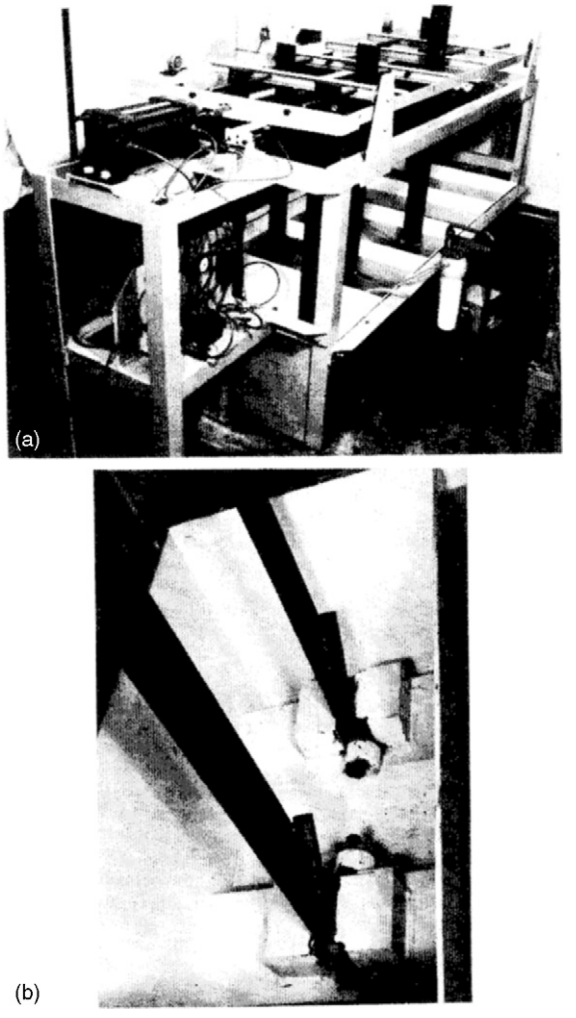


Figure 24 (a) Rig for the laboratory study of corrosion fatigue of welded joints in seawater and (b) view of the test pieces showing welded joint. Reproduced from Jarman, R. A.; Smith, S.; Williams, R. A. *Br. Corros. J.* **1978**, *13*, 195.

Hoeppner[232] pointed out that until the early 1970s, most investigators conducted fatigue tests utilizing rotating bending, flat-plate bending or torsion-type loading configurations, which have the disadvantage that tests at positive or negative mean-stress values are difficult to achieve. In addition, the rotating bending and flat-plate bending tests create complex stress states upon crack initiation, for example, a shifting neutral axis. For these reasons, axial load fatigue machines, as recommended by the ASTM Committee E9, are preferred.

The results obtained from the tests described earlier are presented in the form of the conventional S–N curve, where S is the stress and N the number of cycles to cause fracture. Curves of this type are

obtained for the metal in air and for the metal in the corrodent, and comparison provides information on the effect of the corrosive environment on the fatigue life. Hoeppner points out that even though the S–N curve for either notched or unnotched specimens may be useful for certain applications it cannot always be employed to evaluate the effect of the environment on the fatigue life. This is because in some materials the inherent metallurgical and fabrication discontinuities, which may be undetectable by nondestructive testing will be so large that the only factor of engineering significance will be the rate of propagation of a crack from the initial defect, that is, the fatigue-crack propagation rate may play the dominant rôle in the useful life of the component. For this reason, it is important to conduct fatigue crack growth tests on precracked specimens, and the data are then presented in the form of curves showing crack growth rate, da/dN, versus stress intensity factor range, ΔK.

The NACE publication Corrosion Fatigue[232] gives a comprehensive account of all aspects of the subject, and in this work, a review of the application of fracture mechanics for studying the phenomenon has been presented by McEvily and Wei,[233] while Kitagawa[234] has given a detailed account of crack propagation in unnotched steel specimens. This work should be consulted for details of testing and interpretation of results.

Special requirements for fatigue testing in aqueous environments are addressed in the Annexe to ASTM E 647 'Standard Test Method for Measurement of Fatigue Crack Growth Rates.'

2.34.14 Cavitation-Erosion

In considering these tests, it should be remembered that the phenomenon of cavitation-erosion is often accompained by corrosion effects and that a synergistic effect may operate between the mechanically and chemically induced forms of attack. In fact the term cavitation-erosion-corrosion may often be more applicable in describing the requirements of a test procedure. The subject has been discussed by Wood and Fry.[235]

The methods used have been classified by Lichtman, *et al.*[236] as follows:

1. High-velocity flow.
   a. Venturi tubes.
   b. Rotating discs.
   c. Ducts containing specimens in throat sections.
2. High-frequency vibratory devices.
   a. Magnetostriction devices.
   b. Piezoelectric devices.
3. Impinging jet.
   a. Rotating specimens pass through continuous, stationary jets or droplets.
   b. Stationary specimens exposed to high-speed jet or droplet impact.

All tests are designed to provide high erosion rates on small specimens so that the test can be conducted in a reasonable time, and although vibratory and high-velocity jet methods may not simulate flow conditions, they give rise to high-intensity erosion and can be used, therefore, for screening materials.

The essential component of many high-velocity flow rigs is a venturi-type section in which cavitation occurs in the low-pressure high-velocity region created by the venturi throat. Typical of this type is the double-weir arrangement used by Schroter,[237] but since this technique requires very large volumes of water, it is not readily adaptable to laboratory use. Hobbs[238] and others have used a uniform-area, rectangular-cross-section duct in which a cylinder of small diameter is inserted; cavitation occurs in the wake of the cylinder, which may be used as the test specimen or the specimen may be set in the side wall of the duct near the cylinder. The cavitation intensity will be dependent on the configuration of the test section and the velocity, pressure, temperature, viscosity, surface tension, corrosivity, gas content, and density of the liquid.

Devices in which cavitation is achieved by vibrating a test specimen at high frequencies are often used. The original apparatus was developed by Gaines[239] and was adapted for cavitation-erosion studies by Hunsaker and Peters, as described in the paper by Kerr,[240] and it has been used also by Beeching,[241] Rheingans,[242] and Leith and Thompson.[243] In this method, cavitation is produced by attaching the specimen to the vibrating source or by means of a partially immersed probe vibrating axially at a high velocity and low amplitude and placed close to the test specimen. Although originally magnetostriction oscillators were used,[239] these have now been largely superseded by piezoelectric oscillators, which are more efficient. The apparatus consists basically of a conventional ultrasonic generator, a piezoelectric transducer and a resonating horn or probe, and tests are typically carried out at a frequency of 20 kHz.

Originally, the test specimen was fastened to the end of the ultrasonic probe, and this is still specified

in ASTM D 2809 which describes a method of testing aluminum in antifreeze solution. However, this arrangement also subjects the test-piece to high alternating stresses as a result of the high accelerations associated with vibration at ultrasonic frequencies, which may be overcome by using a stationary test-piece and locating it immediately below a dummy tip placed on the end of the ultrasonic probe.

Vibratory test apparatuses are relatively cheap to build and run and have low power consumption, while flow rigs are bulky, expensive to build and run, and have high power consumptions but have the advantage that they simulate more closely practical conditions of hydrodynamic cavitation. On the other hand, the damage rate is higher in the vibratory tests than in the flow test, although whether this is advantageous depends on the objectives of the test. A further criticism of the vibratory test is that the mechanical component is overemphasized in relation to the effect produced by corrosion. For this reason Plesset[244] uses a technique in which cavitation is intermittent with short bursts of vibration followed by longer static periods, which significantly increases the erosion rate of materials with poor corrosion resistance but has little effect on materials with good corrosion resistance. Tests of this type have distinguished readily between materials having the same hardness but different resistances to corrosion, and between corrosive and noncorrosive solutions.

Figure 25 shows an apparatus for studying cavitation-corrosion using the magnetostriction principle for vibration. A nickel tube is made the core of a magnetic field tuned to the natural frequency of the tube assembly, and since nickel changes its length as it is magnetized and demagnetized, it will vibrate with the frequency of the magnetizing current. The specimen under test vibrates with the nickel tube, and a commonly used frequency is 6500 Hz with an amplitude of 0.008–0.009 cm. Damage is increased by the amplitude of vibration, and the more resistant the material the greater the amplitude to achieve substantial attack. Increase in temperature decreases damage by increasing the vapor pressure within the cavitation bubbles, thus reducing the force of their collapse, but in opposition to this effect is the increased damage resulting from the lower solubility of gases which cushion the collapse of the cavitation bubbles. Consequently, under many circumstances damage reaches a maximum at a test temperature of about 46–52 °C.

Assessment of cavitation-erosion is based on mass loss, and the results are expressed as curves showing cumulative mass (or volume) loss versus the time of the test. Eisenberg and Preiser[245] have expressed the cumulative mass-loss plot on the basis of the rate versus time curve as follows:

1. Incubation zone (little or no mass loss).
2. Accumulation zone (increasing rate to a maximum).
3. Attenuation zone (decreasing loss rate to a steady-state value).
4. Steady-state zone (loss rate at a constant value).

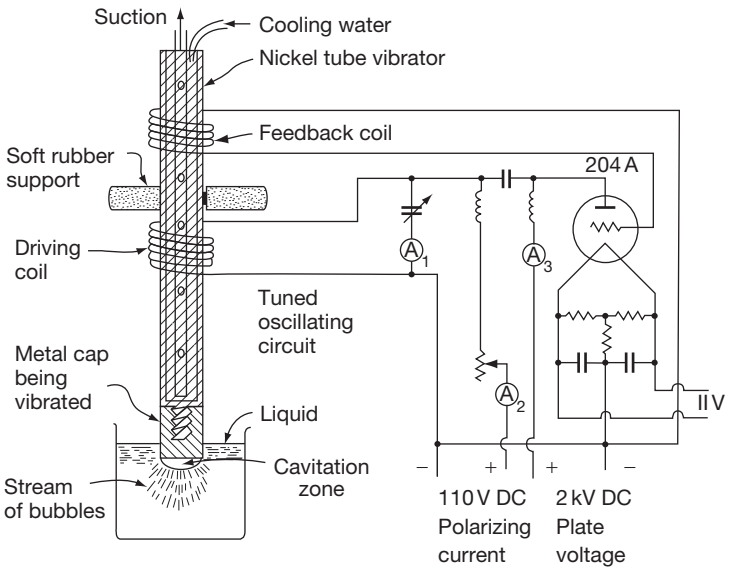


Figure 25 Vibratory cavitation-erosion test using magnetostriction.

It has been proposed that evaluation of the resistance of materials, or the study of experimental variables, should be based on the results obtained for the attenuation zone. Other methods of assessment have been proposed by Hobbs,[246] and by Plesset and Devine.[247]

Examples of various vibratory test procedures for studying cavitation-erosion of metals in inhibited engine coolants have been given in an ASTM Special Technical Publication.[248–250]

2.34.15 Fretting Corrosion

The deterioration of surfaces that occurs when parts supposedly tightly fitted together nevertheless move slightly relative to each other in some sort of cycle under load is called fretting corrosion. With ferrous materials, the characteristic corrosion product is a finely divided cocoa-colored oxide. The general state of knowledge of the subject was reviewed in a symposium on fretting corrosion held by the ASTM in 1952[251] and more recently by Waterhouse.[252,253]

Several techniques for reproducing fretting corrosion have been used. All involve some means for controlling contact pressure, and for achieving and measuring small-amplitude cyclic motion or slip between the contacting surfaces; some control of the environment, particularly moisture, which has a considerable effect on the extent of damage, is also desirable. Fink[254] used an Amsler wear machine. Another early series of tests on fretting corrosion arose from a study of the bottom bearings of electricity meters by Shotter.[255] Tomlinson *et al.*[256] adapted a Haigh alternating-stress machine by which annular specimens were pressed together under load while being subjected to vibration to achieve the required slip. These investigators also used apparatus in which a specimen having a spherical surface was moved cyclically through a small amplitude while in contact under load with a plane surface. A further modification involved an upper specimen machined to provide an annulus which was oscillated under load in contact with a lower plane specimen. A similar technique was used by Wright.[257] The area of damage was measured optically, and the maximum depth of damage was calculated by carefully lapping the lower surface and determining the change in mass. In addition, the amount of oxidized debris was determined chemically.

Uhlig *et al.*[258] measured fretting damage by mass loss of recessed 25.4-mm diameter steel cylinders subjected to radial oscillating motion. The specimens were loaded pneumatically, frequency was varied, and slip was adjusted up to 0.020 mm. Mass loss was determined after debris had been removed by pickling the specimens in inhibited acid.

McDowell[259] used a setup that took advantage of the elastic modulus of one of the test materials to provide a definite deflection subject to control. A rotating-beam fatigue-testing machine was used to produce an alternating compressive and tensile deflection on the surface of the rotating specimen. A sliding specimen slipped back and forth on the rotating specimen as the outer fibers were strained alternately in tension and compression in proportion to the extent of deflection of the rotating specimen.

Horger[251] undertook rotating-beam fatigue tests of press-fitted assemblies using specimens as large as 305-mm diameter shafts.

Warlow-Davies[260] used a technique in which specimens were subjected to fretting corrosion and then tested in fatigue to show the effect of fretting damage in lowering resistance to fatigue.

Herbeck and Strohecker[251] used machines designed particularly for comparing the merits of lubricants in preventing fretting corrosion of antifriction bearings. One provided for both oscillating conditions and combination radial and thrust loads to simulate service. Another was concerned primarily with thrust bearings and correlated satisfactorily with the radial load tester.

An interesting approach involved microscopic observation of fretting corrosion; a glass slide mounted on the stage of a microscope was used for the bearing surface which pressed against a spherical specimen being vibrated by a solenoid.[261]

Other testing machines and techniques have been described by Gray and Jenny,[262] de Villemeur,[263] Wright,[264,265] Barwell and Wright,[266] Field and Waters,[267] and Waterhouse.[268]

2.34.16 Corrosion Testing in Liquid Metals and Fused Salts

Liquid metals have high heat capacities and heat transfer coefficients, and these and other properties make them attractive as coolants for high temperature nuclear reactors and as heat-transfer and working fluids in power-generation systems that operate in conjunction with nuclear reactors. However, austenitic cladding and ferritic structural steels can suffer rapid corrosion when exposed to liquid metals at high temperatures (e.g., in liquid sodium at temperatures above 600 °C or in liquid Pb-17 at %-Li eutectic alloy at

temperatures above 500 °C). Similar corrosion processes affect numerous solid–liquid metal systems, including molybdenum in liquid sodium or lithium, stainless steel in liquid aluminum, platinum in liquid sodium, and carbon steel in liquid zinc.

Corrosion by liquid metals is usually controlled by diffusion processes in the solid and liquid phases and, unlike aqueous corrosion, does not generally involve galvanic effects, and, even where electrochemical phenomena are known to occur, it has not, in general, been demonstrated that they have been responsible for a significant portion of the corrosion observed.[269] In fused salts, there is evidence that electrochemical factors are involved.[270,271] Nevertheless, the corrosion process in relation to liquid metals and fused salts may conveniently be considered under one of the following processes, which do not directly include electrochemical factors: (1) chemical reaction, (2) simple solution, (3) mass transfer, and (4) impurity reactions. Several of these processes may be involved in a single corrosion reaction, but for simplicity, they will be treated separately.

2.34.16.1 Chemical Reaction

This involves the formation of distinct compounds by reaction between the solid metal and the fused metal or salt. If such compounds form an adherent, continuous layer at the interface, they tend to inhibit continuation of the reaction. If, however, they are nonadherent or soluble in the molten phase, no protection will be offered. In some instances, the compounds form in the matrix of the alloy, for example, as grain-boundary intermetallic compound, and result in harmful liquid metal embrittlement (LME), although no corrosion loss can be observed.

2.34.16.2 Simple Solution

The liquid phase may simply dissolve the solid metal, or the liquid may go into solid solution with the metal to form a new phase. In some instances, only a single constituent of an alloy will dissolve in the liquid phase; in this case, a network of voids extending into the metal will result, with obvious deleterious effects.

2.34.16.3 Mass Transfer

This phenomenon manifests itself as the physical transport of a metal from one portion of the system to another, and may occur when there is an alloy compositional difference or a temperature gradient between parts of the unit joined by the flowing liquid phase. An exceedingly small solubility of the metal component or corrosion product in the molten metal or salt appears sufficient to permit mass transfer to proceed at a fairly rapid pace.

2.34.16.4 Impurity Reactions

Small amounts of impurities in the liquid phase or on the surface of the solid metal may result in the initiation of attack or in increased severity of attack by one of the mechanisms just outlined.

In general, it is fair to state that one of the major difficulties in interpreting, and consequently in establishing definitive tests of, corrosion phenomena in fused metal or salt environments is the large influence of very small, and therefore not easily controlled, variations in solubility, impurity concentration, temperature gradient, etc.[272] For example, the solubility of iron in liquid mercury is of the order of 5×10^{-5} at 649 °C, and static tests show iron and steel to be practically unaltered by exposure to mercury. Nevertheless, in mercury boiler service, severe operating difficulties were encountered, owing to the mass transfer of iron from the hot to the cold portions of the unit. Another minute variation was found substantially to alleviate the problem: the presence of 10 ppm of titanium in the mercury reduced the rate of attack to an inappreciable value at 650 °C; as little as 1 ppm of titanium was similarly effective at 454 °C.[273]

In the case of the alkali metals, impurities, such as oxygen and carbon, can have a significant effect on the corrosion of steel and refractory metals. Borgstedt and Frees[274] have shown that for the corrosion of both stabilized and unstabilized austenitic stainless steels in flowing liquid sodium at 700 °C, there is an almost linear dependence of the corrosion constant, k, on the oxygen content of the sodium, as follows:

$$\log k = -5.6637 + 0.919 \log[\mathrm{O}] \qquad [17]$$

where, k is in mg cm^{-2} h^{-1} and [O] is in ppm. Barker *et al.*[275] have demonstrated that oxygen exerts a similarly deleterious effect on the corrosion of AISI Type 316 austenitic stainless steel in liquid Pb–17Li eutectic by increasing the depth of the ferritic corrosion layer and the extent of chromium depletion within the layer.

The effect of carbon on the corrosion of stainless steels in liquid sodium depends upon the test conditions and the composition of the steels.[274] Stabilized stainless steels tend to pick up carbon from sodium,

leading to a degree of carburization that corresponds to the carbon activity in the liquid metal. Conversely, unstabilized stainless steels suffer slight decarburization when exposed to very pure sodium. The decarburization may promote corrosion in the surface region of the material[276] and, under creep rupture conditions, can lead to cavity formation at the grain boundaries and decreased strength.

2.34.16.5 Testing

As in all corrosion testing, the procedure that most nearly duplicates the conditions anticipated in service will provide the most satisfactory and useful information for those aspects of corrosion under consideration here. In fact, in view of the extraordinary sensitivity of fused metal and salt corrosion phenomena to minute variations in operating conditions and purity of components, as already discussed, failure to reproduce these conditions with considerable accuracy may well make any test results completely unrealistic and worthless. In all of the following, then, it should be understood, if not explicitly stated, that all extraneous matter must be carefully excluded from the system and that only materials closely simulating those to be employed in service (including prior history and surface preparation of the metals) should be used. Other factors affecting the corrosion in liquid metals and fused salts include the heat flux of the corroding surface, the volume of liquid to the surface area of the solid, and the liquid flow rate. If, however, screening tests to establish the compatibility of a relatively large number of metals with a given molten metal or salt are to be run, it is often useful to commence with static tests even though the ultimate application involves a dynamic system. This is desirable because static tests are comparatively simple to conduct and interpret, and considerably more economical to operate, and because experience has shown that a metal that fails a static test is not likely to survive the more severe dynamic test.[277] Static tests have been used by Grabner *et al.*[278] to investigate the compatibility of metals and alloys in liquid Pb–Li eutectic at temperatures up to 650°C.

2.34.16.5.1 Static tests

Ideally, a static test would consist of immersing a test sample in the liquid medium held in an inert container under isothermal conditions. Tests in mercury, for example, may be contained in glass at temperatures of several hundred degrees.[279] Unfortunately, at the higher temperatures and with the aggressive metals and salts of interest, there are few readily available inert container materials, and results will often vary according to the nature of the container. The most satisfactory solution is to make the container of the same material as the test sample or, even in some cases, to let it be the sample. Klueh used small capsules for determining the effect of oxygen on the compatibility of Nb and Ta with sodium[280] and potassium.[281] For the Nb–K tests, the Nb specimen was approximately 2.5 × 1.4 × 0.1 cm and was contained in a Nb capsule surrounded by another capsule of welded Type 304 stainless steel. It was demonstrated that the oxygen concentration, added as K_2O, markedly increased the solubility of the Nb in the molten K. DiStefano[282] studied the interaction of Type 316 stainless steel with Nb (or Nb–1Zr) in Na and Na–K by exposing tensile specimens of Nb (Nb–1Zr) to the liquid metal in a stainless steel container. Carbon and nitrogen from the stainless steel were transferred to the Nb, resulting in carbide-nitride at the surface and diffusion of nitrogen into the metal, thus producing an increase in tensile strength and a decrease in ductility. Close control of temperature is also essential if reproducible results are to be obtained, because of differential solubility as a function of temperature. For example, the corrosion rate for Cu–Bi at 500 ± 5.0 °C is several times its rate at 500 ± 0.5 °C.[283]

2.34.16.5.2 Refluxing capsules

In systems in which a liquid metal is used as the working fluid, the liquid is converted to vapor in one part of the system, while the converse takes place in another, and the effect of a boiling-condensing metal on the container materials is most readily studied in a refluxing capsule. DiStefano and De Van[284] used a system in which the lower part of the capsule was surrounded by a heating coil, while the upper part was water cooled. Specimens were inserted in the upper part of the capsule and thus exposed to the condensing vapor, the rate of condensation being controlled by the water flow rate.

When close control of purity is essential, it may be necessary to assemble the test specimens in a dry box under an inert atmosphere and to weld the containers shut under inert gas or vacuum before placing on test. With some environments, even the small amount of oxygen and moisture adsorbed on the component surfaces will significantly affect the test results. In one laboratory, this problem was eliminated by maintaining within the dry box a container of molten sodium at 250 °C[277]; it is a rather cumbersome procedure, but one that emphasises again the importance of purity.

Static test results may be evaluated by measurement of change of mass or section thickness, but metallographic and X-ray examinations to determine the nature and extent of attack are of greater value because difficulty can be encountered in removing adherent layers of solidified corrodent from the surface of the specimen on completion of the exposure, particularly where irregular attack has occurred. Changes in the corrodent, ascertained by chemical analysis, are often of considerable value also. In view of the low solubility of many construction materials in liquid metals and salts, changes in mass or section thickness should be evaluated cautiously. A limited volume of liquid metal could become saturated early in the test, and the reaction would thus be stifled when only a small corrosion loss has occurred, whereas with a larger volume, the reaction would continue to destruction.[283,285]

2.34.16.5.3 Dynamic tests

Various tests have been devised to study the effects of dynamic conditions, and one of the simplest tests is to use a closed capsule that contains a sample at each end and is partially filled with the liquid metal or salt.[286] A temperature gradient is maintained over the length of the tube, and the capsule is rocked slowly so that the liquid metal passes from one end to the other. After the test, the extent of mass transfer is determined from the two specimens placed at each end of the capsule. Tests of this type are useful to establish whether thermal-gradient mass transfer (or concentration-gradient mass transfer if dissimilar metals are incorporated in the system) will occur, but although the method is useful for screening purposes, the dynamic nature of the heating and cooling cycles prevents a rigorous analysis of mass transfer in terms of time and temperature.

High velocity effects can also be studied in spin tests using cylindrical specimens of the solid metal and rotating them at high velocities in an isothermal-metal bath. Although, strictly speaking, only a single alloy should be tested at a time, it is generally satisfactory to include a variety of alloys since the velocity effects become manifest at considerably shorter times than does mass transfer.

Kassner[287] used a rotating disc, for which the hydrodynamic conditions are well defined, to study the dissolution kinetics of Type 304 stainless steel in liquid Bi–Sn eutectic. He established a temperature and velocity dependence of the dissolution rate that was consistent with liquid diffusion control with a transition to reaction control at 860 °C when the speed of the disc was increased. The rotating disc technique has also been used to investigate the corrosion stability of both alloy and stainless steels in molten iron sulfide and a copper/65% calcium melt at 1220 °C.[288] The dissolution rate of the steels tested was two orders of magnitude higher in the molten sulfide than in the metal melt.

2.34.16.5.4 Loop tests

Loop test installations vary widely in size and complexity, but they may be divided into two major categories: (1) thermal-convection loops and (2) forced-convection loops. In both types, the liquid medium flows through a continuous loop or harp mounted vertically, one leg being heated, while the other is cooled to maintain a constant temperature across the system. In the former type, flow is induced by thermal convection, and the flow rate is dependent on the relative heights of the heated and cooled sections, on the temperature gradient and on the physical properties of the liquid. The principle of the thermal convective loop is illustrated in **Figure 26**. This method was used by De Van and Sessions[289] to study mass transfer of niobium-based alloys in flowing lithium, and by De Van and Jansen[290] to determine the

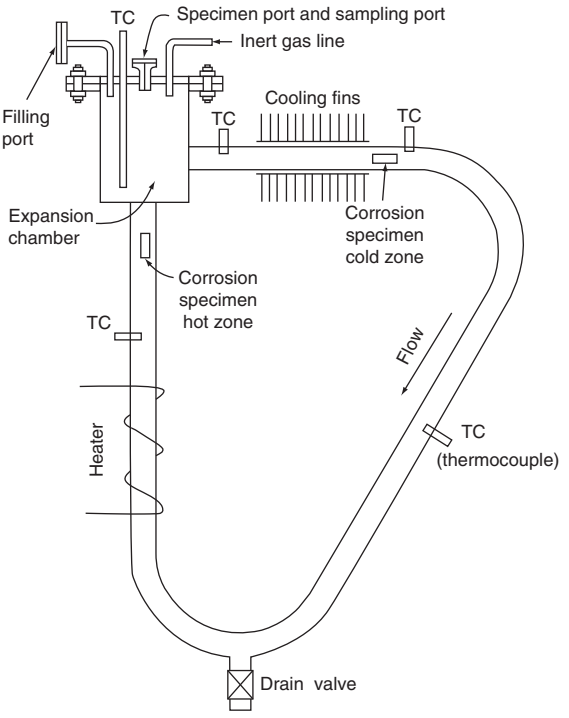


Figure 26 Loop test for studying corrosion in molten metals or salts.

transport rates of nitrogen and carbon between vanadium alloys and stainless steels in liquid sodium.

The thermal-convection loops are limited to flow velocities up to about 6 cm s^{-1}. Where higher velocities are required, the liquid must be pumped, either mechanically or electromagnetically; the latter is usually preferred as it avoids the problem of leakage at the pump seal. Basically, these forced-convection systems[290–293] consist of (1) a hot leg, where the liquid metal is heated to the maximum temperature, (2) an economiser or regenerative heat exchanger, and (3) a cold leg, where the liquid is cooled to its minimum temperature. The economiser consists of concentric tubes with the hotter liquid flowing through the inner tube, while the cooler liquid flows in the opposite direction through the annulus between the two tubes, thus minimizing power requirements. The material under test may be used for constructing all parts of the loop, and the loop is then destructively examined after a given period of test. However, this is costly, and it is now a usual practice to use the loop as a permanent testing facility and to test specimens that are generally placed in the hot leg. Assessment of corrosion is based on changes in weight, dimensions, composition, mechanical properties, and microstructure.

The final stage in a testing program is the design, construction, and testing of loops that simulate the type of system for which data are required.

Since sodium, which is liquid between about 100 °C and 881 °C, has excellent properties as a heat-transfer medium, with a viscosity comparable with that of water and superior heat conductivity,[294] much attention has been paid to liquid sodium corrosion testing of metal and alloys. Indeed, there was an ASTM Standard Practice, which was used for the determination of the corrosion of ferrous alloys, austenitic stainless steels, high nickel alloys, and refractory metals in pumped flowing sodium (ASTM G 68, now withdrawn). This included guidance on the monitoring and control of impurity levels in liquid sodium. The oxygen content of the liquid sodium can be measured continuously by an electrochemical oxygen meter.[295] Similar electrochemical sensors have been used to monitor the carbon content of liquid sodium[296] and the oxygen content of liquid Li–17Bi eutectic.[297] The purity of the liquid metal can be maintained by means of a cold trap through which a small part of the flow is continuously bypassed, the purity level being determined by the temperature of the trap. The ASTM Standard Practice gave the following relationship between the cold trap temperature and oxygen content of the liquid sodium:

$$\log_{10}\mathrm{C}(\text{ppm oxygen content}) = 7.0058 - 2820/T(k) \quad [18]$$

and recommended that the oxygen level of liquid sodium be lowered to 2.85 ppm or less, corresponding to a cold trap temperature of about 150 °C. Borgstedt and Frees[274] found that a cold trap operating at 125 °C further reduced the oxygen content of liquid sodium to 1–2 ppm and acted as a sink for carbon, reducing the level of this element to about 0.01 ppm. The maintenance of low impurity levels in the liquid metal is facilitated if the inert cover gas in the expansion chamber is of high purity (e.g., ≥99.996% argon).

Although the thermal loop test approximates to the conditions that obtain in a dynamic heat-transfer system, in evaluating the results, it is necessary to be aware of those aspects in which the test differs from the full-scale unit, as otherwise unwarranted confidence may be placed in the data. Assuming that adequate attention has been paid to the purity and condition of components, etc., the following factors will, according to ASTM G 68, influence the observed corrosion behavior:

1. liquid metal temperature;
2. degree of nonisothermality of the liquid metal system;
3. liquid metal flow rate;
4. heat flux at the corroding surface;
5. surface-area/volume ratio of solid metal/liquid metal;
6. relative sizes of dissimilar metal surface areas exposed to the liquid metal at the various system temperatures.

The relation between corrosion, and maximum temperature and temperature gradient is obvious, since solubility varies as a function of temperature. If the results are to be useful, these factors should match those anticipated in service. Erratic temperature cycling should be avoided as this can also be modify the corrosion behavior. The effect of surface-to-volume ratio will be more pronounced in thermal convection than in pump loops. It can readily be seen that if a relatively small volume of liquid passes through a given isothermal segment of loop per unit time, it will become saturated quickly, and the corrosion rate will appear lower than would be the case if a substantially larger volume of liquid were passing at the same velocity. In a pumped loop, the velocity can be maintained sufficiently high to prevent the attaining of equilibrium between the solid and liquid

phases, and the rate of dissolution of the solid will be the controlling step. The flow velocity, or Reynolds number, will affect this step too, in that increased velocity will decrease the stagnant or lamellar layer adjacent to the tube wall and decrease the diffusion path that particles must negotiate to enter the rapidly moving stream.[283] The turbulence of the flow may also be modified by the manner in which test specimens are inserted in the loop, and this should also be considered carefully in designing a test unit.

The corrosion rates of the materials of construction are always of importance, but it has been found that, the uniform removal of metal from the hot leg may not impair the load-carrying ability of the container, whereas the deposition of metal in the cold leg can cause the cessation of flow, and the measure of the suitability of an alloy is often the time, under given conditions, that it takes for plugging to occur. Again, the flow velocity and the cross-sectional area are of primary importance in relating test results to operating conditions.

The ultimate test, short of constructing a full-scale unit, is to build a small-scale system in which each item to be incorporated in the final device is represented. Such programs are too specialized to warrant discussion here, and are fully described in the literature.[298–301]

2.34.16.6 Liquid–Metal Embrittlement

Metals have sometimes been observed to crack almost instantaneously when wetted by certain molten metals and subjected to plastic strain at temperatures far below those at which the diffusion-ruled processes involved in liquid–metal corrosion attain significance.[302] The fracture appears to be more brittle than in the absence of the liquid metal, leading to decreased elongation and reduction of area values and, in severe cases, brittle intergranular fracture. Like other forms of environmental cracking, liquid–metal embrittlement is highly specific according to alloy and environment. For example, molten zinc can cause liquid–metal embrittlement of stainless steel if the oxide film is damaged, and because of this molten zinc from associated galvanized parts poses the greatest hazard in welding stainless steel equipment.[303] Other well known examples of liquid–metal embrittlement include the effects of solder on copper alloys and carbon steels and those of mercury on aluminum and nickel alloys. It is generally accepted that most cases of liquid–metal embrittlement arise from the effects of chemisorption of liquid–metal atoms and the consequent reduction of the tensile strength of interatomic bonds at the crack tip since rates of crack growth (up to 10 cm s^{-1}) are usually rapid compared with rates of diffusion of embrittling atoms ahead of cracks or dissolution of the solid in the liquid metal.[304] However, there are a few cases in which diffusion of embrittling atoms ahead of cracks or selective dissolution of a particular phase of an alloy can produce degradation of materials in liquid–metal environments.[305]

Prerequisites for liquid–metal embrittlement are that a solid metal should be subjected to tensile plastic strain while wetted by a liquid metal in which it has low solubility. It has been suggested that such embrittlement may be a general phenomenon occurring under appropriate conditions and to varying degrees between all solid–metal/liquid–metal couples and that a single mechanism may be responsible for all liquid metal embrittlement failures.[306] The occurrence and severity of the embrittlement are governed by:

1. the particular solid–metal/liquid–metal combination;
2. the temperature;
3. the strain rate;
4. the initial mechanical and metallurgical state of the solid metal.

The most commonly used method for assessing liquid–metal embrittlement is by tensile deformation at a slow strain rate. During testing, the specimen should be immersed in the liquid metal in a sealed autoclave to avoid contamination by atmospheric gases.[302] Electrochemical probes similar to those employed in liquid metal corrosion testing can be used to monitor the purity of the liquid metal. Susceptibility to liquid–metal embrittlement can be assessed in terms of the uniform elongation, reduction in area, and fracture appearance of the specimen relative to that determined under similar testing conditions in an inert environment at the same temperature. Where information on crack-propagation behavior is required, use can be made of precracked specimens. These can be tested under static or cyclic loading conditions to determine threshold stress intensity factors and crack growth rates.[307]

2.34.17 Tests in Plant

Although laboratory tests (NACE TMO 169, and Thompson)[308] are obviously of value in selecting materials, they cannot simulate conditions that

occur in practice, and although an initial sorting may be made on the basis of these tests, ultimate selection must be based on tests in the plant. This is particularly important where the process streams may contain small concentrations of unknown corrosive species whose influence cannot be assessed by laboratory trials. Testing is also important for monitoring various phenomena such as embrittlement, hydrogen uptake, corrosion rates, etc.

2.34.17.1 Corrosion Racks

Exposure of coupons or specimens to the process stream cannot be achieved satisfactorily unless they are rigidly supported in a rack, although in some cases, it may be possible to simply hang them in by means of a wire. Methods of exposure of coupons are described in ASTM Method G4.

In the birdcage rack, disc specimens are mounted on a central rod and are insulated from each other and from the rod by insulating spacers and an insulating tube, respectively. PTFE has been found to be suitable for this purpose in aggressive media, particularly at high temperatures. Plates at the end of the rack act as bumpers to prevent the specimens touching the side walls, and the assembly is constructed from a corrosion resistant material such as Monel. Advantages of this method are the following: (1) electrical insulation avoids galvanic effects, and (2) the method of holding the specimen at the centre avoids losses because of corrosion around the point of support. The disadvantages are as follows: (1) specimens are not subjected to either heating or cooling effects and thus will not disclose 'hot-wall' effects and may also escape corrosive condensates when the specimens are in a vapor stream above the dew point, and (2) the corrosivity of the environment may be affected by the presence of corrosion products of the construction material of the racks or by corrosion products of adjacent specimens. A further disadvantage is that because of its size and shape, it must be inserted into the process stream when the plant is out of service. Special devices are required for mounting specimens within pipelines so that they will be subjected to velocity effects.

The insert rack is designed for easy installation and removal through an unused nozzle. The supporting rods (one for each specimen) are welded to a single support plate that is of a width that enables it to be introduced through the nozzle. However, this too cannot be inserted unless the equipment is out of service, although its introduction does not require removal of gas.

A slip-in rack (**Figure 27**) that is designed to be inserted and removed during the operation of the plant through a full-port gate valve attached to a nozzle of suitable diameter (3.8–5.1 cm) is described by Dillon *et al.*[309]. It consists of a short length of pipe flanged at one end to match the gate valve and has a backing-gland arrangement at the other. The coupons are mounted on a rod of small diameter welded to a long heavier rod. The valve is opened, and the support rod is pushed through the packing gland so that the specimen is introduced into the process stream. The specimens are removed by withdrawing the rod until they are again within the pipe section, the gate valve is closed, and the rack removed from the valve.

Access fittings that enable specimens to be introduced into plant that is operating at high pressures, but can also be used for ambient pressures are available. In some instances, it is possible to secure valuable information by substituting experimental materials for parts

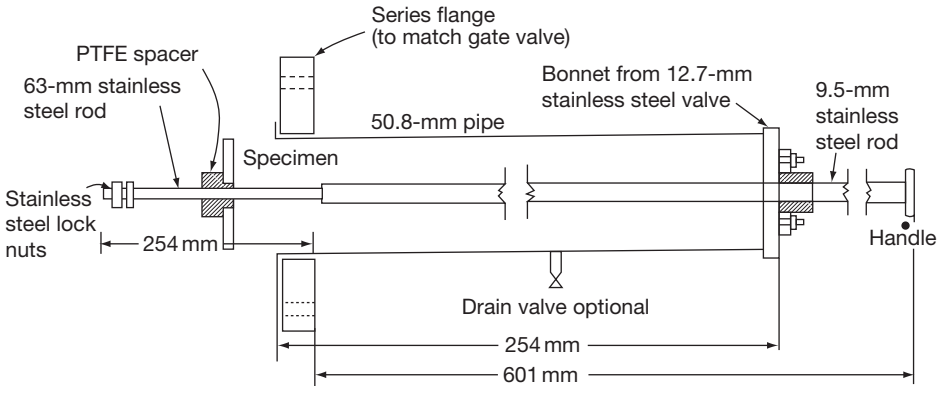


Figure 27 Slip-in corrosion test rack. Reproduced from Dillon, C. P.; Krisher, A. S.; Wissenburg, H. In: Ailor, W. H., Ed.; *Handbook on Corrosion Testing and Evaluation*; John Wiley, 1971.

of the operating equipment, a practice that is used most frequently with condenser tubes, evaporators or other heat exchangers or sections of piping systems.

The prediction of materials performance in plant conditions using modeling and corrosion test methods has been discussed by Strutt and Nichols.[310]

2.34.17.2 Specimens

A convenient size for a circular coupon is 3.8-cm diameter, a thickness of 0.32 cm, and a central hole of 1.1 cm. Although inherent in the philosophy of corrosion testing, the use of coupons with surfaces that simulate those in service has been found to be unsatisfactory owing to irreproducibility, and the standard procedure normally adopted is to abrade down to 120 grit. ASTM Method G4 gives details of preparation of specimens, evaluation of replicate exposures, and the application of statistical methods.

2.34.18 Atmospheric Tests

More or less standardized techniques have been developed for the exposure of specimens to atmospheric weathering. ASTM G 50 and ISO 8565 provide guidance on conducting atmospheric corrosion tests on metals, alloys, and metallic coatings. Procedures for recording data from atmospheric corrosion tests on metallic-coated steel specimens are given in ASTM G 33. The usual practice in the USA[311] is to mount bare specimens on racks that slope 30° from the horizontal and painted specimens on racks that slope 45° from the horizontal. The usual orientation is to have the specimens face south. In coastal exposures, it is not uncommon to have the specimens face the ocean. Steel specimens exposed vertically have been found to corrode about 25% more than similar specimens exposed at the 30° angle.[312] Vertical exposure was used in the large-scale tests of nonferrous metals undertaken by Subcommittee VI of ASTM Committee B-3.[313] Vertical exposure is also favored by Hudson.[314]

A typical test installation uses a frame to support racks on which the specimens are mounted by means of porcelain or plastics insulators. The insulators may be spaced to take specimens varying in size from 10.1 × 13.4 cm to 10.1 × 32 cm and even larger specimens may be used for certain tests. Special types of exposure have been devised to take into account important effects of partial shelter and accumulation of pools of water, as in the case of the specimen and method of support used by Pilling and Wesley[315] to compare steels for roofing.

Copson[316] has described in considerable detail the several factors that require attention in studying atmospheric corrosion, particularly of steels.

Several sizes and shapes of specimen have been used in addition to the common ones already mentioned. In the long-time test of bare and zinc-coated steels undertaken by ASTM Committee A-5 on Corrosion of Iron and Steel, full-size sheets were used.[317] This Committee has also exposed specimens in the form of hardware[318] and wire and fencing.[319]

The extent of deterioration may be measured by one or more of the following methods: visual examination, change in weight, or change in tensile properties. Visual inspection was depended primarily upon the A-5 tests of steel sheets.[319] Here, visible perforation more than 6 mm from an edge was the criterion of failure. This leaves much to be desired for close comparisons because of the frequency with which perforations may be obscured by heavy coats of rust.[312] Other shortcomings of the use of time to visible perforation as the criterion of corrosion resistance are as follows.

1. The removal of rust films or other corrosion products to facilitate inspection for perforation prior to termination of the exposure will change the natural performance of the material and is therefore not tolerable.
2. The recording of a perforation establishes only the time to failure and provides no idea of the progress of corrosion up to the point of failure.
3. The time to perforation may be influenced considerably by the random occurrence of pits that happen to meet after starting from opposite sides of a sheet. This chance meeting of pits may be determined only to a slight extent by the composition of the material and, therefore, will interfere with observations of the effects of composition.

Where changes in appearance are of paramount interest, as in the case of metallic and organic coatings on steel or other metals, visual examination is most desirable. To facilitate ratings on such a basis, photographic standards have been employed, for example, in tests on chromium-plated steel undertaken by ASTM Committee B-8 on Electrodeposited Metallic Coatings.[320] These ratings are supplemented by a shorthand description of the nature of the deterioration observed.

Similarly, photographic standards are recommended for rating organic coatings with respect to different modes of deterioration in ASTM D 610.

The most precise measurements of corrosion resistance require the use of specimens that can be

weighed accurately after careful removal of corrosion products by the techniques described earlier.

A sufficient number of specimens should be exposed initially to permit their withdrawal from test in appropriate groups, for example, 3 to 5 duplicates after at least three time intervals. For long-time tests, a suitable schedule would call for removals after 1, 2, 5, 10, and 20 years. (Editor's note: As with all weight loss tests, it is desirable to retain some specimens in an inert environment to provide a control for the evaluation of systematic errors (such as those due to changes in balance calibration).)

It is good practice to determine depths of pitting as well as mass loss.

As is the case with other types of corrosion testing, mass-loss determinations may fail to indicate the actual damage suffered by specimens that are attacked intergranularly or in such a manner as dezincification. In such cases, mechanical tests or microscopic examination will be required as discussed already in the section on evaluation techniques.

It is desirable for reporting of atmospheric corrosion tests to include a precise description of the climatic conditions that prevailed at the test site during the test so that the weather factors can be tied in with the results of exposure. Progress toward this aim has been made with the development of international standards that provide guidance on evaluating the corrosivity of atmospheric environments. Atmospheric corrosivity can be expressed in terms of environmental factors, the most important of which have to do with contaminants of the atmosphere and the time that the specimens are actually wetted by condensed moisture. In the case of organic coatings, the interacting effects of sunlight and moisture, and their sequence, complicate this problem even more.[321] Methods of measuring pollution (deposition rates of sulfur compounds and chlorides) are provided in ISO 9225. ISO 9223 defines five different categories of atmospheric corrosivity based on time of wetness and pollution. An alternative approach is to express atmospheric corrosivity in terms of the corrosion rate of standard materials, including carbon steel, weathering steel, zinc, copper, and brass. Methods of determining the corrosion rates for this purpose are given in ISO 9226, while ISO 9224 provides a classification of atmospheric corrosivity based on both the average and the steady-state corrosion rates of the standard metals.

Sereda[322] has described a method of determining time of wetness in which a strip of platinum foil (0.8×7 cm) is mounted on a zinc panel (10.1×13.4 cm) on both the skyward and groundward face. Condensed moisture from dew, or rain or snow, results in a galvanic cell, the potential of which is monitored on a recorder, thus giving the time of wetness. Guttman and Sereda[323] found that if the SO_2 content remained essentially constant, the corrosion rate of zinc was related to time of wetness; furthermore, the dew detector registered the presence of moisture on the panel when the relative humidity ranged between 82% and 89%, thus providing a means of estimating from long-term weather data, such as temperature and relative humidity, the time a specimen is likely to be wet.

2.34.19 Atmospheric Galvanic Tests

Studies of galvanic corrosion in the atmosphere are experimentally simpler than those conducted in solution in the laboratory. The environment is taken as it comes, and the relatively high electrical resistance of the rain and moisture films that serve as electrolytes restricts the distance through which the galvanic action can extend and thus limits the relative area effects that complicate galvanic corrosion in solutions of high conductivity. Standard test methods for assessing galvanic corrosion caused by the atmosphere are given in ASTM G 104 and in ISO 7441.

Because of the limited proportion of the areas of a couple that actually participates in the galvanic action, it is difficult to make quantitative measurements that separate the galvanic action from the total effects of exposure. Thus, many of the observations are likely to be qualitative ones, and often no more than what can be determined by visual inspection or measurements of changes in strength, etc. as a result of any localized galvanic action.

An idea of the distribution of galvanic corrosion in the atmosphere is provided by the location of the corrosion of magnesium exposed in intimate contact with steel in the assembly shown in **Figure 28** after exposure in the salt atmosphere 25 m from the ocean at Kure Beach, North Carolina, for 9 years. Except where ledges or crevices may serve to trap unusual amounts of electrolyte, it may be assumed that, even with the most incompatible metals, simple galvanic effects will not extend more than about 4–5 mm from the line of contact of the metals in the couple.

The extent of galvanic action in atmospheric exposure may also be restricted by the development of corrosion products of high electrical resistance between the contacting surfaces: this is especially likely to occur if one of the metals in the couple is an iron or steel that will rust. In long-time tests, such

possible interruptions in the galvanic circuit should be checked by resistance measurements from time to time so as to determine the actual periods in which galvanic effects could operate.

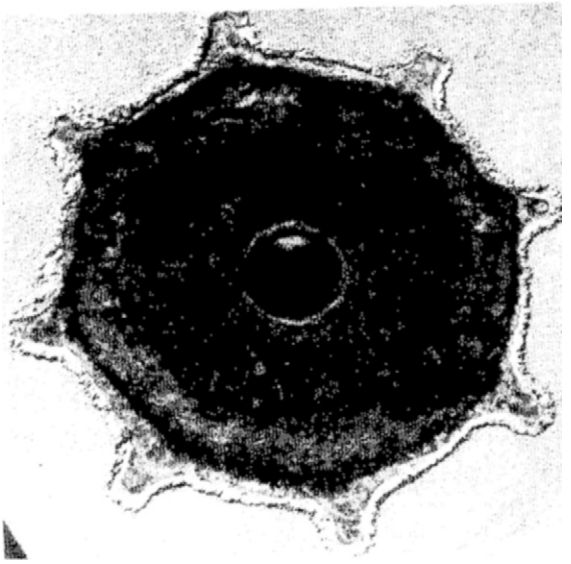


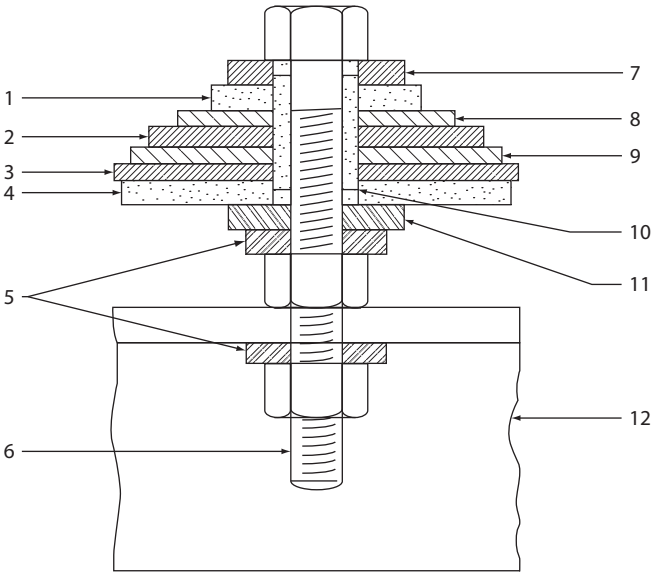
Figure 28 Distribution of galvanic effects around contact of a magnesium casting and a steel core.

The test assembly used originally by Subcommittee VIII of ASTM Committee B-3 in its comprehensive studies of atmospheric galvanic corrosion[324] had the disadvantage that it depended on paint coatings to confine corrosion to the surfaces in actual contact with each other. In interpreting the results, it was frequently difficult to decide how much corrosion was due to galvanic action and how much to a variable amount of normal corrosion through failure of the paint system.

These difficulties were overcome in a design developed by Subcommittee VIII of ASTM Committee B-3[325] (**Figure 29**). In this assembly, each of the two middle specimens has a specimen of the other metal each side of it, and only these middle specimens are considered in appraising the results.

A fairly direct way of observing galvanic effects, which also permits changes in mechanical properties to be measured, involves the preparation of a composite specimen formed by attaching a strip, or strips, of one metal to a panel of another one. Tensile test specimens that include the areas of galvanic action can be cut from these panels after exposure, as shown in **Figure 30**.

A modification of the specimen shown in **Figure 30** may be made simply by lapping a panel of one material over a panel of another one. The greatest effects may



Key
(1) Bakelite washer 19.0 × 3.2 mm. (2) Metal B disc 30 × 1.6 mm. (3) Metal B disc 36.6 × 1.6 mm.
(4) Bakelite washer 35.5 × 3.2 mm. (5) Stainless stell lock washer. (6) Stainless steel bolt 4.8 × 38.1 mm.
(7) Stainless steel washer 15.9 mm o.d. (8) Metal A disc 25.4 × 1.6 mm. (9) Metal A disc 35.5 × 1.6 mm. (10) 11.1 mm bakelite bushing, 5.2 mm i.d. × 7.9 mm o.d. (11) Stainless steel washer 15.9 mm o.d. (12) Galvanised angle support.

Figure 29 Atmospheric galvanic couple test assembly.

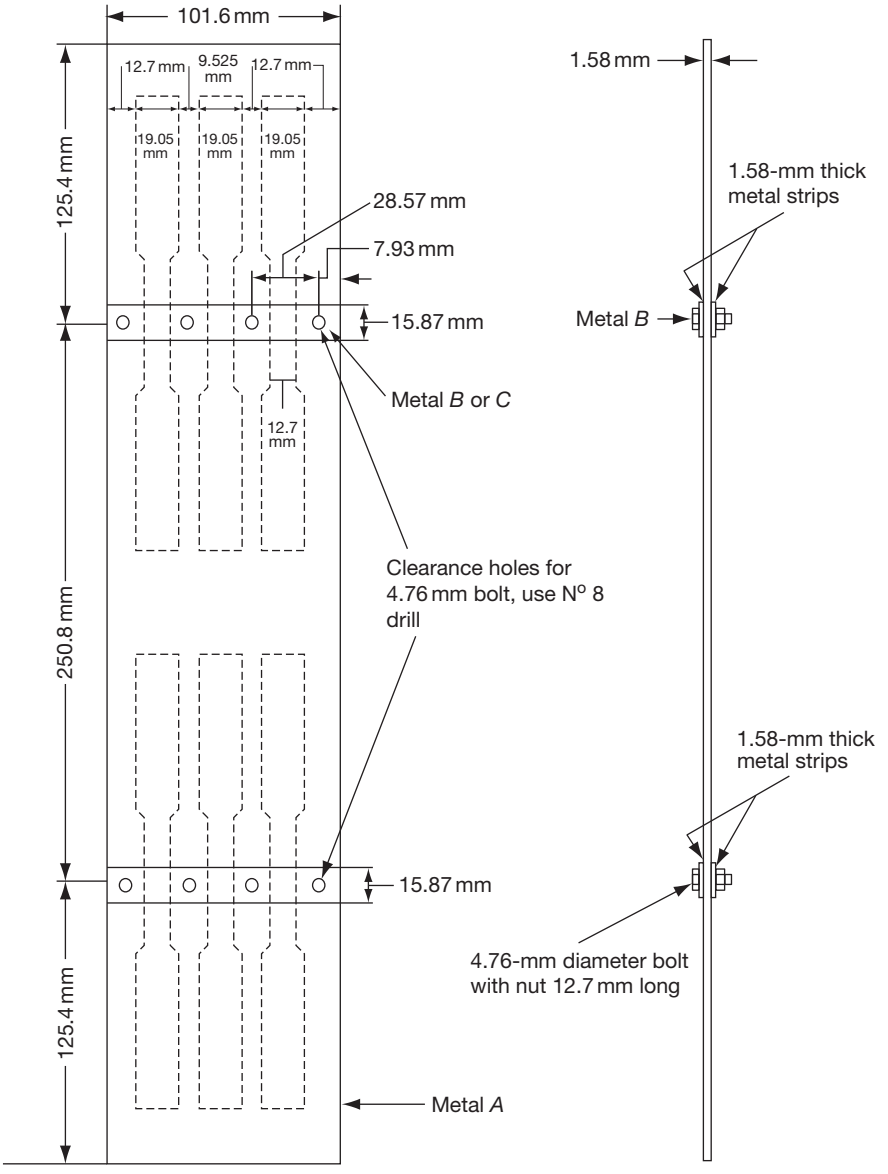


Figure 30 Plate and fastening type galvanic couple test specimen.

be observed when such panels are exposed with the laps facing up so as to favor retention of corrosive liquids along the line of contact. To permit observations of secondary effects of corrosion products, or exhaustion of corrosive constituents, the relative positions of the dissimilar metals should be changed from top to bottom in duplicate test assemblies.

Where the practical interest is in possible galvanic effects of fastenings, it is simple to make up specimens to include such couple assemblies as illustrated in Figure 31.

A type of assembly calculated to favor maximum galvanic action was developed by the Bell Telephone Laboratories and is illustrated in Figure 32. Here, the less noble metal is in the form of a wire wound in the grooves of a threaded specimen of the metal believed to be more noble. Good electrical contact is achieved by means of set screws covered with a protective coating. This assembly favors accumulation of corrosive liquids around the wire in the thread grooves. Corrosive damage is also favored by the high ratio of surface to mass in the wire specimens.

To determine whether a protective metallic coating will retard or accelerate corrosion of a basis metal, and to what distance either effect will extend, specimens in which strips of various widths are left bare or made bare have been used by Subcommittee 11 of ASTM Committee B-8.[325] The extent of corrosion in and near the bare strips as compared with that on a completely bare or completely coated specimen will provide a measure of the extent of galvanic action and the distance through which the effect is able to extend from the edges of the bare strips.

2.34.20 Tests in Natural Waters

The corrosion testing of metals in natural waters is usually conducted in field or service tests since the conditions of flow are important and often rate-determining. Testing will be concerned with mains water (potable water), river-water and seawater, or combinations of these as in estuarine conditions. Test specimens of various geometries will be used, for example, in the form of wires, plates, tubes, etc., and certain general precautions should be followed.

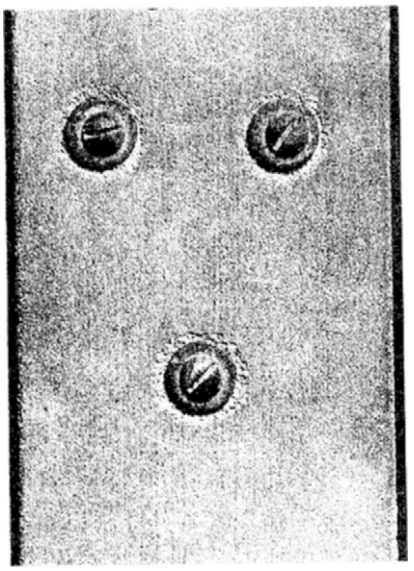


Figure 31 Specimen for studying galvanic corrosion resulting from fasteners.

1. The specimens should be mounted so that they are insulated from their supporting racks and from each other. Such insulation can be achieved by the use of fastening assemblies, such as those illustrated in **Figure 33**. Occasional difficulties have been encountered with this sort of assembly for tests of copper and high-copper alloys because of deposition of copper from corrosion products along the surfaces of the insulating tubes, which provided a metallic bridge between the specimens and the rack and introduced undesired galvanic effects. The required insulation and support can be provided by use of porcelain or plastic knob insulators in much the same manner as used on atmospheric test racks. A modified design has the advantage of offering less resistance to the flow of water and is less likely to serve as a form of screen to catch debris floating on, or suspended in, the water. Additional details of rack design may be found in the section on seawater tests in The Corrosion Handbook (Uhlig, Selective Bibliography).
2. In the case of corrosion tests in the sea or in other large volumes of water, that is, as opposed to tests in waters flowing within pipes, all the specimens to be compared should be suspended at the same depth or should pass through the same range of depths. Isolated specimens exposed at different depths will not be corroded in the same way as continuous specimens that extend through the total range of depth to be studied. This is especially the case with specimens exposed to seawater

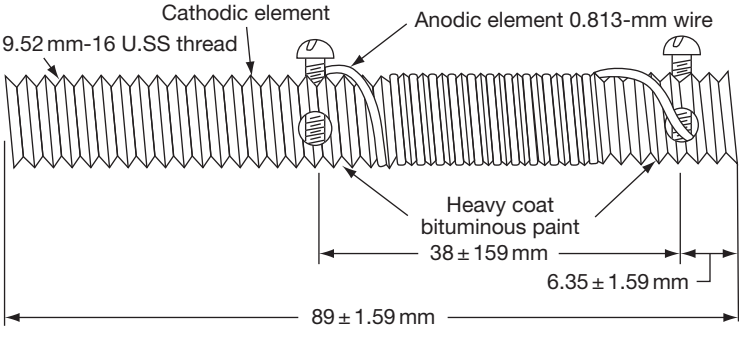


Figure 32 Bolt and wire type atmospheric galvanic couple test specimen.

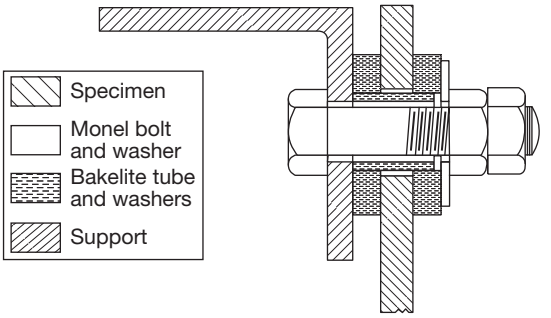


Figure 33 Scheme for insulating specimens from metal test racks.

from above high tide to below low tide. Where the behavior of structures, such as piling, that pass through these zones is to be investigated, the test specimens must be continuous and large enough to extend through the total range in order to take into account differential aeration and other possible concentration cells that may have such a tremendous effect on the results secured.[326] For example, in seawater exposure, isolated specimens of steel exposed in the tidal zone have corroded 10 times as fast as portions of continuous specimens of the same steel in the same zone that extended also below low-tide level.[312]

3. The specimens should be oriented so that their flat surfaces are parallel to the direction of water flow and so that one specimen will neither shield an adjacent specimen from effects of water velocity nor create any considerable extra turbulence upstream of it.
4. In tests in seawater where accumulations of marine organisms are likely, specimens exposed parallel to each other should be spaced far enough apart to ensure that the space between specimens will not become completely clogged by fouling organisms. A minimum spacing of 100 mm is suggested.
5. Wooden racks used in seawater tests are likely to be subject to severe damage by marine borers. The wood used, therefore, must be treated with an effective preservative, for example, creosote applied under pressure, if the test is to extend for several years. Organic copper compound preservatives may suffice for shorter tests, for example, 2 or 3 years. Since the leaching of such preservatives may have some effects on corrosion, metal racks fitted with porcelain or plastic insulators have an advantage over wooden racks.
6. Where constant depth of immersion is desired in spite of tidal action, it is necessary to support the test racks from a float or raft.

Recommended methods for assessing the corrosivity of waters, including flowing potable waters, are described in ASTM D 2688. Three procedures are described in which test specimens in the form of wires, sheets, or tubes are placed in pipes, tanks, or other equipment. The test assembly for the first of these consists of three helical wire coils mounted in series on, and electrically insulated from, a supporting frame. The assembly must be installed so that flow is not disturbed and turbulence and high velocities, for example, of more than $1.53\,m\,s^{-1}$, are avoided. A minimum test period of 30 days is recommended. Procedures for the other specimen forms are given in the standard.

An extensive study of the corrosion of metals in tropical environments has been carried out by Southwell.[327] Tests have included atmospheric exposure, and exposure in seawater under mean tide and fully immersed conditions for a range of ferrous and nonferrous metals and alloys.

The Marine Corrosion Working Party of the European Federation of Corrosion has published valuable advice on corrosion testing in service.[328]

2.34.21 Field Tests in Soil

The precautions generally applicable to the preparation, exposure, cleaning, and assessment of metal test specimens in tests in other environments will also apply in the case of field tests in the soil, but there will be additional precautions because of the nature of this environment. In the case of aqueous, particularly seawater, and atmospheric environment, the physical and chemical characteristics will be reasonably constant over distances covering individual test sites, whereas this will not necessarily be the case in soils, which will almost inevitably be of a less homogeneous nature. The principal factors responsible for the corrosive nature of soils are the presence of bacteria, the chemistry (pH and salt content), the redox potential, electrical resistance, stray currents, and the formation of concentration cells. Several of these factors are interrelated.

These considerations will significantly affect the location of test specimens in field testing. It is clearly important to ensure that the conditions of exposure are accurately known so that the corrosion test results may be interpreted with respect to the end-use requirements.

Two civil engineering operations require particular attention when soil corrosion tests in the field are

required. These are (1) the use of reinforced earth structures in which the corrosion conditions will differ from those at the site from which the soil has been taken and which may take some time to come to equilibrium in the new site, and (2) the use of reclaimed or contaminated land where unusual corrosive agents may be present in irregular distribution. In both these situations, considerable thought should be given to the corrosion test procedures.

Soil burial tests are popular despite the precautions that are needed. It is also important that a sufficient number of specimens are exposed so that statistical treatment of the results may be applied to compensate for some of the inevitable variations in the exposure conditions. Certain precautions originally set out in 1937[329] are still valid, and are as follows:

1. A sufficient number of specimens to yield a reliable coverage should be included.
2. The test site should be typical of the type of soil to be investigated.
3. The depth of burial should be that which will be occupied by the structure of interest. Specimens to be compared should be buried at the same depth. Ideally, tests for structures, such as piling, that will extend through several horizons would require the use of test specimens long enough to extend to the same depth.
4. Specimens should be separate so that they will not affect the corrosion of each other. A minimum spacing of two diameters was proposed.
5. Cylindrical specimens should be laid horizontally.
6. Sheet or plate specimens should be placed on edge.
7. The ends of pipe specimens should be closed to prevent internal corrosion.
8. Sufficient specimens should be provided to allow withdrawals after several time intervals so as to permit observations of changes in corrosion rates with time.
9. A portion of the original surface should be protected so as to provide a datum line for the measurement of pit depths.
10. In applying results of tests on small specimens to estimating corrosion, particularly by pitting on large structures, the effect of the increased area in increasing the depth of pitting must be taken into account.[330]

2.34.21.1 Other Tests

Other tests to determine bacterial-notably sulfate reducing-activity, soil resistivity, pH, redox potential, etc. will provide valuable data to supplement the results obtained with test specimens. A useful account of some of these was given in.[331] A scheme for assessment of corrosivity of soils based on some of these parameters has been given by Tiller.[332]

A number of standards exist for the determination of some of these parameters. BS 1377: Part 3 refers to methods of tests for soils for civil engineering purposes, and Part 9 refers to these and corrosivity tests in situ. It is significant that the standard draws attention to the fact that the results of the tests that are described should be interpreted by a specialist. ASTM tests for pH and resistivity of soil used for corrosion testing are covered by G51 and G57, respectively.

2.34.22 Corrosion Testing of Organic Coatings

Programs to evaluate the corrosion protection by organic coatings on metals are intended to establish relationships between coating properties and performance. Such knowledge is essential to the most effective use of organic coating systems in corrosion control. Depending on the detail with which such studies are performed, light may be shed on the mechanism of coating deterioration as well.

If valid and useful relationships are to be established, it is essential that the factors affecting performance be recognized and form part of the test record. Since the performance is determined by interactions between the coating, the substrate and the surrounding environment against which protection is sought, significant factors and their interrelationships will vary with the nature of the service.

Care in designing and conducting the test in no way reduces the need for discrimination on the part of the person using the test data in the selection of a coating for a particular purpose. Test environments must reflect the deteriorating influences of the service for which they are applicable. A coating system cannot reliably be selected for service in a chemical plant on the basis of performance determined in a rural atmosphere.

Thus, both the proper conduct of the testing program and the valid use of the data depend on an understanding of the nature of organic coatings and of the forces through which they are degraded.

2.34.22.1 Behavior of Organic Coatings

An organic coating provides corrosion protection through the interposition of a continuous, adherent,

high-resistance film between the metal surface and its environment. In principle, its function is the mechanical exclusion of the environment from the metal surface. It seldom, if ever, succeeds practically in achieving this since all continuous organic films are permeable to some degree to moisture, and many coatings either have occasional physical defects or acquire them in service. Surface conversion treatments, such as phosphate and chromate dips, are used to supplement the physical protective properties of coatings, as are chemically inhibiting primers and wash primers. When such treatments are used, they must be included in the record as constituting a part of the coating system.

Critical parts of the test program are the preparation of test specimens, the selection of the exposure conditions (both in laboratory and field tests), and the selection of significant coating properties to be evaluated as a measure of deterioration with time.

2.34.22.2 Preparation

Specimens will normally be flat panels, large enough to avoid any effects caused by nearby edges of the specimen. Edges and backs are usually coated unless the effect of uncoated edges is an intended test variable. Panels may include the structural features of plates, channels, welds, sharp edges, pits, or depressions, depending on the service for which the data are to be applicable.

The composition of the basis metal has been found to influence the performance of organic finishes in many cases. Thus, composition is a significant test variable and must be considered in comparing test data.

It is particularly important that surface roughness and cleanness, which greatly affect adhesion, should be carefully controlled and that the procedures used to achieve them be a part of the test record. A high degree of cleanliness is normally sought. If, however, the data are to be applicable to the painting of outdoor structures, a certain amount of outdoor weathering becomes a part of the specimen preparation prior to coating. Specimens again will be of the basis metal appropriate for the related service application.

The thickness of a coating plays an important part in determining its physical characteristics. Uniformity of thickness among specimens is therefore necessary, particularly when coating deterioration is to be assessed by changes in such properties. For the preparation of reproducible specimens, methods of applying coatings in uniform thicknesses are available, as are methods for accurately measuring film thickness.

2.34.22.3 Exposure Conditions

In considering exposure tests, whether in the form of laboratory, field, or service tests, it is important to consider the purpose of the test and the relevance of the data to the anticorrosion function of the coating. Thus, in the case of paint coatings, factors such as gloss deterioration, chalking, and color retention are of considerable importance in some industries, for example, the automotive industry, but perhaps of minor importance in the painting of structural steelwork. These assessment factors can nevertheless be of significance, since they may be the precursors of corrosion of the basis metal.

2.34.22.3.1 Laboratory tests

Laboratory tests are often conducted with the purpose of providing an accelerated test procedure and if intelligently used, that is, with proper respect for their limitations, they are of value in determining the probable order of durability, and hence, by implication, corrosion protection of a group of paints. They can also be of value in assessing the quality of a range of similar compositions where there is already some knowledge of the performance of the general composition. Although continuing attention is given to the correlation of accelerated tests with field trials and service performance, caution must always be exercised in attempting to predict the type of failure likely to occur under conditions of natural exposure. Certainly, an approach based on 'the rougher the treatment, the better the test' cannot be justified. Three main classes of laboratory test can be identified and may be conveniently classified under the headings of (1) electrochemical, (2) coating adherence, and (3) exposure cabinets – including weatherometers.

Electrochemical tests

This group includes the various electrochemical tests that have been proposed and used over the last fifty years or so. These tests include a number of techniques ranging from the measurement of potential–time curves, electrical resistance, and capacitance to the more complex electrochemical impedance methods. The various methods have been reviewed by Walter.[333] As the complexity of the technique increases, that is, in the mentioned order, the data

that are produced will provide more types of information for the metal-paint system. Thus, the impedance techniques can provide information on the water uptake, barrier action, damaged area, and delamination of the coating as well as the corrosion rate and corroded area of the metal. However, it must be emphasized that the more comprehensive the technique, the greater the difficulties that will arise in interpretation and in reproducibility.

Adherence tests

This group of techniques involves the testing of the metal-to-paint adherence. These techniques are covered by descriptions such as prohesion,[334] blister,[335] pull off (BS EN 24624), and cross-cut (BS EN ISO 2409). Detailed descriptions of these techniques will be found in the appropriate references.

Exposure cabinets

This group of laboratory tests include the so-called exposure cabinets, salt-spray, and weatherometer tests in which the paint-coated panels are subjected to various cycles of wetting and exposure to ultraviolet light to simulate atmospheric conditions of exposure. BS 3900: Part F3 describes a weatherometer consisting of a 1.2 m diameter drum that rotates at 1 rev/20 min, and has facilities for spraying the panels (100 × 150 mm) periodically during a 24 h cycle and exposing them to ultraviolet light by means of an enclosed carbon arc. Spraying with distilled water is effected by means of an atomizer and fan using the following 24-h cycle: 4 h off, 2 h on, 10 h off, 2 h on and 1 h off; the final 1 h is used for checking the arc. The test is continued for 7 days, at the end of which the panels are examined visually for change of color, loss of gloss and blistering, and for checking, cracking, and chalking by means of a lens (×25).

An appraisal of artificial weathering methods was given in a report by Hoey and Hipwood[336] who described the effectiveness of various weatherometer tests such as those described in BS 3900: Part F3 and ASTM G 152/52. Although these tests simulate atmospheric exposure, it is not possible to obtain a direct correlation owing to variation in outdoor exposure conditions from place to place, but they serve a very useful purpose in providing a preliminary sorting of paints that can then be tested in the field.

2.34.22.3.2 Field and plant tests

Field exposure of test panels offers the benefit of a high degree of control over surface preparation and application. Moreover, through standardized exposure conditions, broader comparisons between both paint systems and locations are possible. More importantly, since replicates may be removed and laboratory tested periodically, changes in properties can be followed in considerable detail. At least four replicates should be examined for each exposure period to minimize the effects of atypical specimens.

The exposure site is selected according to the service for which the data are to be applicable. For atmospheric service, such factors as marine and industrial contaminants, sunlight, dew, and sand abrasion must be considered. Atmospheric specimens are normally mounted at 45°, facing south. This has been shown to provide about a 2:1 acceleration of failure compared with a vertical exposure. Whether this or other standardized positions are used, the details of the exposure are an important part of the test record.

The degree of deterioration experienced over a given test period varies with climatic conditions. Since these differ significantly from one season to another, a standard specimen, the performance of which is well known, should be included with each exposure to increase the validity of coating comparisons.

For other environments, such as in seawater or in chemical plants, exposure conditions that most nearly duplicate those of the related service, and are at the same time reproducible, are used. Impingement by water or water carrying entrained solids, thermal effects, and physical abuse are among the factors to be considered.

2.34.22.4 Coating Evaluation

The performance of organic finishes on test is evaluated by visual observation and by physical tests made upon coated specimens that have been exposed for various periods of time to natural or accelerated weathering conditions. Electrical tests are sometimes used on immersed specimens.

Inspection of test panels at the test sites consists of visual observations of blistering (see ASTM D 714) and the appearance of rust (see ASTM D 610). For these, photographs showing various degrees of degradation that serve as observational standards greatly reduce variations between observers. Results of consecutive observations entered on charts provide visual records of the trend of these features with time.

These more serious evidences of degradation, however, are preceded by invisible physical alterations within the coatings, which can be detected readily and quantitatively by suitable physical tests of replicate specimens removed from tests at periodic

intervals. The use of such tests to reveal incipient coating changes is, in a sense, a means of accelerating the test program without distortion of the test environment. This approach is especially dependent on uniformity of properties among replicates, hence on reproducible application techniques. Moreover, since many coating properties are highly sensitive to changes in temperature and relative humidity, equilibrium of the specimens during testing is necessary. Testing conditions are commonly 25 °C and 50% RH.

Physical tests appropriate for this type of evaluation are not necessarily limited to those properties that the coating may be called upon to display in service. A coating that shows a decrease in distensibility from 20% to 10% is still quite capable of withstanding the expansion and contraction of the substrate in atmospheric temperature cycling, yet such a coating can be expected to fail in service earlier than one that shows no decrease. Thus, the properties of value are those that have been established as reliable indices of deterioration.

Besides providing early comparative data on coating performance, physical tests in dealing with intrinsic coating properties provide much-needed quantitative information on the relationships between the several factors affecting the ageing of organic films. The tests cited as follows are those that have been shown to indicate reliably significant changes in the condition of coatings on tests.

2.34.22.4.1 Distensibility

This property is very sensitive to chemical changes within the coating. Its measurement thus shows the beginning of normal ageing or of deterioration through reaction with the environment. Distensibility is generally determined by bending the best panels over a conical mandrel of known radius and calculating the % elongation at first rupture.

2.34.22.4.2 Abrasion tests

In these tests, the end point is normally taken as the amount of abrasion required to penetrate the coating. The results thus reflect the strength of the coating, its cohesion, and in some cases, its adhesion to the basis metal as well as resistance to abrasion.

2.34.22.4.3 Hardness

Coating hardness is related to the method of measurement. Results reflect the resistance to scratching as well as to indentation.

2.34.22.4.4 Impact tests

Such tests reveal the resistance of coatings to deformation and destruction by concentrated sudden stresses. They thus throw considerable light on the integrity of the metal-coating bond. Changes in adhesion through chemical reaction at the paint/metal interface will be reflected in the impact-test values.

These tests for characterizing coating properties necessarily continue to involve a certain amount of empiricism. The intelligent use of these tests, however, has shown that wide variations of physical and electrochemical characteristics of coatings as a function of composition may be obtained, and further, that significant changes in these characteristics, that can be measured before the usual evidence of failure appears, occur upon natural and accelerated ageing.

2.34.23 Test Methods for Corrosion Inhibitors

2.34.23.1 Immersed Conditions

Since corrosion inhibitors are used in a wide range of applications, no universal test method exists. Recognized methods tend to relate to a product or process in which the inhibitor forms a part rather than to the inhibitor *per se*. Thus, tests exist for inhibited coolants, cooling waters, cutting oils, pickling liquids, etc.

The considerations applicable to corrosion test methods also apply to tests for inhibited products. The metals and alloys used, their surface preparation, the temperature, flow rate, composition of the test medium, the presence of heat transfer, and so on must all be relevant to the proposed use of the inhibited product. As with other test methods, there are those tests that have been developed in particular laboratories for the development of inhibitors for particular purposes and those that have acquired national or international recognition by appropriate standards-writing bodies.

The three types of test procedure discussed in this chapter may often be identified in testing of inhibitors or inhibited products. The testing of inhibited engine coolants provides a suitable example.

Laboratory tests used in the development of inhibitors can be of various types and are often associated with a particular laboratory. Thus, in one case, simple test specimens, either alone or as bimetallic couples, are immersed in inhibited solutions in a relatively simple apparatus, as illustrated in **Figure 34**. Sometimes, the test may involve heat transfer, and a simple test arrangement is shown in **Figure 35**. Tests of

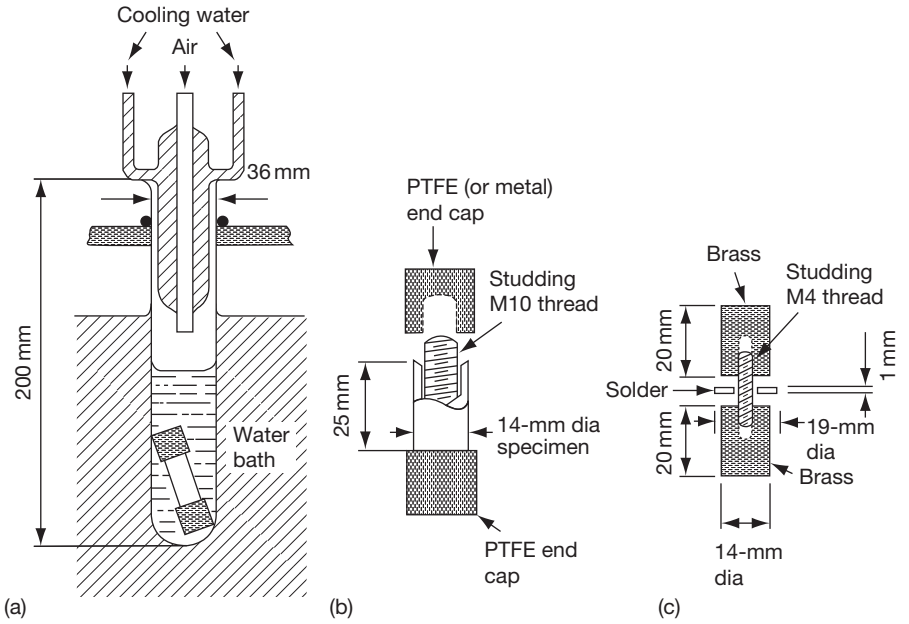


Figure 34 Experimental arrangements for corrosion tests without heat transfer. (a) Test vessel and specimen, (b) assembly for single or bimetallic specimens, and (c) assembly for brass/solder/brass specimens.

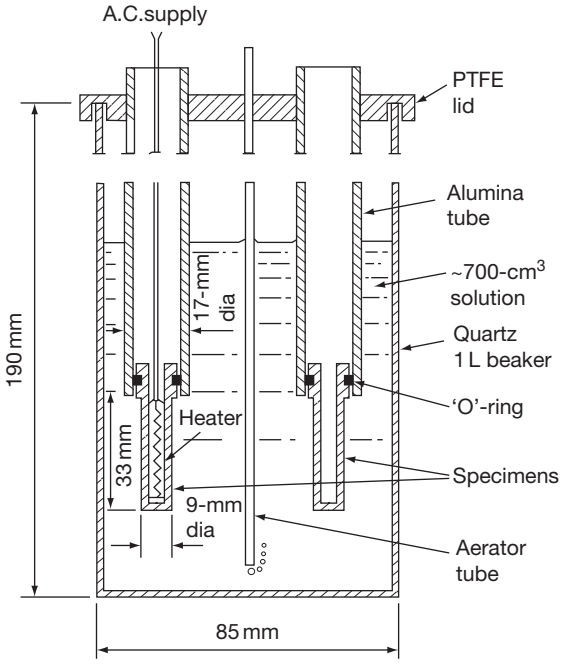


Figure 35 Apparatus for testing inhibitors with metal-to-coolant heat transfer.

these types have been described in the literature.[337,338] However, national standards also exist for this type of test approach. BSI and ASTM documents describe laboratory test procedures and, in some cases, provide recommended pass or fail criteria (BS 5117: Part 2: Section 2.2; BS 6580; ASTM D 1384). Laboratory testing may involve a recirculating rig test in which the intention is to assess the performance of an inhibited coolant in the simulated flow conditions of an engine cooling system. Although test procedures have been developed (BS 5177: Part 2: Section 2.3, ASTM D 2570), problems of reproducibility and repeatability exist, and it is difficult to quote numerical pass or fail criteria.

These laboratory tests may be followed by engine dynamometer tests (BS5117: Section 2.4) and finally by road tests in working vehicles (BS 5117: Part 2: Section 2.5), thus completing the sequence of laboratory, field, service testing.

The problems that have been experienced in the recirculating rig test are indicative of those often met in performance testing. Attempts to reproduce the service conditions in a laboratory test inevitably involve attempting to reproduce each of the controlling conditions that exist in the real situation. Variations, which may be relatively small, in these simulations can lead to significant differences in test results. There is therefore much to be said for keeping test conditions as simple as possible rather than attempting to reproduce accurately the conditions in practice. A balance between reproducibility and realism has to be struck.

The testing of inhibitors for use in oil and gas production, transport, and processing normally involves two-phase oil–water fluids with, sometimes, a solid

phase, for example, entrained sand particles. Tests are usually of the dynamic variety with continuous movement so that test specimens contact all phases present. A well known laboratory test procedure is the so-called wheel test in which bottles of about 200-cm^3 volume containing weighed test specimens and a two-phase fluid saturated with an appropriate gas (CO_2 or H_2S) rotate inside a temperature-controlled chamber.[339] For many applications in these technologies, data are required for high temperature high-pressure conditions, and the use of autoclaves then becomes essential.[340] Two reviews on test procedures for corrosion inhibitors have been published.[341,342]

2.34.23.2 Vapor Phase Conditions

The testing of vapor phase inhibitors, usually referred to as volatile corrosion inhibitors, is essentially a matter of placing a test specimen in the vapor space of a closed vessel containing an aggressive atmosphere – frequently water vapor, perhaps with SO_2 present – and a quantity of the inhibitor. Variations on the basic technique include provision for circulation of the vapor, the use of paper impregnated with inhibitor, provision for temperature cycling, etc.

In the early 1950s, Wachter *et al.*,[343] in the USA, described a humidity cabinet test in which metal specimens were supported inside inverted glass tubes containing a slip of inhibitor-impregnated paper in the lower end. The test was conducted at 37.7 °C and 100% RH. In the UK, Stroud and Vernon[344] described two types of test: (1) with a single test specimen suspended from a cork in the neck of a 250-cm^3 conical flask containing 25 cm^3 of water with 5 mass% of the inhibitor, which was held at 35 °C during the day and at room temperature overnight; and (2) with specimens suspended in the upper part of glass tubes containing water and inhibitor with the lower part of the tubes immersed in a thermostatted bath so that condensation occurred in the upper part of the tubes. Other forms of test using a climatic cabinet with tropical or industrial atmospheres have also been described.[345]

Acknowledgements

Acknowledgements are made to Dr. R. Francis, Dr. T. G. Gooch, Prof. J. S. Llewelyn Leach, Dr. T. C. Lindley, Mr. G. O. Lloyd, Dr. D. Mills, Mr. T. E. Such, Mr. A. K. Tiller, and Mr. P. J. Trant for their helpful comments in connection with the updating of this section, which is based on the original by Dr. L. L. Shreir and Dr. F. L. LaQue.

Appendix A: Methods of Removal of Corrosion Product

This appendix provides information on chemical and electrochemical treatments, which have been recommended for the removal of corrosion products. In using these methods the following points need to be borne in mind:

1. The duration of chemical or electrochemical treatment should be kept to the minimum necessary to remove the corrosion product. Loosely adherent material should be removed beforehand by suitable mechanical means, for example, scrubbing.
2. The combined action of chemical (or electrochemical) treatment and scrubbing is often more effective than either method alone. It is frequently advantageous to alternate short periods of immersion with scrubbing to remove any corrosion product that has become loosened by the action of the chemical reagent.
3. The rate of attack of the chemical reagent on sound metal should be determined on a separate uncorroded sample of the material being cleaned, and if necessary a correction should be applied to the loss in weight of the corroded specimen. However, where a metal, and particularly an alloy, is heavily corroded, thus exposing a surface structure different from that of an uncorroded surface, it will be necessary to check the reliability of the cleaning method (Mercer, A. D.; Butler, G.; Warren, G. M. *Br. Corrosion J.* **1977**, *12*(2), 122–126). A procedure for obtaining more accurate weight loss data in these circumstances has been described [ISO 8407].
4. The possibility of redeposition of metal from the dissolved corrosion product or, if electrochemical treatment is employed, from the anode material should always be kept in mind. If there is reason to believe this has occurred during removal of the corrosion product, further treatment to remove the redeposited metal will be necessary before the weight loss due to corrosion is measured.

A.1 Procedures for Removing Corrosion Products

The removal of corrosion products from metal specimens is described in Reference 1 and in ASTM G 1 and ISO 8407 and certain of these procedures are described as follows.

A.1.1 Electrolytic cathodic cleaning

After scrubbing to remove loosely attached corrosion products, cathodically polarize in hot dilute sulfuric acid under the following conditions:

Electrolyte-sulfuric acid (5 wt. %) plus an inhibitor (0.5 kg m^{-3}) such as diorthotolyl thiourea, quinoline ethiodide, or β-naphthol quinoline. The temperature should be 75°C, the cathode current density 2000 Am^{-2}, and the time of cathodic polarization 3 min. The anode should be carbon or lead. If lead anodes are used, lead may deposit on the specimens and cause an error in the weight loss. If the specimen is resistant to nitric acid the lead may be removed by a flash dip in 1:1 nitric acid. Except for this possible source of error, lead is preferred as an anode, as it gives more efficient corrosion product removal.

After the electrolytic treatment, scrub the specimen with a brush, rinse thoroughly and dry.

Electric treatment may result in the redeposition of a metal, such as copper, from reducible corrosion products and thus decrease the apparent weight loss.

A.1.2 Chemical Cleaning

A.1.2.1 Copper and nickel alloys

Dip for 1–3 min in 1:1 HCl or 1:10 H_2SO_4 at room temperature. Scrub lightly with bristle brush under running water, using fine scouring powder if needed.

A.1.2.2 Aluminum alloys

Dip for 5–10 min in an aqueous solution containing 2 wt. % chromic acid (CrO_3) plus 5 vol. % orthophosphoric acid (H_3PO_4, 85%) maintained at 80°C. Ultrasonic agitation will facilitate this procedure.

Rinse in water to remove the acid, brush very lightly with a soft bristle brush to remove any loose film, and rinse again. If film remains, immerse for 1 min in concentrated nitric acid and repeat previous steps. Nitric acid may be used alone if there are no deposits. (See comments on this method when used for corroded specimens in the paper by Mercer, A. D.; Butler, G.; Warren, G. M. *Br. Corrosion J.* **1977**, *12*, 122.)

A.1.2.3 Tin alloys

Dip for 10 min in boiling trisodium phosphate solution (15%). Scrub lightly with a bristle brush under running water, and dry.

A.1.2.4 Lead alloys

Preferably use the electrolytic cleaning procedure just described. Alternatively, immerse for 5 min in boiling 1% acetic acid. Rinse in water to remove the acid and brush very gently with a soft bristle brush to remove any loosened matter.

Alternatively, immerse for 5 min in hot 5% ammonium acetate solution, rinse and scrub lightly. This removes PbO and $PbSO_4$.

A.1.2.5 Zinc

Immerse the specimens in warm (60–80°C) 10% NH_4Cl for several minutes. Then rinse in water and scrub with a soft brush. Then immerse the specimens for 15–20 s in a boiling solution containing 5% chromic acid and 1% silver nitrate. Rinse in hot water and dry.

Note: in making up the chromic acid solution it is advisable to dissolve the silver nitrate separately and add it to the boiling chromic acid to prevent excessive crystallization of the silver chromate. The chromic acid must be free from sulfate to avoid attack on the zinc. Immerse each specimen for 15 s in a 6% solution of hydriodic acid at room temperature to remove the remaining corrosion products. Immediately after immersion in the acid bath, wash the samples first in tap water and then in absolute methanol, and dry in air. This procedure removes a little of the zinc and a correction may be necessary.

A.1.2.6 Magnesium alloys

Dip for approximately 1 min in boiling 15% chromic acid to which 1% silver chromate solution has been added with agitation.

A.1.2.7 Iron and steel

Preferably use the electrolytic cleaning procedure, or else immerse in Clark's solution (hydrochloric acid 100 parts, antimonious oxide 2 parts, stannous chloride 5 parts) for up to 25 min. The solution may be cold but it should be vigorously stirred. Remove scales formed under oxidizing conditions on steel in 15 vol. % concentrated phosphoric acid containing 0.15 vol. % of an organic inhibitor at room temperature.

A.1.2.8 Stainless steels

Clean stainless steels in 20% nitric acid at 60°C for 20 min. In place of chemical cleaning, use a brass scraper

or brass bristle brush or both, followed by scrubbing with a wet bristle brush and fine scouring powder.

Other methods of cleaning iron and steel include immersion in molten sodium hydride and cathodic treatment in molten caustic soda. These methods may be hazardous to personnel, and should not be carried out by the uninitiated, or without professional supervision.

A.1.3 General Note

Whatever cleaning method is used, the possibility of removal of solid metal is present. This will result in error in the determination of the corrosion rate. One or more cleaned and weighed specimens should be recleaned by the same method and reweighed. Loss due to this second treatment may be used as a correction to that indicated by the first weighing.

Appendix B: Standards

B.1 ISO Standards

ISO standards are managed by the International Organization for Standardization, known as ISO for short (the short name is taken from the Greek 'iso' meaning equal, rather than being an abbreviation of the name of the organization, which would, of course be different in different languages). ISO standards have a simple naming convention, ISO nnnn-pp: yyyy, where nnnn is the standard number, pp the optional part number and yyyy the year of publication. For standards published in collaboration with national bodies, ISO may be supplemented with the abbreviation of the body (as in ISO/ASTM . . .). In the main text we have just given the standard number and part number; here we list the current version at the time of writing.

ISO 1456:2003 Metallic coatings – Electrodeposited coatings of nickel plus chromium and of copper plus nickel plus chromium
ISO 2160:1998 Petroleum products – Corrosiveness to copper – Copper strip test
ISO 2810:2004 Paints and varnishes – Natural weathering of coatings – Exposure and assessment
ISO 3651–1:1998 Determination of resistance to intergranular corrosion of stainless steels – Part 1: Austenitic and ferritic-austenitic (duplex) stainless steels – Corrosion test in nitric acid medium by measurement of loss in mass (Huey test)
ISO 3651–2:1998 Determination of resistance to intergranular corrosion of stainless steels – Part 2: Ferritic, austenitic and ferritic-austenitic (duplex) stainless steels – Corrosion test in media containing sulfuric acid
ISO 4536:1985 Metallic and nonorganic coatings on metallic substrates – Saline droplets corrosion test (SD test)
ISO 4538:1978 Metallic coatings – Thioacetamide corrosion test (TAA test)
ISO 4539:1980 Electrodeposited chromium coatings – Electrolytic corrosion testing (EC test)
ISO 4541:1978 Metallic and other nonorganic coatings – Corrodkote corrosion test (CORR test)
ISO 4543:1981 Metallic and other nonorganic coatings – General rules for corrosion tests applicable for storage conditions
ISO 4623–1:2000 Paints and varnishes – Determination of resistance to filiform corrosion – Part 1: Steel substrates
ISO 4623–2:2003 Paints and varnishes – Determination of resistance to filiform corrosion – Part 2: Aluminum substrates
ISO 6251:1996 Liquefied petroleum gases – Corrosiveness to copper – Copper strip test
ISO 6314:1980 Road vehicles – Brake linings – Resistance to water, saline solution, oil and brake fluid – Test procedure
ISO 6315:1980 Road vehicles – Brake linings – Seizure to ferrous mating surface due to corrosion – Test procedure
ISO 6505:2005 Rubber, vulcanized or thermoplastic – Determination of tendency to adhere to and corrode metals
ISO 6509:1981 Corrosion of metals and alloys – Determination of dezincification resistance of brass
ISO 6988:1985 Metallic and other nonorganic coatings – Sulfur dioxide test with general condensation of moisture
ISO 7120:1987 Petroleum products and lubricants – Petroleum oils and other fluids – Determination of rust-preventing characteristics in the presence of water
ISO 7384:1986 Corrosion tests in artificial atmosphere – General requirements
ISO 7441:1984 Corrosion of metals and alloys – Determination of bimetallic corrosion in outdoor exposure corrosion tests
ISO 7539–1:1987 Corrosion of metals and alloys – Stress corrosion testing,

Part 1: General guidance on testing procedures
ISO 7539–2:1989 Corrosion of metals and alloys – Stress corrosion testing,
Part 2: Preparation and use of bent-beam specimens
ISO 7539–3:1989 Corrosion of metals and alloys – Stress corrosion testing,
Part 3: Preparation and use of U-bend specimens
ISO 7539–4:1989 Corrosion of metals and alloys – Stress corrosion testing,
Part 4: Preparation and use of uniaxially loaded tension specimens
ISO 7539–5:1989 Corrosion of metals and alloys – Stress corrosion testing,
Part 5: Preparation and use of C-ring specimens
ISO 7539–6:2003 Corrosion of metals and alloys – Stress corrosion testing,
Part 6: Preparation and use of precracked specimens for tests under constant load or constant displacement
ISO 7539–7:2005 Corrosion of metals and alloys – Stress corrosion testing,
Part 7: Slow strain rate testing
ISO 7539–8:2000 Corrosion of metals and alloys – Stress corrosion testing, Part 8: Preparation and use of specimens to evaluate weldments
ISO 7539–9:2003 Corrosion of metals and alloys – Stress corrosion testing,
Part 9: Preparation and use of pre-cracked specimens for tests under rising load or rising displacement
ISO 8044 1999 Corrosion of metals and alloys – Basic terms and definitions
ISO 8407:1991 Corrosion of metals and alloys – Removal of corrosion products from corrosion test specimens
ISO 8565:1992 Metals and alloys – Atmospheric corrosion testing – General requirements for field tests
ISO 9223:1992 Corrosion of metals and alloys – Classification of corrosivity of atmospheres
ISO 9224:1992 Corrosion of metals and alloys – Guiding values for the corrosivity categories of atmospheres
ISO 9225:1992 Corrosion of metals and alloys – Corrosivity of atmospheres – Methods of measurement of pollution
ISO 9226:1992 Corrosion of metals and alloys – Corrosivity of atmospheres – Methods of determination of corrosion rate of standard specimens for the evaluation of corrosivity
ISO 9227:2006 Corrosion tests in artificial atmospheres – Salt spray tests
ISO 9400:1990 Corrosion of metals and alloys – Nickel-based alloys – Determination of resistance to intergranular corrosion
ISO 9591:2004 Corrosion of aluminum alloys – Determination of resistance to stress corrosion cracking
ISO 10062:2006 Corrosion tests in artificial atmosphere at very low concentrations of polluting gas(es)
ISO 10270:1995 Corrosion of metals and alloys – Aqueous corrosion testing of zirconium alloys for use in nuclear power reactors
ISO 10270:1995 Corrigendum 1:1997
ISO 11130:1999 Corrosion of metals and alloys – Alternate immersion test in salt solution
ISO 11303:2002 Corrosion of metals and alloys – Guidelines for selection of protection methods against atmospheric corrosion
ISO 11306:1998 Corrosion of metals and alloys – Guidelines for exposing and evaluating metals and alloys in surface sea water
ISO 11463:1995 Corrosion of metals and alloys – Evaluation of pitting corrosion
ISO 11474:1998 Corrosion of metals and alloys – Determination of corrosion resistance through accelerated outdoor atmosphere testing
ISO 11782–1:1998 Corrosion of metals and alloys – Corrosion fatigue tests, Part 1: Cycles to failure testing
ISO 11782–2:1998 Corrosion of metals and alloys – Corrosion fatigue tests, Part 2: Crack propagation testing
ISO 11844–1:2006 Corrosion of metals and alloys – Classification of low corrosivity in indoor atmospheres – Part 1: Determination and estimation of indoor corrosivity
ISO 11844–2:2006 Corrosion of metals and alloys – Classification of low corrosivity of indoor atmospheres – Part 2: Determination of corrosion attack in indoor atmospheres
ISO 11844–3:2006 Corrosion of metals and alloys – Classification of low corrosivity of indoor atmospheres – Part 3: Measurement of environmental parameters affecting indoor corrosivity
ISO 11845:1995 Corrosion of metals and alloys – General principles for corrosion testing
ISO 11846:1995 Corrosion of metals and alloys – Determination of resistance to intergranular corrosion of solution heat treatable aluminum alloys
ISO 11881:1999 Corrosion of metals and alloys – Determination of resistance to exfoliation corrosion of high strength aluminum alloys

ISO 11881:1999 Corrigendum 1:1999
ISO 12732:2006 Corrosion of metals and alloys – Electrochemical potentiokinetic reactivation measurement using the double loop method (based on Cihal's method)
ISO 14993:2001 Corrosion of metals and alloys – Accelerated testing involving cyclic exposure to salt-mist, 'dry' and 'wet' conditions
ISO 15324:2000 Corrosion of metals and alloys – Evaluation of stress corrosion cracking by the drop evaporation test
ISO 15329:2006 Corrosion of metals and alloys – Anodic test for evaluation of intergranular corrosion susceptibility of heat-treatable aluminum alloys
ISO 16151:2005 Corrosion of metals and alloys – Accelerated cyclic tests with exposure to acidified salt-spray, 'dry' and 'wet' conditions
ISO 16701:2003 Corrosion of metals and alloys – Corrosion in artificial atmosphere – Accelerated corrosion test involving exposure under controlled conditions of humidity cycling and intermittent spraying of a salt solution
ISO 16784–1:2006 Corrosion of metals and alloys – Corrosion and fouling in industrial cooling water systems – Part 1: Guidelines for conducting pilot-scale evaluation of corrosion and fouling control additives for open recirculating cooling water systems
ISO 16784–2:2006 Corrosion of metals and alloys – Corrosion and fouling in industrial cooling water systems – Part 2: Evaluation of the performance of cooling water treatment programs using a pilot-scale test rig
ISO 17081:2004 Method of measurement of hydrogen permeation and determination of hydrogen uptake and transport in metals by an electrochemical technique
ISO 17475:2005 Corrosion of metals and alloys – Electrochemical test methods – Guidelines for conducting potentiostatic and potentiodynamic polarization measurements
ISO 17864:2005 Corrosion of metals and alloys – Determination of the critical pitting temperature under potientiostatic control
ISO 21207:2004 Corrosion tests in artificial atmospheres – Accelerated corrosion tests involving alternate exposure to corrosion-promoting gases, neutral salt-spray and drying
ISO 21610:2009 Corrosion of metals and alloys – Accelerated corrosion test for intergranular corrosion susceptibility of austenitic stainless steels.

B. 2 IEC Standards

IEC standards are produced by the International Electrotechnical Commission. They use a naming convention similar to ISO.

IEC 60068–2–11: Edition 3, 1981: Environmental testing – Part 2: Tests. Test – Ka: Salt mist
IEC 60068–2–42: Edition 3, 2003: Environmental testing – Part 2–42: Tests – Test Kc: Sulfur dioxide test for contacts and connections
IEC 60068–2–43: Edition 2, 2003: Environmental testing – Part 2–43: Tests – Test Kd: Hydrogen sulfide test for contacts and connections
IEC 60068–2–46: Edition 1, 1982 Environmental testing – Part 2: Tests. Guidance to Test – Kd: Hydrogen sulfide test for contacts and connections
IEC 60068–2–60: Edition 2, 1995: Environmental testing – Part 2: Tests – Test Ke: Flowing mixed gas corrosion test
IEC 60068–2–49: Edition 1, 1983: Environmental testing – Part 2: Tests. Guidance to Test – Kc: Sulfur dioxide test for contacts and connections
IEC 60068–2–52: Edition 2, 1996: Environmental testing – Part 2: Tests – Test Kb: Salt mist, cyclic (sodium chloride solution).

B. 3 ASTM Standards

ASTM standards use names of the form Xnnn-yy (yyyy), where X is a letter that indicates the general area of the standard (most corrosion standards start with G), nnn is the standard number, yy is the last two digits of the year in which the standard was last revised and yyyy (if present) is the year in which the standard was last reapproved. In the main text we have given only the standard letter and number; here we list the version of the standard that is current at the time of writing (May 2009).

A262–02a (2008) Standard Practices for detecting susceptibility to intergranular attack in austenitic stainless steels
A380–06 Standard Practice for Cleaning, Descaling, and Passivation of Stainless Steel Parts, Equipment, and Systems
A763–93(2009) Standard practices for detecting susceptibility to intergranular attack in ferritic stainless steels
B117–07a Standard practice for operating salt-spray (fog) apparatus
B154–05 Standard test method for mercurous nitrate test for copper alloys

B368–97(2003)e1 Standard method for copper-accelerated acetic acid-salt-spray (fog) testing (CASS test)
B380–97(2008)e1 Standard test method for corrosion testing of decorative electrodeposited coatings by the Corrodkote procedure
B457–67(2008)e1 Standard test method for measurement of impedance of anodic coatings on aluminum.
B537–70(2007) Standard practice for rating of electroplated panels subjected to atmospheric exposure
B735–06 Standard test method for porosity in gold coatings on metal substrates by nitric acid vapor
D130–04e1 Standard test method for corrosiveness to copper from petroleum products by copper strip test
D610–08 Standard practice for evaluating degree of rusting on painted steel surfaces
D714–02e1 Standard test method for evaluating degree of blistering of paints
D807–05 Standard practice for assessing the tendency of industrial boiler waters to cause embrittlement (USBM Embrittlement Detector Method)
D849–05 Standard test method for copper strip corrosion by industrial aromatic hydrocarbons
D1014–09 Standard practice for conducting exterior exposure tests of paints and coatings on metal substrates
D1280–00(2007) Standard guide for total immersion corrosion test for soak tank metal cleaners
D1384–00(2005) Standard test method for corrosion produced by leather in contact with metal
D1654–08 Standard test method for evaluation of painted or coated specimens subjected to corrosive environments
D1743–05ae1 Standard test method for determining corrosion preventive properties of lubricating greases
D1748–02(2008) Standard test method for rust protection by metal preservatives in the humidity cabinet
D1838–07 Standard test method for copper strip corrosion by liquefied petroleum (LP) gases
D2059–03 Standard test method for resistance of zippers to salt spray (fog)
D2247–02 Standard practice for testing water resistance of coatings in 100% relative humidity
D2251–96(2004) Standard test method for metal corrosion by halogenated organic solvents and their admixtures
D2570–08 Standard test method for simulated service corrosion testing of engine coolants
D2688–05 Standard test methods for corrosivity of water in the absence of heat transfer (weight loss methods)
D2758–94(2003) Standard test method for engine coolants by engine dynamometer
D2803–03 Standard guide for testing filiform corrosion resistance of organic coatings on metal
D2809–04e2 Standard test method for cavitation corrosion and erosion-corrosion characteristics of aluminum pumps with engine coolants
D2847–07 Standard practice for testing engine coolants in car and light truck service
D2943–02(2007) Standard test method for aluminum scratch of 1,1,1-trichloroethane to determine stability
D3310–00(2006) Standard test method for determining corrosivity of adhesive materials
D3843–00(2008) Standard practice for quality assurance for protective coatings applied to nuclear facilities
D3911–08 Standard test method for evaluating coatings used in light-water nuclear power plants at simulated design basis accident (DBA) conditions
D3912–95(2001) Standard test method for chemical resistance of coatings used in light-water nuclear power plants
D3929–03 Standard test method for evaluating stress cracking of plastics by adhesives using the bent-beam method
D4340–96(2007) Standard test method for corrosion of cast aluminum alloys in engine coolants under heat-rejecting conditions
D4350–00(2005) Standard test method for corrosivity index of plastics and fillers
D4627–92(2007) Standard test method for iron chip corrosion for water-dilutable metalworking fluids
D4798–08 Standard practice for accelerated weathering test conditions and procedures for bituminous materials (Xenon-Arc method)
D5144–08 Standard guide for use of protective coating standards in nuclear power plants
E647–08 Standard test method for measurement of fatigue crack growth rates
F363–99(2004) Standard test method for corrosion testing of gaskets
F482–03 Standard test method for corrosion of aircraft metals by total immersion in maintenance chemicals
F483–08 Standard test method for total immersion corrosion test for aircraft maintenance chemicals

F981–04 Standard practice for assessment of compatibility of biomaterials for surgical implants with respect to effect of materials on muscle and bone
F1110–08 Standard test method for sandwich corrosion test
G1–03 Standard practice for preparing, cleaning, and evaluating corrosion test specimens
G2/G2M-06 Standard test method for corrosion testing of products of zirconium, hafnium, and their alloys in water at 680°F [360°C]or in steam at 750°F [400°C]
G3–89 (2004) Standard practice for conventions applicable to electrochemical measurements in corrosion testing
G4–01 (2008) Standard guide for conducting corrosion tests in field applications
G5–94 (2004) Standard reference test method for making potentiostatic and potentiodynamic anodic polarization measurements
G7–05 Standard practice for atmospheric environmental exposure testing of nonmetallic materials
G11–04 Standard test method for effects of outdoor weathering on pipeline coatings
G15–08 Standard terminology relating to corrosion and corrosion testing
G16–95 (2004) Standard guide for applying statistics to analysis of corrosion data
G28–02 (2008) Standard test methods for detecting susceptibility to intergranular corrosion in wrought nickel-rich, chromium-bearing alloys
G30–97 (2003) Standard practice for making and using U-bend stress corrosion test specimens
G31–72 (2004) Standard practice for laboratory immersion corrosion testing of metals
G32–06 Standard test method for cavitation erosion using vibratory apparatus
G33–99 (2004) Standard practice for recording data from atmospheric corrosion tests of metallic-coated steel specimens
G34–01 (2007) Standard test method for exfoliation corrosion susceptibility in 2XXX and 7XXX series aluminum alloys (EXCO test)
G35–98 (2004) Standard practice for determining the susceptibility of stainless steels and related nickel–chromium–iron alloys to stress corrosion cracking in polythionic acids
G36–94 (2006) Standard practice for evaluating stress corrosion cracking resistance of metals and alloys in a boiling magnesium chloride solution
G37–98 (2004) Standard practice for use of Mattsson's solution of pH 7.2 to evaluate the stress corrosion cracking susceptibility of copper-zinc alloys
G38–01 (2007) Standard practice for making and using C-ring stress corrosion test specimens
G39–99 (2005) Standard practice for preparation and use of bent-beam stress corrosion test specimens
G41–90 (2006) Standard practice for determining cracking susceptibility of metals exposed under stress to a hot salt environment
G44–99 (2005 Standard practice for exposure of metals and alloys by alternate immersion in neutral 3.5% sodium chloride solution
G46–94 (2005) Standard guide for examination and evaluation of pitting corrosion
G47–98 (2004) Standard test method for determining susceptibility to stress corrosion cracking of 2XXX and 7XXX aluminum alloy products
G48–03 Standard test method for pitting and crevice corrosion resistance of stainless steels and related alloys by the use of ferric chloride solution
G49–85 (2005) Standard practice for preparation and use of direct tension stress corrosion test specimens
G50–76 (2003) Standard practice for conducting atmospheric stress corrosion tests on metals
G51–95 (2005) Standard test method for measuring pH of soil for use in corrosion testing
G52–00 (2006) Standard practice for exposing and evaluating metals and alloys in surface seawater
G57–06 Standard test method for field measurement of soil resistivity using the Wenner four-electrode method
G58–85 (2005) Standard practice for the preparation of stress corrosion test specimen for weldments
G59–97 (2003) Standard test method for conducting potentiodynamic polarization resistance measurements
G60–01 (2007) Standard practice for conducting cyclic humidity exposures
G61–86 (2003)e1 Standard test method for conducting cyclic potentiodynamic polarization measurements for localized corrosion susceptibility of iron-, nickel-, or cobalt-based alloys
G64–99 (2005) Standard classification of resistance to stress corrosion cracking of heat treatable aluminum alloys
G66–99 (2005)e1 Standard test method for visual assessment of exfoliation corrosion susceptibility of 5XXX series aluminum alloys (ASSET test)
G67–04 Standard test method for determining the susceptibility to intergranular corrosion of 5XXX series aluminum alloys by mass loss after exposure to nitric acid (NAMLT test)

G69–97 (2003) Standard test method for measurement of corrosion potentials of aluminum alloys
G71–81 (2003) Standard guide for conducting and evaluating galvanic corrosion tests in electrolytes
G 73–04 Standard practice for liquid impingement erosion testing
G78–01 (2007) Standard guide for crevice corrosion testing of iron base and nickel base stainless steels in seawater and other chloride-containing aqueous environments
G82–98 (2003) Standard guide for development and use of a galvanic series for predicting galvanic corrosion performance
G84–89 (2005) Standard practice for measurement of time-of-wetness on surfaces exposed to wetting conditions as in atmospheric corrosion testing
G85–02e1 Standard practice for modified salt spray (fog) testing
G87–02 (2007) Standard practice for conducting moist SO_2 test
G90–05 Standard practice for performing accelerated outdoor weathering of nonmetallic materials using concentrated natural sunlight
G91–97 (2004) Standard practice for monitoring atmospheric SO_2 using sulfation plate technique
G92–86 (2003) Standard practice for characterization of atmospheric test sites
G96–90(2008) Standard guide for online monitoring of corrosion in plant equipment (electrical and electrochemical methods)
G97–97(2007) Standard test method for laboratory evaluation of magnesium sacrificial anode test specimens for underground applications
G100–89(2004) Standard test method for conducting cyclic galvanostaircase polarization
G101–04 Standard guide for estimating the atmospheric corrosion resistance of low-alloy steels
G102–89 (2004)e1 Standard practice for calculation of corrosion rates and related information from electrochemical measurements
G103–97 (2005) Standard practice for evaluating stress-corrosion cracking resistance of low copper 7XXX series Al-Zn-Mg-Cu alloys in boiling 6% sodium chloride solution
G104:1989 Test method for assessing galvanic corrosion caused by the atmosphere
G106–89(2004) Standard practice for verification of algorithm and equipment for electrochemical impedance measurements
G107–95(2008) Standard guide for formats for collection and compilation of corrosion data for metals for computerized database input
G108–94(2004)e1 Standard test method for electrochemical reactivation (epr) for detecting sensitization of AISI type 304 and 304L stainless steels
G109–07 Standard test method for determining effects of chemical admixtures on corrosion of embedded steel reinforcement in concrete exposed to chloride environments
G110–92(2003)e1 Standard practice for evaluating intergranular corrosion resistance of heat treatable aluminum alloys by immersion in sodium chloride + hydrogen peroxide solution
G111–97(2006) Standard guide for corrosion tests in high temperature or high pressure environment, or both
G112–92(2003) Standard guide for conducting exfoliation corrosion tests in aluminum alloys
G116–99(2004) Standard practice for conducting wire-on-bolt test for atmospheric galvanic corrosion
G123–00(2005) Standard test method for evaluating stress-corrosion cracking of stainless alloys with different nickel content in boiling acidified sodium chloride solution
G129–00(2006) Standard practice for slow strain rate testing to evaluate the susceptibility of metallic materials to environmentally assisted cracking
G135–95(2007) Standard guide for computerized exchange of corrosion data for metals
G139–05 Standard Test method for determining stress-corrosion cracking resistance of heat-treatable aluminum alloy products using breaking load method
G140–02(2008) Standard test method for determining atmospheric chloride deposition rate by wet candle method
G142–98(2004) Standard test method for determination of susceptibility of metals to embrittlement in hydrogen containing environments at high pressure, high temperature, or both
G146–01(2007) Standard practice for evaluation of disbonding of bimetallic stainless alloy/steel plate for use in high-pressure, high temperature refinery hydrogen service
G148–97(2003) Standard practice for evaluation of hydrogen uptake, permeation, and transport in metals by an electrochemical technique
G150–99(2004) Standard test method for electrochemical critical pitting temperature testing of stainless steels
G152–06 Standard practice for operating open flame carbon arc light apparatus for exposure of nonmetallic materials

G153–04 Standard practice for operating enclosed carbon arc light apparatus for exposure of nonmetallic materials
G157–98(2005) Standard guide for evaluating the corrosion properties of wrought iron- and nickel-based corrosion resistant alloys for the chemical process industries
G158–98(2004) Standard guide for three methods of assessing buried steel tanks
G162–99(2004) Standard practice for conducting and evaluating laboratory corrosions tests in soils
G165–99(2005) Standard practice for determining rail-to-earth resistance
G168–00(2006) Standard practice for making and using precracked double beam stress corrosion specimens
G170–06 Standard guide for evaluating and qualifying oilfield and refinery corrosion inhibitors in the laboratory
G180–07 Standard test method for initial screening of corrosion inhibiting admixtures for steel in concrete
G184–06 Standard practice for evaluating and qualifying oil field and refinery corrosion inhibitors using rotating cage
G185–06 Standard practice for evaluating and qualifying oil field and refinery corrosion inhibitors using the rotating cylinder electrode
G186–05 Standard test method for determining whether gas-leak-detector fluid solutions can cause stress corrosion cracking of brass alloys
G187–05 Standard test method for measurement of soil resistivity using the two-electrode soil box method
G188–05 Standard specification for leak detector solutions intended for use on brasses and other copper alloys
G189–07 Standard guide for laboratory simulation of corrosion under insulation
G192–08 Standard test method for determining the crevice repassivation potential of corrosion-resistant alloys using a potentiodynamic-galvanostatic-potentiostatic technique.

B. 4 NACE Standards

NACE standards are produced by NACE International (formerly the National Association of Corrosion Engineers). They use the naming convention NACE XXnnnn-yyyy, where XX is the type of standard, nnnn the standard number and yyyy the year of publication. Here we list only the TM (Test Method) standards; the full list can be found on the NACE website (www.nace.org).

TM0101–2001 Measurement techniques related to criteria for cathodic protection on underground or submerged metallic tank systems
TM0102–2002 Measurement of protective coating electrical conductance on underground pipelines
TM0103–2003 Laboratory test procedures for evaluation of SOHIC resistance of plate steels used in wet H_2S service
TM0104–2004 Offshore platform ballast water tank coating system evaluation
TM0105–2005 Test procedures for organic-based conductive coating anodes for use on concrete structures
TM0106–2006 Detection, testing, and evaluation of microbiologically influenced corrosion (MIC) on external surfaces of buried pipelines
TM0108–2008 Testing of catalyzed titanium anodes for use in soils or natural waters
TM0109–2009 Above ground survey techniques for the evaluation of underground pipeline coating condition
TM0169–2000 Laboratory corrosion testing of metals
TM0172–2001 Determining corrosive properties of cargoes in petroleum product pipelines
TM0173–2005 Methods for determining quality of subsurface injection water using membrane filters
TM0174–2002 Laboratory methods for the evaluation of protective coatings and lining materials on metallic substrates in immersion service
TM0177–2005 Laboratory testing of metals for resistance to sulfide stress cracking and stress corrosion cracking in H_2S environments
TM0183–2006 Evaluation of internal plastic coatings for corrosion control of tubular goods in an aqueous flowing environment
TM0185–2006 Evaluation of internal plastic coatings for corrosion control of tubular goods by autoclave testing
TM0186–2002 Holiday detection of internal tubular coatings of 250- to 760-μm (10 to 30 mils) dry-film thickness
TM0187–2003 Evaluating elastomeric materials in sour gas environments
TM0190–2006 Impressed current laboratory testing of aluminum alloy anodes

TM0192–2003 Evaluating elastomeric materials in carbon dioxide decompression environments
TM0193–2000 Laboratory corrosion testing of metals in static chemical cleaning solutions at temperatures below 93°C (200°F)
TM0194–2004 Field monitoring of bacterial growth in oil and gas systems
TM0197–2002 Laboratory screening test to determine the ability of scale inhibitors to prevent the precipitation of barium sulfate and/or strontium sulfate from solution (for oil and gas production systems)
TM0198–2004 Slow strain-rate test method for screening corrosion-resistant alloys (CRAs) for stress corrosion cracking in sour oilfield service
TM0199–2006 Standard test method for measuring deposit mass loading (deposit weight density) values for boiler tubes by the glass-bead-blasting technique
TM0204–2004 Exterior protective coatings for seawater immersion service
TM0208–2008 Laboratory test to evaluate the vapor-inhibiting ability of volatile corrosion inhibitor materials for temporary protection of ferrous metal surfaces
TM0284–2003 Evaluation of pipeline and pressure vessel steels for resistance to hydrogen-induced cracking
TM0286–2001 Cooling water test unit incorporating heat transfer surfaces
TM0294–2007 Testing of embeddable impressed current anodes for use in cathodic protection of atmospherically exposed steel-reinforced concrete
TM0296–2002 Evaluating elastomeric materials in sour liquid environments
TM0297–2008 Effects of high temperature, high-pressure carbon dioxide decompression on elastomeric materials
TM0298–2003 Evaluating the compatibility of FRP pipe and tubulars with oilfield environments
TM0304–2004 Offshore platform atmospheric and splash zone maintenance coating system evaluation
TM0374–2007 Laboratory screening tests to determine the ability of scale inhibitors to prevent the precipitation of calcium sulfate and calcium carbonate from solution (for oil and gas production systems)
TM0384–2002 Holiday detection of internal tubular coatings of less than 250-μm (10 mils) dry-film thickness
TM0397–2002 Screening tests for evaluating the effectiveness of gypsum scale removers
TM0399–2005 Standard test method for phosphonate in brine
TM0404–2004 Offshore platform atmospheric and splash-zone new construction coating system evaluation
TM0497–2002 Measurement techniques related to criteria for cathodic protection on underground or submerged metallic piping systems
TM0498–2006 Evaluation of the carburization of alloy tubes used for ethylene manufacture
TM0499–99 Immersion corrosion testing of ceramic materials.

B. 5 Other National Standards

Most other national standards have been subsumed into ISO standards. Thus, the British Standards Institute (BSI) now markets standards as BS EN ISO xxxx (for British Standard, European Norm, ISO xxxx).

References

1. Champion, F. A. *Corrosion Testing Procedures*, Chapman and Hall: London, 1964.
2. Ailor, W. H. Ed.; *Handbook on Corrosion Testing and Evaluation*; Wiley: New York, 1971.
3. Gilroy, D.; Mayne, J. E. O. *Corros. Sci.* **1965**, *5*, 55.
4. Wesley, W. A. *Proc. Am. Soc. Test. Mater.* **1943**, *43*, 649.
5. Bayer, R. O.; Kachik, E. A. *Corrosion* **1949**, *5*, 308.
6. Renshaw, W. G.; Ferree, J. A. *Corrosion* **1951**, *7*, 353.
7. Gulbransen, E. A. *Miner. Metall. N.Y.* **1944**, *25*, 172.
8. Kruger, J. J. *Electrochem. Soc.* **1959**, *106*, 847; Kruger, J.; Hayfield, P. C. S. In *Handbook on Corrosion Testing and Evaluation*; Ailor, W. H., Ed.; Wiley: New York, 1971.
9. Rutherford, J. J. B.; Aborn, R. H. *Trans. Am. Inst. Miner. (Metall.) Eng.* **1932**, *100*, 293.
10. Burns, R. M.; Campbell, W. E. *Trans. Electrochem. Soc.* **1929**, *55*, 271.
11. Hudson, J. C. *Proc. Phys. Soc. Lond.* **1928**, *40*, 107.
12. Mercer, A. D.; Butler, G.; Warren, G. M. *Br. Corrosion J.* **1977**, *12*, 122.
13. Finkeldey, W. H. *Proc. Am. Soc. Test. Mater.* **1932**, *32*, 226.
14. Kenworthy, L.; Waldram, J. M. *J. Inst. Met.* **1934**, *55*, 247.
15. Tronstad, L. *Trans. Faraday Soc.* **1933**, *29*, 502.
16. Kruger, J. *J. Electrochem. Soc.* **1959**, *106*, 847, and **1961**, *108*, 504.; Kruger, J. In *Symp. Proc.* Bashara, N. M., Buckman, A. B., Hall, A. C., Eds.; University of Nebraska: Lincoln, 1969; p 16.
17. Bengough, G. D.; Stuart, J. M.; Lee, A. R. *Proc. Roy. Soc. A.* **1928**, *116*, 425; *121*, 89.
18. Shipley, J. W.; McHaffie, I. R.; Clare, N. D. *Ind. Eng. Chem.* **1925**, *17*, 381.
19. Bloom, M. C.; Krulfeld, M. *J. Electrochem. Soc.* **1957**, *104*, 264.
20. Vaurio, V. W.; Clark, B. S.; Lueck, R. H. *Ind. Eng. Chem. (Anal.)* **1938**, *10*, 368.
21. Hudson, R. M.; Stragland, G. L. *Corrosion* **1959**, *15*, 135t.

22. Willey, A. R.; Krickl, J. L.; Hartwell, R. R. *Corrosion* **1956**, *12*, 433t.
23. Burns, R. M. *J. Appl. Phys.* **1937**, *8*, 398.
24. Dravnieks, A.; Cataldi, H. A. *Corrosion* **1954**, *10*, 224.
25. Marsh, G. A.; Schaschl, E. *Corrosion* **1958**, *14*, 155t.
26. Mylius, F. *Z. Metallkd.* **1922**, *14*, 233.
27. El-Kot, A. M.; Al-Suhybani, *Br. Corros. J.* **1987**, *22*, 29.
28. Am. Soc. Test. Mater. Spec. Tech. Publ. No. 32, 1937.
29. Knapp, B. B. In *The Corrosion Handbook*; Uhlig, H. H., Ed.; Wiley/Chapman and Hall: New York/London, 1948; p 1077.
30. Teeple, H. O. *Am. Soc. Test. Mater. Spec. Tech. Publ*. No. 175, 1956, 89.
31. Am. Soc. Test. Mater. Spec. Tech. Publ. No. 970, 1988.
32. Gallagher, P.; Malpas, R. E.; Shone, E. B. *Br. Corros. J.* **1988**, *23*, 229.
33. Mercer, A. D.; Brook, G. M. *La Tribune de Cebedeau.* **1978**, *299*, 417–418.
34. Cheung, W. K.; Thomas, J. G. N. In *The Use of Synthetic Environments for Corrosion Testing;* ASTM STP; ASTM: Philadelphia, 1988; Vol. 970, p 190.
35. LaQue, F. L.; Stewart, W. C. *Métaux Corros.* **1948**, *23*, 147.
36. Vernon, W. H. J. *J. Soc. Chem. Ind. Lond.* **1947**, *66*, 137; *Corrosion* **1948**, *4*, 141.
37. Bengough, G. D.; May, R. *J. Inst. Met.* **1924**, *32*, 81.
38. May, R.; Stacpoole, R. W. de Vere, *J. Inst. Met.* **1950**, *77*, 331.
39. Brownsdon, H. W.; Bannister, L. C. *J. Inst. Met.* **1932**, *49*, 123.
40. Freeman, J. R., Jr.; Tracy, A. W. *Corrosion* **1949**, *5*, 245.
41. Butler, G.; Ison, H. C. K. *J. Appl. Chem.* **1960**, *10*, 80.
42. Fontana, M. G. *Ind. Eng. Chem.* **1947**, *39*, 87A.
43. Wagner, H. A.; Decker, J. M.; Marsh, J. C. *Trans. Am. Soc. Mech. Eng.* **1947**, *69*, 389.
44. Trembler, H. A.; Wesley, W. A.; LaQue, F. L. *Ind. Eng. Chem.* **1932**, *24*, 339.
45. Brennert, S. *J. Iron Steel Inst.* **1937**, *135*, 101P.
46. Smith, H. A. *Metal Prog.* **1938**, *33*, 596.
47. Hanawalt, J. D.; Nelson, C. E.; Peloubet, J. A. *Trans. Am. Inst. Miner. (Metall.) Eng.* **1942**, *147*, 273.
48. Benedicks, C. *Trans. Am. Inst. Miner. (Metall.) Eng.* **1925**, *71*, 597.
49. McAdams, D. J. *Heat Transmission*, 3rd ed.; McGraw-Hill: New York, 1954; p 370.
50. Groves, N. D.; Eisenbrown, C. M. *Metal Prog.* **1959**, *75*, 78.
51. Fisher, A. O.; Whitney, F. L., Jr. *Corrosion* **1959**, *15*, 257t.
52. Fisher, A. O. *Corrosion* **1961**, *17*, 215t.
53. Groves, N. D.; Eisenbrown, C. M.; Scharfstein, L. R. *Corrosion* **1961**, *17*, 173t.
54. Hart, R. J. In *Handbook on Corrosion Testing and Evaluation*; Ailor, W. H., Ed.; Wiley: New York, 1971; p 367.
55. LaQue, F. L. *J. Electrochem. Soc.* **1969**, *116*, 73C.
56. Stern, M.; Geary, A. L. *J. Electrochem. Soc.* **1957**, *104*, 56.
57. Stern, M. *Corrosion* **1958**, *14*, 440t.
58. Pourbaix, M. *Lectures on Electrochemical Corrosion Plenum Press*; 1973; see also references to potential-pH diagrams given in Section.
59. Armstrong, R. D.; Henthorne, M.; Thirsk, H. R. *J. Electroanal. Chem.* **1972**, *35*, 119.
60. Epelboin, I.; Keddam, M.; Takenouti, H. *J. App. Electrochem.* **1972**, *2*, 71.
61. Sathyaharayana, S. *Electroanal. Chem. Interfacial Electrochem.* **1974**, *50*, 411.
62. Macdonald, D. D. *Corrosion* **1990**, *46*, 229.
63. Kendig, M. W.; Mansfeld, F. Proc. Fall Meeting, Detroit, Electrochem. Soc., 82–2, **1982**, p 105.
64. Ferreira, M. G. S.; Dawson, J. L. In *Passivity of Metals and Semiconductors*; Froment, M., Ed.; **1980**, p 359.
65. Ferreira, M. G. S.; Dawson, J. L. *J. Electrochem. Soc.* **1983**, *132*, 760.
66. Keddam, M.; Oltra, R.; Colson, J. C.; Desestret, A. *Corros. Sci.* **1983**, *23*, 441.
67. Isaacs, H. S.; Kendig, M. W. *Corrosion* **1980**, *36*, 269.
68. Park, J. K.; Macdonald, D. D. *Corros. Sci.* **1983**, *23*, 293.
69. Epelboin, I.; Gabrielli, C.; Keddam, M.; Takenouti, H. *Z. Physik. Chem.* **1975**, *98*, 215.
70. Oltra, R.; Keddam, M. *Corros. Sci.* **1988**, *28*, 1.
71. Sato, N. *J. Electrochem. Soc.* **1976**, *123*, 1197.
72. Shibata, T.; Takeyama, T. *Corrosion* **1977**, *33*, 243.
73. Williams, D. E. Proc. In Electrochemical Corrosion Testing, Ferrara. 10–14 September 1985 DECHEMA 1986.
74. Williams, D. E.; Wescott, C.; Fleischmann, M. *J. Electrochem. Soc.* **1985**, *132*, 1796.
75. Williams, D. E.; Wescott, C.; Fleischmann, M. *J. Electrochem. Soc.* **1985**, *132*, 1804.
76. Gabrielli, C.; Huet, F.; Keddam, M.; Oltra, R. *Corrosion* **1990**, *46*, 266.
77. Ives, D. J. G.; Janz, G. J. *Reference Electrodes*; Academic Press: London, 1961.
78. Compton, K. G. *Mater. Res. Standards* **1970**, *10*, 13.
79. Covington, A. K. In *Electrochemistry*; The Chemical Society: London, 1970; Vol. 1, p 56.
80. Meites, L.; Moros, S. A. *Analyt. Chem.* **1959**, *31*, 25.
81. European Federation of Corrosion Publication No. 4, A WP report Guidelines on Electrochemical Corrosion Measurements; The Institute of Metals: London, 1990.
82. Berzins, T.; Delahay, P. *J. Am. Chem. Soc.* **1955**, *77*, 6448.
83. Pouli, D.; Huff, J. R.; Pearson, J. C. *Anal. Chem.* **1966**, *38*, 382.
84. Kooijman, D. J.; Sluyters, J. H. *Electrochim. Acta.* **1966**, *11*, 1147.
85. Bewick, J. *Electrochim. Acta.* **1968**, *13*, 825.
86. Piontelli, R.; Bianchi, G. Proc. 2nd. Meeting C.I.T.C.E., Milan 1951.
87. Piontelli, R. Proc. 4th Meeting C.I.T.C.E., London and Cambridge, 1952.
88. von Fraunhofer, J. A.; Banks, C. A. *Potentiostat and its Applications*; Butterworths: London, 1972.
89. Stern, M.; Makrides, A. C. *J. Electrochem. Soc.* **1960**, *107*, 782.
90. Greene, N. D.; France, W. D.; Wilde, B. E. *Corrosion* **1965**, *21*, 275.
91. Greene, N. D.; Acello, S. J.; Greif, A. J. *J. Electrochem. Soc.* **1962**, *109*, 1001.
92. Cleary, H. J.; Greene, N. D. *Electrochim. Acta.* **1965**, *10*, 1107.
93. France, W. D., Jr. *J. Electrochem. Soc.* **1967**, *114*, 818.
94. Wilde, B. E. *Corrosion* **1967**, *23*, 331.
95. Smith, L. W.; Pingel, V. J. *J. Electrochem. Soc.* **1951**, *98*, 48.
96. Budd, M. K.; Booth, F. F. *Metallurgia* **1962**, *66*, 245.
97. Cleary, H. J. *Corrosion* **1968**, *24*, 159.
98. Doig, P.; Edington, J. W. *Br. Corros. J.* **1974**, *9*, 88.
99. Davis, J. A. Proc. Conf. on Localised Corrosion; NACE: Williamsburg, 1971; p 168.
100. May, R. *J. Inst. Metals* **1928**, *40*, 141.
101. Hines, J. private communication.
102. Hoar, T. P.; Hines, J. G. *J. Iron Steel Inst.* **1965**, *182*(124), 156.
103. Hoar, T. P.; West, J. M. *Proc. Roy. Soc.* **1962**, *A268*, 304.
104. Horst, R. L., Jr.; Hollingsworth, E. H.; King, W. *Corrosion* **1969**, *25*, 199.

105. Pearson, J. M. *Trans. Electrochem. Soc.* **1942**, *81*, 485.
106. Schwerdtfeger, W. J. *Corrosion* **1963**, *19*, 17t.
107. Schwerdtfeger, W. J.; Manuele, R. J. *Corrosion* **1963**, *19*, 59t.
108. Skold, R. V.; Larson, T. E. *Corrosion* **1957**, *13*, 139t.
109. Stern, M.; Weisert, E. D. *Proc. Am. Soc. Test. Mater.* **1959**, *59*, 1280.
110. Barnarrt, S. *Corrosion* **1971**, *27*, 467.
111. Barnarrt, S. *Corrosion Sci.* **1969**, *9*, 145.
112. Leroy, R. L. *Corrosion* **1973**, *29*, 272.
113. Oldham, K. B.; Mansfeld, F. *Corrosion* **1971**, *27*, 434.
114. Oldham, K. B.; Mansfeld, F. *Corrosion Sci.* **1973**, *13*, 811.
115. Hickling, J. Ph.D. Thesis, University of Cambridge, 1974.
116. Hoar, T. P. *Corrosion Sci.* **1967**, *7*, 455.
117. Mansfeld, F. *Corrosion* **1973**, *29*, 397.
118. Mansfeld, F. *Corrosion* **1974**, *30*, 92.
119. Mansfeld, F. *J. Electrochem. Soc.* **1971**, *118*, 545.
120. Mansfeld, F. *J. Electrochem. Soc.* **1973**, *120*, 515.
121. Makrides, A. C. *Corrosion* **1969**, *25*, 455.
122. Legault, R. A.; Walker, M. S. *Corrosion* **1963**, *19*, 222.
123. Walker, M. S.; France, W. D. *Mat. Protect.* **1969**, *8*, 47.
124. Jones, D. A.; Greene, N. B. *Corrosion* **1969**, *25*, 367.
125. Wilde, B. E. *Corrosion* **1967**, *23*, 379.
126. Jones, D. A. *Corros. Sci.* **1968**, *8*, 19.
127. Bureau, M. 9th Fatipec Congress 1968, *79.*
128. Mikhailovskii, Y. N.; Leonev, V. V.; Tomashov, N. D. *Korrozya Metallov i Splavov Sbornik 2*; Metallurgizdat: Moscow, 1965.
129. Butler, T. J.; Carter, P. R. *Electrochem. Tech.* **1963**, *1*, 22.
130. Walpole, J. F. *Bull. Inacol.* **1972**, *23*, 22.
131. Bird, D. W. *Bull. Inacol.* **1971**, *22*, 149.
132. Rowlands, J. C.; Bentley, M. N. *Br. Corros. J.* **1972**, *7*, 42.
133. Wesley, W. A. *Trans. Electrochem. Soc.* **1938**, *73*, 539.
134. Copson, H. R. *Ind. Eng. Chem.* **1945**, *37*, 721.
135. Rowe, L. C. *J. Mater.* **1970**, *5*, 323.
136. Denison, I. A. *J. Res. Nat. Bur. Stand.* **1936**, *17*, 363.
137. Ewing, S. P. *Am. Gas. Assoc. Mon.* **1932**, *14*, 356.
138. Schwerdtfeger, W. J. *J. Res. Nat. Bur. Stand.* **1953**, *50*, 329.
139. Schwerdtfeger, W. J. *J. Res. Nat. Bur. Stand.* **1954**, *52*, 265.
140. Schwerdtfeger, W. J. *J. Res. Nat. Bur. Stand.* **1957**, *58*, 145.
141. Schwerdtfeger, W. J. *J. Res. Nat. Bur. Stand.* **1961**, *65C*, 271.
142. Logan, R. H.; Ewing, S. P.; Denison, I. A. *Am. Soc. Test. Mater. Spec. Tech. Publ.*; No. 32, 1937, 95.
143. Jones, D. A.; Lowe, T. A. *J. Mater.* **1969**, *4*, 600.
144. Fuller, T. S. *Proc. Am. Soc. Test. Mater.* **1924**, *27*, 281.
145. Lathrop, E. C. *Proc. Am. Soc. Test. Mater.* **1924**, *24*, 281.
146. Streicher, M. A. *ASTM Bull* No. 188, **1953**, p 35.
147. Streicher, M. A. *ASTM Bull* No. 195, **1954**, p 63.
148. Streicher, M. A. *J. Electrochem. Soc.* **1959**, *106*, 161.
149. Streicher, M. A. *Corrosion* **1963**, *19*, 272t.
150. Saur, R. L.; Basco, R. P. *Plating* **1966**, *53*, 33.
151. Saur, R. L.; Basco, R. P. *Plating* **1966**, *53*, 320.
152. Saur, R. L.; Basco, R. P. *Plating* **1966**, *53*, 981.
153. Saur, R. L. *Plating* **1966**, *54*, 393.
154. Englehart, E. T.; George, D. J. *Mater. Prot.* **1964**, *3*, No. 11, 25.
155. Capp, J. A. *Proc. Am. Soc. Test. Mater.* **1914**, *14*, 474.
156. LaQue, F. L. *Proc. Am. Soc. Test. Mater.* **1951**, *51*, 495.
157. LaQue, F. L. *Mater. Meth.* **1952**, *35*, No. 2, 77.
158. Sample, C. H. *Bull. Am. Soc. Test. Mater.* , **1943**, *123*, 19.
159. May, T. P.; Alexander, A. L. *Proc. Am. Soc. Test. Mater.* **1950**, *50*, 1131.
160. Darsey, V. M.; Cavanaugh, W. R. *Proc. Am. Soc. Test. Mater.* **1948**, *48*, 153.
161. Nixon, C. F. *Mon. Ren. Am. Electropl. Soc.* **1945**, *32*, 1105.
162. Pinner, W. L. *Plating* **1957**, *44*, 763.
163. Nixon, C. F.; Thomas, J. D.; Hardesty, D. W. *46th Ann. Tech. Proc. Am. Electropl. Soc.* **1959**, 159.
164. Thomas, J. D.; Hardesty, D. W.; Nixon, C. F. *47th Ann. Tech. Proc. Am. Electropl. Soc.* **1960**, 90.
165. LaQue, F. L. 46th Ann. *Tech. Proc. Am. Electropl. Soc.* **1959**, 141.
166. Bigge, D. M. 46th Ann. *Tech. Proc. Am. Electropl. Soc.* **1959**, 149.
167. Edwards, J. 46th Ann. *Tech. Proc. Am. Electropl. Soc.* **1959**, 154.
168. Edwards, J. *Trans. Inst. Metal Finish.* **1958**, *35*, 55.
169. Kesternich, W. *Stahl u. Eisen* **1951**, *71*, 587.
170. Dix, E. H., Jr.; Bowman, J. J. *Am. Soc. Test. Mater. Spec. Tech. Publ.* No. 32, **1937**, 57.
171. Carter, V. E., Ed. *Corrosion Testing for Metal Finishing*; Butterworths: London, 1982.
172. Swinden, T.; Stevenson, W. W. *J. Iron St. Inst.* **1940**, *142*, 165P.
173. Lloyd, T. E. *J. Metals NY* **1950**, *188*, 1092.
174. Evans, U. R.; Britton, S. C. *J. Iron St. Inst. Spec. Rept. No.* **1931**, *1*, 139.
175. Dennis, J. K.; Such, T. E. *Trans. Inst. Metal Finish.* **1963**, *40*, 60.
176. Dennis, J. K.; Such, T. E. *Nickel and Chromium Plating*, 2nd ed.; Butterworths: London, 1986.
177. Chandler, K. A.; Kilcullen, M. B. *Br. Corros. J.* **1970**, *5*, 1.
178. Bromley, A. F.; Kilcullen, M. B.; Stanners, J. F. In 5th European Congress of Corrosion, Paris, September 1973.
179. Pourbaix, M. *CEBELCOR RT*; **1969**, 160.
180. Legault, R. A.; Mori, S.; Leckie, H. P. *Corrosion* **1970**, *26*, 121.
181. Legault, R. A.; Mori, S.; Leckie, H. P. *Corrosion* **1973**, *29*, 169.
182. Okada, H.; Hosio, U.; Naito, H. *Corrosion* **1970**, *26*, 429.
183. Brown, M. H. *Corrosion* **1974**, *30*, 1.
184. Cowan, R. L.; Tedmon, C. S., Jr. In Fontana, M. G., Staehle, R. W., Eds.; *Advances in Corrosion Science and Technology;* Plenum Press: New York, 1973; Vol. 3.
185. Huey, W. R. *Trans. Am. Soc. Steel Treat.* **1930**, *18*, 1126.
186. Henthorne, M. *Corrosion* **1974**, *30*, 39.
187. Strauss, B.; Schottky, H.; Hinnüber, J. *Z. Anorg. Allgem. Chem.* **1930**, *188*, 309.
188. Rocha, H. J. in discussion of paper by Brauns, E., Pier, G., Stahl, U. *Eisen* **1955**, *75*, 579.
189. Scharfstein, L. R.; Eisenbrown, C. M. ASTM STP No. 369, **1963**, 253.
190. Tedmon, C. S., Jr.; Vermilyea, D. A.; Rosolowski, J. H. *J. Electrochem. Soc.* **1971**, *118*, 192.
191. Ebling, H.; Scheil, M. A. *ASTM Special Tech. Publ.* No. 93, **1949**, 121.
192. Warren, D. *ASTM Bulletin*. No. 230, **1958**, 45.
193. Streicher, M. A. *ASTM Bulletin*. No. 188, **1953**, 35.
194. Clerbois, L.; Clerbois, F.; Massart, J. *Electrochem. Acta.* **1959**, *1*, 70.
195. France, W. D.; Greene, N. D. *Corros. Sci.* **1968**, *8*, 9.
196. Streicher, M. A. *Corros. Sci.* **1969**, *9*, 55.
197. France, W. D.; Greene, N. D. *Corros. Sci.* **1970**, *10*, 379.
198. Streicher, M. A. *Corros. Sci.* **1971**, *11*, 275.
199. Cihal, V.; Desestret, A.; Froment, M.; Wagner, G. H. Proc. Conf. European Federation on Corrosion, Paris, France **1973**; p 249.
200. Novak, P.; Stefec, R.; Franz, F. *Corrosion* **1975**, *31*, 344.

201. Kolotyrkin, Ya. M. *Zashch. Met.* **1975**, *11*, 699.
202. Clark, W. L.; Cowan, R. L.; Walker, W. L. *Comparative Methods for Measuring Degree of Sensitisation in Stainless Steel* ASTM STP; ASTM: Philadelphia, 1978; Vol. 656, p 99.
203. Lee, J. B. *Corrosion* **1986**, *42*, 106.
204. Streicher, M. A. *Corrosion* **1974**, *30*, 77.
205. Fontana, M. G.; Greene, N. D. *Corrosion Engineering*; McGraw-Hill: New York, 1967.
206. Wilde, B. E. *Corrosion* **1972**, *28*, 283.
207. France, W. D.; Greene, N. D. *Corrosion* **1970**, *26*, 1.
208. Wilde, B. E.; Williams, E. *J. Electrochem. Soc.* **1970**, *117*, 775.
209. Wilde, B. E.; Greene, N. D. *Corrosion* **1969**, *25*, 300.
210. Henry, W. D.; Wilde, B. E. *Corrosion* **1969**, *25*, 515.
211. Wilde, B. E.; Williams, E. *J. Electrochem. Soc.* **1971**, *118*, 1058.
212. May, R.; Stacpoole, R. W. de V. *J. Inst. Met.* **1950**, *77*, 331.
213. Efird, K. D. *Corrosion* **1977**, *33*, 347.
214. Brownsdon, H. W.; Bannister, L. C. *J. Inst. Met.* **1932**, *49*, 123.
215. Breckon, C.; Gilbert, P. T. Ist Int. Congress on Met. Corrosion, Butterworths, London, **1962**, p 624.
216. Bem, R. S.; Campbell, H. S. Ist Int. Congress on Met. Corrosion, Butterworths, London, **1962**, p 630.
217. Campbell, H. S. MP577, BNFMRA, 1973.
218. LaQue, F. L. *Marine Corrosion*; Wiley: New York, 1975; p 62.
219. Grant, A. A.; Phillips, L. *The Application of Advanced Materials Technology in Fluid Engineering*; Institute of Mechanical Engineering: London, 1990.
220. Kohley, T.; Heitz, E. In *The Use of Synthetic Environments for Corrosion Testing* ASTM STP; Francis. P. E., Lee, T. S., Eds., ASTM: Philadelphia, 1988; Vol. 970, pp 235–245.
221. Sprowls, D. O. In *ASM Metals Handbook*, 9th ed. 1987; Vol. 13, pp 291–302.
222. Haigh, B. P. *J. Inst. Metals* **1971**, *18*, 55.
223. Huddle, A. U.; Evans, U. R. *J. Iron Steel Inst.* **1944**, *149*, 109P.
224. Inglis, N.; Lake, G. F. *Trans. Faraday Soc.* **1931**, *17*, 803.
225. Rawdon, H. S. *Proc. Am. Soc. Test. Mater.* **1929**, *19*, 314.
226. Kenyon, J. N. *Proc. Am. Soc. Test. Mater.* **1940**, *40*, 705.
227. Gould, A. J.; Evans, U. R. *Iron Steel Inst. Spec. Report* No. 24, **1939**, 235.
228. Evans, U. R.; Simnad, M. T. *Proc. Roy. Soc.* **1947**, *A188*, 372.
229. Gough, H. J.; Sopwith, D. G. *J. Iron Steel Inst.* **1933**, *127*, 301.
230. Gough, H. J.; Sopwith, D. G. *Engineering* **1933**, *136*, 75.
231. Jarman, R. A.; Smith, S.; Williams, R. A. *Br. Corros. J.* **1978**, *13*, 195.
232. Hoeppner, D. W. Corrosion Fatigue, NACE-2 University of Connecticut, 1972; p 3.
233. McEvily, A. J.; Wei, R. P. Corrosion Fatigue, NACE-2 University of Connecticut, 1972; p 381.
234. Kitigawa, H. Corrosion Fatigue, NACE-2 University of Connecticut, 1972; p 521.
235. Wood, R. J. K.; Fry, S. A. *J. Fluids Eng.* **1989**, *111*, 271.
236. Lichtman, J. Z.; Kallas, D. H.; Rufola, A. In *Handbook on Corrosion Testing and Evaluation*; Ailor, W. H., Ed.; Wiley: New York, 1971; p 453.
237. Schroter, H. *Z. Ver. Dtsch. Ing.* **1934**, *78*, 349.
238. Hobbs, J. M. *Proc. Cavitation Forum, ASME* **1966**, 1.
239. Gaines, N. *Physics* **1932**, *3*, 209.
240. Kerr, S. L. *Trans. Am. Soc. Mech. Eng.* **1937**, *59*, 373.
241. Beeching, R. *Trans. Inst. Eng. Shipb. Scot.* **1946**, *90*, 203.
242. Rheingans, W. J. In *Engineering Approach to Surface Damage*; Lipson, C., Colwell, L. V., Eds.; University of Michigan: Ann Arbor, MI, 1958; p 249.
243. Leith, W. C.; Thompson, A. L. *Trans Am. Soc. Mech. Eng. J. Basic Eng.* D **1960**, *82*, 795.
244. Plesset, M. S. *Trans. ASME Series D, J. Basic. Eng.* **1963**, *85*, 360.
245. Eisenberg, P.; Preiser, H. S.; Thiruyengadam, A. *Trans. SNAME* **1965**, *73*, 241.
246. Hobbs, J. M. In *ASTM STP*; 1967, Vol. 408; p 159.
247. Plesset, M. S.; Devine, R. D. *J. Basic Eng. Trans. ASME*; **1966**, *692*, Dec.
248. Schulmeister, R.; Speckhardt, H. In *ASTM STP* **1980**; Vol. 705, p 81.
249. Hudgens, R. D.; Carver, D. P.; Hercamp, R. D.; Lauterback, J. In *ASTM STP*; 1980, Vol. 233.
250. Chance, R. L. In *ASTM STP*; 1980, Vol. 270.
251. In *ASTM STP*; 1953, Vol. 144.
252. Waterhouse, R. B. *Fretting Corrosion*; Pergamon Press: New York, 1972.
253. Waterhouse, R. B. Proc. 10th International Conference on Metallic Corrosion, Madras, India, 7–11 November 1987, 5, Oxford and IBH Publishing Co. New Delhi, 1987; p 63.
254. Fink, M. *Trans. Am. Soc. Steel Treat.* **1930**, *18*, 1026.
255. Shotter, G. F. *J. Inst. Elec. Eng.* **1934**, *75*, 755.
256. Tomlinson, G. A.; Thorpe, P. L.; Gough, H. J. *J. Inst. Mech. Eng.* **1939**, *141*, 233.
257. Wright, K. H. R. *Proc. Instn. Mech. Eng.* **1952–1953**, *1B*, 556.
258. Uhlig, H. H.; Tierney, W. D.; McClellan, A. In *ASTM STP* **1953**; Vol. 144, p 71.
259. McDowell, J. R. In *ASTM STP*; 1953, Vol, 24.
260. Warlow-Davies, E. J. *J. Inst. Mech. Eng.* **1941**, *146*, 32.
261. Godfrey, D. *Tech. Note 2039, Natl. Advisory Comm. Aeronaut* 1950.
262. Gray, A. C.; Jenny, R. W. *S.A.E.J.* **1944**, *52*, 511.
263. de Villemeur, Y. *Metaux. Paris* **1959**, *34*, 413.
264. Wright, K. H. R. *Proc. Instn. Mech. Eng.* **1952–1953**, *1B*, 556.
265. Wright, K. H. R. *Proc. Instn. Mech. Eng*. **1966–1967**, *181* (30), 256.
266. Barwell, F. T.; Wright, K. H. R. *J. Res. Br. Cast Iron. Ass.* **1958**, *7*, 190.
267. Field, J. E.; Waters, D. M. *N.E.L. Rep. No.* **1967**, 275.
268. Waterhouse, R. B. *J.I.S.I.* **1961**, *197*, 301.
269. Epstein, L. F. *Proc. Int. Conf. Peaceful Uses of Atomic Energy, New York* **1956**, *9*, 311.
270. Bakish, R.; Kern, F. *Corrosion* **1960**, *9*, 533t.
271. Edeleanu, C.; Gibson, J. G. *J. Inst. Met.* **1960**, *88*, 321.
272. Brasunas, A.; de, S. *Corrosion* **1953**, *9*, 78.
273. Miller, E. C. In *Liquid Metals Handbook*; Lyon, R. N., Ed.; Atomic Energy Comm. and Dept. of the Navy: Washington DC, 1952; p 144.
274. Borgstedt, H. U.; Frees, G. Proc. 10th Int. Cong. on Metallic Corrosion, Madras, India, 7–11 Nov. 1987, 3, **1988**; p 1843.
275. Barker, M. G.; Coen, V.; Kolbe, H.; Lees, J. A.; Orecchia, L.; Sample, T. *J. Nucl. Mater.* **1988**, *155–157B*, 732.
276. Borgstedt, H. U.; Frees, G. *Werkst. Korros.* **1990**, *41*, 1.
277. Vreeland, D. C.; Hoffman, E. E.; Manly, W. D. *Nucleonics* **1953**, *11*, 36.
278. Grabner, H.; Feurstein, H.; Oschinski, J. *J. Nucl. Mater.* **1988**, *155–157B*, 702.
279. Strachan, J. F.; Harris, N. L. *J. Inst. Met.* **1956–1957**, *85*, 17.
280. Klueh, R. L. Proc. Int. Conf. Sodium Technol. Large Fast Reactor Design, Nov. 7–9 ANL-7520, Pt. 1, 171, Argonne National Laboratory, 1968.
281. Klueh, R. L. *Corrosion* **1969**, *25*, 416.
282. DiStefano, J. R. ORNL-4028 Oak Ridge Laboratory, 1966.
283. Manly, W. D. *Corrosion* **1956**, *12*, 336t.

284. DiStefano, J. R.; De Van, J. H. *Nuclear Appl. Tech.* **1970**, *8*, 29.
285. Koenig, R. F.; Vandenberg, S. R. *Metal. Prog.* **1952**, *61*, 71.
286. Hoffman, E. E. *Corrosion of Materials by Lithium at Elevated Temperatures, ORNL-2924*; Oak Ridge National Laboratory, 1960.
287. Kassner, T. F. *AIME Met. Soc. Trans.* **1967**, *239*, 1643.
288. Shibanova, L. N.; Vostryakov, A. A.; Lepinskikh, B. M. *Zashch. Met.* **1986**, *22*, 124.
289. De Van, J. H.; Sessions, C. E. *Nucl. Appl.* **1967**, *3*, 102.
290. De Van, J. H.; Jansen, D. H. Progr. Report 30 September 1968, ORNL-4350, Oak Ridge National Laboratory, p 91.
291. Bonilla, C. F. In *Reactor Handbook*; Mclain, S. ED.; Interscience: New York, 1964; Vol. IV, p 107.
292. Romana, A. J.; Fleitman, A. H.; Klamut, C. J. Proc. AEC-NASA Liquid Metals Inform. Meeting, CONF-650411 1965.
293. Fuller, L. C.; MacPherson, R. E. ORNL-TM-2595 Oak Ridge National Laboratory, 1967.
294. Borgstedt, H. U.; Frees, G. *Werkst. Korros.* **1987**, *38*, 732.
295. RDT Standard C8–5T, Electrochemical Oxygen Meter for Service In *Liquid Sodium*, RDT Standards Office, Oak Ridge National Laboratory, Tennessee.
296. Pillai, S. R.; Mathews, C. K. *J. Nucl. Mater.* **1986**, *137*, 107.
297. Schutter, F.; De Dekeyser, J. A.; Tas, H.; Burbure, S. de. *J. Nucl. Mater.* **1988**, *155*, 744.
298. Roy, P.; Wozaldo, G. P.; Comprelli, F. A. In *Proc. Int. Conf. Sodium Technolog. Large Fast Reactor Design, 7–9 November 1968; ANL-7520, Pt. 1, Argonne National Laboratory, p 131.*
299. Roy, P.; Gebhardt, M. F. GEAP-13548 General Electric Company, 1969.
300. Hoffmann, E. E.; Harrison, R. W. *Metallurgy and Technology of Refractory Metal Alloys*; Plenum Press: New York, 1969; p 251.
301. Harrison, R. W. GESP-258. General Electric Company, 1969.
302. Borgstedt, H. U.; Grundman, M. Preprints of Eurocorr 87th conference, Karlsruhe, Germany, 6–10 April 1987, DECHEMA, Frankfurt **1987**; p 141.
303. Dillon, C. P. *Mater. Perform.* **1990**, *29*(11), 54.
304. Lynch, S. P. Proceedings of 2nd International Conference on Environmental Degradation of Engineering Materials, Blacksburg, Virginia, September 31–23 1981 Virginia Polytechnic Institute, 1981; p 229.
305. Old, C. F. *Metal Science* **1980**, *14*, 433.
306. Preece, C. M. In Proceedings of International Conference on Stress Corrosion Cracking and Hydrogen Embrittlement of Iron Base Alloys, Unieux-Firminy, France, June 12–16 1973 NACE: Houston, 1977; p 625.
307. Kapp, J. A.; Duquette, D.; Kamdar, M. H. *J. Eng. Mater. Technol.* **1986**, *108*, 37.
308. Thompson, D.H. in Ref. 2
309. Dillon, C. P.; Krisher, A. S.; Wissenburg, H. *Ref.* **1971**, *2*, 599.
310. Strutt, J. E.; Nichols, J. R. Eds. *Plant Corrosion: Prediction of Materials Performance*; Ellis Horwood: Chichester, 1987.
311. Rawdon, H. S. *ASTM STP* **1937**, *32*, 36.
312. LaQue, F. L. *Proc. Am. Soc. Test. Mater.* **1951**, *51*, 495.
313. Finkeldey, W. H. *Proc. Am. Soc. Test. Mater.* **1932**, *32*, 226.
314. Hudson, J. C. *J. Iron Steel. Inst.* **1943**, *161P*, 148.
315. Pilling, N. B.; Wesley, W. A. *Proc. Am. Soc. Test. Mater.* **1940**, *40*, 643.
316. Copson, H. R. *Proc. Am. Soc. Test. Mater.* **1948**, *48*, 591.
317. Gibboney, J. H. *Proc. Am. Soc. Test. Mater.* **1919**, *19*, 181.
318. Mendizza, A. *Proc. Am. Soc. Test. Mater.* **1950**, *50*, 114.
319. Passano, R. F. *Proc. Am. Soc. Test. Mater.* **1934**, *34*, 159.
320. Pinner, W. L. *Proc. Am. Soc. Test. Mater.* **1953**, *53*, 256.
321. Wirshing, R. J.; McMaster, W. D. *Paint Varn. Prod.* **1951**, *41*, 13.
322. Sereda, P. J. *Bull. Am. Soc. Test. Mater. No.* **1958**, *228*, 53.
323. Guttman, H.; Sereda, P. *J. ASTM STP* **1968**, 435.
324. Gorman, L. J. *Proc. Am. Soc. Test. Mater.* **1939**, *39*, 247.
325. Pray, H. A. *Proc. Am. Soc. Test. Mater.* **1944**, *44*, 280.
326. Humble, H. A. *Corrosion* **1949**, *5*, 292.
327. Southwell, C. R. NRL Reports, Naval Research Laboratory, Washington, D.C.
328. Ijsseling, F. P. *Br. Corros. J.* **1989**, *24*, 55.
329. Logan, R. H.; Ewing, S. P.; Denison, I. A. *ASTM STP* **1937**, *32*, 95.
330. Scott, G. N. *Proc. Am. Petrol. Inst.* **1937**, 95.
331. Escalante, E. Ed. *Underground Corrosion* ASTM STP; ASTM: Philadelphia, 1981; Vol 741.
332. Tiller, A. K. *Biocorrosion in Civil Engineering*; Cranfield Institute of Technology, 1990.
333. Walter, G. W. *Corros. Sci.* **1986**, *26*, 681.
334. Timmins, F. D. *J. Oil Col. Chem. Assoc.* **1979**, *62*, 131.
335. Ali Elbasir; Scantlebury, J. D.; Callow, L. M. *J. Oil Col. Chem. Assoc.* **1985**, *67*, 282.
336. Hoey, C. E.; Hipwood, H. A. *J. Oil Col. Chem. Assoc.* **1974**, *57*, 151.
337. Butler, G.; Mercer, A. D.; Warren, G. M. In Eurocorr 77, 6th Eur. Congr. Metall. Corros., Soc. Chem. Ind., London **1977**; pp 349–355.
338. Mercer, A. D. In *Engine Coolant Testing: State of the Art 1979*, ASTM STP; Ailor, W. H., Ed., ASTM: Philadelphia, 1980; Vol. 705, pp 53–80.
339. NACE Publication. *Mater. Perform.* **1982**, *21*(12), 45.
340. Schmitt, G.; Bruckhoff, W. In Proc. 5th European Conference on Corrosion Inhibitors (5 SEIC) University of Ferrara, 1980; p 323.
341. Mercer, A. D. *Br. Corros. J.* **1985**, *20*, 61.
342. Mercer, A. D. In Proc. 6th European Symposium on Corrosion Inhibitors, (6 SEIC), University of Ferrara, N.S. Sez. V, Suppl. N8 1985.
343. Wachter, A.; Skei, T.; Stillman, N. *Corrosion* **1951**, *7*, 284.
344. Stroud, E. G.; Vernon, W. H. J. *J. Appl. Chem.* **1982**, *2*, 178.
345. Levin, S. Z.; Gintzberg, S. A.; Dinner, S. M.; Kuchinsky, V. N. 2nd Ferrara Conference on Corrosion Inhibitors (2SEIC) 1965 University of Ferrara, 1966; p 765.
346. Romeo, A. J.; Skrinde, R. T.; Eliassen, R. *Proc. Am. Soc. Civ. Eng.* **1958**, *84*(SA4).
347. Gleekman, L. W.; Swandby, R. K. *Corrosion* **1961**, *17*, 144t.
348. Makar, D. R.; Francis, H. T. *J. Electrochem. Soc.* **1955**, *102*, 669.
349. Todt, F. *Z. Elektrochem.* **1928**, *34*, 586.
350. Todt, F. *Z. Ver. Dtsch. Zuckerind.* **1929**, *79*, 680.

2.35 Environmentally Assisted Cracking Test Methods

R. N. Parkins
Department of Metallurgy and Engineering Materials, Newcastle Univeristy, Newcastle-upon-Tyne NE1 7RU, UK

This article is a revision of the Third Edition article 8.10 by R. N. Parkins, volume 1, pp 8:215–8:242, © 2010 Elsevier B.V.

Abbreviations

ASTM American Society for Testing and Materials
EAC Environmentally assisted cracking
UTS Ultimate tensile strength

Symbols

E Modulus of elasticity
E_{corr} Corrosion potential
h Height
I Current density
K_{ISCC} Threshold stress intensity factor for the onset of stress corrosion cracking
t Specimen thickness
t_f Time to failure
V Potential
σ Maximum tensile stress
σ_{th} Threshold stress for the onset of stress corrosion cracking
σ_y Lower yield strength

2.35.1 Introduction

2.35.1.1 Stressing Systems

Many different methods[1–4] have been used for the stressing of specimens (or testpieces), from which it may be reasonably assumed that there is no single method that is markedly superior to all the others. Each method may have its peculiar advantages in a given situation, but ideally, a test method should not be so severe that it leads to the condemnation of a material that would prove adequate for service or so trifling as to permit the use of materials in circumstances where rapid failure ensues. Methods of stressing testpieces, whether initially plain, notched, or precracked, can be conveniently grouped according to whether they involve:

1. a constant total strain or deflection;
2. a constant load;
3. an imposed strain or deflection rate.

Constant deflection tests usually have the attraction of employing simple and, therefore, often cheap

specimens and straining frames and of simulating the fabrication stresses that are most frequently associated with stress corrosion failure. Constant load tests may simulate failure from applied or working stresses more closely. Although tests involving the application of a constant deflection rate (strain rate) are commonly used, their relevance to service failures continues to be debated.

2.35.1.2 Constant Total-Deflection Tests

Prismatic beams stressed by bending offer a simple means of testing sheet or plate material, typical arrangements being shown in **Figure 1(a)–1(e)**. Below the elastic limit, the stresses may be calculated[1,5] or determined from the response of strain gauges attached to the surface at an appropriate position.

Plastic bending of strip specimens to produce a 'U'-bend, **Figure 1(d)** and **1(e)**, will usually allow the use of a lighter restraining system, although some of the effects of the plastic deformation, if not removed by subsequent heat treatment, may be to influence cracking response, and the stress obtained in the outer fibers of the specimen is usually less reproducible than with more sophisticated specimens. Tubular material may be tested in the form of 'C'- or 'O'-rings, the former being stressed by partial closing of the gap, **Figure 1(c)**, and the latter by the forced insertion of a plug that is appropriately oversized for the bore. The circumferential stress at the outer

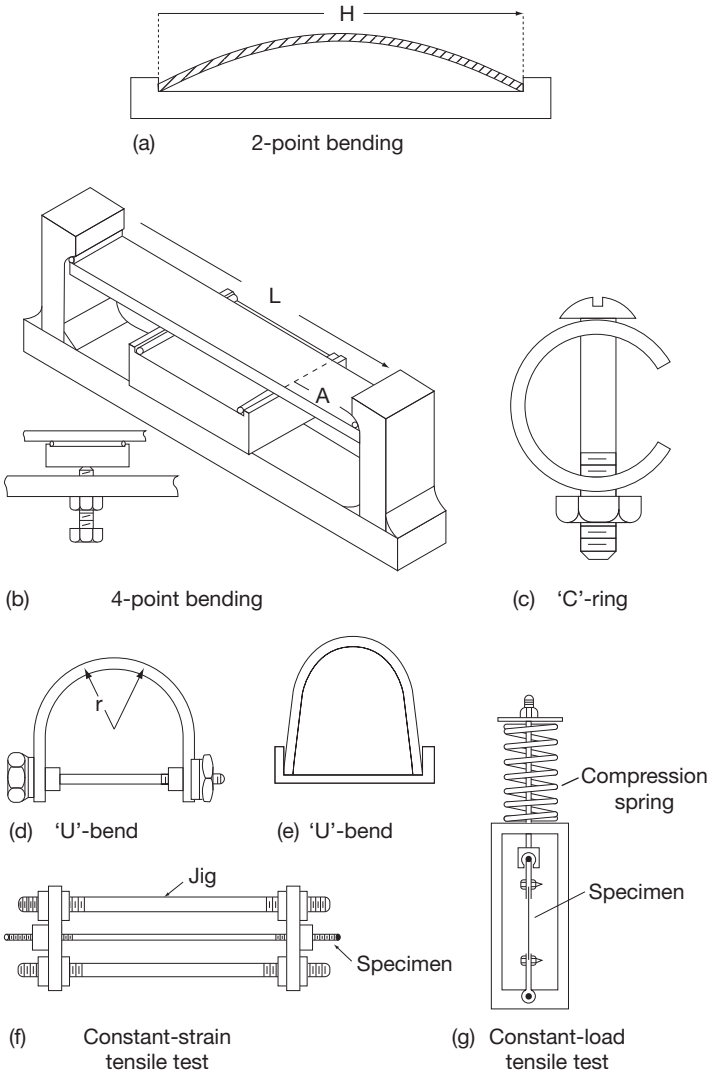


Figure 1 Stressing systems for stress corrosion test specimens; (a) to (f) constant strain, (g) constant load.

surface of a 'C'-ring is maximal midway between the bolt holes, but for the 'O'-ring, it is constant over the periphery, the stresses being readily calculated in terms of measured deflections.[1,2]

Constant-deflection tensile tests, Figures 1(f) and 1(g), are sometimes preferred to bend tests, but for similar cross sections, require a more massive restraining frame. In principle, this problem may be surmounted by the use of internally stressed specimens containing residual stresses as the result of inhomogeneous deformation. The latter may be introduced by plastic bending, for example, by producing a bulge in sheet or plate material, or by welding, but such tests cause problems in the systematic variation of the initial stress, which will usually be in the region of the yield stress. Moreover, elastic spring-back, in introducing residual stresses by bulging plate or partially flattening tube, may introduce problems, and where welding is involved, the structural modifications may raise difficulties unless the test is simulative of a practical situation.

At least as important as the choice of methods of stressing is the realization of the limitations of the various methods, these having been considered in a review of stress corrosion test methods.[6] The stiffness of the stressing frame in constant-deflection tests may influence results because of relaxation in the specimen during the initial loading stage and during subsequent crack propagation. Especially in testing ductile materials, the initial elastic strain is converted in part to plastic strain, even if the total deflection remains constant during cracking. This is because, as the crack propagates, the stress increases on the remaining uncracked portion of the specimen section beyond the crack, eventually reaching the effective yield stress. Yielding will then occur, accompanied by yawning of the crack and frequently, with the propagation of a Lüders band that results in a sharp load drop, which is sometimes mistaken as an indication of the crack having advanced by a burst of mechanical fracture. Once load relaxation has been initiated, the extent to which it proceeds can vary from specimen to specimen. Thus, Figure 2 shows load relaxation curves for two specimens of the same maraging steel in the same stressing frame, which had a facility for load recording throughout the test. The specimens differed in the extent to which they showed load relaxation prior to sudden fracture, this difference being related to the number of cracks that developed in the specimens. Marked load relaxation was associated with the development of many cracks in the specimen and little relaxation, with only a few cracks.

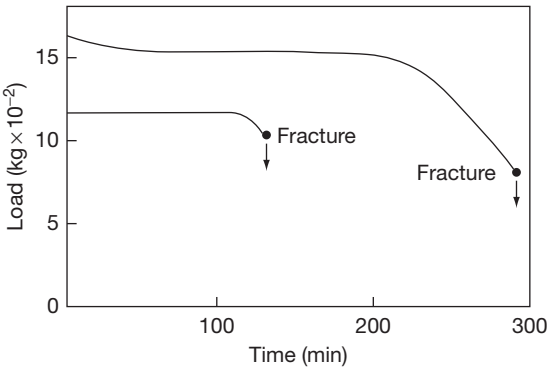


Figure 2 Load relaxation curves for a maraging steel stress corroded in 0.6 M NaCl at pH 2. Reproduced from Parkins, R. N.; Haney, E. G. *Trans. Metall. Soc., AIME* **1968**, *242*, 1943.

This can influence the time to failure, as is apparent from Figure 2, where the specimen stressed at the initially higher load took longer to fail than that at the initially lower one. This is because, when only a single stress corrosion crack develops, it will not need to grow to large dimensions before sudden, final failure occurs, since the applied load remains high, whereas with the marked load relaxation associated with the multicracked specimen, one of the cracks will need to propagate much further before it reaches the size for sudden fracture at the reduced load. Such an explanation conforms to the observations[7] that the load at fracture is related to the area of stress corrosion cracking upon the final fracture surface and to the number of cracks initiated.

This type of result will depend upon the nature of the stress corrosion system being studied, that is, upon such properties as the fracture toughness of the material and even upon the aggressiveness of the environment employed. It will also vary according to the stiffness of the restraining jig employed, since the stiffer the frame, the less the elastic strain that is likely to remain in the specimen after the propagation of a Lüders band, so that a stress-corrosion crack may cease to propagate in some circumstances, especially if the initial stress is in the vicinity of the threshold stress. This indicates some of the dangers inherent in comparing stress corrosion resistances in terms of times to failure at a given initial stress, an approach that is often practiced but can be misleading. Figure 3 shows the results from some tests in which the time to failure of specimens previously cold worked in varying amounts is plotted against initial stress. Comparison of the effects of different amounts of cold work by tests at an initial

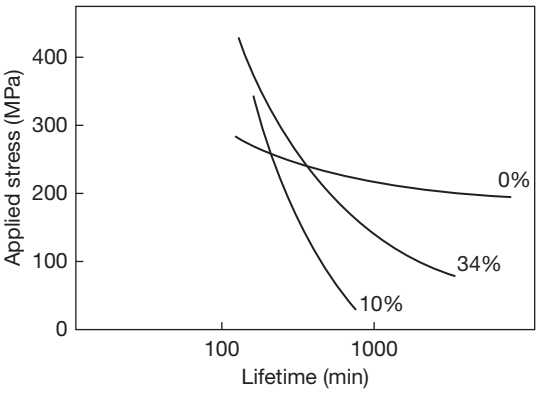


Figure 3 Effects of different amounts of prior cold work (0, 10 and 34%) on the stress corrosion of a 0.07% C steel in boiling 4 M NH_4NO_3.

Table 1 Relative susceptibilities to cracking of a mild steel in boiling 4 M NH_4NO_3 after various amounts of cold work

Initial stress	*Susceptibility of different cold-worked conditions*		
	Most (%)	*Intermediate (%)*	*Least (%)*
280 MPa	0	10	34
155 MPa	10	34	0
100% of yield stress	34	10	0
30% of yield stress	10	34	0

stress of 280 or 155 MPa gives different orders of susceptibility, as shown in Table 1. It could be argued that neither of these results is correct because the prior cold work would result in different yield strengths being developed in the three different conditions and that the results should be rationalized by making the comparison a function of the respective yield strengths. Here again, however, the order of susceptibility varies according to the rationalized stress at which the comparison is made, as the results in Table 1 show. It is difficult to escape the conclusion that a more satisfactory basis of comparison is the threshold, but even the latter may not be a basis for comparison of results obtained using different restraining frames.

The simplicity of the rigs used in the constant-strain tests is an advantage in the application of the corrosive solution. Thus, in the case of two-point bending, Figure 1(a), several specimens may be strained in the same rig, which can be constructed of plastic and immersed in a tank containing the test solution.

2.35.1.3 Constant-Load Tests

Dead-weight loading (with or without the assistance of levers to reduce the load requirements) of tensile specimens has the advantage of avoiding some of the difficulties already discussed, not the least in allowing accurate determination of the stress if the specimen is uniaxially loaded. The relatively massive machinery usually required for such tests upon specimens of appreciable cross section is sometimes circumvented by the use of a compression spring, Figure 1(g), chosen with characteristics that ensure that it does not change significantly in length during testing, thereby approximating to a constant-load application. For immersion tests, the frame may be coated in polymer and the specimen insulated from the shackles by plastic sleeves and washers to avoid bimetallic effects; alternatively, the specimen may be enclosed in a glass cell containing the test solution. The alternative approach of minimizing the size of the loading system by reducing the cross section of the specimen to the dimensions of a wire is dangerous unless failure by stress corrosion cracking is confirmed by, say, metallography. This is because failure may result from pitting and an attendant increase in the effective stress to the ultimate tensile strength (UTS) in some stress corrosion environments. Indeed, there is evidence for some systems that before stress corrosion cracking proper can begin, a pit must form wherein certain chemical or electrochemical conditions are established that permit cracks to be initiated, and in such systems, the use of fine wires has obvious pitfalls.

The load relaxation that accompanies some, if not all, constant-deflection tests is replaced in constant-load tests by an increasing stress condition, since the effective cross section of the test piece is reduced by crack propagation. This suggests that it will be less likely that cracks will cease to propagate once initiated, as may happen with constant deflection tests at initial stresses in the region of the threshold stress, and therefore, the threshold stresses are likely to be lower when determined under constant-load conditions than under conditions of constant deflection. Some results attributed to Brenner and Gruhl[8] for an aluminum alloy, Figure 4, confirm this expectation. These results also show shorter times to failure for the same initial stress with constant load testing and, as already indicated for constant-deflection tests, raise queries as to the significance of time to failure, the parameter so frequently used in assessing cracking susceptibility.

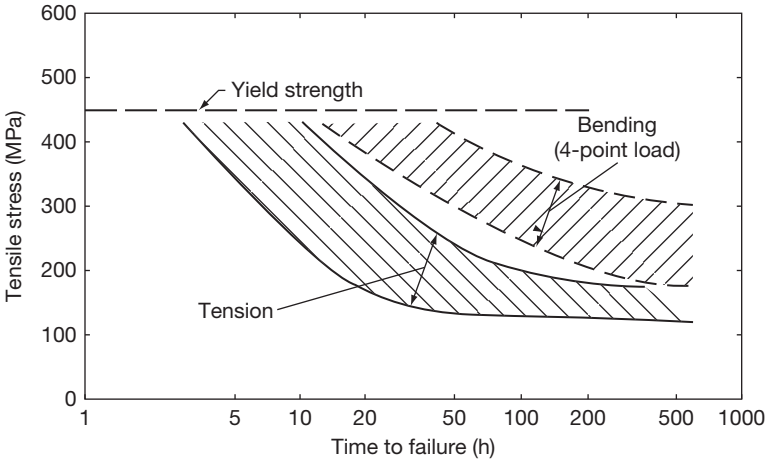


Figure 4 Comparison of test results from bend and tension tests upon Al–Zn–Mg alloy in 3% NaCl plus 0.1% H_2O_2. Reproduced from Brenner, P.; Gruhl, W. *Z. Metall.* **1961**, *52*, 599.

2.35.1.4 Slow Strain-Rate Tests

While this method of testing has been in use in some laboratories for two decades or more and has increased in use considerably in very recent years, there remain some skepticism and unfamiliarity with the method. In essence, it involves the application of a relatively slow strain or deflection rate ($\sim 10^{-6}\,s^{-1}$) to a specimen[9] subjected to appropriate electrochemical conditions. It should be emphasized that the strain rates employed are very much lower than those involved in straining electrode experiments where the object, the measurement of current transients, is totally different. In slow strain-rate corrosion tests, the object is to produce stress corrosion cracks that are metallographically indistinguishable from those produced in constant-load or constant-deflection experiments. The object in all these laboratory tests is normally to obtain data in a relatively shorter period of time, and this is frequently achieved by adopting an approach that increases the severity of the test. In stress corrosion testing, this usually takes the form of increasing the aggressiveness of the environment by changing its composition, temperature, or pressure, stimulating the corrosion reactions (galvanostatic or potentiostatic), increasing the susceptibility of the alloy through changes in structure, or increasing the severity of the stress by the introduction of a notch or precrack. The application of dynamic straining to a stress corrosion test specimen also comes into this last category, and, like all of the other accelerating approaches, its justification will vary according to the circumstances in which it is used.

Most stress corrosion crack velocities fall in the range from 10^{-3} to $10^{-6}\,mm\,s^{-1}$, which implies that failures in laboratory test specimens of usual dimensions occur in not more than a few days. This is found to be so if the system is one in which stress corrosion cracks are readily initiated, but it is common experience to find that some testpieces do not fail even after extended periods of testing, which are then terminated at some arbitrarily selected time. The consequences are that considerable scatter may be associated with replicate tests, and the arbitrary termination of the test leaves an element of doubt concerning what the outcome would have been if it had been allowed to continue for a longer time. Just as the use of precracked specimens assists in stress corrosion crack initiation, so does the application of slow dynamic strain, which has the further advantage that the test is not terminated after some arbitrary time, since the conclusion is always achieved by the specimen fracturing, and the criterion of cracking susceptibility is then related to the mode of fracture. Thus, in the form in which it is normally employed, the slow strain-rate method will result in failure in not more than $\sim$2 days, either by ductile fracture or by stress corrosion cracking, according to the susceptibility towards the latter, and metallographic or other parameters may then be assigned in assessing the cracking response. The fact that the test concludes in this positive manner in a relatively shorter period of time constitutes one of its main attractions.

Early use of the test was in providing data whereby the effects of such variables as alloy

composition and structure or inhibitive additions to cracking environments could be compared, and also for promoting stress corrosion cracking in combinations of alloy and environment that could not be caused to fail in the laboratory under conditions of constant load or constant strain. Thus, they constitute a relatively severe type of test in the sense that they frequently promote stress corrosion failure in the laboratory where other modes of stressing plain specimens do not promote cracking, and in this sense, they are in a category similar to tests on precracked specimens. In recent years, an understanding of the implications of dynamic strain testing has developed, and it now appears that this type of test may have much more relevance and significance than just that of an effective and rapid sorting test. It may, at first sight, be argued that laboratory tests involving the pulling of specimens to failure at a slow strain-rate show little relation to the reality of service failures. In point of fact, in constant-strain and constant-load tests, crack propagation also occurs under conditions of slow dynamic strain, to a greater or lesser extent depending upon the initial value of stress, the point in time during the test at which a stress corrosion crack is initiated, and various metallurgical parameters that govern creep in the specimen. Moreover, there is an increasing amount of evidence for some systems which suggests that the function of stress in stress corrosion cracking is to promote a strain rate which, rather than stress *per se*, is the parameter that really governs crack initiation or propagation. In these cases, the minimum creep rate for cracking is as much an engineering design parameter as is the threshold stress or stress–intensity factor obtained from constant-load tests on plain or precracked specimens.

The point may be illustrated by data for a ferritic steel exposed to a carbonate–bicarbonate solution as fatigue precracked cantilever beams subjected to constant loads. Deformation in the plastic zone associated with the precrack is time dependent following load applications and can be measured, and the threshold conditions for stress corrosion cracking defined, in terms of a limiting average creep rate over a specific time interval. That limiting creep rate may then be used in subsequent experiments to calculate the threshold stress from creep data determined independently, these calculated threshold stresses then being compared with values determined experimentally. The creep properties of ferritic steels may be varied by prior strain aging, following different amounts of cold work, and **Figure 5** shows the observed and calculated threshold stresses from tests on specimens subjected to various strain aging treatments. Clearly,

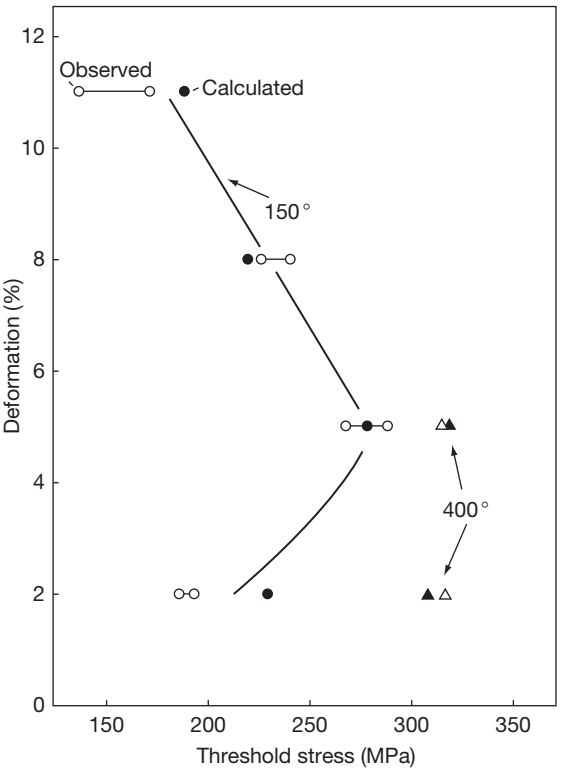


Figure 5 Observed and calculated values of the threshold net section stresses for stress corrosion cracking of a C–Mn steel after various prior deformations and aging treatments.

the general trend of the experimentally determined curve showing the effects of the amount of prior deformation is reflected in the calculated results.

The equipment required for slow strain-rate testing is simply a device that permits a selection of deflection rates while being powerful enough to cope with the loads generated. Plain or precracked specimens in tension may be used, but if the cross section of these needs to be large or the loads high for any reason, cantilever bend specimens with the beam deflected at appropriate rates may be used. It is important to appreciate that the same deflection rate does not produce the same response in all systems and that the rate has to be chosen in relation to the particular system studied.

The representation of the results from slow strain-rate tests may be through the usual ductility parameters, such as reduction in area, the maximum load achieved, the crack velocity, or even the time to failure, although as with all tests, metallographic or fractographic examination, while not readily quantifiable, should also be involved. Since stress corrosion failures are usually associated with relatively little

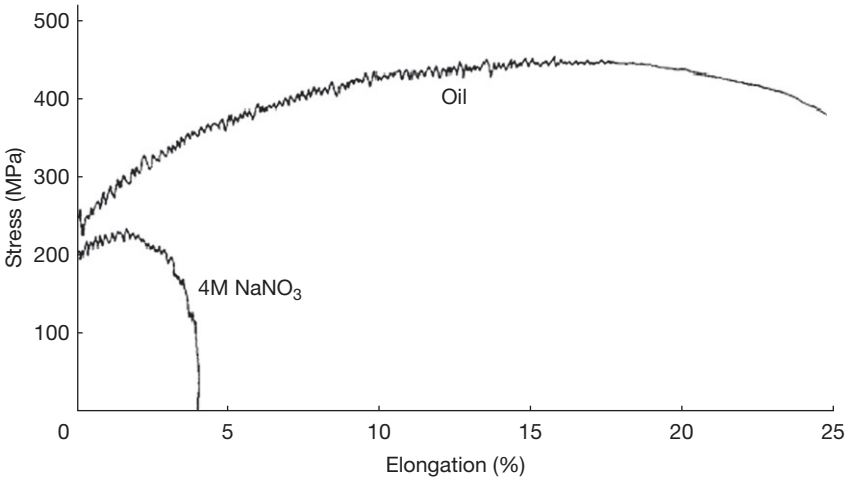


Figure 6 Nominal stress–extension curves for mild steel in oil giving ductile failure, and in 4 M $NaNO_3$ producing stress corrosion failure, at the same test temperature (104 °C).

plastic deformation, the ductility of specimens will be variable according to the extent to which stress corrosion contributes to the fracture process. This will also influence the shape of the load–extension curve that may be obtained by continuous monitoring of the response of a load cell incorporated in the system; Figure 6 shows the forms of curves obtained with and without attendant stress corrosion. It is apparent from these curves that not only is the extension to fracture dependent upon the presence or otherwise of stress corrosion cracks, but so also is the maximum load achieved. The latter may be used for expressing cracking susceptibility in some systems, as also may the area bounded by the load–extension curve. However, the variations in maximum load achieved in slow strain-rate tests in circumstances of varying cracking severity are not always large enough for significant distinctions to be made. Even measurements of ductility, such as reduction in area, are readily invariably not made, if only because the final fracture of the specimen does not always follow a simple path, and the fitting of the two broken pieces together is not easy. Probably, the easiest quantity to measure with reasonable accuracy is the time to failure, which has as much significance in a slow strain-rate test as it does in constant-load or constant-deflection tests. Indeed, the time to failure in slow strain-rate tests is simply related to ductility parameters, a not very surprising result when it is remembered that the less the intensity of stress corrosion cracking, the greater will be the ductility to fracture and therefore, the greater the time to failure for a given strain rate.

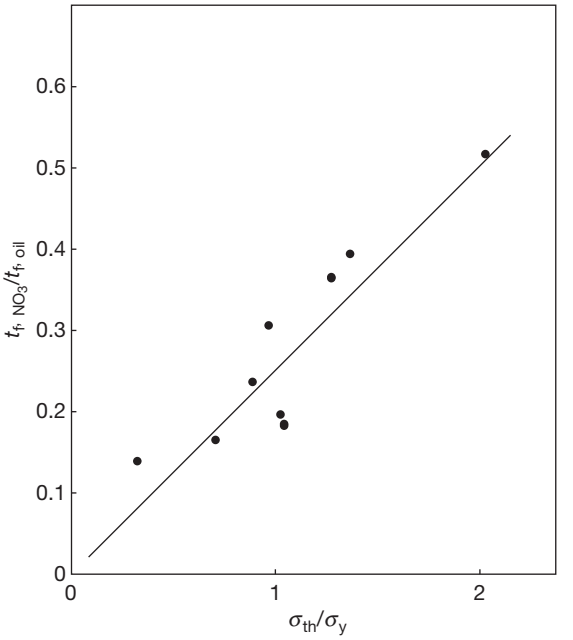


Figure 7 Time to failure ratios from constant-deflection rate tests and normalized threshold stresses σ_{th}/σ_y obtained from constant-strain tests for a series of low-alloy ferriic steels in boiling 4 M NH_4NO_3.

Clearly, for slow strain-rate testing to have credence, it should give results that are comparable with those obtained by other methods. Figure 7 shows some results for tests upon low-alloy ferritic steels in boiling 4 M NH_4NO_3, the various alloying elements producing a range of cracking susceptibilities as measured by the threshold stresses obtained from constant-strain tests. These results have been

normalized by dividing the threshold stress σ_{th} by the lower yield strength σ_y for each steel, while the slow strain test results have been normalized by dividing the time to failure in the 4 M NH_4NO_3 by the time to failure in oil at the same temperature, so that increasing departure from unity indicates increasing cracking susceptibility. The general trend of the results in **Figure 7** is clear in indicating reasonable agreement between the two types of tests in placing the steels in essentially the same order of merit.

Although slow strain-rate tests are most frequently taken to total failure in order to produce a 'go/no-go' type of result in which threshold stresses are not defined, they can be conducted in a manner that allows such definition. Specimens are preloaded to various initial stresses in the absence of the cracking environment or at a potential that prevents cracking, after which they are allowed to creep until the latter falls below the strain rate to be applied. The applied straining is continued for a sufficient time only to allow cracks to grow to a measurable size. During straining, the stress upon the specimen varies in a manner dependent upon the magnitude of the applied strain-rate, hence the importance of restricting the test time to no longer than that necessary to produce measurable cracks. The cracks are probably most conveniently measured by microscopy on longitudinal sections of the gauge lengths, the length of the deepest detectable crack divided by the test time giving an average crack velocity. **Figure 8** shows some results from tests upon a cast nickel–aluminum bronze exposed to seawater[10] and clearly it is possible to define a threshold stress below which cracking is not observed. However, that threshold stress depends upon the strain-rate applied, as is to be expected. Another approach to defining threshold stresses in slow strain-rate tests that may sometimes be useful is to use tapered specimens, with the taper angle minimized to avoid complications by resolved components of the tensile load.[11] Applied to the cracking of α-brass exposed to sodium nitrite solutions, a single tapered specimen gave threshold stresses close to those obtained by the use of a number of plain specimens loaded at a given strain rate to various stress levels.

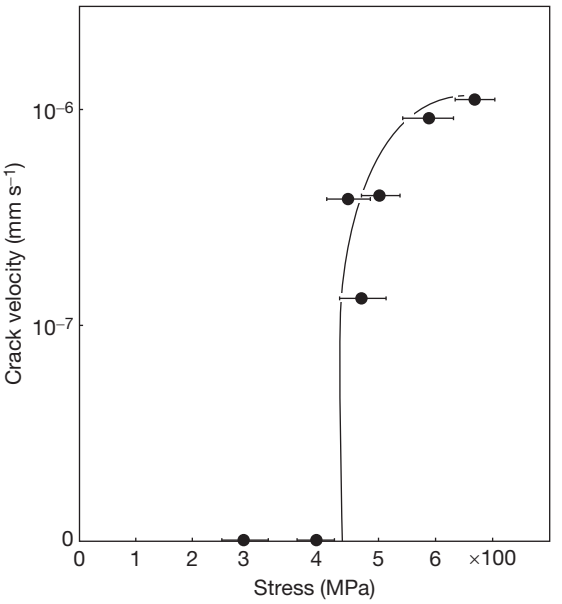


Figure 8 Average stress corrosion crack velocity from monotonic slow strain rate tests at $1.5 \times 10^{-7}\,s^{-1}$ conducted over various restricted ranges of stress on a cast Ni–Al bronze in seawater at 0.15 V(SCE). The stress range traversed in each test is shown by the length of the bar. Reproduced from Parkins, R. N.; Suzuki, Y. *Corros. Sci.* **1983**, *23*, 577.

2.35.2 Testpiece

2.35.2.1 Precracked Samples

The literature contains many references to the use of notched, as opposed to precracked or plain specimens in laboratory studies of stress corrosion, for reasons of improved reproducibility, inability to crack plain specimens under otherwise identical conditions or ease of measuring some parameters, such as crack growth rate, when the crack location is predetermined. However, the developments in fracture mechanics have resulted in a whole new field of stress corrosion testing involving the use of specimens containing a sharp precrack, usually produced from a notch by subjecting the specimen to fatigue. The application of fracture mechanics to stress corrosion cracking is the subject of an admirable review by Brown,[12] and various aspects of the method are considered in papers presented at an Advisory Group for Aerospace Research and Development (AGARD) conference.[13]

The problems associated with the choice of plain specimens for assessing stress corrosion resistance may, at first sight, appear equally large in relation to precracked specimens in the sense that in the relatively shorter time during which such tests have been in use, a large number of specimen types have been used (**Figure 9**). However, the differing specimen geometries are rationalized through the stress–intensity factor, with the result that data from different testpieces are comparable, provided appropriate precautions are taken in

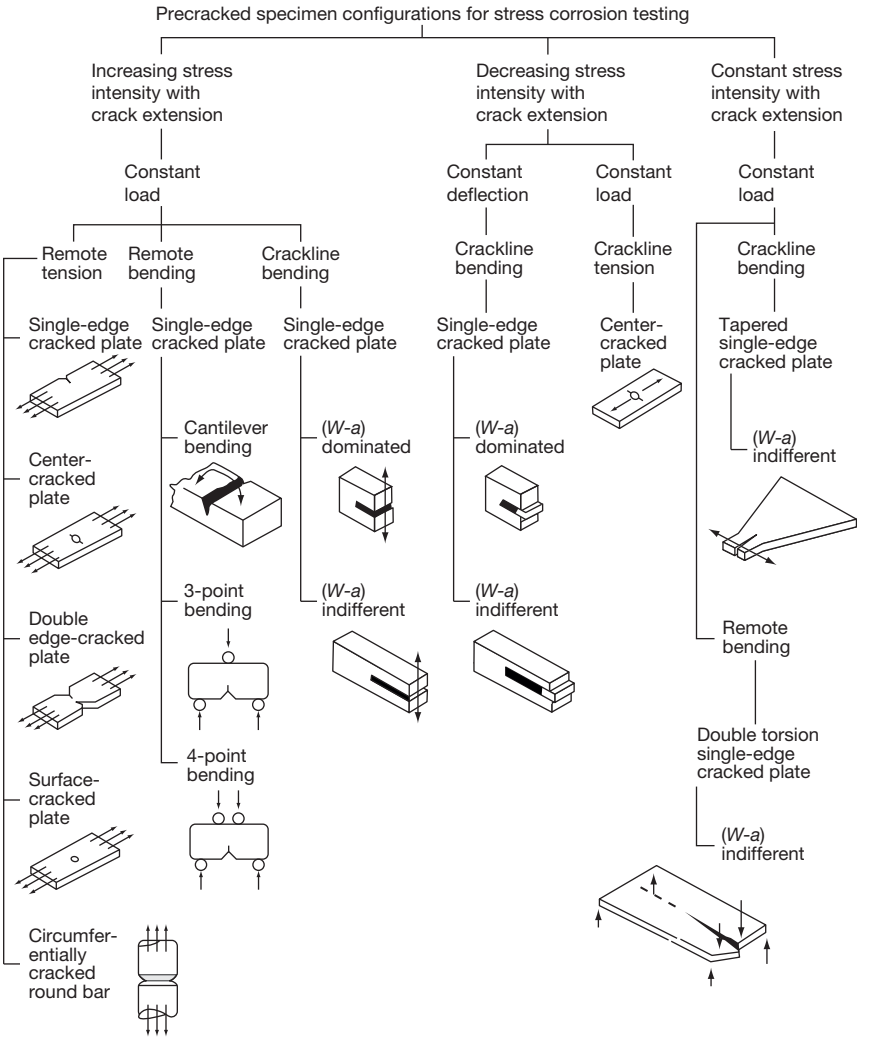


Figure 9 Classification of precracked specimens for stress corrosion testing. Reproduced from Smith, H. R.; Piper, D. E. Stress Corrosion Testing with Precracked Specimens, The Boeing Co, D6–24872, ARPA 878, June, 1970.

specimen preparation. The biggest single difficulty is in relation to the large size of specimen that is necessary for highly ductile materials if the concepts of linear elastic analysis are to be applicable. Since it is probable that most service stress corrosion failures occur in highly ductile materials in relatively thin sections, it is clear that there are problems here, but the use of precracked specimens that do not conform dimensionally to the requirements for linear elastic analysis to be strictly valid is still worthwhile in some instances, and in any case, developments in fracture mechanics over the last decade or so allow alternative approaches than that of linear elastic analysis.

Precracked specimens are sometimes useful for other reasons than the analysis that they afford in relation to stress–intensity factors. Such applications may be associated with the simulation of service situations, the relative ease with which stress-corrosion cracks can be initiated at precracks, or the advantages that sometimes accrue from the propagation of a single crack. The claim that has sometimes been made of precracked specimen tests – that they circumvent the initiation stage of cracking in plain specimens, erroneously assumed invariably to be related to the creation of a corrosion pit that provides a measure of stress concentration approaching that achieved at the outset with a precrack – is rarely entirely valid. Thus, the geometries of a pit, notch, or precrack are often as important for electrochemical reasons as they are for any reason associated with their influences upon stress distribution. This is because a geometrical discontinuity may be necessary

to provide the localized electrochemical conditions, in terms of environment composition or electrode potential, that are necessary for stress corrosion crack propagation. The objections that have sometimes been made against the use of precracked specimens, for example, to the validity of introducing a transgranular precrack into a specimen that suffers intergranular stress corrosion cracking, or of the necessity of incurring a considerable expense to produce a very sharp crack when the introduction of a corrosive environment may blunt the crack by the dissolution, miss the point that such sharp discontinuities do indeed exist in some real materials. Indeed, one of the major attractions of precracked specimen testing is that it can provide data that allow the designation of the maximum allowable defect sizes in structures for the latter to remain in a safe condition.

In view of the significance of strain rate in stress corrosion cracking, mentioned earlier, it is good to remember that its significance is as applicable to precracked specimens as it is to initially plain specimens, in relevant systems. This has a number of implications, not the least of which is the possible influence of time delay between loading precracked specimens and exposing them to the test environment. Moreover, the limiting stress–intensity factor K_{ISCC}, above which cracks grow relatively rapidly (**Figure 10**) may well depend upon the conditions under which it is determined, and it should not be regarded as some property of the material equivalent to, say, a yield stress. There is now a considerable volume of data that shows how relatively small fluctuating stresses may reduce the threshold stresses or stress-intensity factors for stress corrosion cracking, and some of these effects are probably related to cyclic loading sustaining creep-related effects. Crack-tip strain rates have consequently become a topic of interest, and expressions are available for cyclically loaded precracked specimens[14,15] and also for multicracked specimens[16] of the form that initially plain specimens take during slow strain-rate tests.

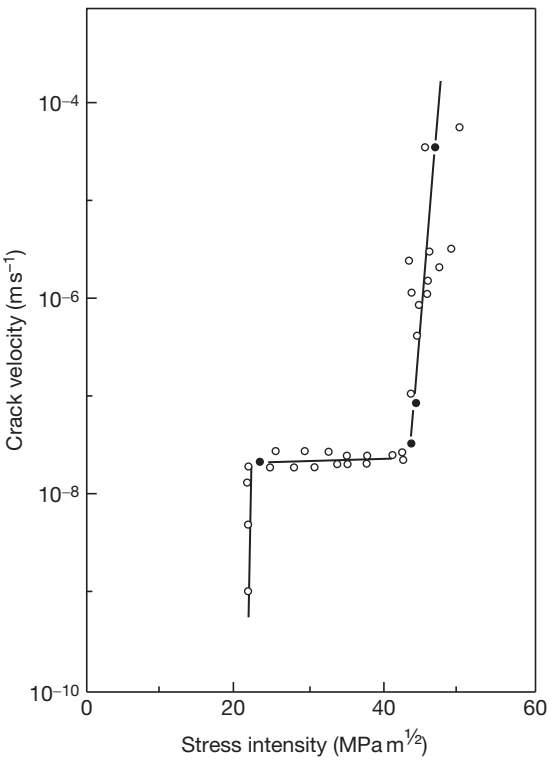


Figure 10 Effect of applied stress intensity upon crack velocity for high-strength (1800 MPa UTS) quenched and tempered steel (AFC 77) in distilled water. Reproduced form Spiedel, M. O.; In Conference on *Hydrogen in Metals*; National Association of Corrosion Engineers, Houston, TX, 1975.

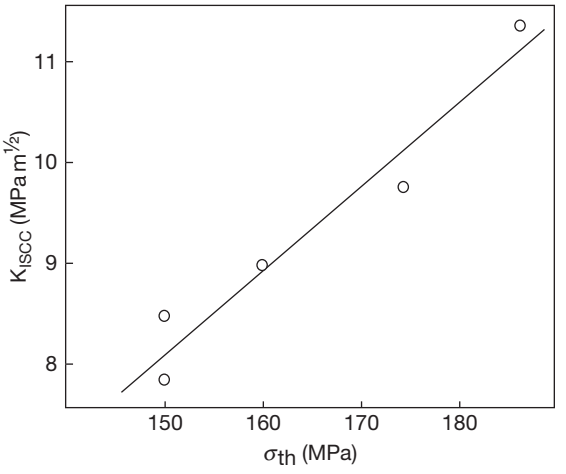


Figure 11 Threshold stress intensities K_{ISCC} from precracked specimen tests, and threshold stresses σ_{th} from plain specimen tests, for a Mg–7Al alloy in various structural conditions tested in chromate–chloride solution. Reproduced from Wearmouth, W. R.; Dean, G. P.; Parkins, R. N. *Corrosion* **1973**, *29*, 251.

2.35.2.2 Comparison of the Results from Plain and Precracked Specimens

It is clear that an initially plain specimen that develops a stress corrosion crack may, if the geometry is appropriate, conform to the conditions obtaining in an initially precracked specimen. This raises a question, despite the opposing views of the protagonists of the two types of testpieces, as to the comparability of the result from each. **Figure 11** shows the results[17]

obtained from stress corrosion tests upon a Mg–7Al alloy exposed to a chromate–chloride solution, the cracking susceptibility of the alloy being varied by different heat treatments. The implication of Figure 11 is that the threshold stress σ_{th}, determined upon initially plain specimens of small cross section, is related to the threshold stress intensity K_{ISCC} obtained from precracked specimens of relatively large section. Since K_{ISCC} represents the stress intensity below which an existing crack does not propagate, it would appear that the threshold stress given by tests on plain specimens corresponds to values below which cracks do not propagate to give total failure, that is, the threshold stress for plain specimens is not necessarily the stress below which cracks do not form. Examination of plain specimens stressed below the threshold stress revealed the presence of small stress corrosion cracks that had ceased to propagate and moreover, the maximum sizes of the cracks that did not propagate to total failure were quadratically related to the threshold stress, as would be expected if the concepts of fracture mechanics were applicable to these initially plain specimens. Nonpropagating cracks have also been observed at stresses below the threshold stress in other systems, such as low-alloy ferritic steels exposed to various environments, and would therefore, support the suggestion that what are being measured in tests upon plain and precracked specimens are not as different as has sometimes been suggested.

2.35.2.3 Crack Velocity Measurements

In mechanistic studies of stress corrosion and also in the collection of data for remaining-life predictions for plants there is need for stress corrosion crack velocity measurements to be made. In the simplest way, these can be made by microscopic measurements at the conclusion of tests, the assumption being that the velocity is constant throughout the period of exposure, or, if the crack is visible during the test, *in situ* measurements may be made by visual observation, the difficulty then being that it is assumed that the crack visible at a surface is representative of the behavior below the surface. Indirect measurements must frequently be resorted to, and these have involved observation of the elongation of the specimen, crack-opening displacement, changes in the electrical resistivity of the specimen, and acoustic emissions that sometimes accompany crack extension.

Measurement of the elongation of the specimen is probably the least satisfactory of all these, even though it is often the simplest, requiring a transducer that responds to dimensional change. Multiple cracking of initially plain specimens raises problems in interpreting the data in terms of crack velocities, and so the technique has been frequently used for attempting to determine the point in time when cracking was initiated, the preceding time during which the transducer shows no response being equated to an incubation period for cracking. However, such results can be completely misleading because the sensitivity of most transducers is such that they will only detect change when the specimen undergoes some plastic deformation, resulting from the propagation of a crack to the size where the remaining uncracked portion of the section beyond the crack is raised to its yield strength. Consequently, crack propagation can occur during the (so-called) incubation period when the stress is insufficient to cause the propagation of a deformation band. The latter frequently occurs suddenly, producing a sharp response by the transducer, which has sometimes been interpreted as evidence of a burst of fast mechanical fracture but which may, in fact, be nothing of the sort. It is much more satisfactory to use a crack-opening displacement gauge[18] located across a precrack. These gauges usually take the form of two thin cantilever beams to which strain gauges are attached, the beams being located at opposite sides of the extremity of the precrack. As crack extension occurs and the sides of the crack undergo relative displacement, the strain gauges respond to the unbending of the beams.

Changes in the electrical resistivity of a specimen containing a propagating crack[19] depend upon applying a high constant direct current at each end of the specimen and measuring the potential difference across electrical leads situated at the opposite sides of the crack. The potential field in the region of the crack is disturbed by the presence of the latter and as the crack extends, the potential difference between the leads on opposite sides increases, the total current remains constant. This requires a reliable constant current source, and the technique is dependent in some degree upon the exact positioning of the leads and gives less reproducible results if crack branching occurs. The initial thought that the application of DC to the specimen may influence the electrochemistry of the stress corrosion reactions is not sustained in practice, and the technique can provide reliable data.

High-frequency stress waves are generated when stress corrosion cracks propagate in some materials, especially the high-strength steels when these undergo

hydrogen-induced cracking. The detection of these acoustic signals, which are filtered from lower amplitude background noise, affords a means of studying crack propagation.[20] While the technique involves the use of sophisticated and relatively costly equipment if it is to be correctly practiced, it has been suggested that it may also offer a means of distinguishing between active paths and hydrogen-embrittlement mechanisms of cracking.[21] However, that is not universally accepted, and the data from acoustic signals need treating with caution.[22]

2.35.2.4 Effects of Surface Finish

It is hardly surprising that the preparation of surfaces of plain specimens for stress corrosion tests can sometimes exert a marked influence upon the results. Heat treatments carried out on specimens after their preparation is otherwise completed can produce barely perceptible changes in surface composition, for example, decarburization of steels or dezincification of brasses, that promote quite dramatic changes in stress corrosion resistance. Similarly, oxide films, especially if formed at high temperatures during heat treatment or working, may influence results, especially through their effects upon the corrosion potential.

However, quite apart from these chemical changes at surfaces occasioned by the method of specimen preparation, physical effects may be important. Paxton and Procter[23] have prepared a review of what little is known about the effects of machining and grinding upon stress corrosion susceptibility, the most obvious effects being related to surface topography and the introduction of residual stresses into the surface layers. The former is more likely to be important in the higher strength notch-sensitive materials, while surface compressive stresses are likely to have the general effect of delaying or preventing failure.

2.35.3 Choice of Environment

Although the list of environments reported as promoting stress corrosion cracking in any alloy continues to grow with time, the concept of solution specificity remains in that not all corrosive environments will initiate or sustain stress corrosion cracking in all alloys. While it is inevitable that the environment will always remain as one of the variables that may need to be assessed by stress corrosion tests, nevertheless certain solutions, by their widespread use over many years, have tended to become standard test solutions for certain types of alloys. Thus, boiling $MgCl_2$ solution for stainless steels, boiling nitrate solutions for carbon steels, and 3.5% NaCl for aluminum alloys, to mention but a few, have been extensively used, for example, in comparing the effects of metallurgical variables upon cracking propensities. Such approaches raise two questions: the first concerned with the extent to which 'standard' solutions prepared in different laboratories may be regarded as identical, and the second with the extent to which degrees of susceptibility of a range of alloys to cracking in one environment are related to cracking in a different environment.

While the relatively small differences that may be expected to occur between laboratories preparing a solution to the same specification often will not influence stress corrosion test results, there are situations where relatively small changes in environment can promote marked changes in cracking response. Thus, Streicher and Casale[24] have pointed to the potential problems associated with the use of nominally 42% boiling $MgCl_2$ in testing stainless steels. Since the hydrate of $MgCl_2$ is hygroscopic, solution preparation by weighing may lead to appreciable differences in boiling point and hence, times to failure in a stress corrosion test, so that it is preferable to prepare the solution by adding water to the hydrate to achieve a particular boiling point.

Similarly, pH variations resulting from either the initial preparation or from changes during a stress corrosion test may exert a marked influence upon results in some systems. Thus, the cracking of carbon steels in nitrates is markedly pH sensitive and, depending upon the volume of solution and the surface area of the specimen exposed, as well as upon the time involved in making the test, significant pH rises can occur, and cracking can cease as a result. Moreover, if tests are carried out with anodic stimulation, these effects may be aggravated, especially if the counter electrode is immersed in the test cell. In other cases, for example, the medium and higher strength steels, the initiation and maintenance of cracking frequently requires localized pH changes within the confines of the crack, and these can only occur if the initial conditions of exposure are appropriate.

The oxygen concentration of the solution, as in many instances of corrosion, can also be critical in stress corrosion cracking tests. Instances are available in the literature that show very markedly different test results according to the oxygen concentration in systems as widely different as austenitic steels immersed in chloride-containing phosphate-treated boiler water[25] and aluminum alloys[26] immersed in 3% NaCl.

The assumption that the relative cracking responses of a series of alloys will be the same irrespective of the environment to which they are exposed can be extremely dangerous. Many examples could be quoted of the dangers of drawing conclusions from tests in a given environment and applying these to a different situation, but some results by Lifka and Sprowls[27] will suffice. The results for the relative cracking susceptibilities of three aluminum alloys subjected to different exposure conditions are shown in Figure 12, which indicate that an intermittent spray test using acidified 5% NaCl solution gives the same order of susceptibility for the three alloys as was observed in outdoor exposure tests at three different locations. On the other hand, an alternate immersion test in 3.5% NaCl, widely used for testing aluminum alloys, places the alloys in a completely different order of susceptibility. This single example will suffice to indicate the necessity for simulating service conditions as closely as possible where laboratory data are to be used for selection or design in relation to industrial equipment.

While the dangers inherent in using standardized environments in relation to environment-sensitive fracture are readily indicated by many examples that can be quoted, there remains a problem in relation to alloy development where possible service environments may not always be identifiable at the time of the development program. In such circumstances, it appears desirable that an alloy should be assessed in a range of environments, but even then, it is necessary for realism to be injected into the program if an excessively large number of test environments are not to be involved. The potential dependence of cracking, with its implications for dissolution and filming reactions or the discharge of hydrogen, suggests that the solution pH is also likely to exert significant influence upon cracking. Plots of cracking domains on potential–pH diagrams sometimes indicate correlations with certain reactions, and this may be useful in guiding a testing program,[28] that is, involving solutions of different pH values and exploring the cracking propensity as a function of potential.

The importance of potential cannot be over emphasized, and some aspects of this part of stress corrosion testing may be conveniently discussed in the context of Figure 13. This shows the different potential ranges for the cracking of ferritic steel according to the environment in tests involving potentiostatic control. Also shown are the free corrosion potentials for that steel in the different solutions. These indicate that while failure would occur in the nitrate at the free corrosion potentials, this would not be so in the hydroxide or carbonate–bicarbonate solution. This does not mean that carbon steels will never fail by stress corrosion in these two environments at the free corrosion potential, since the latter is, of course, dependent upon the composition of the steel, its surface condition, the composition of the environment, and other factors. It is possible, therefore, that as the result of, say, small additions to the environments, added intentionally or present as impurities, the corrosion potential can be caused to be within the cracking range, or that as a result of small additions to a steel, the corrosion potential may fall outside the cracking range. Quite small changes in

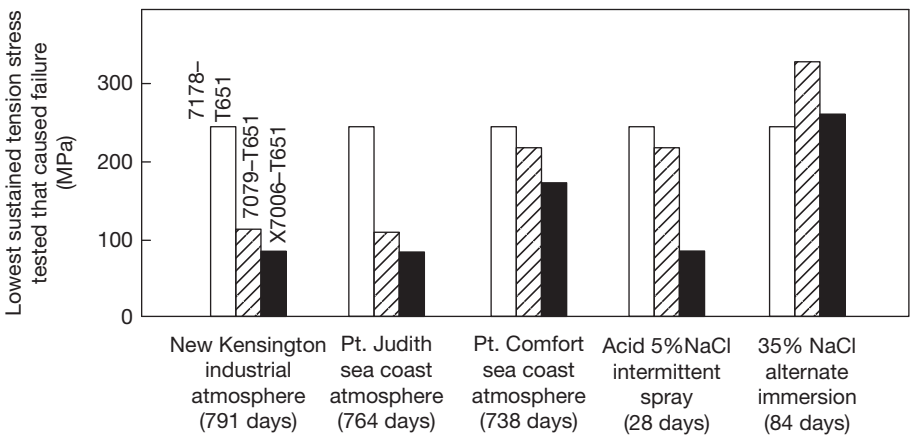


Figure 12 Relative resistance to stress corrosion cracking of three aluminum alloys subjected to different environments. The stress levels employed corresponded to 75, 50, and 25% of the respective transverse yield strengths. Reproduced from Lifka, B. W.; Sprowls, D. O. Stress Corrosion Testing, ASTM STP No. 425, 1966; p. 342.

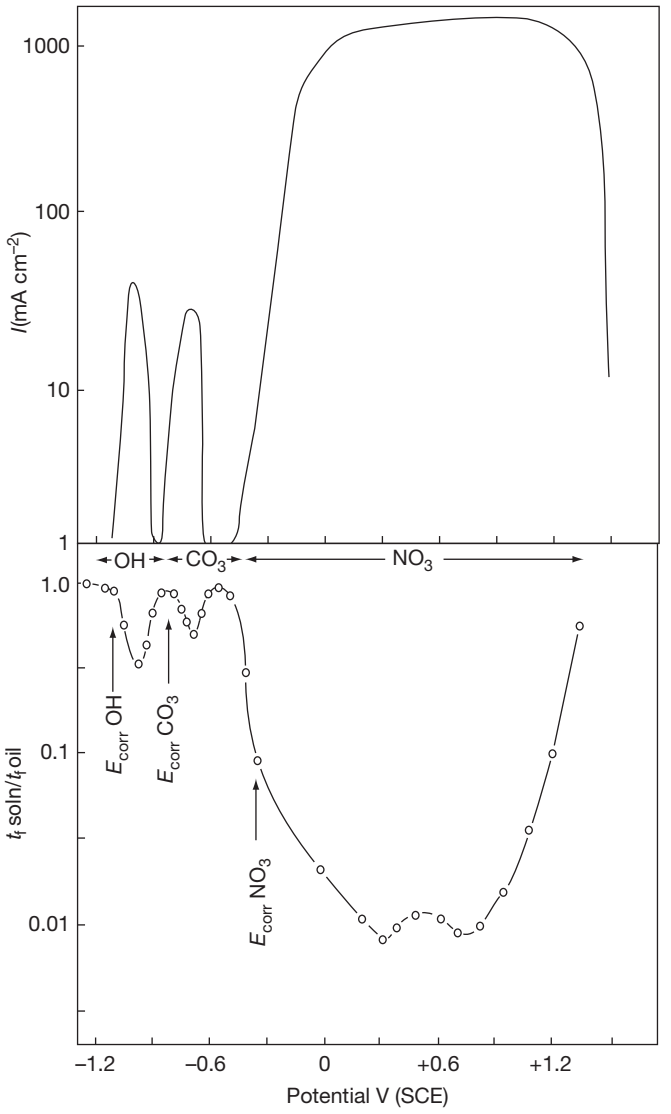


Figure 13 Current density differences between fast and slow sweep rate polarisation curves and stress corrosion cracking susceptibility as a function of potential for C–Mn steel in nitrate, hydroxide, and carbonate–bicarbonate solutions.

potential, often only a few tens of millivolts, can therefore produce dramatic changes in cracking response and point to the necessity, especially in laboratory tests attempting to simulate a service failure, of reproducing the potential with precision.

2.35.4 Stress Corrosion Test Methods

2.35.4.1 Types of Test Cells

The cells that contain the specimen and environment for stress corrosion tests often need to be more than a vessel made in some substance, usually glass, that is inert to the environment and which produces no electrical response upon the test specimen. Where cracking is initiated at surfaces through which heat transfer occurs, it may be necessary to design a cell in which such an effect is incorporated, since the concentration of substances in solution that may occur at an interface through which heat passes, may play a significant role in promoting cracking, especially if surface deposits allow concentration by evaporation while preventing mixing with the bulk or the environment. The cracking of riveted mild-steel boilers and the concentration of carbonate–bicarbonate

solutions under pipeline coatings to produce cracking in high-pressure gas transmissions lines are significant examples. Dana[29] has developed a method for simulating the conditions for cracking of stainless steels in contact with thermal insulating materials, while concentration in leaking boiler seams is simulated in the 'embrittlement detector' developed by Schroeder and Berk.[30]

Such test cells involve, among other things, a crevice, the essence of which is that the volume of solution that it contains is relatively small compared with the area of exposed metal, a ratio that may influence stress corrosion test results determined in more conventional cells where crevices do not exist. The experiments of Pugh *et al.*,[31] on the stress corrosion of 70–30 brass in ammoniacal solutions of various volumes are particularly instructive in indicating how this ratio may influence results, the time to failure varying by about an order of magnitude for a similar change in solution volume. Changes in the surface area of exposed specimens, apart from the effects already implied, may influence the cracking response for other reasons, as shown by the results of Farmery and Evans[32] for an Al–7Mg alloy immersed in a chloride solution. They found that coupling unstressed to stressed specimens of varying area ratio influences failure times, relatively shorter times being obtained when the area of unstressed to stressed specimen was large.

2.35.4.2 Initiation of Stress Corrosion Tests

It may be felt that the initiation of a stress corrosion test involves no more than bringing the environment into contact with the specimen in which a stress is generated, but the order in which these steps are carried out may influence the results obtained, as may certain other actions at the start of the test. Thus, in outdoor exposure tests, the time of the year at which the test is initiated can have a marked effect upon the time to failure,[26] as can the orientation of the specimen, that is, according to whether the tension surface in bend specimens is horizontal upwards or downwards or at some other angle. However, even in laboratory tests, the time at which the stress is applied in relation to the time at which the specimen is exposed to the environment may influence results. Figure 14 shows the effects of exposure for 3 h at the applied stress before the solution was introduced to the cell, upon the failure of a magnesium alloy immersed in a chromate–chloride solution. Clearly, such prior creep extends the lifetime of specimens and raises the threshold stress very considerably, and since other metals are known to be strain-rate sensitive in their cracking response, it is likely that the type of result apparent in Figure 14 is more widely applicable.

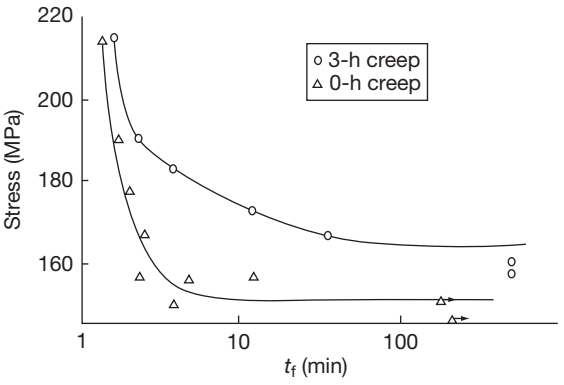


Figure 14 Effect of delay period between application of load and introduction of solution to test cell in the failure of a Mg–7Al alloy exposed to a chromate–chloride solution. Reproduced from Wearmouth, W. R.; Dean, G. P.; Parkins, R. N. *Corrosion* **1973**, *29*, 251.

2.35.4.3 Hydrogen Embrittlement Tests

The absorption of hydrogen by various materials, including high-strength steels, results in loss of ductility, which in turn can result in cracking and fracture when the metal is subjected to a sustained tensile stress. Hydrogen may be introduced into these various alloys from the gas phase (during manufacture or welding), or from aqueous solution during surface treatment (pickling, plating, and phosphating) or from the environment during a spontaneous corrosion process in which the development of acidity within the crack results in hydrogen evolution and absorption. Various test methods[33] may be used to evaluate the effect of hydrogen on the properties of alloys, including some *ad hoc* tests that were specifically developed for high strength steels.

Although similar constant-load test rigs are used for both active-path corrosion and hydrogen stress cracking, there is one fundamental difference in the test procedure. In the case of active-path corrosion testing, it is always carried out in the presence of the corrosive environment, but in the case of hydrogen-related cracking, testing may be carried out after hydrogen has been introduced into the alloy either deliberately by gaseous or cathodic charging, or following processes such as welding, pickling, or electroplating. However, with precharged specimens, loss

of hydrogen may occur[34] when they are removed from the environment, resulting in the entry of that substance and so, sustained-load tests are also carried out in the presence of an environment (gaseous or aqueous) so that hydrogen is introduced into the test piece during the application of the tensile stress.

2.35.4.4 Dynamic Tests

All of the properties evaluated by the conventional tensile test (yield strength, tensile strength, elongation, and reduction in area) are affected by the presence of hydrogen, but in the case of the tensile strength and yield strength, the effect is significant only when the steel has a very high tensile strength and has been severely embrittled. On the other hand, the reduction in area, and, to a lesser extent, the elongation may be used for detecting embrittlement. Hobson and Sykes[35] found that with low-carbon steels there was an almost linear relationship between reduction in area and hydrogen content of the steel. Slow strain-rate tests are sometimes employed in testing materials (and not only steels) after preexposure to a source of hydrogen. The strain rate may be critical in that not only can it be too high but also, if it is too low the hydrogen may diffuse out of the specimen before cracking occurs.

Various types of bend tests have been used to evaluate embrittlement. Beck *et al.*[36] used thin strip specimens and determined the decrease in height, Δh, at fracture when the specimen was bent by compressing it at a constant rate in a tensile testing machine (Figure 15). The decrease in height, Δh, gives a measure of the embrittlement, the maximum elongation of the outside fiber of the specimen being calculated from the radius of curvature at maximum bending. In general, the ductility is found to increase with the rate of straining, and for this reason, high-strain-rate tests, such as impact tests, are insensitive to hydrogen embrittlement. Where the material is available only in the form of tubing, semicircular specimens may be used in place of flat strips in the compression bend test. The total cross-head travel from the unstressed height along the diameter to the point of fracture gives a measure of embrittlement, which may be compared with that obtained from an unembrittled specimen of the same steel.

A constant-rate bend-test machine, which provides an effective method for testing highly embrittled steel wires of high-tensile strength, was designed by Zapffe and Haslem[37] (Figure 16). The motor *A* pulls a chord attached to the traveling arm *D* that rotates about a

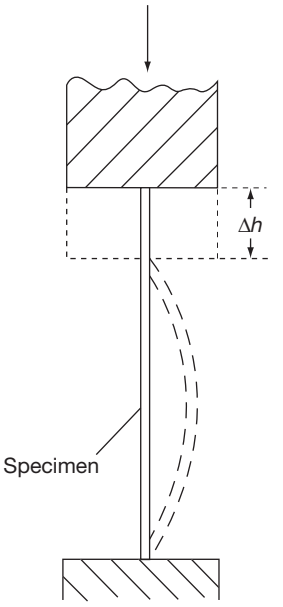


Figure 15 Bend test using a tensile testing machine. Reproduced from Beck, W.; Klier, E. P.; Sachs, G. *Trans. AIME* **1956**, *206*, 1263.

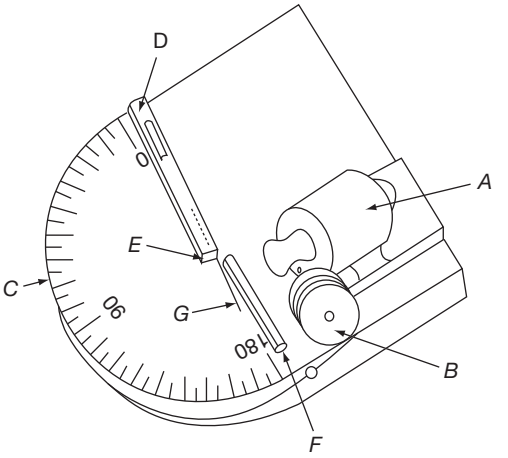


Figure 16 Constant-rate bend-test for determining hydrogen embrittlement of wires. *A*, drive unit; *B*, pulley; *C*, semicircular base; *D*, traveling arm; *E*, axial pin; *F*, fixed arm; *G*, wire specimen. Reproduced from Zapffe, C. A.; Haslem, M. E. *Trans. AIME* **1946**, *167*, 281.

pivot pin. The wire specimen *G* (1.6 × 100 mm) is inserted in *D* and is supported by the fixed arm *F*, the arrangement being so designed that tensile or torsional stresses are avoided. The specimen is thus bent around the pivot pan *E* (radius 1.6 mm) at a constant rate, the angle of bend to cause fracture giving a measure of its ductility. Since ductility increases with rate of straining, the bending rate must be slow and (4%) is

considered to be suitable for detecting embrittlement. A similar machine has been used for studying the embrittlement of spring steel strip after hydrogen has been introduced by cadmium plating.

An alternative procedure is the reverse-bend test in which the specimen in the form of wire or strip is bent repeatedly backward and forward over a mandrel until it fractures, the number of bends indicating the degree of embrittlement. However, this method is considered to be less sensitive than the single-bend test.

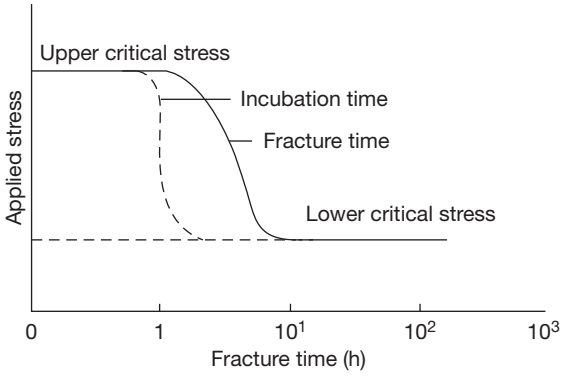


Figure 17 Schematic representation of delayed failure characteristics of a hydrogenated high-strength steel.

2.35.4.5 Static Tests

Whereas ductile materials, such as iron and mild steel, are often considered not to crack when charged with hydrogen and subjected to a tensile stress below the yield stress, the position is different with high-strength ferrous alloys where, depending on the strength of the steel and the hydrogen content, failure may occur well below the yield stress. However, the fracture process is not instantaneous, and there is a time delay before cracks are initiated; for this reason, the phenomenon is sometimes referred to as 'delayed failure.'

In the majority of cases, the tests are conducted using a dead-weight lever-arm stress-rupture rig with an electric timer to determine the moment of fracture, but a variety of test rigs similar to those shown in Figure 1(g) are also used. The evaluation of embrittlement may be based on a delayed-failure diagram in which the applied nominal stress versus time to

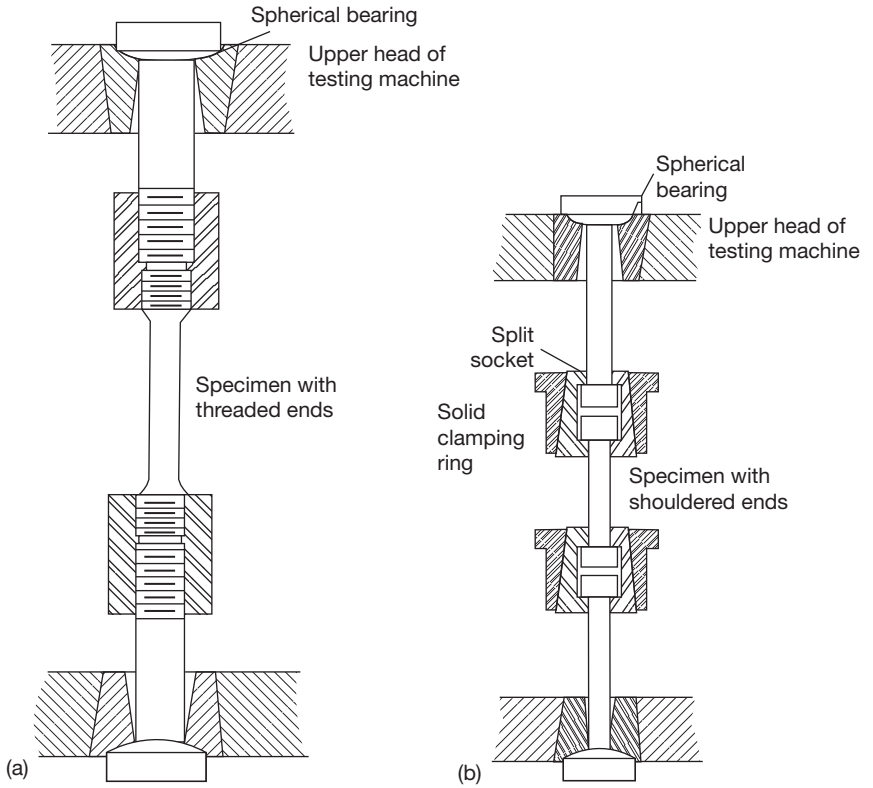


Figure 18 Methods of gripping specimens in order to avoid bending stresses: (a) Device for threaded end specimens, and (b) device for shouldered-end specimens. Reproduced from Stress Corrosion Testing, ASTM STP No. 425, 1967.

failure is plotted, Figure 17, or the specimen may be stressed to a predetermined value (say 75% of the ultimate notched tensile strength) and is considered not to be embrittled if it shows no evidence of cracking within a predetermined time (say 500 h). Troiano[38] considers that the nature of delayed fracture failure can be described by four parameters (see Figure 17):

1. the upper critical stress corresponding to the fracture stress of the unembrittled notched specimen;
2. the lower critical stress, which is the applied stress below which failure does not occur;
3. the incubation period or the time required for the formation of the first crack;
4. the failure time or the time for specimen failure at a given applied stress; in the intermediate stress range, this includes a period of relatively slow crack growth.

During the constant-load test, it is essential that only axial tensile stresses are applied, as any bending stresses that are introduced will result in a higher true stress than that calculated. For this reason, the ends of the specimens and the grips must be designed to avoid bending stresses. The ASTM Standard E8 (Standard Test Methods for Tension Testing of Metallic Materials) specifies that in the case of specimens with threaded ends the grips should be attached to the heads of the testing machine through properly lubricated spherical-seated bearings and that the distance between the bearings should be as great as is feasible (Figure 18).

In order to simplify the test procedure, a number of investigators have designed test rigs in which the bulky lever arm is replaced by a loading nut, the stress in the specimen being determined by means of strain gauges; these rigs are similar in principle to those shown in Figure 1(g). Figure 19 shows a spring-loaded rig that was used by Cavett and van Ness[39] to study the embrittlement produced by hydrogen gas at high pressures, in which the tensile load is applied by compressing a heavy-duty spring.

Raring and Rinebold[40] have devised a method in which the specimen is supported along the diameter of a steel loading ring, Figure 20, and the stress is applied by tightening the bottom nut until the diameter corresponds with the required load. The sudden release of elastic energy stored within the ring when the specimen fractures results in displacement of the tightening nut, and this is used to actuate a microswitch and timer. Williams *et al.*[41] have used notched 'C'-rings, the stress being applied by tightening the nut of a calibrated loading bolt that passes through the diameter of the 'C' ring, Figure 21. The strain gauges attached to the bolt form two arms of a Wheatstone bridge circuit and to compensate for temperature

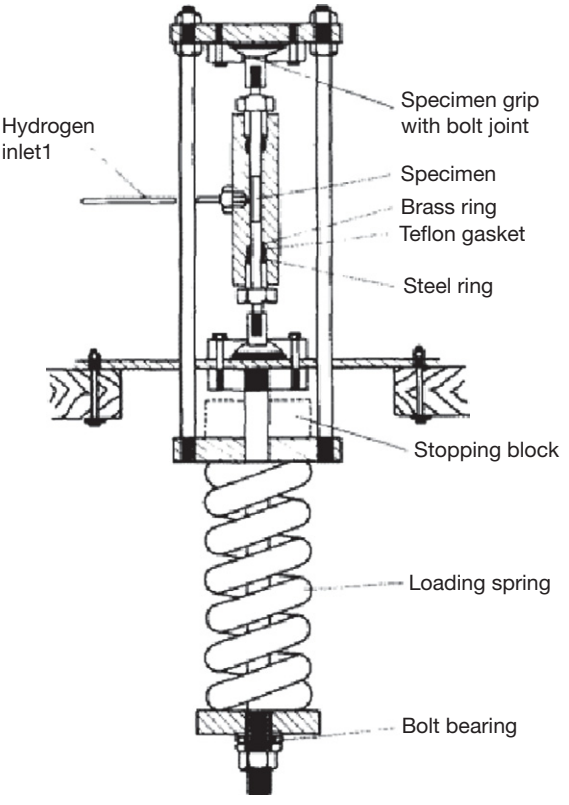


Figure 19 Spring-loaded rig for sustained load testing of a steel specimen in gaseous hydrogen at high pressure. Reproduced from Cavett, R. H.; van Ness, H. C. *Welding J.* (research supplement) **1967**, *42*, 317.

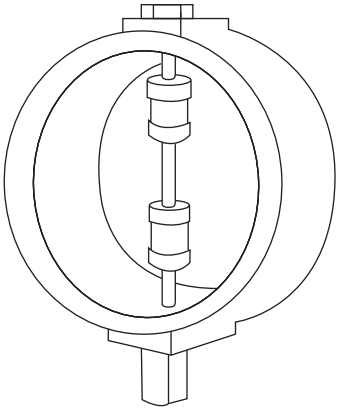


Figure 20 Loading-ring method of stressing a specimen. Reproduced from Raring, R. H.; Rinebold, J. A. ASTM Bulletin No. 213, 1956.

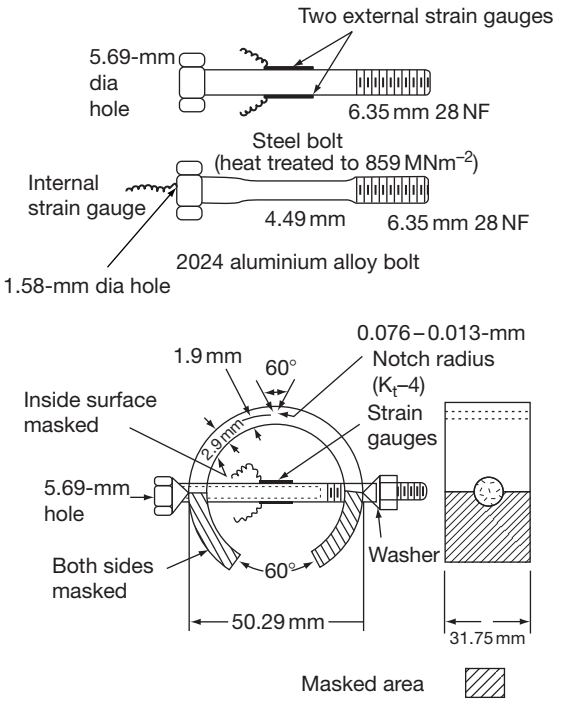


Figure 21 Notched 'C'-ring specimen with attached strain gauges.

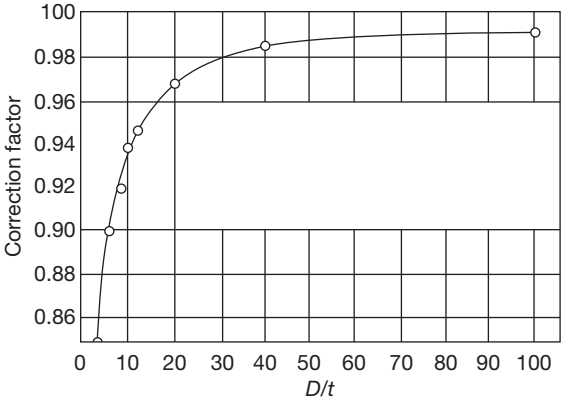


Figure 22 Correction factor Z for calculating stress in 'C'-ring tests.

changes, the other two arms consist of two identical strain gauges attached to a similar unstrained bolt.

Appendix A: Stresses in Bent Specimens

In each of the following equations: σ = maximum tensile stress, E = modulus of elasticity, and t = specimen thickness.

A.1 Two-point bending, Figure 1(a):

$$L = \frac{k.t.E}{\sigma}\sin^{-1}\left(\frac{H.\sigma}{k.t.E}\right)$$

where L = specimen length, k = constant (1.280), and H = holder span.

A.2 Four-point bending, Figure 1(b):

$$\sigma = \frac{12E.t.y}{3L^2 - 4A^2}$$

where y = maximum deflection, L = distance between outer supports, and A = distance between inner and outer supports.

A.3 'U'-bends, Figures 1(d) and 1(e):

Applied strain $\varepsilon = t/2r$, when $t < r$, where r = radius of curvature at section of interest.

A.4 'C'-rings, Figure 1(c):

$$\sigma = \frac{4E.t.Z.\Delta}{\pi D^2}$$

where $\Delta = OD_f - OD_i$, OD_f = final outside diameter of stressed 'C'-ring, OD_i = initial outside diameter of unstressed 'C'-ring, D = mean diameter, that is, $(OD - t)$, and Z = a correction factor, related to D/t as indicated in Figure 22.

References

1. Ailor, W. H. Ed.; *Handbook on Corrosion Testing and Evaluation*; Wiley: New York, 1971.
2. *Stress Corrosion Testing*, ASTM STP No. 425, 1967.
3. BS 6980, Stress corrosion testing, Parts 1–7, (ISO 7539–1/7) BSI, Milton Keynes.
4. Logan, H. L. *The Stress Corrosion of Metals*; Wiley: New York, 1966; p 273.
5. Haaijer, G.; Loginow, A. W. *Corrosion* **1965**, *21*, 105.
6. Parkins, R. N.; Mazza, F.; Royuela, J. J.; Scully, J. C. *Br. Corr. J.* **1972**, *7*, 154.
7. Parkins, R. N.; Haney, E. G. *Trans. Metall. Soc., AIME* **1968**, *242*, 1943.
8. Brenner, P.; Gruhl, W. *Z. Metall.* **1961**, *52*, 599.
9. Ugianski, G. M.; Payer, J. H. Eds., *Stress Corrosion Cracking – The Slow Strain Rate Technique, ASTM STP 665*; ASTM: Philadelphia, 1975.
10. Parkins, R. N.; Suzuki, Y. *Corros. Sci.* **1983**, *23*, 577.
11. Yu, J.; Holroyd, N. J. H.; Parkins, R. N. In *Environment-Sensitive Fracture: Evaluation and Comparison of Test Methods, ASTM STP 821*; Dean, S. W., Pugh, E. N., Ugianski, G. M., Eds.; ASTM: Philadelphia, 1984; p 288.
12. Brown, B. F. *Met. Rev.* **1968**, *13*, 17.
13. Specialists Meeting on Stress Corrosion Testing Methods, AGARD Conference Proceedings, No. 98 NATO, 1972.

14. Lidbury, D. P. G. *Embrittlement by the Localized Crack Environment*; Gangloff, R. P., Ed.; AIME: New York, 1983; p 149.
15. Parkins, R. N. *Corrosion* **1987**, *43*, 130.
16. Congleton, J.; Shoji, T.; Parkins, R. N. *Corros. Sci.* **1985**, *25*, 633.
17. Wearmouth, W. R.; Dean, G. P.; Parkins, R. N. *Corrosion* **1973**, *29*, 251.
18. Fisher, D. M.; Bubsey, R. T.; Srawley, J. E. Design and Use of a Displacement Gauge for Crack Extension Measurements, NASA IN-D 3724, 1966.
19. Barnett, W. J.; Troiano, A. R. *Trans. AIME* **1957**, *209*, 486.
20. Gerberich, W. W.; Hartblower, C. E. Proceedings of the Conference on Fundamental Aspects of Stress Corrosion Cracking NACE, 1986; p 420.
21. Okada, H.; Yukawa, K.; Tamura, H. *Corrosion* **1974**, *30*, 253.
22. Pollock, W. J.; Hardie, D.; Holroyd, N. J. H. *Br. Corros. J.* **1982**, *17*, 103.
23. Paxton, H. W.; Proctor, R. P. M. Paper No. EM68–520 presented at American Society of Tool and Manufacturing Engineers Symposium on Surface Integrity in Machining and Grinding, Pittsburgh 1968.
24. Streicher, M. A.; Casale, I. B. Proceedings of Conference on Fundamental Aspects of Stress Corrosion Cracking NACE: Houston, 1969; p 305.
25. Williams, W. I. *Corrosion* **1957**, *13*, 539.
26. Romans, H. B. *Stress Corrosion Testing, ASTM STP No. 425*, 1966, p 182.
27. Lifka, B. W.; Sprowls, D. O. *Stress Corrosion Testing, ASTM STP No. 425* 1966, p 342.
28. Parkins, R. N. In *The Use of Synthetic Environments for Corrosion Testing, ASTM STP 970*; Francis, P. E., Lee, T. S., Eds.; ASTM: Philadelphia, 1988; p 132.
29. Dana, A. W. *ASTM Bulletin No. 225, TP 196* 1957, p 46.
30. Schroeder, W. C.; Berk, A. A. *Bull. US Bur. Mines* 1941, p 443.
31. Pugh, E. N.; Montague, W. G.; Westwood, A. R. *Trans. Am. Soc. Met.* **1965**, *58*, 665.
32. Farmery, H. K.; Evans, U. R. *J. Inst. Met.* **1956**, *84*, 413.
33. Smialowski, M. *Hydrogen in Steel*; Pergamon Press: London, 1962.
34. Hardie, D.; Holroyd, N. J. H.; Parkins, R. N. *Met. Sci.* **1979**, *13*, 603.
35. Hobson, J. D.; Sykes, C. *J. Iron Steel Inst.* **1951**, *169*, 209.
36. Beck, W.; Klier, E. P.; Sachs, G. *Trans. AIME* **1956**, *206*, 1263.
37. Zapffe, C. A.; Haslem, M. E. *Trans. AIME* **1946**, *167*, 281.
38. Troiano, A. R. *Trans. Am. Soc. Met.* **1960**, *52*, 54.
39. Cavett, R. H.; van Ness, H. C. *Welding J.* (research supplement) **1967**, *42*, 317.
40. Raring, R. H.; Rinebold, J. A. ASTM Bulletin No. 213 1956.
41. Williams, F. S.; Beck, W.; Jankowasky, E. J. *Proc. ASTM* **1960**, *60*, 1192.

2.36 Applications of Statistical Analysis Techniques in Corrosion Experimentation, Testing, Inspection and Monitoring

N. Laycock
Shell Todd Oil Services, Private Bag 2035, New Plymouth, New Zealand
P. Laycock
School of Mathematics, Alan Turning Building, University of Manchester, Oxford Road, Manchester, M13 9PL, UK
P. Scarf
Centre for OR and Applied Statistics, Salford Business School, University of Salford, The Crescent, Manchester, M5 4WT, UK
D. Krouse
Industrial Research Limited, Gracefield Research Centre, PO Box 31-310, Lower Hutt, New Zealand

Glossary

Exceedance The amount by which a data value exceeds a selected threshold.

Extrapolation The process of predicting the value of some quantity of interest over a region (interval of space and/or time) that is outside the region of observation or inspection.

Extreme value The largest (else smallest) number in a data set.

Likelihood The joint probability distribution for the sampled data, considered as a function of the parameters of this distribution.

Operational failure rate The reciprocal of the mean time between operational failures of the function of a system that is subject to preventive maintenance such as inspection and repair.

Outlier A discordant or suspect data value.

Poisson process A sequence of events occurring purely at random in time or over an area at a constant mean rate.

Probability The relative 'long-run' frequency of occurrence of an outcome from many repetitions of a test or experiment, or the degree of one's current belief that a statement is true. Measured on a scale from 0 to 1, where 0 indicates that the outcome of interest cannot occur and 1 indicates that it must occur.

Abbreviations

CDF Cumulative distribution function
EV Extreme value distribution
F dn Fisher's variance-ratio distribution
GEV Generalized extreme value distribution
GPD Generalized pareto distribution
IRIS Internal rotary inspection system
LOTIS Laser optical tube inspection system
MAT Minimum allowable thickness
ML Maximum likelihood
MTBOF Mean time between operational failures
NDT Nondestructive testing (also known as nondestructive examination, NDE)
NWT Nominal wall thickness
PDF Probability density function
RWT Remaining wall thickness
S-IDAP Shell inspection design analysis and plotting package
***t* dn** Student's *t* distribution
TML Thickness measurement location
WT Wall thickness

Symbols

A Extrapolation area
a, b Intercept and slope for data regression line
C_f Cost of failure due to a crack
C_p Cost of inspection for cracks including the repair of crack defects
$C(T)$ Long-run cost per unit time of policy that inspects every T time units
$F_{\max,T}(z)$ Distribution function of the maximum crack depth at inspection at time T
$F_T(z)$ Unconditional distribution of the crack depth at inspection at time T
$F(y)$ Probability model cumulative frequency distribution (cdf, Pr(Y $<$ y))
$f(x, \theta)$, $F(x, \theta)$ Density and distribution function with parameter θ
$F(x)$, $f(x)$, $f(x|\cdot)$ Distribution function, density, and conditional density
$F_n(y)$ Sample cdf for n data values
$G_T(z; T - x)$ The distribution of crack depth at inspection at time T conditional on initiation of a crack at time x
$G(Z; \tau)$ Distribution of crack depth for a crack that has been growing for time τ
N_t Number of cracks at time t
$P_{WT}(\,)$ Density function of initial wall thickness
$p(\,)$ Posterior or prior density function
Pr(A) Probability that event A occurs
Pr($A|B$) Probability that A occurs given that B occurs

Pr(*A* & *B*), Pr(*A*, *B*) Probability that both A and B occur
R Half-width of uniform error range
S(t) Survivor function (Pr(T > t) = 1 − F(t))
S^2 Sample variance of data
T Optimum inspection interval
u Threshold wall loss
$\bar{x}, \bar{y}$ Means or averages for sample data
$X_{(p)}$ The $_p$th quantile of the maximum wall loss distribution
y(i) *i*'th order statistic, the *i*'th sample value in ordered data *y(1)* < *y(2)* <. . ..
z_f Critical crack depth: if $Z >_{z_f}$, then failure occurs
$\alpha\alpha$ Intercept for model regression line or similar *else* scale parameter for Extreme Value distribution
$\beta\beta$ Slope for model regression line or similar
δ, α Location and scale parameters of a GEV or GPD
Δ, $\hat{\Delta}$ Change and estimated change in extreme value mode between inspections
ε Measurement error effect on the mode of a Gumbel distribution
$\gamma\gamma$ Euler constant, 0.57721. . .
γ, α_0, β Parameters of Weibull distribution for $G_T(z;T-x)$
κ Shape parameter of a GEV or GPD
$\lambda\lambda$ Mean event rate for a Poisson process (e.g. pit initiation or crack initiation)
μ Mean of a probability distribution
$\hat{\mu}, \hat{\sigma}$ Maximum likelihood estimates of distribution parameters
$\varphi(\,)$ Density of wall loss measurement error
Θ, θ, θ_1, θ_2 Possible values of the Bernoulli parameter
$\rho\,\rho$ Coefficient of Variation, σ/μ
σ Standard deviation (square-root of variance) *else* scale parameter for a distribution
$\sigma^2, \sigma_c^2, \sigma_e^2$ Variance and variance components of a Normal distribution, $N(\mu, \sigma^2)$
ξ Bound for extreme value distribution

2.36.1 Introduction

Statistical techniques are used extensively across the full range of scientific and engineering disciplines. More often than not, the results of experimental studies are reported in terms of averages, best-fit lines, confidence limits, and distribution functions. The special relevance of statistics to corrosion was identified as early as 1933,[1] and the application of extreme value analysis techniques to some corrosion engineering problems was introduced in the 1950s.[2,3] Statistical approaches were then more widely adopted and developed in Japan, both in the study of localized corrosion processes[4,5] and for remnant life prediction in an engineering context.[6] The 1980s and 1990s saw increasing academic activity,[7–9] and recently there have appeared a small number of guidelines and standards for the practical applications of these techniques.[10–12]

Other chapters in this book describe the state of the art in corrosion modeling, but here we briefly mention a few examples where in statistical techniques have proved particularly useful. In modeling pitting corrosion of stainless steels, it is common to assume that pit initiation is a *Poisson process* occurring randomly in time and at sites distributed randomly across the sample surface,[13–15] although interactions between pits are now being considered more significant.[16,17] Similarly, *percolation theory*[18] has been used to explain the critical Cr concentration for the passivation of stainless steels in dilute acids,[19,20] the degree of grain boundary sensitization required for intergranular SCC of 304 stainless steel in thiosulfate solutions,[21] the mechanism of pit initiation for stainless steels in chloride solutions,[22] and the critical compositional thresholds associated with dealloying processes.[23,24] Of course, *Monte-Carlo (MC) techniques*[25] are frequently applied to enable the computer simulation of these and other corrosion processes; for example, **Figure 1** shows the results of an MC simulation of the passivation of Fe–Cr alloys.

Statistical approaches are frequently applied to the analysis of the results from the experimental studies of pitting corrosion, and the so-called 'electrochemical noise' monitoring is fundamentally dependent on statistical theory. Both these application areas are covered in some detail later. However, the most economically important use of statistics in corrosion is undoubtedly in the analysis of inspection results to estimate the in-service corrosion rate of industrial plants. These results are used to determine the planned replacement and future inspection strategies for plants all over the world, with enormous safety, financial, and environmental consequences potentially arising from poor decisions. In this context, it is natural to consider extreme corrosion events so that extreme value analysis has found wide application. In this chapter, we provide a basic introduction to the underlying mathematics behind the most commonly used techniques in the statistical

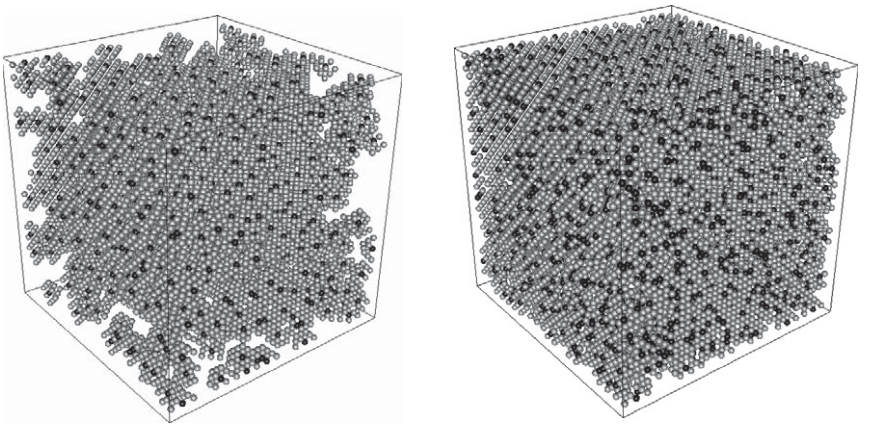


Figure 1 Surface exposed Fe (grey) and Cr (black) atoms after 100 000 time steps in a dissolution simulation involving 17 261 atoms in a bcc lattice: (a) 3 at.% Cr; (b) 18 at.% Cr. This simulation is based on the method described in Qian *et al.*,[20] and it shows passivation when the Cr concentration is above the relevant percolation threshold, $P_c(1,2) = 17.2$ at.%.

evaluation of corrosion data, but we focus, in particular, on those methods that have either been customized to a significant extent for specific applications in corrosion, or seem most likely to be further developed by the corrosion community in the future.

2.36.2 Basic Terminology and Methods

This is an introduction to the relevant statistical terminology and methods discussed in this chapter, as can be found in a fuller form in an introductory text book or at online resources.[26,27] Dedicated statistical software packages, such as SPSS, MINITAB, SAS, or GENSTAT, are also widely available and are usually preferable for implementing statistical techniques. Extensions of these ideas and nomenclature to allow for other probability distributions and contexts are introduced as and when they appear in this chapter.

2.36.2.1 The Normal Distribution and Associated Distributions

Statistical methods are designed to explore the interface between data and models. The simplest standard model is the Normal (or Gaussian) distribution, as illustrated in Figure 2 with an example dataset comprising 75 measurements of crack depth on circumferential welds on piping at various locations inside a nuclear power plant facility.

Note that although the data are clearly not 'a perfect fit' to the model, they can nevertheless be shown to be compatible with these model assumptions.

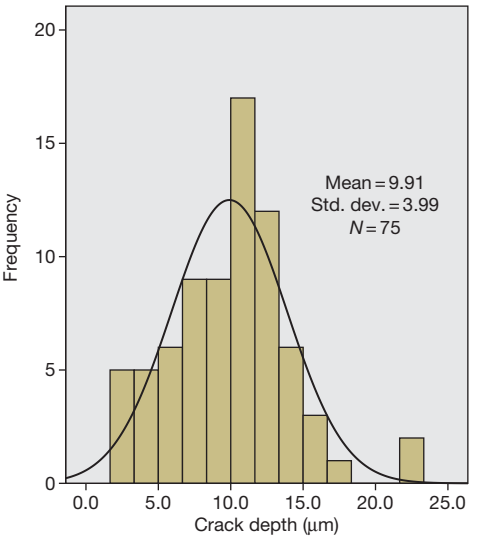


Figure 2 An example of a Normal (or Gaussian) distribution. The dataset comprises 75 measurements of crack depth on circumferential welds on piping inside a nuclear power plant.

We write $Y \sim N(\mu, \sigma^2)$ to be read as 'the random variable Y has a Normal distribution with mean μ and variance σ^2.' However, there are two possible 'outliers' in the right-hand 'tail' of the distribution at a depth of 22 μm, which might warrant a validity check and an assessment of any special circumstances at that measurement location. The model allows us to make statements such as 'There is a 95% probability that the next observation on Y will lie in the interval $\mu \pm 1.96\sigma$.' Note that σ is called the standard deviation of Y and is measured in the same units. Replacing

1.96 (~2) with 1.64 produces a 90% interval size, while 3.29 (~3) produces 99%, and 4.89 (~5) leaves just a one in a million chance of landing outside of the interval. The smooth curve drawn in Figure 2 is called the probability density function (PDF), and for the Normal distribution, this has formula

$$f(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\left(\frac{y-\mu}{\sigma}\right)^2\right] \qquad [1]$$

Probabilities for this or any other random variable are obtained by integrating the relevant PDF over the chosen interval. In particular, the cumulative distribution function (CDF) is defined by

$$\Pr(Y \le y) = F(y) = \int_{-\infty}^{y} f(y)\mathrm{d}y \qquad [2]$$

This probability may or may not have a simple formula. For the Normal distribution, access is required to a table of values for the standard distribution with mean zero and standard deviation one. When the measured variate is time, especially a failure time, then $S(t) = \Pr(T > t) = 1 - F(t)$ is commonly called the Survivor function.

A random sample of observations $y_1, y_2, \ldots, y_n$ on any such random variable Y has sample mean (or average) $\bar{y} = (y_1 + y_2 + \cdots + y_n)/n$, which can be used to estimate the population mean, μ, and we write $\hat{\mu} = \bar{y}$. The distinction between capital letters (Y, standing for the name of the random variable) and lower-case letters (y, for the collected data) is frequently dropped, whereas the distinction between Latin letters (such as y and s for data) and Greek letters (such as μ and σ for model parameters) is always vital and should only be discarded when the context and intent is clear – although for simplicity of presentation, this will usually be assumed to be the case in most of this chapter. By the Central Limit Theorem, almost all sample means, $\bar{y}$, have approximately (although exactly so if Y is itself Normal) a Normal distribution, with mean μ and variance σ^2/n – and hence standard deviation (or 'standard error') equal to $\sigma/\sqrt{n}$, which tends to zero as n increases.

The data will also have their own sample variance, given by

$$s^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (y_i - \bar{y})^2 \qquad [3]$$

which provides an 'unbiased' estimate of the population variance σ^2, where we write $\tilde{\sigma}^2 = s^2$, which then provides an estimate $s/\sqrt{n}$ of the 'standard error of the mean'. The divisor '$n - 1$' is called the 'degrees of freedom' for the numerator and will change as more parameters (currently only one, namely μ) are introduced into the model of the population mean. When Y is Normal, a 95% (say) 'confidence interval' for the true mean, μ, is $\hat{\mu} \pm ts/\sqrt{n}$, where t comes from tables for the t-distribution, with $t = 1.96$ for a large n when this can also be used for other sampled distributions.

Proposed values of μ (such as zero, for example), which lie outside of this interval, are said to be rejected at a 5% significance level. Conversely, the confidence interval could alternatively be derived from the formalities of the significance test.

The mean μ represents the deterministic part of the model and can be expanded – as can the noise (or random) component σ^2 – to incorporate covariate data $x_1, x_2, \ldots, x_n$ collected along with y on a p-vector $\underline{x} = (x_1, x_2, \ldots x_p)^{\mathrm{T}}$. Such information was available for the data on Y, as summarized in Figure 2, namely 'location,' $x_1 = 1$ or 2, and $x_2 =$ 'working hours'; and this was initially absorbed into a more complex model of Y, incorporating parallel linear relationships between Y and x_2. The effect of location (reactor number) was found to be statistically nonsignificant, leaving just the simple regression line through the data displayed in Figure 3. The fitted model is

$$y_i = 4.7609 + 0.1875x_{2i} + e_i, \text{ where}$$
$$e_i \sim \mathrm{N}(0, 0.993^2), i = 1, \ldots, n \qquad [4]$$

The histogram of the fitted residuals $\{e_i\}$ is shown in Figure 4, where it can be seen that the visual fit for a Normal distribution is superior to that for the original y data seen in Figure 2 – as is often the case with the residuals from more complex models, thereby making the assumption of Normality of the random component of the model more plausible. However, the two outliers are still noticeable and may warrant special attention. A typical definition of an outlier is any data point for which the modulus of the associated residual term in the fitted model exceeds 3σ ; refer to the book on outliers by Barnett and Lewis.[28]

Note that when modeling the variance in a corrosion context, it can be important to distinguish between system or corrosion noise, σ_c^2, and measurement error, σ_e^2. These may add to give $\sigma^2 = \sigma_c^2 + \sigma_e^2$, $\sigma^2 = t\sigma_c^2 + \sigma_e^2$, or $\sigma^2 = t^2\sigma_c^2 + \sigma_e^2$, depending on the context or assumptions – where t is the exposure time.

It is often useful to reorder the data by size, with notation $y_{(1)} \le y_{(2)} \le \cdots \le y_{(n)}$, implying that $y_{(1)}$ is the data minimum, $y_{(n)}$ is the data maximum, and more generally, $y_{(i)}$ is called 'the ith'-order statistic. Plotting $y_{(i)}$ against i/n gives the sample cumulative distribution function, F_n, where

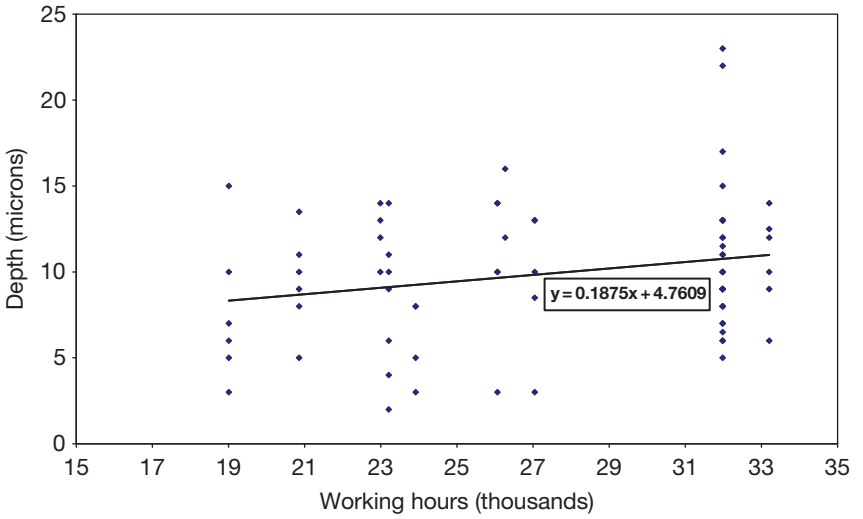


Figure 3 For the data shown in **Figure 2**, the effect of location on crack depth was found to be statistically nonsignificant, leaving just the simple regression line shown here for the effect of 'working hours.'

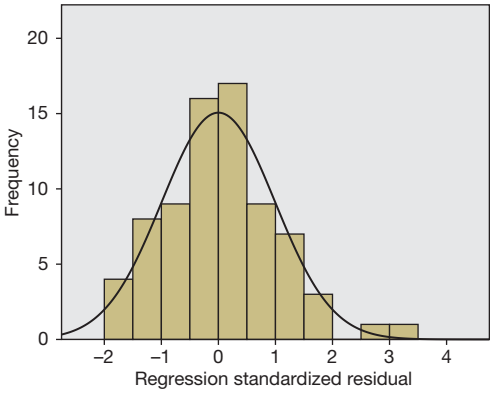


Figure 4 The histogram for the fitted residuals e_i from the linear regression of **Figure 3**. The visual fit for a Normal distribution is superior to that for the original *crack depth* data seen in **Figure 2**. However, the two outliers are still noticeable and may warrant special attention.

$$F_n(y) = i/n, y_{(i)} < y \le y_{(i+1)} i = 0, 1, 2, \ldots, n,$$
$$\text{with } y_{(0)} = -\infty, y_{(n+1)} = \infty$$

Alternatively, since $E[F(y_{(i)})] = i/(n+1), i = 1, 2, \ldots, n$, plotting $y_{(i)}$ against $i/(n+1)$ and connecting points directly by straight lines or a smooth curve will provide an approximately unbiased estimate $\hat{F}$ of the population cumulative distribution function, F, with $\hat{S} = 1 - \hat{F}$ for the Survivor function. Also, plotting $y_{(i)}$ against $F^{-1}(i/(n+1))$, when F is known (perhaps only up to location and scale), should give an approximately straight line if the assumed distributional form is correct. The intercept and slope of this line can provide estimates of the unknown location and scale parameters. This technique can be implemented algorithmically or *via* the appropriate probability paper.

Discrete probability distributions can be used for counting pits, failures, or other corrosion-induced events. A standard model is the Binomial distribution

$$p_r = \Pr(X = r) = {}^nC_r p^r q^{n-r},$$
$$q = 1 - p, r = 0, 1, \ldots, n \qquad [5]$$

for the number of 'successes' X in n independent trials with a constant probability of success, p. This distribution has mean $\mu = np$ and variance $\sigma^2 = npq$, and for a large n, this can be approximated by a Normal distribution with the same mean and variance by replacing 'r' with the interval $r \pm \frac{1}{2}$ – except for 'rare events' with $\mu < 5$ when a Poisson distribution can be used. For event counts in random samples of n objects from a finite population of N objects, such as a boiler with m 'failed tubes' out of N, we can use the hypergeometric distribution

$$p_r = {}^mC_r^{N-m}C_{n-r}/{}^NC_n, \max\{0, m+n-N\}$$
$$\le r \le \min\{m, n\} \qquad [6]$$

This distribution has mean $\mu = mn/N$ and variance $\sigma^2 = npq[(N-n)/(N-1)]$, where $p = (m/N)$, $q = 1 - p$, and this can be approximated by a binomial – and hence by Normal or Poisson distributions as appropriate – for a large enough N.

2.36.2.2 Linear Regression

The statistical methods in 'linear regression' cover a very wide range of models though the default assumption is a simple 'straight line regression,' which has the form $y = \alpha + \beta x$. The textbook by Freund and Wilson[29] provides an introduction to this topic, with a more extensive treatment in Neter *et al.*[30] The standard *straight line regression* model can be written as

$$y_i = \alpha + \beta x_i + e_i, \text{where } e_i \sim N(0, \sigma^2), \quad i = 1, \ldots, n \qquad [7]$$

The least-squares estimates a and b of the intercept α and slope β and the unbiased estimate of the residual variance σ^2 are given by

$$b = \sum (x_i - \bar{x})(y_i - \bar{y}) / \sum (x_i - \bar{x})^2;$$
$$a = \bar{y} - b\bar{x}; \tilde{\sigma}^2 = \sum (y_i - \hat{y}_i)^2 / (n - 2) \qquad [8]$$

where $\hat{y}_i = a + bx_i$ is the predicted value of y_i from the fitted model. Confidence limits about the line and prediction limits for future observations can be constructed. Formulae with further explanatory discussion are supplied in an example of straight line regression analysis of weight loss data, as is given later in Section 2.36.3.2.1. Variants on this model are parallel regressions $y_i = \alpha_j + \beta x_i$, $j = 1, 2$ or broken-line regressions

$$y = \alpha_1 + \beta_1 x \text{ for } x < x^* \text{ then } y = \alpha_2 + \beta_2 x \text{ for } x \geq x^* = (\alpha_2 - \alpha_1)/(\beta_1 - \beta_2) \qquad [9]$$

When duplicate observations are available for at least one fixed value of x and there are more than two distinct x values, then this will provide an independent estimate of the residual variance, enabling 'analysis of variance' techniques to be used to test the 'goodness-of-fit' of a straight line to the data. This method can be extended to more complex models.[30]

The *polynomial regression* model is given by

$$y_i = \alpha + \beta_1 x_i + \cdots + \beta_k x_i^k + e_i \ (i = 1, \ldots, n) \qquad [10]$$

In this model, x is called a *genuine* variable, while its powers, $x^2, \ldots, x^k$ are called *carrier* variables. This model can be used when searching for a local optimum in a dataset, and the most important special case is *quadratic regression*, for which $k = 2$.

Finally, we have *general linear regression*, for which

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} \cdots + \beta_k x_{ki} + e_i \ (i = 1, \ldots, n) \qquad [11]$$

where some of the x_j' s may be *genuine* physically distinct variables (such as temperature or pressure) and some may be *carriers* (such as powers or interaction terms such as temperature.pressure) or *indicators* of qualitative variates (such as location or manufacturer). When all the x_j' s ($j = 1, \ldots, k$) are genuine, this model is called *multiple linear regression.*

These models are best fitted using a standard statistical package, which will supply the fitted parameters along with their standard errors, confidence intervals, and tests for goodness of fit. Note that in a model-fitting context, a nonlinear model is one where the unknown parameters (conventionally denoted by Greek letters) occur in a nonlinear fashion, as, for example, in the models $y = \alpha e^{\beta x}$ or $y = \alpha x/(\beta + x)$.

2.36.2.3 Poisson Process Modeling

If 'events,' such as pits or cracks, occur 'purely at random' along a line, or in time, or over a surface – but at a constant mean rate, say λ – then they are said to follow a (homogeneous) 'Poisson process.' The count, say X, of events in any fixed interval/area of size t then follows a Poisson distribution:

$$\Pr(X = r) = \mu^r \exp(-\mu)/r!, \quad r = 0, 1, \ldots; \mu = \lambda t \qquad [12]$$

This distribution has mean and variance equal to μ and can be approximated by a Normal distribution for a large μ. The gaps, t, between such events along a single axis will follow an exponential distribution with PDF

$$f(t) = \lambda e^{-\lambda t}, t > 0 \qquad [13]$$

This distribution has mean and standard deviation equal to $1/\lambda$. If the mean event rate λ varies, then this is called a nonhomogeneous Poisson process, in which case, if the variation is known, the (time) axis can be transformed to produce a homogeneous process (i.e., by squeezing and stretching the axis appropriately). This type of approach was used in Laycock *et al.*[31] to model the initiation of crevice corrosion of stainless steels *via* metastable pitting events in open-circuit conditions.

A standard all-purpose technique for estimating the unknown parameters in these and more extensive models is the method of maximum likelihood (ML), which directly enables the production of standard errors, confidence intervals, and prediction intervals for estimated parameters and predictions about future values for Y. An application of all these techniques in a corrosion context can be found elsewhere.[32]

2.36.2.4 Extreme Value Distributions

Extreme value (EV) distributions are typically applied in a corrosion context to maximum pit depths measured on a set of test coupons or from sampled areas on a large metal structure. They may have been recorded at one moment in time or at several distinct times, and interest may center on extrapolation in time and/or over larger areas. Relevant environmental parameters may have been measured along with the measured maxima. The models can be applied to single maxima or extended to include the largest two or largest three *et cetera* on each coupon or sampled area. An introductory text for the application of these methods can be seen elsewhere.[33]

Gnedenko[34] fully characterized the limit (implying a large amount of unrecorded data, such as smaller pits) distribution of the maximum (or minimum) as being one of three types: I, II, or III. One textbook[6] suggests that the maximum pit-depth is often described by the Type I distribution, although Type III with a finite upper bound has been found elsewhere.[35] Gnedenko's limit theorems on the maximum are analogous to the celebrated 'Central Limit theorem' on the mean. These distributions reproduce their form under maximization of further data in the same way that the Normal distribution reproduces itself under summation. The Type I distribution is more commonly known as the Gumbel or, simply, the EV distribution. The Type II is known as a Fréchet distribution, and the Type III for maxima can be described as a reversed three parameter Weibull distribution. The distribution function of the EV with location parameter δ and scale parameter α – EV(δ, α) – is

$$F(y) = \exp[-\exp[-(y-\delta)/\alpha]], \quad -\infty < y < \infty, \alpha > 0 \qquad [14]$$

The mode is δ, and the mean is $\mu = \delta + \gamma\alpha$, where γ is the Euler constant, 0.57721..., while the standard deviation is $\sigma = \alpha\pi/\sqrt{6}$.

The generalized extreme value distribution (GEV) subsumes the Type I, II, and III distributions into a single three parameter distribution – GEV (δ,α,κ) – with distribution function (for maxima)

$$F(y) = \exp\{-[1-\kappa(y-\delta)/\alpha]^{1/\kappa}\}, \quad \alpha > 0, \kappa y \le \kappa\xi = \alpha + \kappa\delta \qquad [15]$$

and shape parameter $\kappa < 0$ for Type II with lower bound ξ, $\kappa > 0$ for Type III with upper bound ξ, and when the κ parameter reaches zero, the GEV becomes the unbounded EV. The mapping $y \to -y$ recovers the form for minima. The mean is

$$\mu = \delta + \frac{\alpha}{\kappa}[1-\Gamma(1+\kappa)] = \xi - \frac{\alpha}{\kappa}\Gamma(1+\kappa), \text{provided } \kappa > -1, \qquad [16]$$

and the standard deviation is

$$\sigma = \frac{\alpha}{\kappa}[\Gamma(1+2\kappa) - \Gamma(1+\kappa)^2]^{1/2}, \quad \text{provided } \kappa > -\frac{1}{2} \qquad [17]$$

Note that some references use ξ or $\gamma = -\kappa$ for the shape parameter. Most, but not all, continuous distributions have a limiting form that is a member of the GEV family. In particular, both tails of the Normal distribution (i.e., maximum or minimum values) have a limiting Type I distribution, as have maxima from a Weibull distribution, while minima from a Weibull distribution have exactly a Type III distribution.

The joint generalized extreme value (JGEV) multivariate density function for the r largest order statistics $y^*_{(1)} > y^*_{(2)} > \ldots > y^*_{(r)}$ (e.g., from a single coupon at one moment in time) is

$$f\left(y^*_{(1)}, y^*_{(2)}, \ldots, y^*_{(r)}\right) = \alpha^{-r} \exp\left\{-(1-\kappa(y^*_{(r)}-\delta)/\alpha)^{1/\kappa} - \left(\frac{1}{\kappa}-1\right)\sum_{j=1}^{r}\log\left(1-\kappa(y^*_{(j)}-\delta)/\alpha\right)\right\} \qquad [18]$$

with bounds as for the GEV and a limiting form when $\kappa = 0$, but the joint cumulative distribution function does not have a simple closed form expression.

These distributions can be fitted to data using ML techniques. For the EV distribution, suitable starting values for iterative ML can be obtained by plotting the order statistics $y_{(1)} \le y_{(2)} \le \cdots \le y_{(n)}$ against $F^{-1}(i/(n+1)) = -\ln(-\ln(i/(n+1)))$, $i = 1, 2, \ldots, n$, which can be done using EV plotting paper. The intercept and slope of the line then provide estimates of δ and α, respectively, while deviations downwards from the line at the upper end may indicate a Type III distribution with $\kappa > 0$ and hence an absolute upper bound. By extending the EV fitted line, extrapolated values for percentiles such as

$$\hat{y}_N = \hat{F}^{-1}(N/(N+1)) = \hat{\delta} - \hat{\alpha}\ln(-\ln(N/(N+1))) \approx \hat{\delta} + \hat{\alpha}\ln(N) \qquad [19]$$

can be read directly off the graph. In hydrology and meteorology, where data are typically annual maxima, '$N = T$ years' is described as 'the return period,' and setting $T = 100$ years leads to the so-called 'hundred year storm' prediction. In corrosion, where data are more typically from sampled areas of size a, say, $\hat{y}_N$ becomes an extrapolation over the extrapolation area $A = Na$. The Nth return level $\hat{y}_N$ is that value which would be exceeded, on average, once every N time or area units. Alternatively, since y_N is the maximum of a set of maxima, it will also have a distribution from the GEV family. This will be GEV(δ_N, α_N, κ) where

$$\delta_N = \delta + \alpha(1 - N^{-\kappa})/\kappa, \alpha_N = \alpha N^{-\kappa}$$
$$\text{and bound } \xi = \delta + \alpha/\kappa, \kappa \neq 0, \text{unchanged} \quad [20]$$

The mean of this distribution is a best estimator for y_N in the usual mean-squared error sense

$$\mu_N = \delta_N + \frac{\alpha_N}{\kappa}[1 - \Gamma(1+\kappa)], \text{ provided } \kappa > -1$$
$$= \delta + \alpha(\gamma + \ln(N)) \text{ when } \kappa = 0$$
$$\to \delta + \alpha \ln(N), \text{ as seen before}$$
$$\text{for the EV 'return period' method} \quad [21]$$

When covariates, such as time, are measured along with each maximum, a regression model for the mean – and possibly the variance or scale parameter – is required. Given a covariate vector $\underline{x} = (x_1, x_2, \ldots, x_p)^{\mathrm{T}}$ at each maximum, a multiple regression model might set $\delta_i = \underline{\beta}^{\mathrm{T}}\underline{x}_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki}, i = 1, 2, \ldots, n$ implying modeled mean

$$E[y_i|\underline{x}_i] = \beta_0 + \underline{\beta}^{\mathrm{T}}\underline{x}_i \text{ with intercept}$$
$$\beta_0 = \frac{\alpha}{\kappa}[1 - \Gamma(1+\kappa)] \quad [22]$$

and constant residual variance

$$\mathrm{Var}[y_i|\underline{x}] = \sigma_e^2$$
$$= (\alpha/\kappa)^2\left[\Gamma(1+2\kappa) - \Gamma(1+\kappa)^2\right] \quad [23]$$

After fitting such a model using ordinary least-squares, the fitted intercept and residual variance could be equated to the above formulae and solved for α and κ, enabling an estimate of the bound ξ to be obtained as a function of the covariates. The alternative of a constant coefficient of variance model can be achieved by additionally setting $\alpha_i = \rho\underline{\beta}^{\mathrm{T}}\underline{x}_i$, where $\rho = \sigma/\mu$, so $100\rho\%$ is the percentage variability in the data, independent of mean size. When the only covariate is time, t, and a fixed 'corrosion rate' parameter, β, is desired, then we must set $\delta_i = \beta t_i$. Alternatively, setting $\delta_i = \beta t_i^b$ allows for square-root, cube-root, or other nonlinear dependency on time, as referenced elsewhere in this chapter.

2.36.2.5 Threshold Techniques

These techniques can be applied in situations wherein we are interested in the statistical properties of extremes from a single large sample, such as feature depths from an Intelligent Pigging run or current noise measurements obtained by online corrosion monitoring.

Given a large random sample of measurements on such a y, for a given threshold value u, the exceedances, x, are defined by $x = y - u$ for $y > u$. It can be shown, see for example Smith and Davison,[36] that for sufficiently high thresholds and for a wide variety of initial distributions, the number, n, of these exceedances has asymptotically a Poisson distribution, while their sizes have a generalized pareto distribution (GPD) with CDF

$$F(x) = 1 - (1 - \kappa x/\sigma)^{1/\kappa}, x > 0 \quad [24]$$

The mean is $\sigma/(1+\kappa), \kappa > 1$, and the variance is $\sigma^2/[(1+\kappa)^2(1+2\kappa)], \kappa > 1/2$. Maxima from this distribution have a GEV distribution with shape parameter κ. A suitable value of the minimum threshold, above which the GPD will hold, can be found by plotting the mean exceedances in the data against the increasing threshold. This plot should eventually follow a straight line with slope $-\kappa/(1+\kappa)$ and intercept $\sigma/(1+\kappa)$. A horizontal plot corresponds to $\kappa = 0$ and a simple exponential distribution for this tail area distribution.

2.36.2.6 Wiener Process Modeling

Successive pigging runs along a corroding pipeline or successive surveys over any corroding structure will, in general, exhibit an increasing spread, as measured by range or variance, σ^2. Models with variance proportional to time are a consequence of incrementally additive assumptions for corrosion – which assumption is used, for example, in the theoretical 'shot noise' models for electrochemical noise measurements in pitting and crevice corrosion, as proposed by Cottis and other workers.[37,38] A statistical model incorporating linear growth in the mean and additive corrosion noise plus

instrument noise for tracked feature depths on two successive runs or surveys is

$$y_{ji} = \beta t_j + \omega_{ji} + \varepsilon_{ji}, \text{where } \omega_{ji} \sim W\left(0, t_j \sigma_W^2\right)$$

is a Wiener additive (corrosion)noise process

and $\varepsilon_{ji} : N\left(0, \sigma_M^2\right)$ is measurement noise;

$j = 1, 2$ runs or surveys; $i =$ feature identifier. [25]

A Wiener process has the simple property that its additive increments are normally distributed. It follows that the model for the corrosion increment $y_{(2-1)}$ is

$$\begin{aligned} y_{(2-1)i} &= \beta(t_2 - t_1) + \omega_{(2-1)i} + \varepsilon_{(2-1)i} \\ &\sim N\left(\beta(t_2 - t_1), (t_2 - t_1)\sigma_W^2 + 2\sigma_M^2\right), \\ &\text{where } \omega_{(2-1)i} \sim N\left(0, (t_2 - t_1)\sigma_W^2\right) \text{and } \varepsilon_{(2-1)i} \\ &\sim N(0, 2\sigma_M^2);\ i = \text{feature identifier.} \end{aligned} \quad [26]$$

Since we know $(t_2 - t_1)$, the sample mean for $y_{(2-1)}$ will provide an estimate $\hat{\beta} = \bar{y}_{(2-1)}/(t_2 - t_1)$, which can then be scaled to a yearly rate if the time units are not years.

Given the sample variance $s_{(2-1)}^2$ for $y_{(2-1)}$, the estimated variance for this estimated β will be $\hat{\sigma}_\beta^2 = s_{(2-1)}^2/[N(t_2 - t_1)^2]$, which can then be used to provide a confidence bound on this mean rate. Because the sample count, N, for pigging runs and oil platform or chemical plant surveys is typically large or very large, this particular adjustment may prove minor. However, feature tracking with successive pigging runs is technically and practically difficult, and this may lead to concentration on just a few of the measured features. In particular, discarded data, here, will include features that appear in one run but not in the other. Assuming a constant probability in time for such omissions (note that dependence on size will not matter for the differenced data), discarding these data will clearly not bias the subsequent estimates, although including such discarded data in a full likelihood analysis could improve the precision of such estimates. Given an independent estimate of σ_M – as supplied by the manufacturer of the measuring instrument, say, and the sample variance $s_{(2-1)}^2$ for $y_{(2-1)}$, we can then estimate σ_W^2 to be equal to $(s_{(2-1)}^2 - 2\sigma_M^2)/(t_2 - t_1)$. Assuming the time units are years, a (two-sided) prediction interval for y_δ – the (true, as opposed to measured) corrosion increment after 1 year – will therefore be $y_\delta = \hat{\beta} \pm t\sqrt{(\sigma_\beta^2 + \sigma_w^2)}$, where '$t$' comes from Student's t-distribution on $(N - 1)$ degrees of freedom. Dropping the minus sign converts this into a one-sided (upper bound) prediction interval with an appropriately adjusted probability level. A set of n successive runs or surveys will provide a set of $(n - 1)$ independent increments with means and variances to be pooled appropriately, which can then be used in the formulae mentioned earlier.

2.36.2.7 Sampling Theory and Sample Surveys

The statistical methods in sampling theory are principally designed for sampling large inhomogeneous populations of discrete items, such as motor vehicles, books in libraries, or populations of people. Two key ideas that can be useful when measuring corrosion on large structures such as an oil platform or pipeline are 'stratification' and 'clustering.' The strata of a population arise from dividing the whole population into identifiable groups such that the between-groups variability is likely to be larger than that within the groups. So elbows, joints, verticals, horizontals, and valves might form such a stratification on a large industrial structure. Clusters arise from grouping the whole of each such stratum into convenient and identifiable clumps of items for subsequent ease of sampling within each cluster. This process can be extended in a nested fashion. In a two-stage procedure, one or more clusters are chosen at random from every stratum, and then one or more items are chosen at random from each of the selected clusters. Practical introductions to these ideas, along with their theoretical justification, can be found elsewhere.[39,40]

2.36.2.8 Experimental Design

When used, the statistical principles and methods of experimental design can both help to justify the assumptions made in the standard analyses of experimental results and also improve the efficiency of the information gathering for a given amount of experimental effort. Two reference books on this subject are Cobb[41] and Cox and Reid.[42] Randomization, blocking, factorials, and fractions, Latin Squares, central composite, and optimal design are names of some of the main ideas in this area of statistics. For an illustrative example in corrosion, consider an experiment wherein metal coupons made of a particular alloy are to be corroded in a controlled environment with two flow rates F1 and F2, two temperatures T1 and T2, and two levels of acidity pH1 and pH2. In a full factorial experiment, a 'single replicate' would consist

of one experiment at each of the $2 \times 2 \times 2 = 8$ combinations of the three 'factors' flow, temperature, and acidity. Repeats would consist of exposing a further set of eight coupons to the same conditions. The addition of a base position or 'central point' at F0. T0.pH0 is often useful. This method can be contrasted with the so-called one-at-a-time experimentation, wherein each factor is altered one at a time from the base position, with repeats at each such selection. This implies a basic set of $2 + 2 + 2 = 6$ experimental combinations. For a given number of coupons used, most analyses would show that the factorial system provides a significantly more useful set of results. Alternatively, consider a situation wherein, say for financial reasons, only four of the above eight possible factorial combinations can be attempted (though, as usual, repeats might be more easily available). There are ${}^{8}C_4 = 70$ possible selections of four experiments from the eight possible, of which just two form a 2^{3-1} fractional factorial (which is also a simple Latin Square in this instance). The two experiments are (F1.T1.pH1, F1.T2.pH2, F2.T2.pH1, F2.T1.pH2) and its complement. These two have the unique property that all possible pairings appear. Selecting levels at random would have a chance of just 2/70 of achieving this property. With more factors or more levels, the potential gains increase markedly.

2.36.2.9 Bayes' Theory

The specific relevance of Bayesian methods to engineering applications is the ability to update subjective information with data for decision-making. Although the mathematical theory dates back to 1774, the numerical integration routines required to solve practical problems were not available until the 1980s. Since then, Bayesian techniques have increasingly been used in reliability and risk applications and are likely to become indispensable tools for corrosion risk management. Bolstad[43] is a suitable introductory text for 'Bayesian Statistics,' while Gelman *et al.*[44] provide the computational details necessary to implement the methods to practical problems.

2.36.2.9.1 Bayes' theorem

The essential mathematical content of the theory is 'Bayes' theorem' also known as 'Bayes' rule,' which was discovered in a restricted form by the amateur mathematician the Reverend Thomas Bayes and was subsequently published posthumously in 1763.[45] A more general formula first appeared in a 1774 paper by Pierre-Simon de Laplace.[46] In present-day notation,

$$\Pr(B|A) = \frac{\Pr(A|B)\Pr(B)}{\Pr(A)} \qquad [27]$$

or, in words, the probability that B occurs given that A occurs (i.e., $\Pr(B|A)$) is equal to the probability that A occurs given that B occurs (i.e., $\Pr(A|B)$) times the probability that B occurs (i.e., $\Pr(B)$) divided by the probability that A occurs (i.e., $\Pr(A)$).

This rather simple-looking formula hides a deep mathematical subtlety that must be thoroughly understood if we are to make proper use of Bayesian methods. As a formula in probability calculus, it is mathematically rigorous, yet ever since its first enunciation, there have been divergent interpretations of its meanings, leading to intense philosophical debate and controversy in applications. The problem lies in the interpretation of 'probability,' and we must devote some time and effort to comprehending its various meanings.

2.36.2.9.2 Probability

Bayes' formula can be rearranged to give the symmetrical form

$$\Pr(B|A)\Pr(A) = \Pr(A|B)\Pr(B), \qquad [28]$$

where, in fact, the quantity, $\Pr(B|A)\Pr(A)$, is the probability that both A and B occur, denoted by $\Pr(A\&B)$ or $\Pr(A, B)$.

The generic probabilities appearing in the above formulae have the following special names:

- $\Pr(A)$ – the *total probability* that A occurs
- $\Pr(B|A)$ – the *conditional probability* that B occurs given that A occurs
- $\Pr(A \& B)$ – the *joint probability* that A and B occur

In 'frequentist' statistics, such as that presented in previous sections, these probabilities are all interpreted as long-run frequencies that occur in repeated sampling. Thus, the probability that a coin lands heads is expressed as $\Pr(C = h) = 1/2$, with the informal interpretation that in a 100 tosses of the coin, the frequency of heads will be 'approximately' 50.

The alternative, 'Bayesian,' interpretation is that probabilities can express degrees of belief. Thus, $\Pr(C = h) = 1/2$, quantifies one's current belief that the coin is fair – *a statement that can be made even without any data*! Note, however, that an *Empirical Bayesian* would be allowed to use previously collected data to calculate a direct estimate of such a probability. The chief use of Bayes' Theorem in the context of *Bayesian*

methods is to revise such beliefs given data. We will illustrate this procedure, assuming that the coin has landed tails-up in a single toss; that is, the event $C = t$ has occurred.

2.36.2.9.3 Prior and posterior probabilities

The probability of tails is given by the Bernoulli data model, $\Pr(C = t|\Theta = \theta) = 1 - \theta$, where Θ is the probability of heads. Notice that we use a capital Greek letter for the Bernoulli parameter, which is, according to the Bayesian paradigm, *a random variable* and not a *fixed unknown value*. This is an essential point of difference between Bayesian and frequentist data modeling. Most statisticians currently use both methods according to convenience, but *true Bayesians* will never use frequentist methods.

To avoid unnecessary technical details, we assume that it is known that Θ has only two possible values θ_1 and θ_2, which are known *a priori*. Then, applying Bayes' rule,

$$\begin{aligned}\Pr(\Theta = \theta|C = t) &= \frac{\Pr(C = t|\Theta = \theta)\Pr(\Theta = \theta)}{\Pr(C = t)} \\ &= \frac{(1-\theta)}{\Pr(C = t)}\Pr(\Theta = \theta)\end{aligned} \quad [29]$$

To calculate the probability in the denominator, we use *the total probability formula*

$$\begin{aligned}\Pr(C = t) = &\Pr(C = t|\Theta = \theta_1)\Pr(\Theta = \theta_1) \\ &+ \Pr(C = t|\Theta = \theta_2)\Pr(\Theta = \theta_2)\end{aligned} \quad [30]$$

which gives

$$\begin{aligned}&\Pr(\Theta = \theta|C = t) \\ &= \frac{(1-\theta)}{(1-\theta_1)\Pr(\Theta = \theta_1) + (1-\theta_2)\Pr(\Theta = \theta_2)}\Pr(\Theta = \theta)\end{aligned} \quad [31]$$

In Bayesian statistics, the probability $\Pr(\Theta = \theta)$ is known as the *prior probability* and $\Pr(\Theta = \theta|C = t)$ is the *posterior probability* (having observed tails). Many of the problems that arise applying Bayes' rule can be traced to inappropriate specification of the prior. The key point to note here is that it cannot be specified independently of the data model, that is, the Bernoulli model, because in using the total probability formula, it must be true that

$$\Pr(\Theta = \theta_1) + \Pr(\Theta = \theta_2) = 1 \quad [32]$$

Hence, the prior itself cannot be entirely subjective nor can it be an arbitrary quantification of degrees of belief as some Bayesian analysts insist.

2.36.2.9.4 Specifying the prior

We are now faced with the problem of specifying the prior, since, evidently, it cannot be arbitrary. Laplace's own prescription for the prior – now known as an *uninformative prior* – has $\Pr(\Theta = \theta_1) = \Pr(\Theta = \theta_2) = 1/2$, giving the posterior

$$\Pr(\Theta = \theta|C = t) = \frac{1-\theta}{2-\theta_1-\theta_2} \quad [33]$$

and the *prior predictive probability*

$$\begin{aligned}\Pr(C = h) = &\Pr(C = h|\Theta = \theta_1)\Pr(\Theta = \theta_1) \\ &+ \Pr(C = h|\Theta = \theta_2)\Pr(\Theta = \theta_2) \\ &= (\theta_1 + \theta_2)/2\end{aligned} \quad [34]$$

In the special case that $\theta_1 + \theta_2 = 1$, the uninformative prior also gives $\Pr(C = h) = 1/2$, and is therefore in accord with our prior belief in a fair coin. However, in all other cases, the uninformative prior gives $\Pr(C = h) \neq 1/2$.

An alternative to the uninformative prior is to use a customized prior that matches both the total probability constraint and our prior predictive constraint. Effectively, with $\Pr(\Theta = \theta_1) = p$, we must take

$$\theta_1 p + \theta_2(1-p) = 1/2 \quad [35]$$

and hence

$$\Pr(\Theta = \theta_1) = p = \begin{cases} 1/2 & \theta_1 = \theta_2 \\ (1/2 - \theta_2)/(\theta_1 - \theta_2) & \theta_1 \neq \theta_2 \end{cases} \quad [36]$$

Such a prior is consistent with the data model and belief in a fair coin even when, *a priori*, it is known that neither θ_1 nor θ_2 is equal to $1/2$. This is because the parameter Θ in the Bernouilli data model is a *random variable* and *not some fixed unknown value*.

Now, however, there is a degenerate case in which one, but not both, of θ_1 or θ_2 is equal to $1/2$. In this case, the prior becomes $\Pr(\Theta = \theta) = \delta(\theta - 1/2)$ (δ is the delta function: $\delta(0) = 1$ and $\delta(x) = 0$ for $x \neq 0$). The prior no longer has a nonzero value on all possible values, and, therefore, there is no possibility of updating the prior, since Bayes' rule gives $\Pr(\Theta = \theta|C = t) = \Pr(\Theta = \theta|C = h) = \Pr(\Theta = \theta)$.

This phenomenon has been called the 'zero preservation problem.' It actually points to an inconsistency between our prior belief and the data model and requires revision of one or more of our models. Clearly, the quantification of even the most uncontroversial beliefs – here, simply that a coin is fair – is

fraught with technicalities, and it is no wonder that Bayesian methods are often controversial.

2.36.2.9.5 Robust Bayes

The coin example reveals how, in the Bayesian paradigm, consistency between prior beliefs and a data model is obtained by randomization over the parameters using the prior. At the limits of this randomization process are degenerate cases in which the prior has zero probability, and there is no possibility of updating prior beliefs by data. Although Bayes' rule is still formally applicable, there is really too great an inconsistency to proceed with the analysis, without a major revision of either our beliefs or the data model.

These issues, especially the 'zero preservation problem,' are especially relevant in applying Bayesian methods to engineering risk, where risks can be seriously underestimated. There is now increasing recognition by practitioners that some form of sensitivity analysis is required to augment the standard Bayesian analysis.

Currently under development are 'robust Bayes' methods that address many of the problems identified. The most promising methods employ *classes of priors*, which are appealing from a mathematical viewpoint, because Bayes' rule maps classes of priors to *multivalued predictive probabilities*. Thus, in the coin example, a prior class such as

$$\{\delta(\theta-\theta_1), \delta(\theta-\theta_2), [\delta(\theta-\theta_1)+\delta(\theta-\theta_2)]/2\} \quad [37]$$

yields the multivalued prior predictive probability

$$\Pr(C=h) \in \{\theta_1, \theta_2, (\theta_1+\theta_2)/2\} \quad [38]$$

Such multivalued probabilities have a clear interpretation in terms of prior belief – here simply, that the probability of heads takes on one of the three values.

2.36.3 Corrosion Science

This section demonstrates the application of some of the above techniques to the analysis of experimental data, primarily from laboratory studies.

2.36.3.1 Pitting Potentials and Induction Times

One common method for the experimental study of pitting corrosion is the measurement of 'pitting potentials.' Typically, multiple repeat experiments are carried out for each given set of experimental conditions, using a freshly prepared sample for each test. This produces several measurements of the pitting potential for the given set of conditions, and these measurements will be spread over a range of potentials that is much wider than that of the inherent measurement error and can be attributed to the mechanistic details of the pit initiation process. Following Shibata and Takeyama,[5] such data are often presented as sample cumulative probability distribution functions, as shown in **Figure 5**.

Similarly, another common experiment is to measure the 'induction time' before stable pitting is observed on a given sample. Results of these experiments on multiple samples can be presented as the (estimated) survival probability function, $S(t)$, by plotting the ith-order statistic against $1 - i/(n+1)$, as described in the basic terminology and methods section. Further examples and discussion on these methods can be found elsewhere.[5,7,14]

2.36.3.2 Coupon Testing

Corrosion coupons are often exposed in a test environment for significant periods of time and then examined to enable the determination of their corrosion rate. Usually, the examination consists of removing solid corrosion products and then measuring both the depth of any pits and the total weight lost from the coupon during the exposure. Sometimes, electrochemical methods may also be used to obtain corrosion rate information at certain times during the exposures. This section describes some examples of statistical approaches to the analysis of the resulting data.

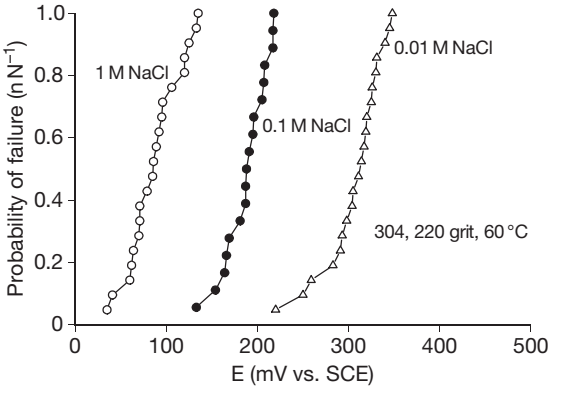


Figure 5 Typical results from potentiodynamic pitting potential measurements for 304 stainless steel with a 220 grit surface finish, in sodium chloride solutions at 60 °C. Each datapoint is the pitting potential of an individual test on a freshly prepared sample.

2.36.3.2.1 Straight-line regression analysis of weight loss data

Figure 6 shows the metal loss from carbon steel coupons exposed at two statistically similar sites along the River Thames over a period of 5 years.[47] The metal loss (mm) was calculated from the weight loss by assuming one-side uniform corrosion and applying Faraday's Second Law, and the slope of the 'best fit' line gives an estimate of the 'corrosion rate' (mm year^{-1}) over the measurement period.

Although a formal approach to corrosion might suggest omitting the intercept, a, to give a simple regression through the origin, that is $y = bt$, it is rarely sensible to do this in practice for the whole of a data set unless an appropriate significance test on a suggests that the intercept can be safely omitted, even though a may have no metallurgical validity at time zero. By forcing y to be zero at time zero, it may extend the formal range of validity of the model, but it will then almost certainly predict a higher (possibly much higher) corrosion rate, and it is the validity of the model for extrapolation into the future, not the past, which is important in practice.

In this case, the calculation of a prediction interval is also appropriate. This has the form

$$y_x = \hat{y}_x \pm t\hat{\sigma}\sqrt{\left(1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right)} \qquad [39]$$

for a future observation on y at x, with t selected appropriately from t-tables with $(n - 2)$ degrees of freedom. A confidence interval about the line (which omits the '1' inside the bracket above – as this term relates not to the line itself, but to future noise in the observations as they appear about the line) and the prediction interval for future observations are demonstrated in Figure 6.

2.36.3.2.2 Extreme value analysis of pit depth measurements

Figure 7 shows the maximum pit depth as a function of immersion time for 316L stainless steel coupons in a 10% ferric chloride solution at 50 °C, as reported by Pierpoline *et al.*[48] Laycock *et al.*[32] have fitted to this data a constant coefficient of variance model with nonlinear time dependence, using ML. This had location parameter $\delta_i = \beta t_i^b$ and scale parameter $\alpha_i = \rho\beta t_i^b$ (where $\rho = \sigma/\mu$), and produced the following parameter estimates (with associated standard errors in brackets)

$$\tilde{\alpha} = 1.004(\pm 0.317), \tilde{\delta} = 6.322(\pm 1.721),$$
$$\tilde{\kappa} = 0.401(\pm 0.162), b = 0.376(\pm 0.047) \qquad [40]$$

This implies the following formula for the extrapolated mean for maxima as a function of extrapolation area A in square inches and time t in days

$$\tilde{\mu} = (29.16 - 14.404/A^{0.401})t^{0.376} \rightarrow$$
$$29.16t^{0.376} = \xi t^{0.376} \text{ for large } A \qquad [41]$$

When the two or three largest (etc.,) maxima are collected at each location, this can afford greater

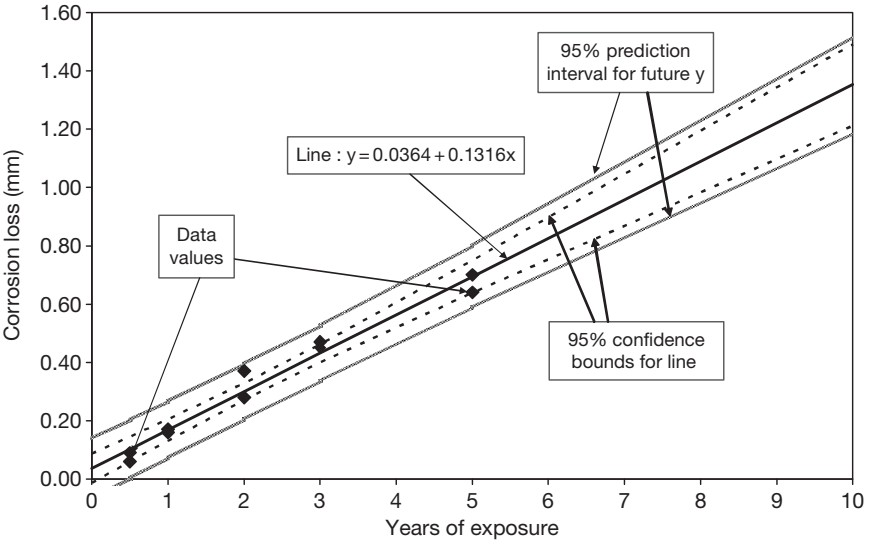


Figure 6 Corrosion loss from carbon steel coupons exposed in the River Thames for 5 years,[47] with the best-fit line obtained by straight-line regression and extrapolated for a further 5 years. Also, shown are the 95% confidence bounds for the fitted line, and the 95% prediction interval for future 'corrosion loss.'

precision for the same estimated parameters, as demonstrated for the earlier data.[49]

Another example is provided by a set of 1800 maximum pit depths extracted from 8×8 grids over experimental coupons of a corrosion-resistant alloy subjected to a complex cyclic testing regime, involving hold periods at 350 °C (for between 1 and 23 h), with and without exposure to salt, for total test durations of 1 year. The statistical package Genstat was used to prepare an analysis of variance table for the data, with the results shown in **Table 1**.

Inspection of a half-normal plot for the residuals showed good linearity, suggesting that the standard model assumptions, for this particular analysis, were sound. It can be seen that the variation of maximum pit depths between days of exposure (3, 4, or 11, treated as a factor with three levels) and temperature cycles (either 1 or 23 h at 350 °C) and salt treatments (Yes or No) was highly significant ($P < 0.001$), but the other two factors (salt replenishment and measurement position) were definitely nonsignificant ($P > 0.5$) although not so small as to cast doubt on the validity of the model or data. Therefore, four experimental conditions, defined by the two temperature cycle schemes and exposure to or not to salt, were found to give statistically different results in terms of maximum pit depths, though no detectable growth of pits could be found for the cycle with 23 h at 350 °C and no salt exposure. Hence, the Fortran program described in Laycock *et al.*[32] was used to fit four-parameter GEV distributions to the three sets of data, which had significant upward time trends. **Figure 8** shows a plot against time of one of these fitted GEV mean functions, which were all Type III, for a selection of extrapolated

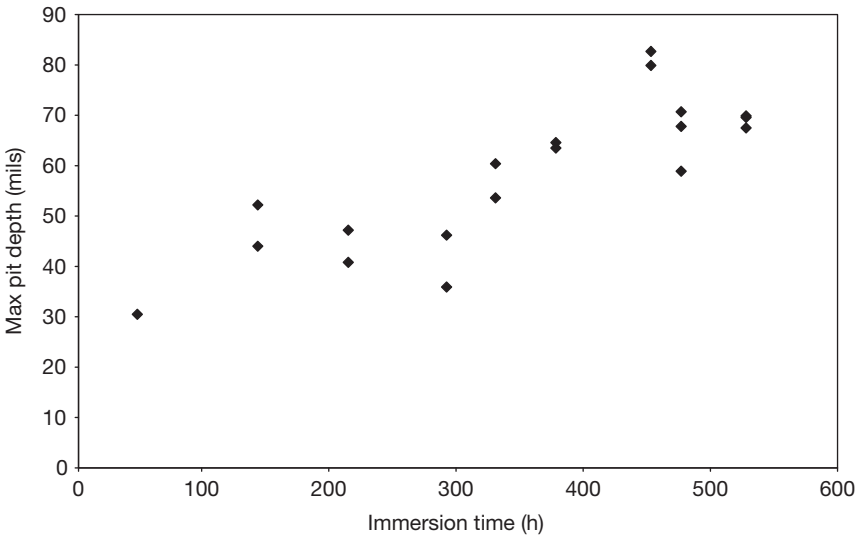


Figure 7 Maximum pit depth as a function of immersion time for 316L stainless steel coupons in a 10% ferric chloride solution at 50 °C. Reproduced from Pierpoline, R.; White, R. J.; Wrong, C.; Cornwell, L.; Griffin, R. In *NACE-9: Advances in Localised Corrosion;* Isaacs, H., Bertocci, U., Kruger, J., Smialowska, S., Eds.; NACE: Houston, TX, 1990; pp 123–126.

Table 1 Analysis of variance for pit depth data from coupons of a corrosion-resistant alloy subjected to a complex cyclic testing regime

Source of variation	*Degrees of freedom*	*Sum-of-squares*	*Mean-square*	*Variance-ratio*	*F dn probability*
Cycles(2)	1	79 644.4	79 644.4	548.74	<.001
Days(3)	2	19 183.6	9 591.8	66.09	<.001
Salt(Y/N)	1	10 339.5	10 339.5	71.24	<.001
Replication	1	0.1	0.1	0.00	0.975
Position	24	2 666.2	111.1	0.77	0.784
Residual	1770	256 898.0	145.1	–	–
Total	1799	368 757.5	–	–	–

1800 maximum pit depths were extracted as internal 5×5 grids (to remove edge effects) from 8×8 grids over experimental coupons. The statistical package Genstat was used to prepare this table.

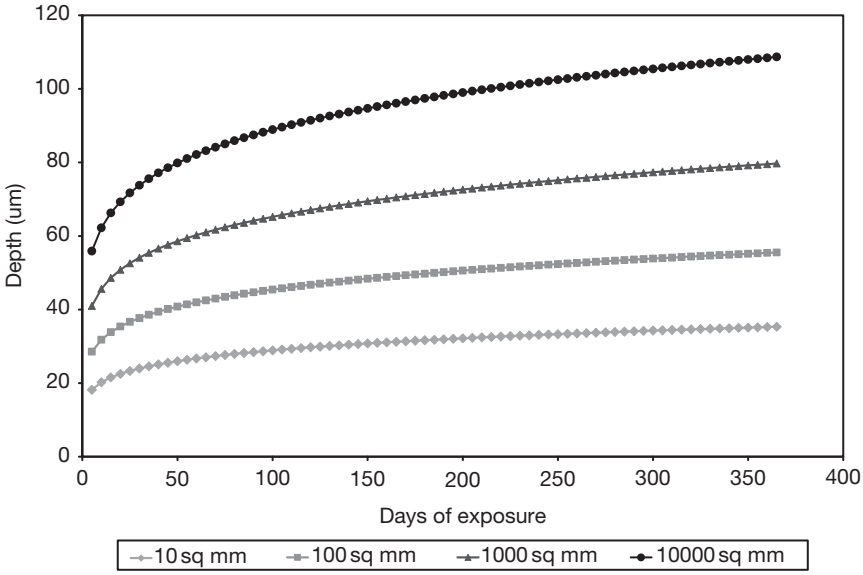


Figure 8 A Fortran program[32] was used to fit four-parameter GEV distributions to pit depth data from cyclic testing of corrosion-resistant alloy coupons. This figure shows a plot against time of one of these fitted GEV mean functions for a selection of exposure areas.

exposure areas related to operational requirements for this alloy, which was to be used inside jet engines operating in a marine environment.

2.36.3.2.3 Response surface regression analysis of LPRM data

Sometimes, experiments are carried out to determine the influence of various parameters (such as alloy composition) on the corrosion rate in a given environment. In one such set of experiments, linear polarization resistance measurements (LPRM) were used to measure corrosion rates for a selection of different carbon steels in a concentrated HF environment. The steels were selected because they had small differences in residual element (nickel, chromium, copper) and carbon content, and it was suspected that such variations were responsible for unexpected failures due to corrosion in HF alkylation plants. To determine the optimum steel composition for corrosion resistance in this environment, the data were analyzed using response surface regression. The most popular version of this is the bivariate second order model:

$$y_i = \beta_{00} + \beta_{10}x_{1i} + \beta_{01}x_{2i} + \beta_{20}x_{1i}^2 + \beta_{02}x_{2i}^2 + \beta_{11}x_{1i}x_{2i} + e_i \quad (i = 1, \ldots, n), \qquad [42]$$

which can be described as the *quadratic* version of *bivariate linear regression*, and it represents a smooth hill, or valley, or saddle-point in the response surface.

The general tth-order, k-variate, model has the form

$$y_i = \sum_{0 \le i_1 + i_2 + \ldots i_k \le t} \beta_{i_1+i_2+\ldots+i_k} x_{1i}^{i_1} x_{2i}^{i_2} \ldots x_{ki}^{i_k} + e_i \quad (i = 1, \ldots, n) \qquad [43]$$

This model has k *genuine* variables, the rest being *carrier* variates.

The fitted model was a second-order response surface for LPRM regressed on carbon, chromium, copper, and nickel content, based on 45 determinations of LPRM. This model gave a good fit to the data, with a squared multiple correlation coefficient of ~80% and 30 degrees of freedom for the residual. A contour plot for the raw data is shown in **Figure 9(a)**, while the corresponding contour plot for the fitted model but controlled for the hidden variates (nickel and copper at their median values) is shown in **Figure 9(b)**. From this saddle-point figure, the qualitative conclusions for this application were that when copper and nickel are at their median levels (0.06% and 0.04%, respectively), a decreasing carbon content is good for corrosion resistance, provided chromium is low, while an increasing carbon content will be good when chromium is high. This analysis supported a recommendation for new guidelines on the residual element and carbon content of steels for concentrated HF service.

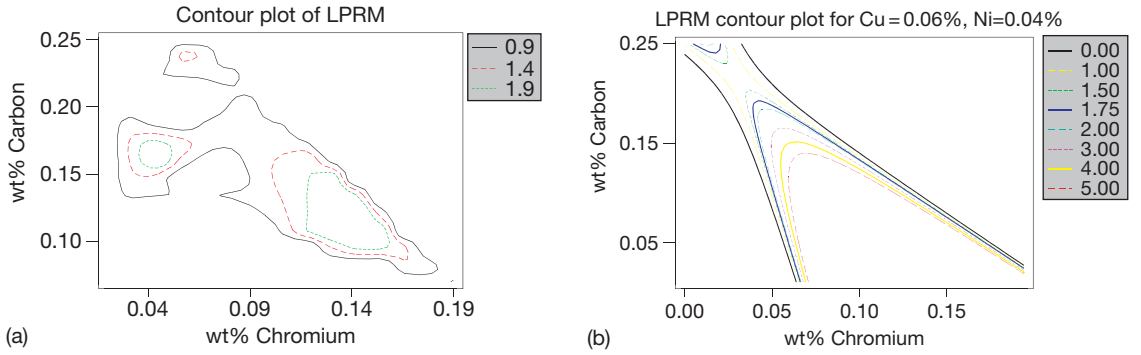


Figure 9 LPRM data for carbon steels of different alloy content in deaerated 60% HF (by volume) at 71 °C, with 5 ppm of predissolved iron. The fitted model is a second-order response surface regressed on carbon, chromium, copper, and nickel content, based on 45 determinations of LPRM. A contour plot for the raw data is shown in (a), while the corresponding contour plot for the fitted model but controlled for the hidden variates (nickel and copper at their median values) is shown in (b). Data produced in a multiclient project at CAPCIS Ltd, Manchester, UK.

2.36.3.3 Crack Depth Modeling

Cracks in engineered structures, whether they are caused by (corrosion) fatigue or stress-corrosion cracking, can be considered, in general, as defects, which initiate and 'grow' according to some random mechanism. Hence, there is much scope for modeling. There is, of course, an extensive literature on crack growth, and, in particular, fracture mechanics has been extensively used to study the principle and nature of fatigue crack growth. Many authors present purely theoretical considerations in an attempt to provide a model to predict crack growth, and such methods have been reviewed.[50,51] Dolinski[52] is skeptical as to whether theoretical considerations alone can provide an adequate model for prediction. On the other hand, when data are presented, they are often experimental in nature,[53] from which it is extremely difficult to infer useful practical models for the prediction of crack growth in real complex systems. For this reason, an empirical model, which makes reasonable assumptions, and which can be evaluated or fitted using data on the nature and extent of cracks observed at inspection may be more useful in practice. This empirical approach will suppose that the physical properties of the system, such as loading history and material characteristics, are inherent in the field data and also suppose that data are available in sufficient quantity and in a form that allows the fitting of a statistical model. This model can then be used to predict fatigue crack growth and to investigate the consequences of various inspection policies. This philosophical contrast between a physical model and an empirical one based on the observation of large scale effects is not uncommon and not easily resolved. Although the model in this section is not based strictly on fracture mechanics, the approach and the assumptions made are simple, and the approach taken is a practical one.

2.36.3.3.1 Two-stage modeling of crack growth

Detailed models of crack growth will consider crack development in stages. Suppose a crack initiates according to one random process, and subsequent to initiation, it grows according to another random process independently of the initiation process. Given a simple model for each stage, the distribution of crack depths at some time T can be formulated. The simplest model that can capture the randomness of the initiation process is the homogeneous Poisson process.[54] This supposes that cracks initiate purely at random but at an average rate of λ crack initiations per unit of time. Then, the number of crack initiation events in the time period $[0,T]$, N_T, will be Poisson distributed with mean λT, that is, $E(N_t) = \lambda T$ and $\Pr(N_t = n) = \mathrm{e}^{-\lambda T}(\lambda T)^n/n!$.

Next, let the distribution of crack depth for a crack that has been growing for some time τ be $G(z;\tau) - G$, which will be specified later. For now, put simply, not all cracks that have been growing for the same time would be expected to be of the same size – some will propagate more quickly than others. Thus, the depth of a crack Z that has been growing for a particular time will be a random variable, and this random variable will have distribution G.

Suppose a structure or an area on a structure is inspected for cracks at some time T, and a number of

cracks are observed. If any of these cracks is selected at random (with no size-related selection bias), then it is a property of the Poisson process of initiation events that the time of initiation of the crack is uniformly distributed over the interval $[0,T]$ – that is, the crack could have initiated in any equal interval of time within $[0,T]$ with equal probability. This can be thought of as the defining property of the Poisson process and one that characterizes events that occur 'purely at random.' With inspection carried out at time T, conditional on initiation of a crack at time x, the distribution of crack depth at time T will be $G_T(z;\, T-x)$. However, we do not observe x, and, therefore, predictions have to be made unconditionally. To do this, we appeal to the law of total probability so that the unconditional distribution of the crack depth at time T (with no regard to the initiation time), $F_T(z)$, is

$$F_T(z) = \int_0^T \Pr(Z_T \le z | x = x) f(x)\mathrm{d}x = \frac{1}{T}\int_0^T G_T(z;\, T-x)\mathrm{d}x, \qquad [44]$$

where f is the probability density function of the initiation time. This expression is interpreted as follows: the distribution of the depth Z is a weighted average of the distribution of depth, given the initiation time with weights given by the probabilities of particular initiation times. We have already stated that for the Poisson process, these initiation times are uniformly distributed on $[0,T]$ so that $f(x) = 1/T$.

For n crack depths $z_1, \ldots, z_n$ observed at time T, the log-likelihood function is given by

$$\sum_{i=1}^{n} \log\left[\frac{\mathrm{d}}{\mathrm{d}z}\left\{\int_0^T G_T(z_i;\, T-x)\mathrm{d}x\right\}\right] \qquad [45]$$

From this, we can estimate the parameters of G and F by implication, and make predictions about Z_t for $t > T$.

It may be natural to focus on the maximum crack depth. Only the maximum crack depth may be recorded. In this case, given n crack initiations in the interval $[0,T]$, the distribution function of the maximum crack depth is $F_{\max,T}(z|N_t = n) = \{F_T(z)\}^n$, given the assumption that the n crack depths are independent and identically distributed. If, in general, the number of cracks at a location at an inspection is unknown, it follows, arguing unconditionally on N_t, that the distribution function of the maximum crack depth is given by $F_{\max,T}(z) = \sum_{n=0}^{\infty}\{F_T(z)\}^n \Pr(N_t = n)$, and the likelihood for an observed maximum crack depth $z_{(1)}$ is $\log[\frac{\mathrm{d}}{\mathrm{d}z} F_{\max,T}(z_{(1)})]$ if $z_{(1)} > 0$ and $\log[\Pr(N_t = 0)] = -\lambda T$ if $z_{(1)} = 0$ (i.e., if no cracks have initiated).

With $F_T(z)$ given by eqn [44], it then follows that

$$F_{\max,T}(z) = \sum_{n=0}^{\infty}\left\{\frac{1}{T}\int_0^T G_T(z;\, T-x)\mathrm{d}x\right\}^n \mathrm{e}^{-\lambda T}(\lambda T)^n/n! = \exp\left\{-\lambda\int_0^T 1 - G_T(z;\, T-x)\mathrm{d}x\right\} \qquad [46]$$

As to the form of the distribution function for $G_T(z;\, T-x)$, for simplicity, the two-parameter Weibull distribution is recommended. This distribution is widely used in reliability for modeling failure times.[55] Taking the scale parameter α of the Weibull to be a power function of the 'growing time,' $\tau = T - x$, G becomes

$$G_T(z;\tau) = 1 - \exp[-\{z/(\alpha_0\tau^\gamma)\}^\beta], \qquad [47]$$

where β is the shape parameter. This implies that the mean depth grows according to $\mu = \alpha_0\Gamma(1 + 1/\beta)\tau^\gamma$, and where $\Gamma(\cdot)$ is the gamma function. This power law function for the scale parameter, while having no immediate physical justification, provides, along with the Weibull distribution, a flexible family of growth laws and also implies power law growth in the mean. The implied distribution of the maximum crack depth is then

$$F_{\max,T}(z;\lambda,\alpha_0,\beta,\gamma) = \exp\left[-\lambda\int_0^T \exp\left[-\left\{\frac{z}{\alpha_0(T-x)^\gamma}\right\}^\beta\right]\mathrm{d}x\right] \qquad [48]$$

Extrapolation in space and time now proceeds as follows. Suppose the estimation of the maximum is carried out using maxima from sampled regions of area A at time T, if we wish to determine the distribution of the maximum over an area $a > A$ at time $t > T$, then it is required to scale up the initiation rate λ to $\lambda(a/A)$, and, therefore, the distribution of the maximum over an area a at time t is then

$$F_{\max,t,a}(z;\lambda,\alpha_0,\beta,\gamma) = \exp\left[-\lambda\left(\frac{a}{A}\right)\int_0^t \exp\left[-\left\{\frac{z}{\alpha_0(t-x)^\gamma}\right\}^\beta\right]\mathrm{d}x\right]. \qquad [49]$$

If X is the time for the maximum crack depth to reach some critical level, z_f, then the expected value of X can be determined from this distribution, since $\Pr(X > t) = \Pr(Z_{\max,t,a} < z_f)$ and $E(X) = \int_0^\infty \Pr(X > t)\mathrm{d}t$.

2.36.4 Corrosion Engineering

Statistical data analysis techniques have long been used in the assessment of corrosion damage to industrial plant. Such techniques can be applied to *condition assessment*, wherein there typically are concerns about the influence of measurement error and the extrapolation of sample inspection results to uninspected areas; and *life prediction*, wherein the goal is to determine the remaining life of the equipment.

2.36.4.1 Sample Inspection of Heat Exchanger Tubes

Heat exchanger tubes are often inspected using nondestructive testing (NDT) methods – such as IRIS (internal rotary inspection system) or LOTIS (laser optical tube inspection system) – and the results are typically reported in terms of the maximum pit depth in each inspected tube. However, in many such cases, no more than 10 or 20% of the tubes are tested, at least initially, and the sample data must be used to estimate the overall condition of the tube bundle. The hypergeometric distribution – as described in the introductory section – can be applied here to determine the required level of inspection for a given detection requirement.

ASTM G46[10] describes a procedure based on the Type-I EV (Gumbel) distribution for extrapolating maximum pit depths over areas larger than the inspected area, and this technique is described in the following example. A carbon steel ammonia condenser with cooling water on the tube-side was inspected after ~2 years service. The exchanger contained 773 nominally identical tubes, of which 54 were inspected, and the maximum measured pit depth from each inspected tube was recorded. The method described in G46 was then applied to this dataset, as shown in Figure 10. In this case, the extent of inspection was ~7%, and the maximum measured pit depth was 0.99 mm, while the maximum predicted pit depth was 1.32 mm.

The previous example provides no validation that the analysis has produced an accurate estimate of the condition for the uninspected tubes. However, the next example concerns another exchanger with cooling water on the tube side, in which 357 out of 410 tubes were inspected (with the remaining 53 tubes being uninspectable due to obstructions). To assess the accuracy of the GEV analysis technique, the full dataset was randomly sampled to create simulated inspection datasets of different sizes from 5 to 100 tubes. For each of these simulated datasets, the GEV technique was used to estimate the maximum pit depth in the entire 410 tube bundle, with 95% confidence limits also determined for each prediction. As shown in Figure 11, for sample sizes over ~50 tubes (or ~25% inspection), the estimated maximum pit depth is both quite stable and in excellent agreement with the maximum measured pit depth from all the inspected tubes.

In the latter example, it is also instructive to consider the measurement error. For the NDT technique used in this case, the stated accuracy on clean tubes was ± 0.05 mm, but this varies for other inspection methods. For a sample size of 10% (40 tubes), Figure 12 shows the impact of taking this effect into account. The ML estimates for the Gumbel parameters were obtained using the Xtremes software of Reiss and Thomas.[56] To allow for small sample calculations, we use the parametric bootstrap method to obtain standard errors and confidence limits.[57] The bootstrap method is implemented in the Xtremes software. Therefore, all we need to do is adjust the ML estimates to allow for the measurement error.

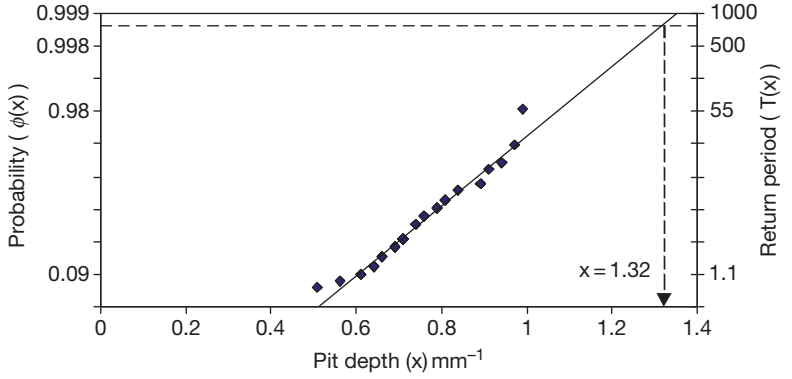


Figure 10 ASTM G46 extrapolation method for maximum pit depth in an ammonia condenser with 54 tubes inspected out of 773 in total. The return period is 773, and the maximum expected pit depth is 1.32 mm.

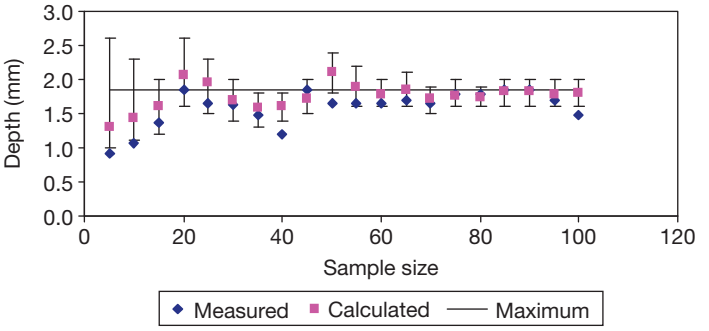


Figure 11 For a cooler with pitting corrosion on the internal surface of the tubes, this figure compares the maximum measured pit depth in samples of varying size (up to 100 sampled tubes); the calculated maximum pit depth based on GEV analysis of the given sample size; and the maximum measured pit depth in the 357 inspected tubes from the 410 tube bundle (1.85 mm).

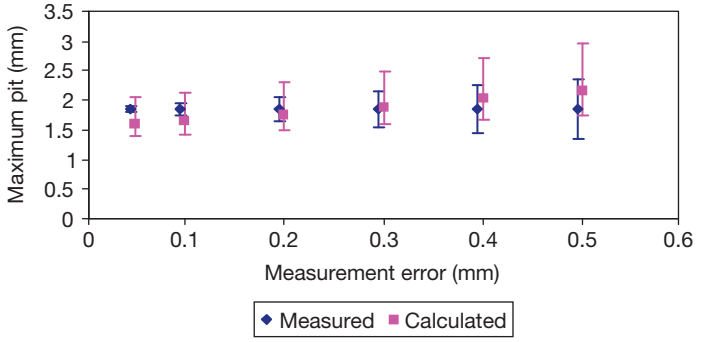


Figure 12 For a cooler with pitting corrosion on the internal surface of the tubes. As a function of the assumed inherent measurement error, this figure compares; the maximum calculated pit depth based on inspection of a random 10% sample of the tubes; and the maximum measured pit depth in the 357 inspected tubes from the 410 tube bundle (1.85 mm).

In the special case that the measurement error is uniform, we use the following method.

The (ideal) genesis of the Gumbel is as follows:

1. The distribution of deep pits follows an exponential distribution, $1 - \exp(-(x-\theta)/\sigma)$.
2. The number of deep pits follows a Poisson distribution with mean λ.

Then, the distribution of the maximum pit depth follows the Gumbel distribution, $\exp[-\exp\{-(x-\theta)/\sigma + \ln(\lambda)\}]$, which has mode $\theta + \sigma\ln(\lambda)$ and scale σ.

If, in fact, the pits are measured with uniform error between $-R$ and R, then the distribution of measured depths x is exponential with distribution function $1 - \exp[-(x-\theta-\varepsilon)/\sigma]$, where

$$\varepsilon = \sigma \ln\left[\frac{\sinh(R/\sigma)}{R/\sigma}\right] \qquad [50]$$

The resulting distribution of the maximum measured pit follows the Gumbel distribution, $\exp[-\exp\{-(x-\theta-\varepsilon)/\sigma + \ln(\lambda)\}]$, which has mode $\theta + \varepsilon + \sigma\ln(\lambda)$ and scale σ. Therefore, the effect of the measurement error is to increase the mode by an amount ε. Hence, given the ML estimates, $\hat{\mu}$ and $\hat{\sigma}$, the appropriate adjustment for uniform error is to *reduce* the mode by an amount

$$\hat{\varepsilon} = \hat{\sigma} \ln\left[\frac{\sinh(R/\hat{\sigma})}{R/\hat{\sigma}}\right] \qquad [51]$$

while leaving the scale estimate as $\hat{\sigma}$. The usual parametric bootstrap can now be applied with the adjusted parameters to obtain the required standard errors and confidence limits. Further discussion on these techniques can be found elsewhere.[58]

2.36.4.2 Thickness Measurement Locations

Corrosion engineers are frequently faced with the situation wherein they have inspection data showing the change in wall thickness (or increase in pit depth) at certain locations on a given piece of equipment

over several months or years, and are now being asked to predict the remaining useful life of this equipment. Here, we ignore the various possible different definitions for 'end-of-useful-life' and assume that there is a defined minimum allowable wall thickness, and that reaching this thickness at any one location constitutes the end-of-life for the given piece of equipment. In-service inspection codes[59,60] describe simple ways of calculating corrosion rates for a single thickness measurement location (TML); for example, the long-term corrosion rate is determined from the wall loss between the first measurement and the most recent measurement, while the short-term corrosion rate is obtained using only the two most recent measurements. However, there are also several different statistical approaches to this problem.

2.36.4.2.1 Straight-line regression

Figure 13 shows the wall thickness data obtained by ultrasonic testing at one particular location on a wet gas flowline over a period of several years.

In this case, the observed erosion–corrosion damage was restricted to the outside of a bend, and the inspections always covered 100% of the damaged area such that there is no concern here about extrapolation of the results to other locations. As in the earlier examples concerning coupon testing data, the slope of this 'best fit' line gives an estimate of the 'corrosion rate' over the measurement period, and a confidence interval about the line and the prediction interval for future observations are also shown. The corresponding graphic produced for these data by the Shell inspection design analysis and plotting (S-IDAP) package[61] is presented in Figure 14, where it is seen to present much the same predictions out to the same time point in 2008.

2.36.4.2.2 Extreme value analysis

In this example, ultrasonic testing has been used to map the extent of internal corrosion damage to a carbon steel Benfield process vessel.[62] Inspections were carried out ~1 year apart, the measurements made using continuous scans of 50 mm × 50 mm grids and the minimum measured wall thickness being reported for each grid. For the worst case (thinnest) locations at the first inspection, the maximum measured wall loss between the two inspections was 0.6 mm, but, on average, there was negligible change (Table 2). It might be assumed, therefore, that the apparent differences between results at specific TMLs are all due to the measurement error (which has a mean of zero). However, for 8 of the 11

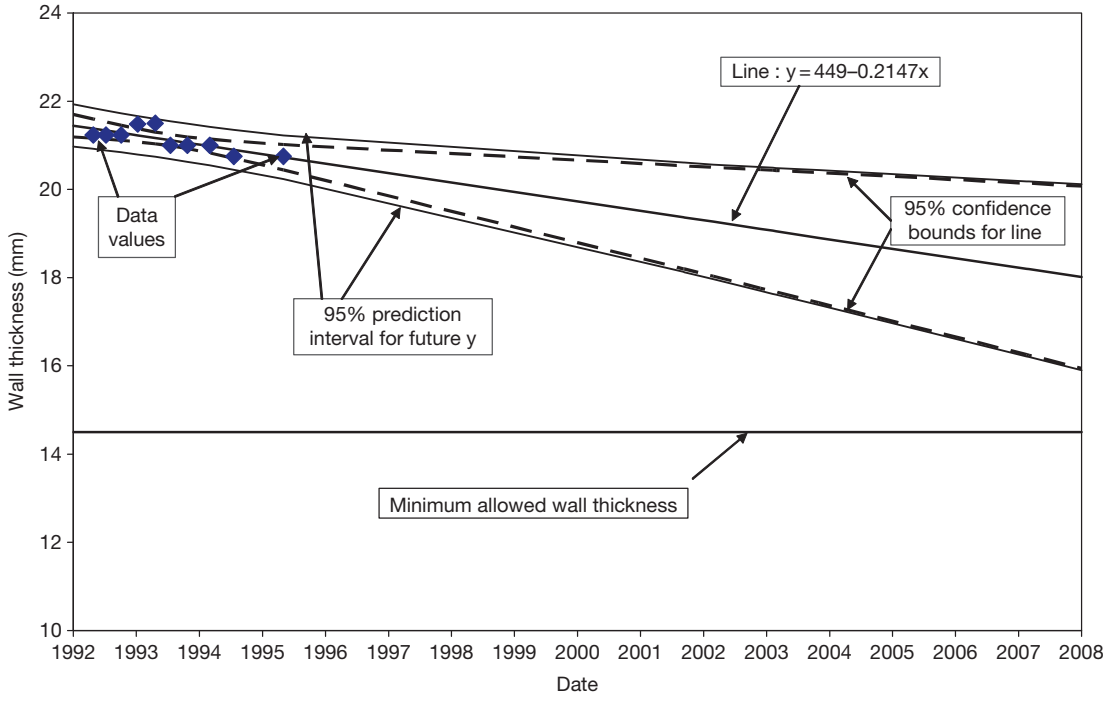


Figure 13 Minimum wall thickness data from ultrasonic testing measurements on the outside of a bend in an offshore gas flowline, with confidence and prediction intervals around the best fitting line.

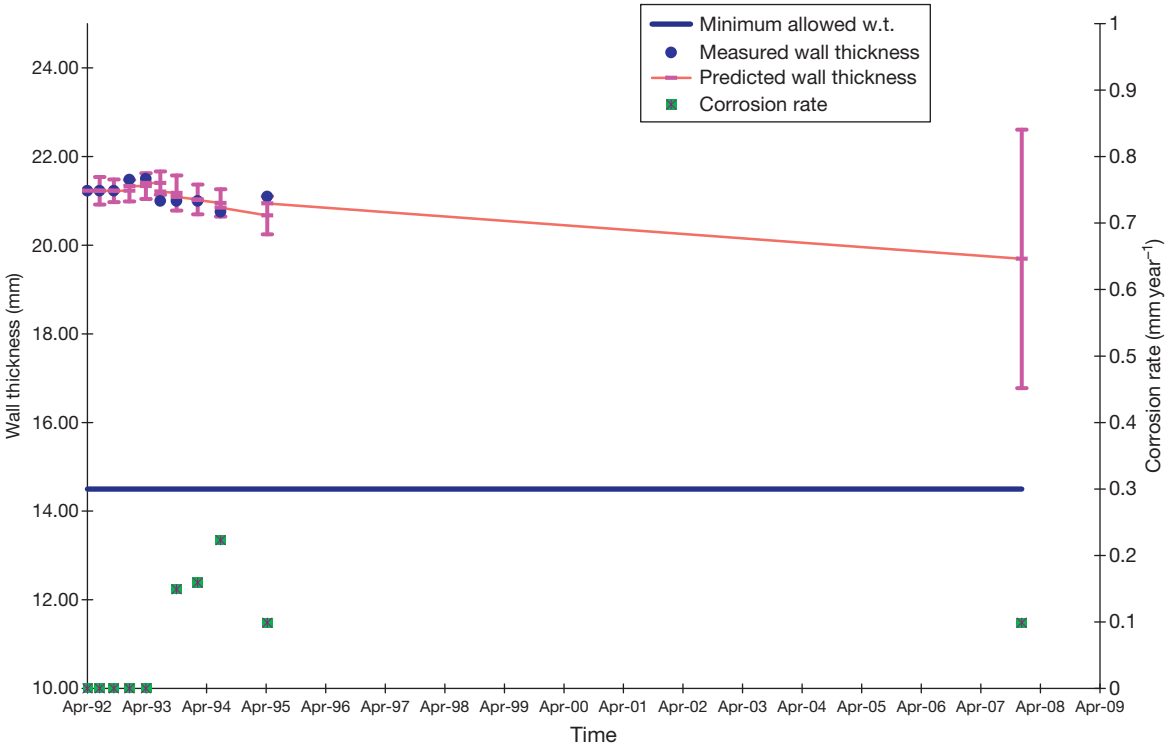


Figure 14 The same measurement data as in **Figure 13**, with the extrapolation in time carried out using the S-IDAP package described by Hoeve *et al.*[61]

Table 2 Ultrasonic testing results from inspection of a Benfield Process vessel at several thickness measurement locations in 2001 and 2002[62]

TML	*Wall loss Aug 01–Nov 02 (mm)*
1	−0.4
2	0
3	−0.9
4	0.1
5	−0.5
6	0.3
7	0.6
8	0.6
9	−0.1
10	0.2
11	0.3
Mean	0.01

The selected locations are those where the minimum wall thicknesses were found in 2001. Note that negative values indicate an apparent increase in wall thickness.

circumferential rows making up the inspected area, the minimum measured thickness was found in a different grid in the two inspections. This happens partly because of inherent measurement errors and partly because the highest corrosion rates are not necessarily experienced at the same locations at all times.

An alternative approach to this problem again takes advantage of extreme value analysis methods. In **Figure 15**, the maximum wall loss in each circumferential row is plotted for each inspection time. This is done in such a way that data conforming to a Type-I EV distribution would fall along a straight line. Hence, it is clear that the data from each inspection follow a Type-I EV distribution, but the entire distribution has shifted to higher wall loss values over time, as might reasonably be expected in a vessel suffering from corrosion. The corresponding corrosion rate can be calculated from the change in the Gumbel mode (assuming the same scale). Thus, we consider $x \sim \mathrm{EV}(\mu, \sigma)$ for the wall loss data from August 2001 and $x \sim \mathrm{EV}(\mu + \Delta, \sigma)$ for the data from November 2002. As the data are based on minima from different locations, we can assume that the wall loss results are statistically independent, yielding two independent samples, each of size 11. Using the ML estimation procedure, we obtain the parameter estimates $\hat{\mu} = 5.05, \hat{\sigma} = 0.36$, and $\hat{\Delta} = 0.43$. In general, ML estimates are biased in small samples. To estimate the bias-corrected estimate and its standard error, we use the two-sample *jackknife* procedure, which gives $\hat{\Delta} = 0.46$ with

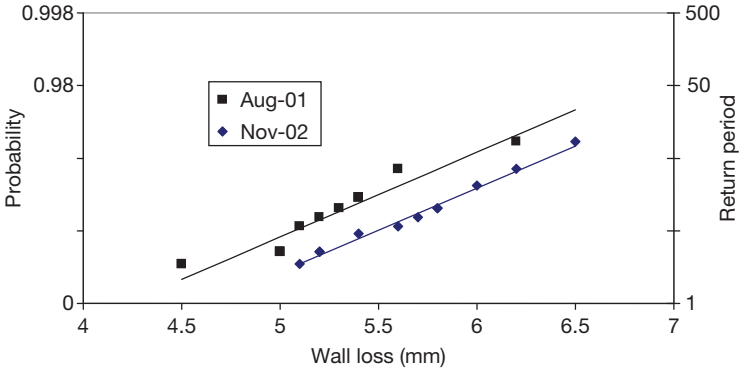


Figure 15 EV analysis of minimum wall thickness data from surveys of a Benfield process vessel in August 2001 and November 2002. The progression of the distribution from left to right shows the rate at which the minimum wall thickness is decreasing, which is a practical definition of the corrosion rate. Reproduced from Wilson, P. T.; Krouse, D. P.; Moss, C. J. Proceedings of Corrosion Control and NDT; ACA, Melbourne, Australia, 23–26 November 2003; ACA: Melbourne, 2003; Paper no. 45.

standard error 0.22. Note that the *large sample* standard error for the ML estimate of the mode is $\sigma\sqrt{\frac{1.11}{n}}$.

In the present case, this is estimated by $0.36\sqrt{1.11/11} = 0.114$. Hence, the standard error for the difference of two independent ML estimates is $0.114\sqrt{2} = 0.16$, which is, as expected, somewhat less than that obtained using the jackknife procedure.[57] Thus, we conclude that a wall loss of 0.46 ± 0.22 mm occurred over a period of 457 days, yielding a corrosion rate estimate of 0.37 ± 0.18 mm year^{-1}.

2.36.4.2.3 Bayesian methods

In many corrosion applications, fitness for purpose is assessed using the remaining wall thickness (RWT). However, the RWT varies with the initial wall thickness, WT, independently of the corrosion process. The primary corrosion damage measure is actually the wall loss, which is independent of the initial wall thickness. For a given wall loss, $X = x$, the determination of the RWT depends on the measurement method. Generically, measurement methods determine either the material remaining or the material lost. In the first case, the RWT is measured directly as

$$y^{(1)} = (WT - x) + e^{(1)} \qquad [52]$$

(for some error $e^{(1)}$), and in the second case, the wall loss is measured directly as $x + e^{(2)}$ (for some error $e^{(2)}$), giving the apparent RWT

$$y^{(2)} = \mathrm{NWT} - (x + e^{(2)}) \qquad [53]$$

where NWT is the nominal wall thickness. However, noting that $y^{(2)} = (\mathrm{WT} - x) + [(\mathrm{NWT} - \mathrm{WT}) - e^{(2)}]$, in general, the RWT is given by $y = (\mathrm{WT} - x) + e$.

Bayesian methods are ideally suited to the problem of varying wall thickness. Wall loss X is assumed to be independent of the initial wall thickness, WT, and to have a continuous distribution function $F(x)$. Hence, conditional on the initial wall thickness, $\mathrm{WT} = w$, the wall loss has the truncated distribution function $F(x)/F(w)$ and differentiation gives the conditional density,

$$f(x|\mathrm{WT} = w) = \frac{f(x)}{F(w)} \qquad [54]$$

Hence, the first application of Bayes' rule is to obtain the posterior as

$$p(\theta, \mathrm{WT} = w|x) = \frac{1}{c(x)}\frac{f(x,\theta)}{F(w,\theta)}p(\theta)p_{\mathrm{WT}}(w) \qquad [55]$$

where the normalizing constant is given by

$$c(x) = \int\left[\int\frac{p_{\mathrm{WT}}(w)}{F(w,\theta)}\mathrm{d}w\right]f(x,\theta)p(\theta)\mathrm{d}\theta \qquad [56]$$

Now the posterior density $p(\theta|x)$ can be obtained by integrating over the initial wall thickness to give

$$p(\theta|x) = \frac{f(x,\theta)}{c(x)}p(\theta)\int\frac{p_{\mathrm{WT}}(w)}{F(w,\theta)}\mathrm{d}w \qquad [57]$$

This formula explicitly quantifies the effect of uncertainty in initial wall thickness. In cases wherein the wall loss is small compared with the initial wall thickness, $F(w,\theta) \cong 1$, and the posterior reduces to

$$p(\theta|x) = \frac{f(x,\theta)p(\theta)}{\int f(x,\theta)p(\theta)\mathrm{d}\theta} \qquad [58]$$

Having obtained the posterior, the posterior predictive probability of wall loss exceeding a

threshold, u, is given by integrating over the posterior, that is,

$$\Pr(X > u|x) = 1 - \int F(u,\theta)p(\theta|x)\mathrm{d}\theta \qquad [59]$$

These results can be extended to the case when the wall loss is measured directly with measurement-error density, $\varphi(z)$, since the likelihood conditional on the wall thickness is now given by

$$f(y|WT = w) = \frac{1}{F(w)}\int_{y-w}^{y} f(y-z)\varphi(z)\mathrm{d}z \qquad [60]$$

The case in which the RWT is measured directly gives a similar result,

$$f(y|WT = w) = \frac{1}{F(w)}\int_{y-w}^{y} f(w-y+z)\varphi(z)\mathrm{d}z \qquad [61]$$

In cases wherein the wall loss can be localized to a 'defect' and there are a large number per component, the problem reduces to considering the maximum wall loss. In case the number of defects is Poisson with mean λ, the maximum wall loss has distribution function $\exp\{\lambda[1-F(x)]\}$.

In many practical situations, the GEV distribution can be used to approximate the distribution of the maximum wall loss. A question of special interest is whether the distribution has an upper bound, which is given by

$$\delta + \frac{\alpha}{\kappa} = X_{(1/e)} + \frac{X_{(0.5)} - X_{(1/e)}}{1 - (\ln 2)^{\kappa}} \qquad [62]$$

where $X_{(p)}$ denotes the p the quantile of the distribution. Evidently, since the quantiles are readily estimated by the sample values, uncertainty in the putative upper bound is dominated by uncertainty in the shape parameter, κ. Indeed, any Bayesian analysis using the GEV is very sensitive to the shape-parameter prior, and more experience is required to assess whether the results are suitably robust for practical application.

Figure 14 shows an analysis of wall thickness data produced by the S-IDAP package described by Hoeve *et al.*[61] S-IDAP makes Bayesian predictions based on a number of alternative models in addition to the linear or constant corrosion rate model. These include a genuine change in corrosion rate as well as other data anomalies such as an outlier or incorrect measurement. Given prior probabilities for the candidate models, Bayes' rule can be used to calculate the posterior probability of each model, thereby indicating the most probable model. Such information is useful for reviewing the data to ensure a robust prediction. The final prediction is a weighted combination of the predictions from each model. According to Bayes' rule, the optimal prediction is obtained by weighting the predictions from each model with the model's posterior probability.

2.36.4.3 Pit Depth Data

For pitting corrosion, analysis of inspection data is generally focused on predicting the time at which the deepest pit is expected to penetrate through the complete thickness of the inspected item. Of course, extreme value techniques are naturally suited to this problem. Furthermore, pits in engineering metals and alloys generally initiate at defects, such as nonmetallic inclusions or intermetallic particles, such that the number of potential pit initiation sites is typically very large. It is also well known that many pits will initiate and propagate for short periods, producing only small pits, while a relatively small number of pits may become 'stable' and grow to significant depths from an integrity perspective. For example, experimental data presented by Bhakta and Solomon[63] show pit densities for low carbon steel in the range 100–350 cm^{-2}. Coupled with the resolution of the measuring instrument, this implies that data sets consisting of all practicably measurable pits at a given site may inherently consist of data incorporating a known and (relative to the majority of the initiated pits) large threshold.

2.36.4.3.1 Extreme value analysis for life prediction

Figure 16 shows all pit depths over 6 μm, as found by inspection of two copper-plated 316L stainless steel college roofs in the south of England. The first roof had been exposed for 1523 days and the second for 883 days. The mean exceedance plots for these data gave a slope that was not significantly different from zero after applying the usual least-squares criteria, implying a simple exponential distribution for these exceedances, and hence a Type-I EV for maxima.

Time dependence can be incorporated into the fitted GPD models by assuming the square root law for growth (see other comments in this chapter), which then had exponential type probability density functions

$$f(y) = 2.893t^{-0.5}\exp[-2.893(y-6)t^{-0.5}],$$
$$y > 6 \text{ Roof } 1 \qquad [63]$$

with mean $0.346t^{0.5}$, and

$$f(y) = 1.062t^{-0.5}\exp[-1.062(y-6)t^{-0.5}],$$
$$y > 6 \text{ Roof } 2 \qquad [64]$$

with mean $0.0.941t^{0.5}$.

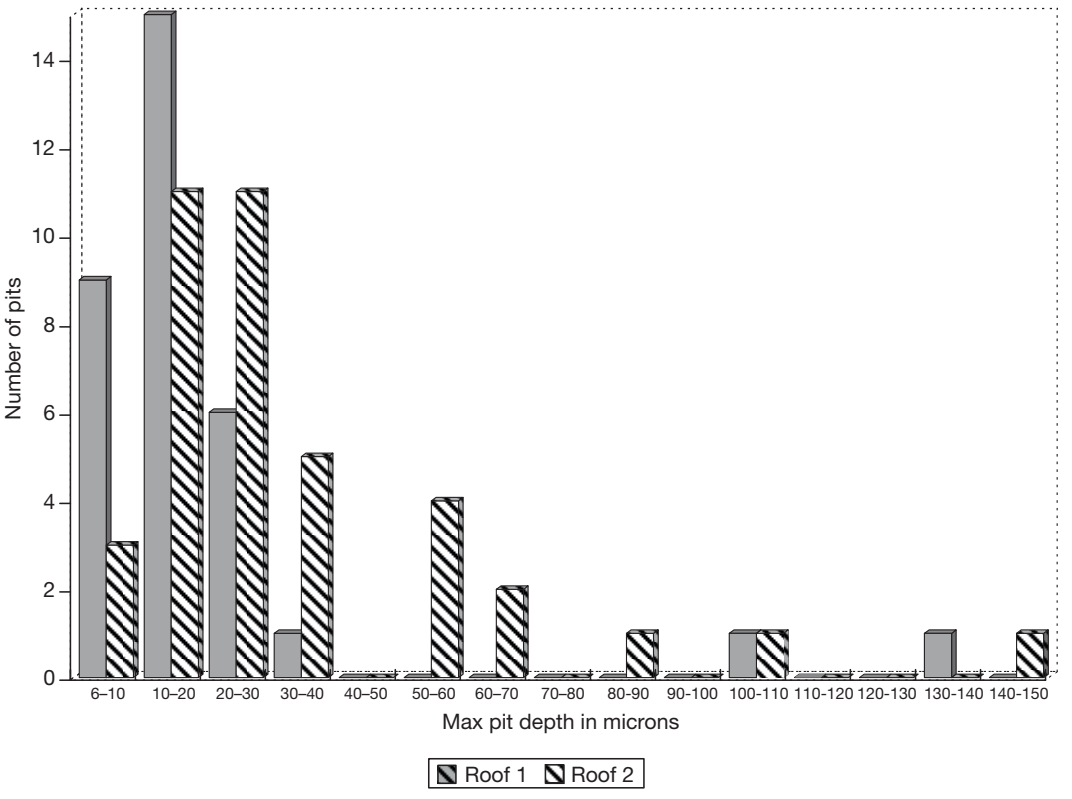


Figure 16 All pit depths over 6 μm, as found by inspection of two copper plated 316L stainless steel college roofs in the south of England. The first roof had been exposed to the weather for 1523 days and the second for 883 days.

The count of exceedances over the specified value will follow a Poisson distribution with mean rates over the whole roof

$$= 4.3A \exp[-2.893(y-6)t^{-0.5}] \text{ for Roof 1}$$
$$= 3.9A \exp[-1.062(y-6)t^{-0.5}] \text{ for Roof 2} \quad [65]$$

where $A = 5\ 000\ 000\ \text{cm}^2$. Given that the roof (cover) thickness is 400 μm, this reaches a mean count of 1 after 12 years for Roof 1, and after 3 years for Roof 2. This analysis gives the expected time to the first penetration for each roof covering.

In the majority of cases, life prediction based on the analysis of historical data involves the key assumption that the environmental conditions in the future can reasonably be considered similar to those in the past. However, even if this assumption is reasonable on average, over long time periods, the environmental conditions usually vary quite significantly over time. For example, Figure 17 shows the variation of the chloride concentration in an open recirculating cooling water system over several years. Despite this variation, which will be not limited to the chloride level but will extend to many other factors such as pH and inhibitor concentration, a Japanese survey of carbon steel heat exchangers in cooling water circuits found that the depth of the deepest pit tends to increase with the square root of time over many years of service.[64] From their survey data, they calculated the variation of EV model parameters over time and used this to produce remaining life estimates for exchanger tube bundles in this type of service. Some further discussion and alternative approaches to this problem are provided elsewhere.[65]

2.36.4.3.2 Life prediction models transformable to linearity

A model with profound implications for the extrapolation of long-term corrosion rates in time is the power-law model $y = at^b$, as reported in Laycock *et al.*,[32] Aziz and Godard,[66] and de la Fuente,[67] and elsewhere, all with $b < 1$ and $b = 0.5$ or 0.67 typical. It is important to note that the corrosion rate, that is, dy/dt, is now time dependent and cannot be reported in isolation as a single number. By assuming a multiplicative error, taking logs of this model produces a straight line regression model, which can be fitted by ordinary least-squares with dependent variable $y' = \log(y)$, intercept $a' = \log(a)$, slope b, and

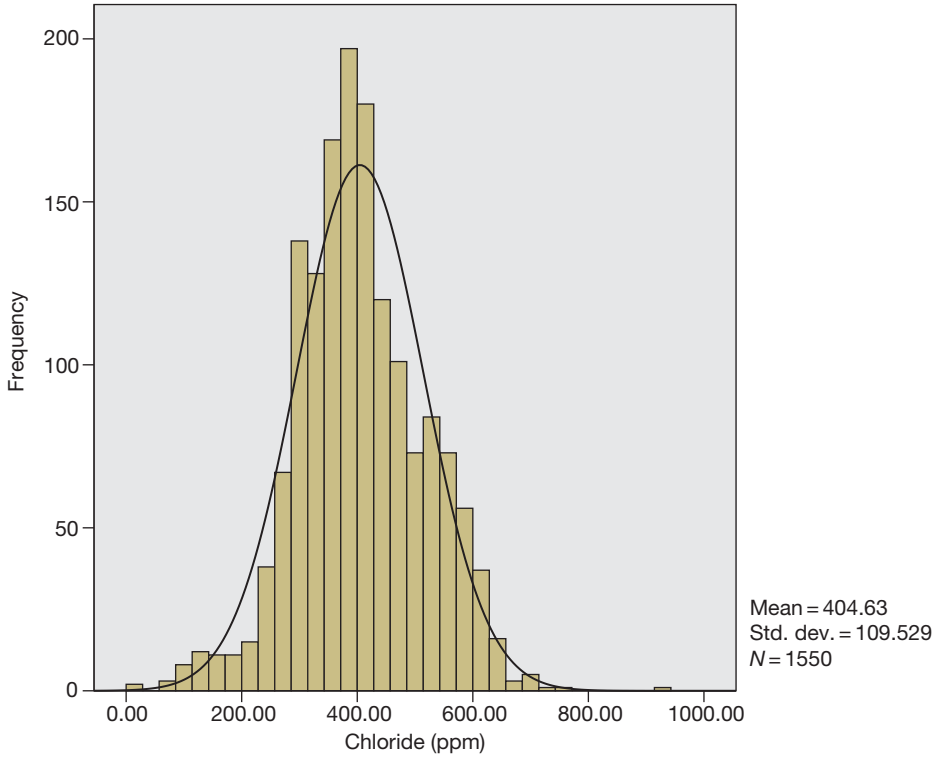


Figure 17 Distribution of measured chloride concentrations (ppm) in an open recirculating cooling water system, obtained from 1550 measurements over 12 years of operation.

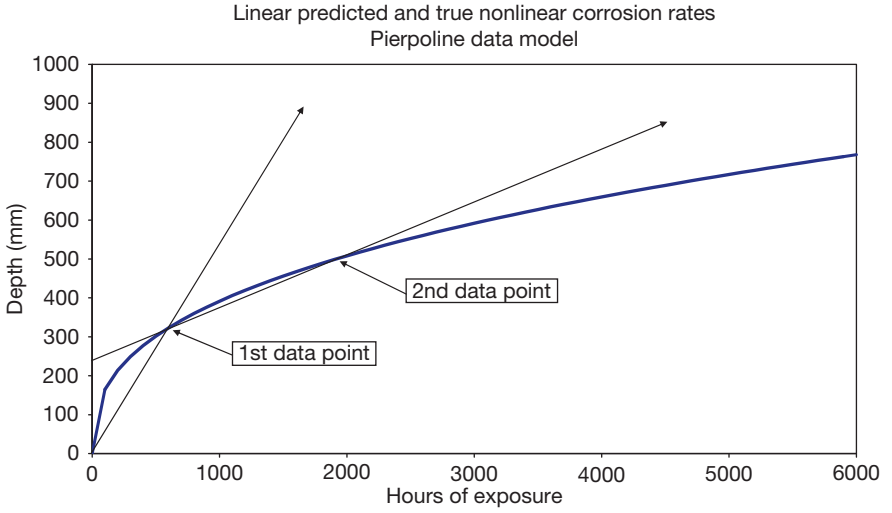


Figure 18 A true power-law growth curve, with decreasing corrosion rate, fitted through the pit depth data of Pierpoline *et al.*,[48] using the method described in Scarf *et al.*[49] This is compared with fitted straight lines based on (a) first data point plus origin; then (b) first and second data points, both extrapolated to show how prediction errors decrease.

covariate $x = \log(t)$, except that b is of course no longer the corrosion rate. Now, suppose that corrosion is progressing according to such a law, as illustrated in Figure 18, but we only have some current data, collected at time $t = t_1$, say. Then, fitting a simple straight line through the origin with these data will produce the straight line superimposed on the curve as shown in Figure 18, and an extrapolated corrosion rate much in excess of future corrosion rates. Now, suppose that we collect some more data, at a later time t_2 and fit a straight line regression through the two time points we now have, as again

illustrated in Figure 18. This time the predicted corrosion rates are much closer to reality for reasonable extrapolations into the future. This suggests that, where it is convenient to have just a single reportable 'corrosion rate,' straight-line regression should be used, with an intercept, utilizing just the latest two or three time points. If more points are available, then the tests for deviations from linearity (as described earlier) can be used to assist in model determination.

2.36.4.4 Intelligent Pigging Data

Pipeline inspections by Intelligent Pigs can produce vast amounts of data, which are typically summarized in lists of 'features' with depths greater than some predefined reporting threshold. When comparing the data from two different inspections of the same pipeline, it is often attempted to match the 'corrosion features' from the two runs and thereby determine a corrosion rate for each feature. However, for various reasons, such as changes in technology over time, it is often difficult to match all the reported features in this way. Here, we describe methods for analysis of both 'tracked' and 'untracked' features.

2.36.4.4.1 Untracked features: Pareto distribution example

Suppose that the results of two pigging runs down a large and actively corroding pipeline are presented as two sets of untracked feature depths $y_{11}, y_{12}, \ldots y_{1n_1}$ at time t_1 and $y_{21}, y_{22}, \ldots, y_{2n_2}$ at time t_2 with $n_2 \geq n_1$. Because the later survey may have new features, the two data mean $\bar{y}_1$ and $\bar{y}_2$ may not provide a sound estimate of the corrosion rate. Indeed, if there are enough fresh features, the second mean could be smaller than the first. Instead, consider the $n < \min(n_2, n_1)$ largest order statistics $y^*_{(11)} > y^*_{(12)} > \cdots > y^*_{(1n)}$ and $y^*_{(21)} > y^*_{(22)} > \cdots > y^*_{(2n)}$ for the two series. Since interest will center on the growth rate of the largest values, it seems sensible to use a mean exceedance plot to decide on a common cutoff point for the two ordered series based on the same count of measurements to maintain rough comparability and equivalence to a fully tracked sequence. A plot of the two fitted GPD distributions for a simulated set of pigging data that has been filtered and processed in this manner can be seen in Figure 19. The modal shift between these two roughly parallel curves is 2.85 mm and the runs were 6 months apart, implying a corrosion rate of 5.7 mm year^{-1}.

2.36.4.4.2 Tracked feature comparison

Figure 20 shows a set of 22 tracked feature depths from two pigging runs of a subsea gas pipeline. The tracked features are all from one pipe spool known to have suffered some internal damage early in service life, many years before these two pigging runs. However, there was also a risk of ongoing internal corrosion due to upsets in dewpoint control. The pigging data are expressed as percentage wall thickness lost, with the difference between the two sets of data converted to an implied wall loss in millimeters – based on a nominal original wall thickness of 15.9 mm. Since 7 of the 22 values are negative, with a maximum absolute value of 1.27 mm, there is substantial measurement error involved. The average

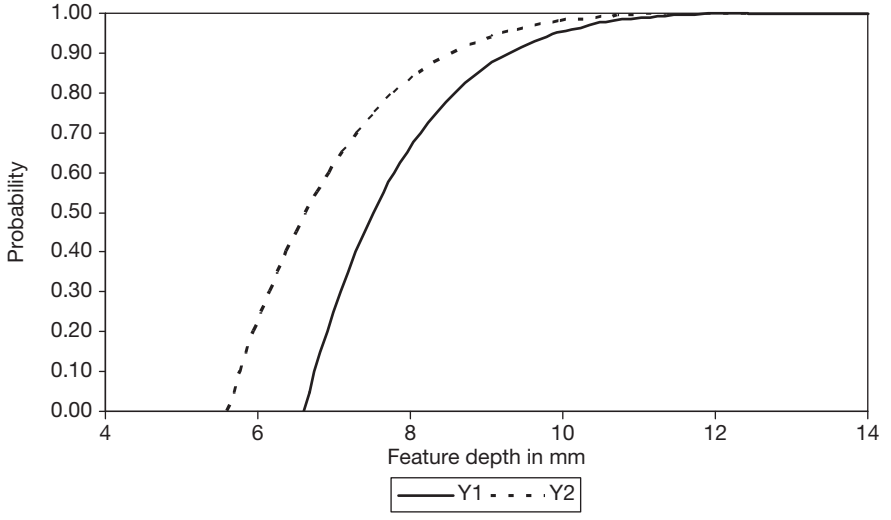


Figure 19 Fitted Pareto distributions for the largest 1500 order statistics from two successive pigging runs (simulated data only). The modal shift between these two roughly parallel curves is 2.85 mm, and the runs were 6 months apart, implying a corrosion rate of 5.7 mm year^{-1}.

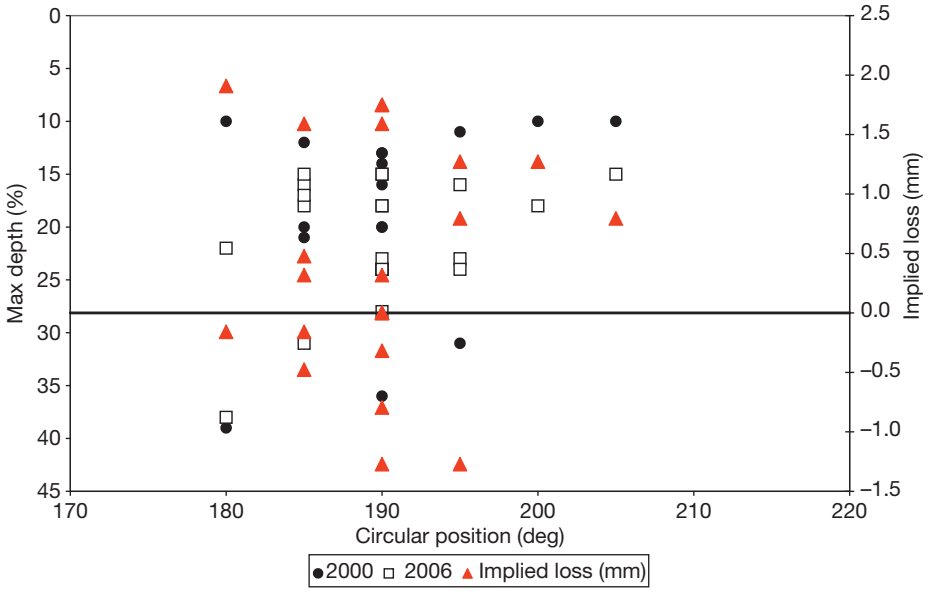


Figure 20 A set of 22 tracked feature depths from two pigging runs of a subsea gas pipeline. These features are all from one pipe spool known to have suffered some internal damage early in service life, many years before these two pigging runs.

implied wall loss is −0.426 mm with a standard error of 0.208 mm (so an implied corrosion rate of −0.065 mm year^{-1} with standard error of 0.032 mm year^{-1}). Note that this average loss is based on all the 22 values. Discarding the seven negative values as 'erroneous' would produce a severely biased estimate of the true corrosion rate – since, without further information, it must be assumed that the unknown measurement errors are as likely to be positive as negative. An inferred tolerance for these measurements of ±1.3 mm implies a measurement error standard deviation of 0.65 mm with variance of 0.4225. The variance of the recorded wall loss data is 0.9520; hence if we assume an additive corrosion process, the implied Wiener noise variance (see **Section 2.36.1**) is $(0.9520 - 2 \times 0.4225) = 0.1070$ over the 2393 days between the pigging runs, or 0.0163 scaled to 1 year – so a standard deviation of 0.13 mm year^{-1}. Since the 95% (one-sided) point from Student's t-distribution on $(22 - 1) = 21$ degrees of freedom is 1.72, and a 95% upper prediction bound on future values for the corrosion rate will be $-0.065 - 1.72 \times \sqrt{(0.032^2 + 0.0163)} = -0.29$ mm year^{-1}, which coincides with the single (so ~5%) maximum observed corrosion rate found in the 22 measured annual corrosion rates.

2.36.4.5 Optimizing Inspection and Repair Intervals

Corrosion damage and cracks detected by nondestructive inspection may be subject to repair. Such maintenance will, in general, need to be carried out within an effective maintenance program.[68–70] Based on assumptions about the rate of corrosion or crack growth, the consequences of failure, the cost of inspection and repair, and data to estimate parameters relating to these quantities, maintenance decision-making can follow. In particular, the optimum time between inspections may be determined.

2.36.4.5.1 *Planning inspections for cracks*

In the simplest case, assume that an inspection is carried out at T time units since the last inspection or failure; inspections are perfect (all crack defects are detected at inspection and repaired) so that they renew the system; on failure, all crack defects are repaired, and the system is renewed; the cost inspection including the repair of crack defects is fixed at C_p; and the cost of failure is C_f. Denote the time to failure by X, then

$$F_X(t) = \Pr(X \le t) = \Pr(Z_{\max,t} > z_f) = 1 - F_{\max,t}(z_f) \quad [66]$$

where $F_{\max,t}(z_f)$ is the distribution of the maximum crack depth at time T. The long-run cost per unit time is given by the renewal-reward theorem[71]:

$$C(T) = E(U)/E(V) \quad [67]$$

where $E(U)$ is the expected cost of an renewal interval, and $E(V) = E\{\min(X, T)\}$ is the expected

length of a renewal interval. The renewal interval is defined as the time between successive renewals. Now

$$E(U) = C_f - (C_f - C_p)F_{\max,T}(z_f) \quad [68]$$

since the cost of a renewal interval is C_p if the renewal interval ends with an inspection that occurs with probability $F_{\max,T}(z_f)$, and C_f if the renewal interval ends with a failure that occurs with probability $1 - F_{\max,T}(z_f)$. Also,

$$E(V) = E\{\min(X, T)\} = \int_0^T \{1 - F_X(t)\}dt = \int_0^T F_{\max,t}(z_f)dt \quad [69]$$

Thus,

$$C(T) = \{C_f - (C_f - C_p)F_{\max,T}(z_f)\} / \int_0^T F_{\max,t}(z_f)dt \quad [70]$$

This is the classic age-based replacement policy of Barlow and Proschan [72]. $C(T)$ is minimized with respect to T to obtain T^*, the cost-optimal age limit for inspection (or more generally, preventive maintenance). For this purpose, $F_{\max,t}(z_f)$, C_p, and C_f must be specified and $C(T)$ minimized. Differentiating eqn [70] with respect to T and setting the derivative equal to zero, we obtain

$$\{f_{\max,T^*}(z_f)/F_{\max,T^*}(z_f)\} \int_0^{T^*} F_{\max,t}(z_f)dt + F_{\max,T^*}(z_f) = C_f/(C_f - C_p) \quad [71]$$

where $f_{\max,t}(z_f)$ is the density function of the maximum crack depth. T^* can be found by solving eqn [71] numerically, or by plotting $C(T)$ against T for various values of T.

If the cost of crack repairs is additional to the cost of inspection, the cracks costing C_r each to repair, then the optimum inspection interval does not depend on C_r, because the expected cost of a renewal interval $E(U)$ increases by $\lambda C_r \times E(V)$, where λ is the rate of crack initiation – a longer inspection interval will result in more cracks developing. Thus, the long-run cost per unit time, $C(T)$, increases by a factor λC_r that does not depend on T.

2.36.4.5.2 Optimizing inspections with unknown cost of failure

Often in practice, it is difficult to specify the cost of failure. Engineers may be happier to prescribe a required reliability. Such a required reliability is best expressed in terms of the operational reliability, which is defined as the distribution of the time between operational failures of the system. Assuming that system renewal intervals are independent of one another, the time of any failure can be taken as time origin, and events up to the first subsequent failure can be considered without loss of generality. Thus, if the time to the first operational failure of the system is denoted by Y_1 and subsequent time to second operational failure by $Y_2, \ldots$, then $Y_1, Y_2, \ldots$ are independent and identically distributed with distribution function $F_Y(y)$, say. It follows that

$$R_Y(y) = R_X(y - nT)\{R_X(T)\}^n, \quad (nT \le y < (n+1)T, n = 0, 1, \ldots) \quad [75]$$

where $R_Y(y) = 1 - F_Y(y)$ is the operational reliability function. This result is given in this form by Lewis.[73] The mean and variance of this distribution can be determined explicitly:

$$E(Y) = \mu_T + T\{1 - F_X(T)\}/F_X(T) = \int_0^T (1 - F_X(x))dx/F_X(T) \quad [76]$$
$$\mathrm{var}(Y) = \sigma_T^2 + T^2\{1 - F_X(T)\}/\{F_X(T)\}^2$$

where $\mu_T = E(X|X < T)$ and $\sigma_T^2 = \mathrm{var}(X|X < T)$. We call $E(Y)$ the mean time between operational failures (MTBOF). For small T, $\mu_T \sim 0$ and $\sigma_T^2 \sim 0$. Therefore, $E(Y) \sim T/F_X(T)$, $\mathrm{var}(Y) \sim \{T/F_X(T)\}^2$, and Y has approximately an exponential distribution with mean $T/F_X(T)$. This approximation is supported by the results of Xie *et al.* [74]. Furthermore,

$$F_X(T)/T = \Pr(Z_{\max,T} > z_f)/T \quad [77]$$

can be interpreted as the operational failure rate.

Thus, given a reliability constraint that is expressed in terms of the MTBOF, the required inspection interval T^* can be obtained from eqn [77] and using the fact that $F_X(t) = 1 - F_{\max,t}(z_f)$. Note that an implied cost of failure, C_f, can be determined by substitution of T^* in eqn [71]. T^* can also be determined by specifying the operational failure rate in eqn [77].

2.36.4.5.3 Crack inspection case study

Maximum crack depths were measured on identical circular welds at locations in two identical plants after varying amounts of operating time (**Figure 21**). A number of welds were subjected to repeated inspection by NDT, but here, only the first inspections are considered. The complete dataset is presented in Scarf *et al.*[75] Where a weld was inspected and no crack found, this was recorded as a zero. At inspection, welds were repaired. It was of interest

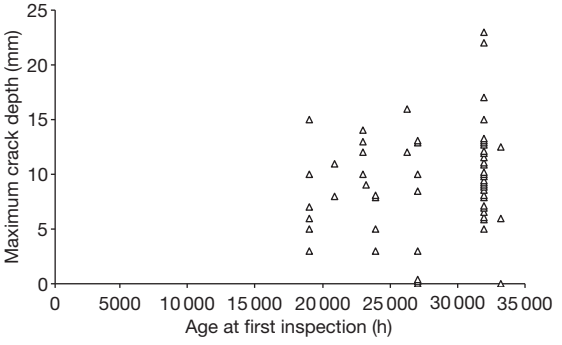


Figure 21 Maximum crack depth against age-at-first-inspection for 59 circular welds from two identical plants – jitter added to depths to distinguish all datapoints.

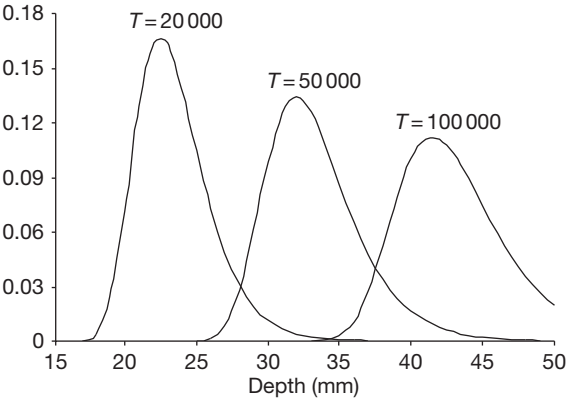


Figure 22 The distribution (probability density) of the maximum crack depth over the system at various times, T, in operating hours.

Table 3 Parameter estimates for the model: Poisson process of crack initiation (rate λ) with distribution of crack depth given by eqn [47]

Parameter	*Estimate*
α_0	7.092
β	1.645
γ	0.00079
λ	0.00013

to the engineers operating the plant to determine if there was some positive trend in crack depth with age (or time since last inspection), and if so, at what rate cracks were 'growing'; and also to determine how often they should carry out costly inspections.

The crack depth data are used here to demonstrate the estimation of the maximum crack depth distribution (eqn [49]). The estimated parameters are shown in **Table 3**. These imply that visible cracks initiate at a rate per weld of one every 7937 operating hours and then subsequently grow at a mean rate given by $\mu = \alpha_0\Gamma(1 + 1/\beta)\tau^{\gamma} = 6.32\tau^{0.000785}$. This growth rate implies that, for example, the mean depth of a crack that has been growing for 10 000 h is 6.37 mm. The estimated growth rate is extremely slow, suggesting that the cracks were in fact not growing at all and were present in the welds from new.

To demonstrate the methods of this section further, the growth rate parameter is taken to be $\gamma = 0.3$, and α_0 is rescaled. Taking $\alpha_0 = 0.4$ implies that the mean depth of a crack that has been growing for 10 000 h is 6.34 mm, and 9.23 mm for one growing for 35 000 h. The distribution of the *maximum* crack depth over the entire system (assuming 60 welds on the system) at time T is then as shown in **Figure 22**.

The probability density is drawn here for various values of T. The mean maximum depth and the standard deviation of the maximum crack depth are increasing as expected.

The probability of observing a maximum crack depth that is at least as big as some specified threshold can be determined. For example, from **Figure 22**, there is a significant risk that a maximum crack depth in a weld is larger than 40 mm after 50 000 h ($P = 0.033$). This probability rises to 0.727 after 100 000 h of operation.

To consider an optimum time between inspections, assuming perfect inspection and repair (weld renewal), the cost of inspection and repair of welds, C_p, can be considered as the unit of cost. Then, the cost of failure, C_f, need only be specified in terms of the cost of inspection and repair. The failure-critical threshold for the maximum crack depth must also be specified. This is taken to be 40 mm. Then, the long-run cost per unit time can be determined for various T. This is shown in **Figure 23(a)** with $C_f = 4C_p$. The optimum inspection interval is ~55 000 h. The sensitivity to C_f and z_f may be investigated simply by determining T^* for various values of C_f and z_f. For example, for a scenario in which the cost of failure is considerably higher than the cost of inspection, say $C_f = 10C_p$, T^* reduces to ~45 000 h. Where an operational reliability is required, this is done most simply by specifying an operational failure rate, eqn [77], and solving for T^* by trial and error. In **Figure 23(b)**, the operational failure rate is plotted against T. Given a required operational reliability, T^* may be read off this graph. Note that this T^* then implies a cost of

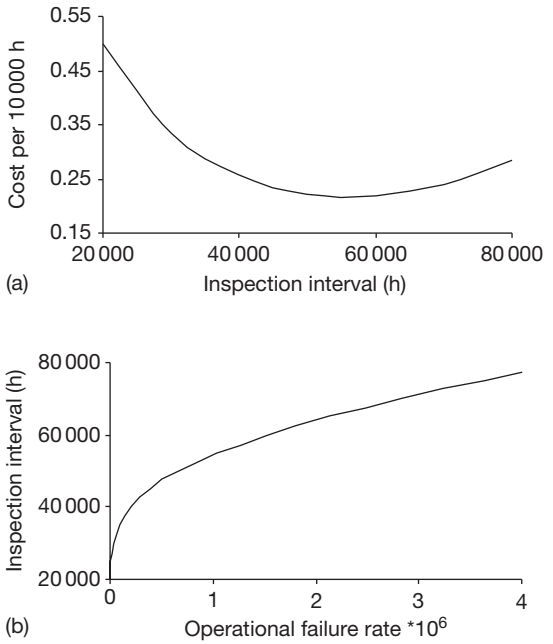


Figure 23 (a) The long-run cost per unit time, $C(T)$, against T, with $C_f = 4C_p$ and $z_f = 40$mm. (b) Time between inspections against operational failure rate, $\Pr(Z_{\max,T} > z_f)/T$ with $z_f = 40$mm.

failure that may be calculated and used to confirm engineering judgment about these quantities. An operational reliability of (at most) one failure for every 5×10^6 operating hours requires that the inspections take place every 40 000 h, which in turn implies that $C_f = 20C_p$ (approximately).

The calculations mentioned earlier may be performed on a spreadsheet. The integral in eqn [49] may be evaluated most simply using Simpson's rule. For the calculation of the long-run cost (eqn [70]), the integral in the denominator can be calculated likewise and noting that $F_{\max,t}(z_f) \approx 1$ for small t (here for $t \leq 20\,000$). These approximations provide sufficient accuracy for optimization purposes, given that the cost function is relatively flat in the region of the optimum.

2.36.5 Corrosion Monitoring

The science of electrochemical noise monitoring began in the 1980s[76,77] and has since extended into commercially available corrosion monitoring techniques. The standard noise analysis techniques are typically built on incrementally additive assumptions for corrosion, as in the theoretical 'shot noise' models for pitting and crevice corrosion proposed by Eden, Cottis, and other workers.[37,38] These result in Poisson models and consequently have variance equal to the mean, which can be approximated by Normal distributions. More generally, statistical analysis methods may ignore the ordering of the measured potential or current values by using sequence-independent parameters such as the mean, standard deviation, skew, and kurtosis, or they may take the sequencing into account by computing the autocorrelation function, power spectral density PSD – ψ_E for potential, or ψ_I for current – or higher-order spectra. When potential and current noise are measured simultaneously, electrochemical noise resistance R_n, electrochemical noise impedance, characteristic charge q, and characteristic frequency f_n can be calculated. The R_n is obtained as the ratio of the standard deviation of potential to the standard deviation of current, justified by appealing to the definition of electrical power for an alternating current. The PSD referred to is that at low frequencies, where the noise is white (hence, the PSD is frequency-independent). Electrochemical noise impedance can be estimated in essentially the same way as the R_n, by dividing the PSD of the potential noise by the PSD of the current noise. Several researchers have analyzed the relationship between R_n and the polarization resistance R_p, and some claim to have proved that the two are equivalent.[78]

2.36.5.1 Threshold Techniques

When this technique is used on a correlated time-series, such as the current noise plot in **Figure 24(a)**, preliminary filtering of the data is required so as to produce a set of statistically independent measured maxima. The most widely adopted method is declustering (see Coles[33], p. 99). In this, an empirical rule is used to define isolated clusters of exceedances, followed by extraction of the maximum within each such cluster. In meteorology, this is sometimes described as 'rainstorm identification'. The empirical spreadsheet formula with a bandwidth of 10 units used to extract the maxima in **Figure 24(b)** from the noise data in **Figure 24(a)** was

```
IF(AND(((MAX(F11:F20)−AVERAGE(F11:F20))<1),
  ((MAX(F10:F20)−AVERAGE(F10:F20))>1)),1,0)
*F20IF(AND((MAX(F11:F20)−AVERAGE
   (F11:F20))<1),
((MAX(F10:F20)−AVERAGE(F10:F20))>1)),1,0)
*F20
```

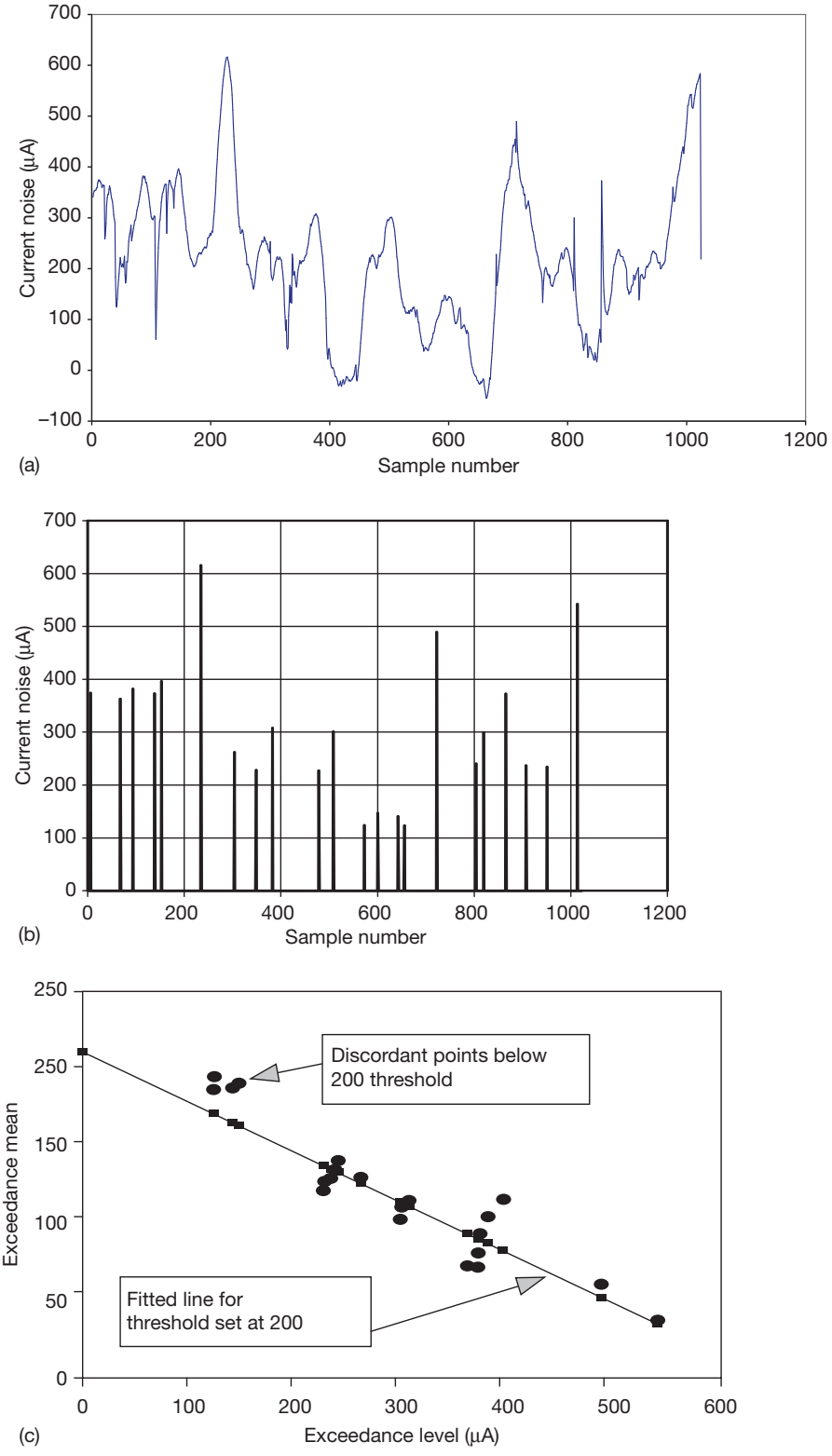


Figure 24 Threshold analysis techniques applied to electrochemical noise measurements. (a) Raw current noise data, (b) extracted peaks, and (c) mean exceedance plot for the current noise measurements, with fitted straight line at selected threshold.

This is for data row 20 updated as usual for other rows. The mean exceedance plot for these data can be seen in Figure 24(c), in which a threshold of 200 μA was used as a cutoff point for fitting the straight line. The statistical package Xtremes, which is supplied with the book by Reiss and Thomas,[56] offers an automated procedure for selecting this threshold.

2.36.6 Summary and Conclusions

Statistical techniques have long been used in both corrosion science and engineering. Rigorous data analysis can help to ensure that the maximum possible benefit is achieved from analysis of corrosion data, be it from laboratory experiments, or plant inspection programs. However, the full range of available methods is not widely used in either field. For corrosion engineering in particular, modern computer technology and readily available statistical software packages present a considerable opportunity for wider use of advanced analysis methods that will make better use of limited inspection data and provide quantitative measures of uncertainty, thereby enabling improved decision making. Key applications include estimating the maximum extent of damage from sample inspections and estimating corrosion rates from wall thickness measurements at different times. While values extrapolated far beyond the sampling region have to be interpreted with care, extreme value analysis provides an established method for carrying out such predictions. Also, when combined with risk-based inspection and reliability engineering methodologies, statistical techniques can be used to guide decisions on the required extent of inspection, especially in relation to piping systems wherein only relatively small sample inspections are practically feasible.

References

1. Evans, U. R.; Mears, R. B.; Queneau, P. E. *Engineering* **1933**, *136*, 689–701.
2. Aziz, P. M. *Corrosion* **1956**, *12*, 495t–506t.
3. Eldredge, G. G. *Corrosion* **1957**, *13*, 51t–60t.
4. Sato, N. *J. Electrochem. Soc.* **1976**, *123*, 1197–1199.
5. Shibata, T.; Takeyama, T. *Corrosion* **1997**, *33*, 243–251.
6. Kowaka, M. Ed. *Introduction to Life Prediction of Industrial Materials: Application of the Extreme Value Statistical Method for Corrosion Analysis*; Allerton Press: New York, 1994. [First published in Japanese in 1984].
7. Gabrielli, C.; Huet, F.; Keddam, M.; Oltra, R. In *NACE-9: Advances in Localised Corrosion*; Isaacs, H., Bertocci, U., Kruger, J., Smialowska, S., Eds.; NACE: Houston, TX, 1990; pp 93–108.
8. Bertocci, U. In *NACE-9: Advances in Localised Corrosion*; Isaacs, H., Bertocci, U., Kruger, J., Smialowska, S., Eds.; NACE: Houston, TX, 1990; pp 127–130.
9. Laycock, P. J.; Scarf, P. A. *Corros. Sci.* **1993**, *35*, 135–145.
10. ASTM G 46-94: Standard Guide for Examination and Evaluation of Pitting Corrosion; American Society for Testing and Materials: West Conshohocken PA, 1994.
11. ASTM G 16-95: Standard Guide for Applying Statistics to Analysis of Corrosion Data; American Society for Testing and Materials: West Conshohocken PA, 1995.
12. Guidelines for Use of Statistics for Analysis of Sample Inspection of Corrosion, Health and Safety Executive, UK, 2001.
13. Williams, D. E.; Westcott, C.; Fleischmann, M. *J. Electrochem. Soc.* **1985**, *132*, 1796–1811.
14. Baroux, B. *Corros. Sci.* **1988**, *28*, 969–986.
15. Laycock, N. J.; Noh, J. S.; White, S. P.; Krouse, D. P. *Corros. Sci.* **2005**, *47*, 3140–3177.
16. Wu, B.; Scully, J. R.; Hudson, J. L.; Mikhailov, A. S. *J. Electrochem. Soc.* **1997**, *144*, 1614–1629.
17. White, S. P.; Krouse, D. P.; Laycock, N. J. *ECS Trans.* **2005**, *1*(16), 37–45.
18. Grimmett, G. *Percolation*, 2nd ed.; Springer-Verlag: Berlin, 1999.
19. Sieradzki, K.; Newman, R. C. *J. Electrochem. Soc* **1986**, *133*, 1979–1980.
20. Qian, S.; Newman, R. C.; Cottis, R. A.; Sieradzki, K. *J. Electrochem. Soc.* **1990**, *137*, 435–439.
21. Wells, D. B.; Stewart, J.; Herbert, A. W.; Scott, P. M.; Williams, D. E. *Corrosion* **1989**, *45*, 649–660.
22. Williams, D. E.; Newman, R. C.; Song, Q.; Kelly, R. G. *Nature* **1991**, *350*, 216–219.
23. Sieradzki, K.; Kim, J. S.; Cole, A. T.; Newman, R. C. *J. Electrochem. Soc.* **1987**, *134*, 1635–1639.
24. Sieradzki, K.; Corderman, R. R.; Shukla, K.; Newman, R. C. *Philos. Mag. A* **1989**, *59*, 713–746.
25. Fishman, G. S. *Monte Carlo: Concepts, Algorithms, and Applications*; Springer Verlag: New York, 1995.
26. http://onlinestatbook.com
27. www.statsoft.com
28. Barnett, V.; Lewis, T. *Outliers in Statistical Data*, 3rd ed.; Wiley Series in Probability & Statistics: New York, 1994.
29. Freund, R. J.; Wilson, W. J. *Regression Analysis*; Academic Press: New York, 1998.
30. Neter, J.; Kutner, M. H.; Wasserman, W.; Nachtsheim, C. J. In *Applied Linear Statistical Models*; 4th ed.; Richard, D., Ed.; Irwin: Illinois, 1996.
31. Laycock, N. J.; Stewart, J.; Newman, R. C. *Corros. Sci.* **1997**, *39*, 1791–1809.
32. Laycock, P. J.; Cottis, R. A.; Scarf, P. A. *J. Electrochem. Soc.* **1990**, *137*, 64–69.
33. Coles, S. *An Introduction to Statistical Modelling of Extreme Values*; Springer-Verlag: London, 2001.
34. Gnedenko, B. V. *Ann. Math.* **1943**, *44*, 423–453.
35. Cottis, R. A.; Laycock, P. J.; Holt, D.; Moir, S. A.; Scarf, P. A. In *NACE-9: Advances in Localised Corrosion*; Isaacs, H., Bertocci, U., Kruger, J., Smialowska, S., Eds.; NACE: Houston, TX, 1990; pp 117–122.
36. Smith, R. L.; Davison, A. C. *J. Roy. Stat. Soc. B.* **1990**, *52*, 393–442.
37. Eden, D. A. *Electrochemical Noise – the First Two Octaves*; CORROSION/98, NACE International: Houston, TX, 1998; Paper no. 386.
38. Cottis, R. A.; Turgoose, S. *Mater. Sci. Forum* **1995**, *192–194*, 663–672.
39. Barnett, V. *Sample Survey: Methods and Principles (Mathematics)*, 3rd ed; Hodder Arnold: London, 2002.

40. Thompson, M. *Theory of Sample Surveys (Monographs on Statistics and Applied Probability)*; Chapman & Hall: London, 1997.
41. Cobb, G. W. *Introduction to Design and Analysis of Experiments*; Springer-Verlag: New York, 1998.
42. Cox, D. R.; Reid, N. *The Theory of the Design of Experiments*; CRC Press: London, 2000.
43. Bolstad, W. M. *Introduction to Bayesian Statistics*, 2nd ed.; Wiley: New York, 2007.
44. Gelman, A.; Carlin, J. B.; Stern, H. S.; Rubin, D. B. *Bayesian Data Analysis*; 2nd ed.; Chapman & Hall/CRC: London, 2003.
45. Bayes, T. *Philos. Trans. R. Soc.* **1763**, 330–418. Reprinted in *Biometrika* **1958**, *45*, 293–315.
46. Laplace, P. S. *Memoires de Mathematique et de Physique* **1774**. Tome Sixieme. English translation by Stigler, S.M. *Statist. Sci.* **1986**, *1*(19), 364–378.
47. Melchers, R. E. *Corros. Sci.* **2007**, *49*, 3149–3167.
48. Pierpoline, R.; White, R. J.; Wong, C.; Cornwell, L.; Griffin, R. In *NACE-9: Advances in Localised Corrosion*; Isaacs, H., Bertocci, U., Kruger, J., Smialowska, S., Eds.; NACE: Houston, TX, 1990; pp 123–126.
49. Scarf, P. A.; Laycock, P. J.; Cottis, R. A. *J. Electrochem. Soc.* **1992**, *9*, 2621–2627.
50. Maddox, S. J. *Int. J. Fract.* **1975**, *11*, 389–408.
51. Hoeppner, D. W.; Krupp, W. E. *Eng. Fract. Mech.* **1974**, *6*, 47–70.
52. Dolinski, K. *Eng. Fract. Mech.* **1992**, *43*, 195–216.
53. To, S.; Lambert, S. B.; Burns, D. J. *Int. J. Fatigue* **1993**, *15*, 333–340.
54. Karlin, H. M.; Taylor, S. *An Introduction to Stochastic Modeling*, revised ed.; Academic Press: San Diego, 1994.
55. Crowder, M. J.; Kimber, A. C.; Smith, R. L.; Sweeting, T. J. *Statistical Analysis of Reliability Data*; Chapman & Hall: London, 1991.
56. Reiss, R. D.; Thomas, M. *Statistical Analysis of Extreme Values with applications to Insurance, Finance, Hydrology and Other Fields*; 2nd ed.; Birkhauser: Berlin, 2001.
57. Efron, B.; Tibshirani, R. J. *An Introduction to the Bootstrap*; Chapman & Hall: London, 1993.
58. Krouse, D.; Laycock, N. J. Proceedings of Corrosion & Protection, Auckland, New Zealand, November 2000; ACA: Melbourne, 2000; Paper # 70.
59. API Publication 510, Pressure Vessel Inspection Code: In-Service Inspection, Rating, Repair, and Alteration, 9th ed.; API: Washington DC, 2006.
60. API Publication 570, Piping Inspection Code: Inspection, Repair, Alteration, and Rerating of In-Service Piping Systems, 2nd ed.; API: Washington DC, 1998.
61. Hoeve, F.; Terpstra, S.; van de Camp, P.; Paleja, R. Proceedings of World Congress on Aset Management, Gold Coast, Australia, 11–14 July, 2006.
62. Wilson, P. T.; Krouse, D. P.; Moss, C. J. Proceedings of Corrosion Control and NDT; ACA, Melbourne, Australia, 23–26 November, 2003; ACA: Melbourne, 2003; Paper # 45.
63. Bhakta, P.; Solomon, A. In *NACE-9: Advances in Localised Corrosion*; Isaacs, H., Bertocci, U., Kruger, J., Smialowska, S., Eds.; NACE: Houston, TX, 1990; pp 445–452.
64. Nakahara, M. Proceedings of the International Symposium on Plant Aging and Life Prediction of Corrodible Structures; NACE: Houston TX, 1997; pp 169–174.
65. Laycock, N. J.; Hodges, S.; Krouse, D.; Keen, D.; Laycock, P. J. Proceedings of Corrosion Science in the 21st Century; UMIST, Manchester, UK, 6–11 July, 2003; *J. Corros. Sci. Eng.* 2003; CO54; www.jcse.org.
66. Aziz, P. M.; Godard, H. P. *Ind. Eng. Chem.* **1952**, 1791–1795.
67. de la Fuente, D.; Otero-Huerta, E.; Morcillo, M. *Corros. Sci.* **2007**, *49*, 3134–3148.
68. Christer, A. H.; Waller, W. M. *J. Opl. Res. Soc.* **1984**, *35*, 401–406.
69. Thoft-Christensen, P.; Sorensen, J. D. *Civil Eng. Sys.* **1987**, *4*, 94–100.
70. Fujimoto, Y.; Swilem, S. A. M. *OMAE Vol II, Safety and Reliability*; Guedes Soares, C., Ed.; ASME: New York, 1992; pp 219–226.
71. Tijms, H. C. *Stochastic Models: An Algorithmic Approach*; Wiley: New York, 1994.
72. Barlow, R.; Proschan, F. *Mathematical Theory of Reliability*; SIAM: Philadelphia, 1996.
73. Lewis, E. E. *Introduction to Reliability Engineering*; Wiley: New York, 1987.
74. Xie, M.; Kong, H.; Goh, T. N. *J. Qual. Maintenance Eng.* **2000**, *6*, 260–268.
75. Scarf, P. A.; Wang, W.; Laycock, P. J. *Reliab. Eng. Syst. Saf.* **1996**, *51*, 331–339.
76. Hladky, K.; Dawson, J. *Corros. Sci.* **1981**, *21*, 317–322.
77. Hladky, K.; Dawson, J. *Corros. Sci.* **1982**, *22*, 231–237.
78. Cottis, R. A. *Corrosion* **2001**, *27*(3), 265–285.

2.37 Introduction to the Modeling of Corrosion

R. A. Cottis
School of Materials, University of Manchester, Manchester M60 1QD, UK

Symbols
V[x,y] Voltage at position *x,y* on a regular 2-D grid

2.37.1 Introduction

For the purposes of this chapter, modeling is defined as the construction of a mathematical-based description of the corrosion process that permits the calculation of expected corrosion behavior for arbitrary conditions within a specified range. While physical models (i.e., practical systems that are expected to have the same behavior as a real corrosion system) have been used (e.g., to study current distribution in cathodic protection), these have received relatively little attention recently. In contrast, the rapidly increasing power of computer hardware and software has led to a huge growth in the use of mathematical models, usually using numerical methods to solve the complex systems of equations that govern most corrosion processes.

2.37.2 Applications of Modeling

There are two distinct reasons for the development of models of corrosion processes:

1. *Prediction of behavior*: Many would expect that the construction of predictive models (i.e., models that can predict corrosion behavior over a range of conditions) would be the most important aspect of corrosion modeling, and this is certainly a common view. For some well-defined problems, such as the distribution of cathodic protection current in a uniform conductive environment such as seawater, predictive models based on 'first-principles' methods can be very accurate and form an important tool in the corrosion engineer's armoury. However, many corrosion processes are very complex and the construction of first-principles predictive models is currently difficult (but see **Chapter 2.38, Modeling of Aqueous Corrosion** and **Chapter 2.39, Predictive Modeling of Corrosion**), and practical predictive models often use relatively simple functions (possibly, but not necessarily, based on theoretically expected dependencies) that are calibrated using measured data. Thus, such models are often concerned more with fitting functions to data than with real understanding of the corrosion process. One such method, the use of neural networks, is discussed in **Chapter 2.40, Neural Network Methods for Corrosion Data Reduction**, while other data-based models may be discussed in the context of the specific corrosion process.
2. *Testing and developing understanding*: While it may not be appreciated by those who have not attempted to develop models, one of the most powerful applications of modeling is to test understanding of mechanisms. One of the fundamental reasons for this is that the construction of a theoretically-based model requires every aspect of a mechanism to be exactly defined and quantified, which is often not done (and may be essentially impossible without constructing a model, due to the complexity of corrosion processes) in the absence of the rigorous analysis required by the modeling process. Then the comparison of the output of the model with

observations provides a challenging test of the precision of understating of the mechanism. Furthermore, the process of creating a model almost inevitably enhances the understanding of the process by those involved in the creation – a challenge for modelers is often to convey their enhanced understanding to others.

2.37.3 Modeling Methods

This section covers a number of modeling methods:

1. Modeling of electrochemical reactions
   Electrochemical kinetic theory is well-developed, and accurate models of the behavior of pure metals in reasonably simple environments can be developed. This is covered in **Chapter 2.38, Modeling of Aqueous Corrosion** and **Chapter 2.39, Predictive Modeling of Corrosion**, which demonstrates the detail required for the construction of accurate first-principles models.
2. Modeling of spatial distribution
   For a number of corrosion processes, such as galvanic corrosion, pitting corrosion, crevice corrosion, and cathodic protection, the challenge is to evaluate the distribution of reaction over the surface of a component. In the simplest cases, such as cathodic protection, it is sufficiently accurate to treat the problem as that of potential and current distribution in a uniform resistive medium. In more challenging cases, such as pitting and crevice corrosion, it is necessary to account for changes in the chemistry of the solution and the consequent changes in reaction kinetics. In this case, the techniques in the previous section become important, although most work in this area tends to use rather simplified models of interfacial chemistry. Such models have been developed for many years, but the advent of general-purpose finite element programs has made these methods much more accessible. **Chapter 2.39, Predictive Modeling of Corrosion** provides examples of the state-of-the art in such modeling, but does not go into the details of the principles underlying the spatial aspects of the models, which are described in outline in the following section.
3. Data-based modeling
   Owing to the complexity of the interactions between metallurgical structure, environment and mechanical stress, many corrosion processes are not yet amenable to theoretically-based modeling, and predictive models must therefore be based on measured data. Such models are then highly dependent on both the quality and the quantity of the measured data on which they are based. We can define the 'problem domain' for a particular set of measured data as being a region in a multidimensional space where each dimension represents one of the variables that have a significant effect on the behavior. For accurate models, it is important that the measured data provide an adequate coverage of the problem domain and that the measured data have well-defined values for all of the dimensions of the multidimensional space. Unfortunately, this is rarely achieved, especially where real service data are concerned. The modeling of the data then becomes a problem of fitting the 'best' function to the measured values. For small data sets, it is usually necessary to make some simplifying assumptions about the behavior, such as assuming linear dependencies and neglecting interactions between variables. If sufficient information is available, more complex functions can be fitted, such as higher-order polynomials. Neural network methods (introduced in **Chapter 2.40, Neural Network Methods for Corrosion Data Reduction**) provide a particularly interesting approach, whereby the functional relationship is effectively determined by the fitting process, rather than being defined in advance.

2.37.4 Approaches to Modeling Spatial and Time Distribution

When we are concerned with the spatial distribution of corrosion (e.g., when attempting to model crevice corrosion or pitting corrosion), either for the final steady state or for incorporating the evolution over time, the problem can almost invariably be expressed as requiring the solution of a set of partial differential equations over the space and time concerned (the 'problem domain') with specified boundary conditions. Occasionally, it is possible to obtain an analytical solution, but usually only in highly simplified cases. Consequently, it is normally necessary to use a numerical method to obtain a solution. As a generalization, many of the numerical methods aim at converting the differential equations into a finite (but usually large) set of simultaneous difference equations. These can then be solved by conventional methods for the solution of simultaneous equations (essentially these write the simultaneous equations as a matrix equation that can be solved by matrix inversion). There are three approaches to this.

2.37.4.1 Finite Difference Methods

In the finite difference method, the problem domain is covered by a regular grid of points, the objective then being to determine the values corresponding to the solution to the system of differential equations at the grid points. These will be points with a constant spacing along a line for a 1-dimensional problem, crossing points of a square grid for a 2-dimensional problem and points on a cubic grid for a 3-dimensional problem. (Some finite difference methods use grid spacings that vary according to a regular pattern, and these may offer advantages for electrochemical problems.[1]) Then the solution can be approximated by replacing the differential equations by the corresponding difference equations. For the detailed mathematics of the process (and an indication of the large variety of methods that have been developed to obtain a solution), the readers are referred to Britz,[1] but we can illustrate the method by a simplified explanation.

We shall consider one of the simplest classes of problems, that of determining the steady-state potential distribution in a uniform 2-dimensional conductor (e.g., a conductive sheet). If we consider a grid point somewhere in a 2-dimensional array (**Figure 1**), it is easy to see that the best estimate of the true value of the potential at a point is the average of the corresponding potentials in the four nearest neighbors. Thus, if $V[x,y]$ is the potential at point x, y, we can write

$$V[x,y] = (V[x-1,y] + V[x+1,y] + V[x,y-1] + V[x,y+1])/4$$

Edge and corner points and internal boundaries will obviously have to be handled specially (and this is where the boundary conditions will be introduced). Writing the appropriate equations for all grid points leads to a set of simultaneous equations, which can be solved by matrix inversion or by a range of alternative techniques (iterative solution methods are usually used, because the arrangement of grid points in the problem domain is normally very regular; these effectively implement the matrix inversion in a very efficient manner by taking advantage of the regularity). The method has the advantage of conceptual simplicity and relatively easy implementation without specialist software. However, it does not handle complex geometries very well and it is inefficient for problems where small regions need to be analyzed in detail in a large problem domain (since this requires a very large number of very closely spaced points).

2.37.4.2 Finite Element Method

The limitations of the finite difference method are, to an extent, a result of the requirement to use an array of elements of constant size and shape. The finite element method provides the ability to use elements of arbitrary (and variable) size and shape, although at the expense of more complex mathematics; we shall not attempt to explain the mathematical basis here (see the Comsol web site[2] for extensive documentation on the method and its application to electrochemical problems). Fortunately several programs, both commercial and open-source, are available to handle finite element problems. Most of these are targeted to well-defined problems such as stress analysis and the analysis of heat flow in uniform media. Heat flow is slightly more complex than current flow, due to the heat capacity of the material, but the steady-state temperature distribution will be identical to the steady-state voltage distribution. Because of the greater generality and greater practical significance, program literature will usually refer to the analysis of thermal rather than electrical problems.

Some programs[3] are able to handle problems involving multiple physical processes (such as diffusion, migration, and reaction) and provide a simple

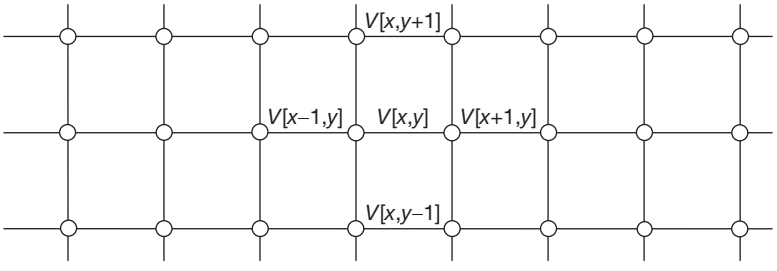


Figure 1 2-Dimensional grid for finite difference analysis of potential distribution.

way to analyze problems associated with the distribution of corrosion processes.

2.37.4.3 Boundary Element Method

Both the finite difference and finite element methods need to fill the problem domain with elements, and consequently, it is often necessary to use very large numbers of elements, with the result that very large matrices need to be inverted (the matrix size required is approximately proportional to the square of the number of elements). The boundary element method is based on analyzing only elements on the boundaries of the problem domain. This leads to a major reduction in the number of elements required, and hence the size of matrix that needs to be inverted. However, the computational benefits are not quite as significant as this would suggest, as the matrix that needs to be inverted is usually much more completely filled than for the other methods, and the inversion is consequently much slower for given matrix dimensions. Furthermore, the boundary element method can only be used for a subset of possible characteristics within the problem domain; fortunately the electrical behavior of a uniform conductive material can be handled. Hence this method is popular for the solution to current distribution problems in cathodic protection. A subsidiary benefit of the method is that it can handle problem domains with no exterior boundary (i.e., the conductive medium can extend to infinity), which is particularly useful for cathodic protection problems. It is beyond the scope of this article to describe the mathematical basis of the boundary element method and the readers are referred to appropriate texts.[4] As with the finite element method, commercial software is available to implement the boundary element method.

References

1. Britz, D. *Digital Simulation in Electrochemistry*, 3rd ed.; Springer: Berlin, 2005.
2. Pepper, D. W.; Heinrich, J. C. *The Finite Element Method: Basic Concepts and Applications*; Hemisphere Publishing: New York, 1992.
3. http://www.comsol.com/ (viewed 5 March 2009).
4. Beer, G. *Programming the Boundary Element Method: An Introduction for Engineers*; Wiley Blackwell, 2002.

2.38 Modeling of Aqueous Corrosion

A. Anderko
OLI Systems Inc., 108 American Road, Morris Plains, NJ 07950, USA

Abbreviations

CCT Critical crevice temperature
CR Corrosion rate
HKF Helgeson–Kirkham–Flowers equation of state
Me Metal
MSA Mean spherical approximation
NRTL Nonrandom two-liquid (equation)
Ox Oxidized form
Re Reduced form
SHE Standard hydrogen electrode
UNIFAC Universal functional activity coefficient (equation)
UNIQUAC Universal quasi-chemical (equation)

Symbols

a_i Activity of species i
A Surface area
A_{ij} Surface interaction coefficient for species i and j
b Tafel coefficient using decimal logarithms

$c_{i,b}$ Bulk molar concentration of species *i*
$c_{i,s}$ Surface molar concentration of species *i*
C_p Heat capacity at constant pressure
d Characteristic dimension
D_i Diffusion coefficientsta of species *i*
D_t Turbulent diffusion coefficient
e^- Electron
E Potential
E_b Passivity breakdown potential
E_{corr} Corrosion potential
E_{crit} Critical potential for localized corrosion
E_{rp} Repassivation potential
E_0 Equilibrium potential
f Friction factor
f_i Fugacity of species *i*
F Faraday constant
G^{ex} Excess Gibbs energy
i Current density
i_a Anodic current density
$i_{a,ct}$ Charge-transfer contribution to the anodic current density
$i_{a,L}$ Limiting anodic current density
i_c Cathodic current density
$i_{c,ct}$ Charge-transfer contribution to the cathodic current density
$i_{c,L}$ Limiting cathodic current density
i_{corr} Corrosion current density
i_p Passive current density
i_{rp} Current density limit for measuring repassivation potential
i_0 Exchange current density
i^* Concentration-independent coefficient in expressions for exchange current density
J_i Flux of species *i*
k_{ads} Adsorption rate constant
k_{des} Desorption rate constant
k_i Reaction rate constant for reaction *i*
$k_{m,i}$ Mass transfer coefficient for species *i*
K Equilibrium constant
K_{ads} Adsorption equilibrium constant
K_{sp} Solubility product
l_i Reaction rate for *i*th reaction
m_i Molality of species *i*
m_0 Standard molality of species *i*
n_i Number of moles of species or electrons
Nu Nusselt number
Pr Prandtl number
R Gas constant
R_k Rate of production or depletion of species *k*
Re Reynolds number
S Supersaturation
Sc Schmidt number
Sh Sherwood number
t Time
T Temperature
u_i Mobility of species *i*
v_i Rate of reaction *i*
V Linear velocity
z Direction perpendicular to the surface
z_i Charge of species *i*
α Thermal diffusivity
α_i Electrochemical transfer coefficient for species *i*
β Tafel coefficient using natural logarithms
γ_i Activity coefficient of species *i*
δ_i Nernst layer thickness for species *i*
$\Delta G_{ads,i}$ Gibbs energy of adsorption for species *i*
$\Delta H^{\neq}$ Enthalpy of activation
$\Delta\Phi$ Potential drop
ϵ Dielectric permittivity
η Dynamic viscosity
θ_i Coverage fraction of species *i*
θ_P Coverage fraction of passive layer
λ Thermal conductivity
μ_i Chemical potential of species *i*
μ_i^0 Standard chemical potential
$\bar{\mu}_i$ Electrochemical potential of species *i*
ν Kinematic viscosity
ν_i Stoichiometric coefficient of species *i*
ρ Density
Φ Electrical potential
ξ_i Electrochemical transfer coefficient for reaction *i*
ω Rotation rate
∇ Vector differential operator
$\equiv$**X** Surface species *X*

2.38.1 Introduction

Aqueous corrosion is an extremely complex physical phenomenon that depends on a multitude of factors including the metallurgy of the corroding metal, the chemistry of the corrosion-inducing aqueous phase, the presence of other – solid, gaseous, or nonaqueous liquid – phases, environmental constraints such as temperature and pressure, fluid flow characteristics, methods of fabrication, geometrical factors, and construction features. This inherent complexity makes the development of realistic physical models very challenging and, at the same time, provides a strong incentive for the development of practical models to understand the corrosion phenomena, and to assist in their mitigation. The need for tools for simulating aqueous corrosion has been recognized in various industries including oil and gas production and

transmission, oil refining, nuclear and fossil power generation, chemical processing, infrastructure maintenance, hazardous waste management, and so on. The past three decades have witnessed the development of increasingly sophisticated modeling tools, which has been made possible by the synergistic combination of improved understanding of corrosion mechanisms and rapid evolution of computational tools.

Corrosion modeling is an interdisciplinary undertaking that requires input from electrolyte thermodynamics, surface electrochemistry, fluid flow and mass transport modeling, and metallurgy. In this chapter, we put particular emphasis on corrosion chemistry by focusing on modeling both the bulk environment and the reactions at the corroding interface. The models that are reviewed in this chapter are intended to answer the following questions:

1. What are the aqueous and solid species that give rise to corrosion in a particular system? What are their thermophysical properties, and what phase behavior can be expected in the system? These questions can be answered by thermodynamic models.
2. What are the reactions that are responsible for corrosion at the interface? How are they influenced by the bulk solution chemistry and by flow conditions? How can passivity and formation of solid corrosion products be related to environmental conditions? How can the interfacial phenomena be related to observable corrosion rates? These questions belong to the realm of electrochemical kinetics and mass transport models.
3. What conditions need to be satisfied for the initiation and long-term occurrence of localized corrosion? This question can be answered by electrochemical models of localized corrosion.

These models can be further used as a foundation for larger-scale models for the spatial and temporal evolution of systems and engineering structures subject to localized and general corrosion. Also, they can be combined with probabilistic and expert system-type models of corrosion. Models of such kinds are, however, outside the scope of this review, and will be discussed in companion chapters.

2.38.2 Thermodynamic Modeling of Aqueous Corrosion

Historically, the first comprehensive approach to modeling aqueous corrosion was introduced by Pourbaix in the 1950s and 1960s on the basis of purely thermodynamic considerations.[1] Pourbaix[1] developed the *E*–pH stability diagrams, which indicate which phases are stable on a two-dimensional plane as a function of the potential and pH. The potential and pH were originally selected because they play a key role as independent variables in electrochemical corrosion. Just as importantly, they made it possible to construct the stability diagrams in a semianalytical way, which was crucial before the advent of computer calculations. Over the past four decades, great progress has been achieved in the thermodynamics of electrolyte systems, in particular for concentrated, mixed-solvent, and high temperature systems. These advances made it possible to improve the accuracy of the stability diagrams and, at the same time, increased the flexibility of thermodynamic analysis so that it can go well beyond the *E*–pH plane.

The basic objective of the thermodynamics of corrosion is to predict the conditions at which a given metal may react with a given environment, leading to the formation of dissolved ions or solid reaction products. Thermodynamics can predict the properties of the system in equilibrium or, if equilibrium is not achieved, it can predict the direction in which the system will move towards an equilibrium state. Thermodynamics does not provide any information on how rapidly the system will approach equilibrium, and, therefore, it cannot give the rate of corrosion.

The general condition of thermodynamic equilibrium is the equality of the electrochemical potential, $\bar{\mu}_i$ in coexisting phases,[2] that is,

$$\bar{\mu}_i = \mu_i + z_i F\phi = \mu_i^0 + RT \ln a_i + z_i F\phi \qquad [1]$$

where μ_i is the chemical potential of species i; μ_i^0, its standard chemical potential; a_i, the activity of the species; z_i, its charge; F, the Faraday constant; and ϕ is the electrical potential. The standard chemical potential is a function of the temperature and, secondarily, pressure. The activity depends on the temperature and solution composition and, to a lesser extent, on pressure. The activity is typically defined in terms of solution molality m_i,

$$a_i = (m_i/m^0)\gamma_i \qquad [2]$$

where m^0 is the standard molality unit (1 mol kg^{-1} H_2O), and γ_i is the activity coefficient. It should be noted that the molality basis for species activity becomes inconvenient for concentrated solutions because molality diverges to infinity as the concentration increases to the pure solute limit. Therefore, the mole fraction basis is more generally applicable for calculating activities.[3] Nevertheless, molality remains the most common concentration unit for aqueous systems and is used here for illustrative purposes.

Computation of μ_i^0 and γ_i is the central subject in electrolyte thermodynamics. Numerous methods, with various ranges of applicability, have been developed over the past several decades for their computation. Several comprehensive reviews of the available models are available (Zemaitis *et al.*,[4] Renon,[5] Pitzer,[6] Rafal *et al.*,[7] Loehe and Donohue,[8] Anderko *et al.*[3]). In the next section, the current status of modeling μ_i^0 and γ_i is outlined as it applies to the thermodynamics of corrosion.

2.38.2.1 Computation of Standard-State Chemical Potentials

The computation of the standard chemical potential μ_i^0 requires the knowledge of thermochemical data including

1. Gibbs energy of formation of species i at reference conditions (298.15 K and 1 bar).
2. Entropy or, alternatively, enthalpy of formation at reference conditions.
3. Heat capacity and volume as a function of temperature and pressure.

For numerous species, these values are available in various compilations (Chase *et al.*,[9] Barin and Platzki,[10] Cox *et al.*,[11] Glushko *et al.*,[12] Gurvich *et al.*,[13] Kelly,[14] Robie *et al.*,[15] Shock and Helgeson,[16] Shock *et al.*,[17,18] Stull *et al.*,[19] Wagman *et al.*,[20] and others). In general, thermochemical data are most abundant at near-ambient conditions, and their availability becomes more limited at elevated temperatures.

In the case of individual solid species, the chemical potential can be computed directly from tabulated thermochemical properties according to the standard thermodynamics.[2] In the case of ions and aqueous neutral species, the thermochemical properties listed above are standard partial molar properties rather than the properties of pure components. The standard partial molar properties are defined at infinite dilution in water. The temperature and pressure dependence of the partial molar heat capacity and the volume of ions and neutral aqueous species are quite complex because they are manifestations of the solvation of species, which is influenced by electrostatic and structural factors. Therefore, the computation of these quantities requires a realistic physical model.

An early approach to calculating the chemical potential of aqueous species as a function of temperature is the entropy correspondence principle of Criss and Cobble.[21] In this approach, heat capacities of various types of ions were correlated with the reference-state entropies of ions, thus making it possible to predict the temperature dependence of the standard chemical potential.

A comprehensive methodology for calculating the standard chemical potential was developed by Helgeson *et al.* (Helgeson *et al.*,[22] Tanger and Helgeson[23]). This methodology is based on a semi-empirical treatment of ion solvation, and results in an equation of state for the temperature and pressure dependence of the standard molal heat capacities and volumes. Subsequently, the heat capacities and volumes are used to arrive at a comprehensive equation of state for standard molal Gibbs energy and, hence, the standard chemical potential. The method is referred to as the HKF (Helgeson–Kirkham–Flowers) equation of state. An important advantage of the HKF equation is the availability of its parameters for a large number of ionic and neutral species (Shock and Helgeson,[16] Shock *et al.*,[18] Sverjensky *et al.*[24]). Also, correlations exist for the estimation of the parameters for species for which little experimental information is available. The HKF equation of state has been implemented both in publicly available codes (Johnson *et al.*[25]) and in commercial programs. A different equation of state for standard-state properties has been developed on the basis of fluctuation solution theory (Sedlbauer *et al.*,[26] Sedlbauer and Majer[27]). This equation offers improvement over HKF for nonionic solutes and in the near-critical region. However, the HKF remains as the most widely accepted model for ionic solutes.

2.38.2.2 Computation of Activity Coefficients

In real solutions, the activity coefficients of species deviate from unity because of a variety of ionic interaction phenomena, including long-range Coulombic interactions, specific ion–ion interactions, solvation phenomena, and short-range interactions between uncharged and charged species. Therefore, a practically-oriented activity coefficient model must represent a certain compromise between physical reality and computational expediency.

The treatment of solution chemistry is a particularly important feature of an electrolyte model. Here, the term 'solution chemistry' encompasses ionic dissociation, ion pair formation, hydrolysis of metal ions, formation of metal–ligand complexes, acid–base reactions, and so on. The available electrolyte models can be grouped in three classes:

1. models that treat electrolytes on an undissociated basis,
2. models that assume complete dissociation of all electrolytes into constituent ions, and

3. speciation-based models, which explicitly treat the solution chemistry.

The models that treat electrolytes as undissociated components are analogous to nonelectrolyte mixture models. They are particularly suitable for supercritical and high temperature systems, in which ion pairs predominate. Although this approach may also be used for phase equilibrium computations at moderate conditions (e.g., Kolker *et al.*[28]), it is not suitable for corrosion modeling because it ignores the existence of ions. The models that assume complete dissociation are the largest class of models for electrolytes at typical conditions. Compared with the models that treat electrolytes as undissociated or completely dissociated, the speciation-based models are more computationally demanding because of the need to solve multiple reactions and phase equilibria. Another fundamental difficulty associated with the use of speciation models lies in the need to define and characterize the species that are likely to exist in the system. In many cases, individual species can be clearly defined and experimentally verified in relatively dilute solutions. At high concentrations, the chemical identity of individual species (e.g., ion pairs or complexes) becomes ambiguous because a given ion has multiple neighbors of opposite sign, and, thus, many species lose their distinct chemical character. Therefore, the application of speciation models to concentrated solutions requires a careful analysis to separate the chemical effects from physical nonideality effects.

It should be noted that, as long as only phase equilibrium computations are of interest, comparable results could be obtained with models that belong to various classes. For example, the overall activity coefficients and vapor–liquid equilibria of many transition metal halide solutions, which show appreciable complexation, can be reasonably reproduced using Pitzer's[29] ion-interaction approach without taking speciation into account. However, it is important to include speciation effects for modeling the thermodynamics of aqueous corrosion. This is due to the fact that the presence of individual hydrolyzed forms, aqueous complexes, and so on is often crucial for the dissolution of metals and metal oxides. It should be noted that activity coefficients of individual species are different in fully speciated models than in models that treat speciation in a simplified way. Therefore, it is important to use activity coefficients that have been determined in a fully consistent way, that is, by assuming the appropriate chemical species in the solution.

The theory of liquid-phase nonideality is well-established for dilute solutions. A limiting law for activity coefficients was developed by Debye and Hückel[30] by considering the long-range electrostatic interactions of ions in a dielectric continuum. The Debye–Hückel theory predicts the activity coefficients as a function of the ionic charge and dielectric constant and density of the solvent. It reflects only electrostatic effects and, therefore, excludes all specific ionic interactions. Therefore, its range of applicability is limited to ~0.01 M for typical systems. Several modifications of the Debye–Hückel theory have been proposed over the past several decades. The most successful modification was developed by Pitzer[29] who considered hard-core effects on electrostatic interactions. A more comprehensive treatment of the long-range electrostatic interactions can be obtained from the mean-spherical approximation (MSA) theory,[31,32] which provides a semianalytical solution for ions of different sizes in a dielectric continuum. The MSA theory results in a better prediction of the long-range contribution to activity coefficients at somewhat higher electrolyte concentrations.

The long-range electrostatic term provides a baseline for constructing models that are valid for electrolytes at concentrations that are important in practice. In most practically-oriented electrolyte models, the solution nonideality is defined by the excess Gibbs energy G^{ex}. The excess Gibbs energy is calculated as a sum of the long-range term and one or more terms that represent ion–ion, ion–molecule, and molecule–molecule interactions:

$$G^{ex} = G^{ex}_{long\text{-}range} + G^{ex}_{specific} + \cdots \quad [3]$$

where the long-range contribution is usually calculated either from the Debye–Hückel or the MSA theory, and the specific interaction term(s) represent(s) all other interactions in an electrolyte solution. Subsequently, the activity coefficients are calculated according to standard thermodynamics[2] as

$$\ln \gamma_i = \frac{1}{RT}\left(\frac{\partial G^{ex}}{\partial n_i}\right)_{T,P,n_{j\neq i}} \quad [4]$$

Table 1 lists a number of activity coefficient models that have been proposed in the literature, and shows the nature of the specific interaction terms that have been adopted. In general, these models can be subdivided into two classes:

1. models for aqueous systems; in special cases, such models can also be used for other solvents as long as the system contains a single solvent;
2. mixed-solvent electrolyte models, which allow multiple solvents as well as multiple solutes.

The aqueous electrolyte models incorporate various ion interaction terms, which are usually defined

Table 1 Summary of representative models for calculating activity coefficients in electrolyte systems

Reference	*Terms*	*Features*
Aqueous (or single-solvent) electrolyte models		
Debye and Hückel[30]	Long-range	Limiting law valid for very dilute solutions
Guggenheim[33,34]	Long-range + ion interaction	Simple ion interaction term to extend the Debye–Hückel limiting law; applicable to fairly dilute solutions
Helgeson[35]	Long-range + ion interaction	Ion interaction term to extend the limiting law; applicable to fairly dilute solutions
Pitzer[29]	Long-range + ion interaction	Ionic strength-dependent virial coefficient-type ion interaction term; revised Debye–Hückel limiting law; applicable to moderately concentrated solutions (~6 m)
Bromley[36]	Long range + ion interaction	Ionic strength dependent ion interaction term; applicable to moderately concentrated solutions (~6 m)
Zemaitis[37]	Long-range + ion interaction	Modification of the model of Bromley[36] to increase applicability range with respect to ionic strength
Meissner[38]	One-parameter correlation as a function of ionic strength	Generalized correlation to calculate activities based on a limited amount of experimental information
Pitzer and Simonson,[39] Clegg and Pitzer[40]	Long-range + ion interaction	Mole fraction-based expansion used for the ion interaction term; applicable to concentrated systems up to the fused salt limit
Mixed-solvent electrolyte models		
Chen *et al*.[41]	Long-range + short-range	Local-composition (NRTL) short-range term
Liu and Watanasiri[42]	Long-range + short-range + electrostatic solvation (Born) + ion interaction	Modification of the model of Chen *et al*.[41] using a Guggenheim-type ion interaction term for systems with two liquid phases
Abovsky *et al*.[43]	Long-range + short-range	Modification of the model of Chen *et al*.[41] using concentration-dependent NRTL parameters to extend applicability range with respect to electrolyte concentration
Chen *et al*.[44]	Long-range + short-range + ion hydration	Modification of the model of Chen *et al*.[41] using an analytical ion hydration term to extend applicability with respect to electrolyte concentration
Chen and Song[45]	Long-range + short-range + electrostatic solvation (Born)	Modification of the model of Chen *et al*.[41] by introducing segment interactions for organic molecules
Sander *et al*.[46]	Long-range + short-range	Local composition model (UNIQUAC) with concentration-dependent parameters used for short-range term
Macedo *et al*.[47]	Long-range + short-range	Local composition model (UNIQUAC) with concentration-dependent parameters used for short-range term
Kikic *et al*.[48]	Long-range + short range	Local composition group contribution model (UNIFAC) used for short-range term
Dahl and Macedo[49]	Short-range	Undissociated basis; no long-range contribution; group contribution model (UNIFAC) used for short-range term
Iliuta *et al*.[50]	Long-range + short-range	Local composition (UNIQUAC) model used for short-range term
Wu and Lee[31]	Extended long-range (MSA)	Mean-spherical approximation (MSA) theory used for the long-range term
Li *et al*.[51]	Long range + ion interaction + short range	Virial-type form used for the ion interaction term; local composition used for short-range term
Yan *et al*.[52]	Long-range + ion interaction + short range	Group contribution models used for both the ion interaction term and short-range term (UNIFAC)
Zerres and Prausnitz[53]	Long-range + short range + ion solvation	Van Laar model used for short-range term; stepwise ion solvation model
Kolker *et al*.[28]	Short-range	Undissociated basis; no long-range contribution
Wang *et al*.[54,55]	Long range + ion interaction + short range	Ionic strength-dependent virial expansion-type ion interaction term; local composition (UNIQUAC) short-range term; detailed treatment of solution chemistry
Papaiconomou *et al*.[32]	Extended long-range (MSA) + short-range	MSA theory used for the long-range term; local composition model (NRTL) used for the short-range term

in the form of virial-type expansions in terms of molality or mole fractions (see Table 1). Among these models, the Pitzer[29] molality-based model has found wide acceptance. Parameters of the Pitzer[29] model are available in the open literature for a large number of systems.[6]

The mixed-solvent electrolyte models are designed to handle a wider variety of chemistries. They invariably use the mole fraction and concentration scales. A common approach in the construction of mixed-solvent models is to use local-composition models for representing short-range interactions. The well-known local-composition models include NRTL, UNIQUAC, and its group-contribution version, UNIFAC (see Prausnitz *et al.*[56] and Malanowski and Anderko[57] for a review of these models). The local composition models are commonly used for nonelectrolyte mixtures and, therefore, it is natural to use them for short-range interactions in electrolyte systems. The combination of the long-range and local-composition terms is typically sufficient for representing the properties of moderately concentrated electrolytes in any combination of solvents. For systems that may reach very high concentrations with respect to electrolyte components (e.g., up to the fused salt limit), more complex approaches have been developed. One viable approach is to explicitly account for hydration and solvation equilibria in addition to using the long-range and short-range local composition terms (Zerres and Prausnitz,[53] Chen *et al.*[44]). A particularly effective approach is based on combining virial-type ion interaction terms with local composition models (Li *et al.*,[51] Yan *et al.*,[52] Wang *et al.*[54,55]). In such combined models, the local-composition term reflects the nonelectrolyte-like short-range interactions, whereas the virial-type ion interaction term represents primarily the specific ion–ion interactions that are not accounted for by the long-range contribution. These and other approaches are summarized in Table 1. Among the models summarized in Table 1, the models of Pitzer,[29] Zemaitis,[37] Chen *et al.*[41] (including their later modifications),[44,45] and Wang *et al.*[54,55] have been implemented in publicly available or commercial simulation programs.

2.38.2.3 Electrochemical Stability Diagrams

The *E*–pH diagrams, commonly referred to as the Pourbaix[1] diagrams, are historically the first, and remain the most important class of electrochemical stability diagrams. They were originally constructed for ideal solutions (i.e., on the assumption that $\gamma_i = 1$), which was the only viable approach at the time when they were introduced.[1] The essence of the procedure for generating the Pourbaix diagrams is analyzing all possible reactions between all – aqueous or solid – species that may exist in the system. The simultaneous analysis of the reactions makes it possible to determine the ranges of potential and pH at which a given species is stable. The reactions can be conveniently subdivided into two classes, that is, chemical and electrochemical reactions. The chemical reactions can be written without electrons, that is

$$\sum \nu_i M_i = 0 \qquad [5]$$

Then, the equilibrium condition for the reaction is given in terms of the chemical potentials of individual species by

$$\sum \nu_i \mu_i = 0 \qquad [6]$$

According to eqn [1], eqn [6] can be further rewritten in terms of species activities as

$$\sum \nu_i \ln a_i = -\frac{\sum \nu_i \mu_i^0}{RT} = \ln K \qquad [7]$$

where the right-hand side of eqn [7] is defined as the equilibrium constant because it does not depend on species concentrations.

In contrast to the chemical reactions, the electrochemical reactions involve electrons, e^-, as well as chemical substances M_i, that is

$$\sum \nu_i M_i + n e^- = 0 \qquad [8]$$

The equilibrium state of an electrochemical reaction is associated with a certain equilibrium potential. Since electrode potential cannot be measured on an absolute basis, it is necessary to choose an arbitrary scale against which the potentials can be calculated. If a reference electrode is selected, the equilibrium state of the reaction that takes place on the reference electrode is given by an equation analogous to eqn [8], that is

$$\sum \nu_{i,\mathrm{ref}} M_{i,\mathrm{ref}} + n e^- = 0 \qquad [9]$$

Then, the equilibrium potential E_0 of eqn [8] is given with respect to the reference electrode as

$$E_0 - E_{0,\mathrm{ref}} = \frac{\sum \nu_i \mu_i - \sum \nu_{I,\mathrm{ref}} \mu_{i,\mathrm{ref}}}{nF} \qquad [10]$$

According to a generally used convention, the standard hydrogen electrode (SHE) (i.e., $H^+(a_{H^+} = 1)/H_2(f_{H_2} = 1)$) is used as a reference. The corresponding reaction that takes place on the SHE is given by

$$H^+ - \frac{1}{2} H_2 + e^- = 0 \qquad [11]$$

Then, eqn [10] becomes

$$E_0 - E_{0,\text{ref}} = \frac{\sum \nu_i \mu_i - n\left(\mu^0_{H^+} - \frac{1}{2}\mu^0_{H_2}\right)}{nF} \quad [12]$$

In eqn [12], the standard chemical potentials μ^0_H and $\mu^0_{H_2}$ as well as the reference potential $E_{0,\text{ref}}$ are equal to zero at $T = 298.15$ K. For practical calculations at temperatures other than 298.15 K, two conventions can be used for the reference electrode. According to a universal convention established by the International Union of Pure and Applied Chemistry (the 'Stockholm convention'), the potential of SHE is arbitrarily defined as zero at all temperatures (i.e., $E_{0,\text{ref}} = 0$). When this convention is employed, eqn [12] is used as a working equation with $E_{0,\text{ref}} = 0$. In an alternate convention, the SHE reference potential is equal to zero only at room temperature, and its value at other temperatures depends on the actual, temperature-dependent values of $\mu^0_{H^+}$ and $\mu^0_{H_2}$. In this case, it is straightforward to show (see Chen and Aral,[58] Chen *et al.*,[59]) that eqn [12] becomes

$$E_0 = \frac{\sum \nu_i \mu_i}{nF} \quad [13]$$

at all temperatures. Equation [13] can be further rewritten in terms of activities as

$$E_0 = \frac{\sum \nu_i \mu^0_i}{nF} + \frac{RT}{nF}\sum \nu_i \ln a_i = E^0_0 + \frac{RT}{nF}\sum \nu_i \ln a_i \quad [14]$$

where E^0_0 is the standard equilibrium potential, which is calculated from the values of the standard chemical potentials μ^0_i.

For a brief outline of the essence of the Pourbaix diagrams, let us consider a generic reaction in which two species, A and B, undergo a transformation. The only other species that participate in the reaction are hydrogen ions and water, that is:

$$aA + cH_2O + ne^- = bB + mH^+ \quad [15]$$

If eqn [15] is a chemical reaction (i.e., if $n = 0$), then its equilibrium condition (eqn [7]) can be rewritten as

$$m\text{pH} = \log\left(\frac{a^a_B}{a^a_A}\right) - \log K - \log a^c_{H_2O} \quad [16]$$

where $\text{pH} = -\log a_{H^+}$, and decimal rather than natural logarithms is used. For dilute solutions, it is appropriate to assume that $a_{H_2O} = 1$, and the last term on the right-hand side of eqn [16] vanishes. If we assume that the components A and B are aqueous dissolved species and their activities are equal, eqn [16] defines the boundary between the predominance areas of species A and B. If one of the species (A or B) is a pure solid and the other is an aqueous species, the activity of the solid is equal to one and the activity of the aqueous species can be set equal to a certain predetermined, typically small, value (e.g., 10^{-6}). Then, eqn [16] represents the boundary between a solid and an aqueous species at a fixed value of the dissolved species activity. Similarly, for a boundary between two pure solid phases, the activities of the species A and B are equal to one. Such a boundary is represented by a vertical line in a potential–pH space, and its location depends on the equilibrium constant according to eqn [16].

If eqn [15] represents an electrochemical reaction (i.e., $n \neq 0$), the equilibrium condition (eqn [14]) becomes

$$E_0 = E^0_0 + \frac{RT}{nF}\ln\frac{a^a_A}{a^b_B} + \frac{RT}{nF}\ln a_{H_2O} + \frac{RT\ln 10}{F}\frac{m}{n}\text{pH} \quad [17]$$

As with the chemical reactions, the term that involves the activity of water vanishes for dilute solutions, and the ratio of the activities of species A and B can be fixed to represent the boundary between the predominance areas of two species. Under such assumptions, the plot of E versus pH is a straight line with a slope determined by the stoichiometric coefficients m and n.

Thus, for each species, boundaries can be established using eqns [16] and [17]. As long as the simplifying assumptions described above are met, the boundaries can be calculated analytically. By considering all possible boundaries, stability regions can be determined using an algorithm described by Pourbaix.[1] Sample E–pH diagrams are shown for iron in **Figure 1**.

One of the main reasons for the usefulness of stability diagrams is the fact that they can illustrate the interplay of various partial processes of oxidation and reduction. The classical E–pH diagrams contain two dashed lines, labeled as 'a' and 'b,' which are superimposed on the diagram of a given metal (see **Figure 1**). The dashed line 'a' represents the conditions of equilibrium between water (or hydrogen ions) and elemental hydrogen at unit hydrogen fugacity, that is

$$H^+ + e^- = 0.5H_2 \text{ (line 'a')} \quad [18]$$

The 'b' line describes the equilibrium between water and oxygen, also at unit fugacity,

$$O_2 + 4H^+ + 4e^- = 2H_2O \text{ (line 'b')} \quad [19]$$

Accordingly, water will be reduced to form hydrogen at potentials below line 'a,' and will be oxidized to form oxygen at potentials above line 'b.' The location

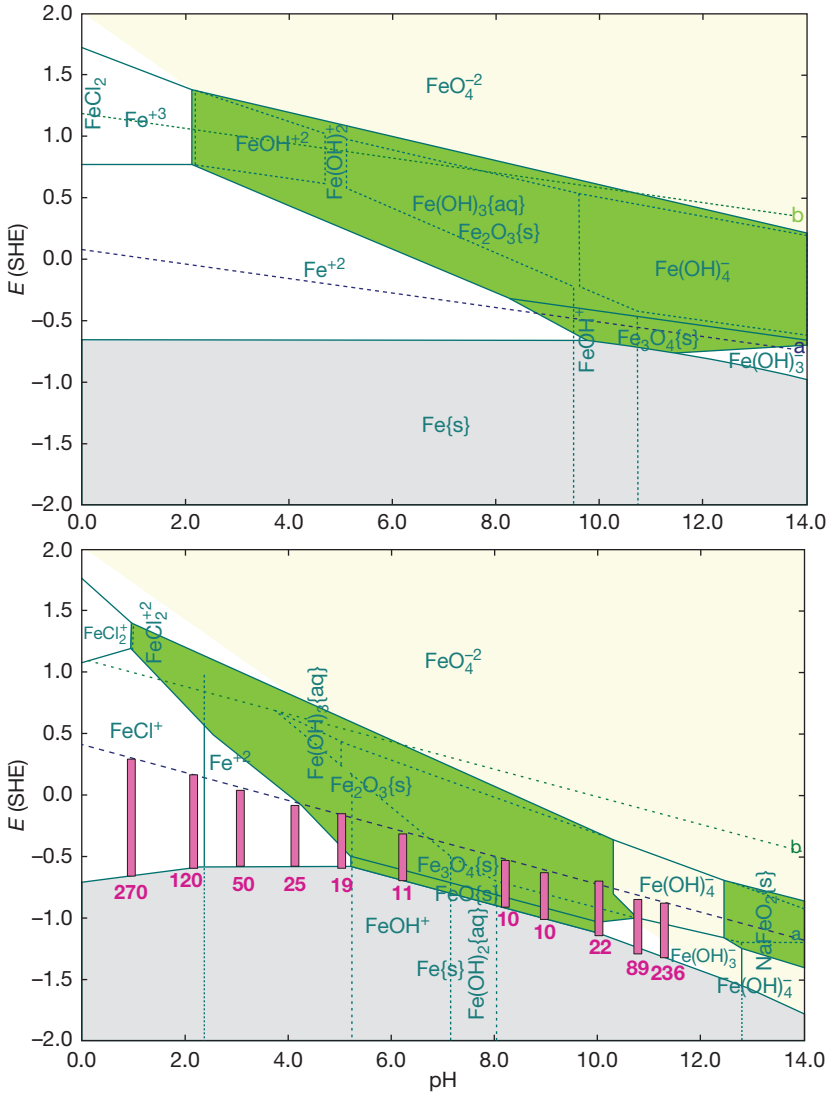


Figure 1 *E*–pH (Pourbaix) diagrams for iron at 25 °C (upper diagram) and 300 °C (lower diagram). At 300 °C, the conditions of the experiments of Partridge and Hall[60] are superimposed on the diagram. The vertical bars show the range between the equilibrium potentials for the reduction of H^+ and oxidation of Fe, thus bracketing the mixed potential in the experiments. The numbers under the bars denote the experimentally determined relative attack. The diagrams have been generated using the Corrosion Analyzer software[61] using the algorithm of Anderko *et al*.[62]

of the 'a' and 'b' lines on the diagram indicates whether a given redox couple is thermodynamically possible. For example, in the region between the line 'a' and the upper edge of the stability region of $Fe_{(s)}$ in Figure 1, the anodic reaction of iron oxidation can be coupled with the cathodic reaction of water or hydrogen ion reduction. In such a case, the measurable open-circuit potential of the corrosion process will establish itself between the line 'a' and the equilibrium potential for the oxidation of iron (i.e., the upper edge of the $Fe_{(s)}$ region). If the potential lies above line 'a,' then water reduction is no longer a viable cathodic process, and the oxidation of iron must be coupled with another reduction process. In the presence of oxygen, reaction eqn [19] can provide such a reaction process. In such a case, the open-circuit (corrosion) potential will establish itself at a higher value for which the upper limit will be defined by line 'b.'

Pourbaix[1] subdivided various regions of the *E*–pH diagrams into three categories, that is, immunity, corrosion, and passivation. The immunity region

encompasses the stability field of elemental metals. The corrosion region corresponds to the stability of dissolved, either ionic or neutral species. Finally, passivation denotes the region in which solid oxides or hydroxides are stable. In **Figure 1**, the immunity and passivation regions are shaded, whereas the corrosion region is not. It should be noted that this classification does not necessarily reflect the actual corrosion behavior of a metal. Only immunity has a strict significance in terms of thermodynamics because in this region the metal cannot corrode regardless of the time of exposure. The stability of dissolved species in the 'corrosion' region does not necessarily mean that the metal rapidly corrodes in this area. In reality, the rate of corrosion in this region may vary markedly because of kinetic reasons. Passivation is also an intrinsically kinetic phenomenon because the protectiveness of a solid layer on the surface of a metal is determined not by its low solubility alone. The presence of a sparingly soluble solid is typically a necessary, but not sufficient condition for passivity.

Although the *E*–pH diagrams indicate only the thermodynamic tendency for the stability of various metals, ions, and solid compounds, they may still provide useful qualitative clues as to the expected trends in corrosion rates. This is illustrated in the lower diagram of **Figure 1**. In this diagram, the vertical bars denote the difference between the equilibrium potentials for the reduction of water and oxidation of iron. Thus, the bars indicate the tendency of the metal to corrode in deaerated aqueous solutions with varying pH. They bracket the location of the corrosion potential and, thus, indicate whether the corrosion potential will establish itself in the 'corrosion' or 'passivation' regions. The numbers associated with the bars represent the experimentally determined relative attack. It is clear that the observed relative attack is substantially greater in the regions where a solid phase is predicted to be stable than in the regions where no solid phase is predicted. Thus, subject to the limitations discussed above, the stability diagrams can be used for the qualitative assessment of the tendency of metals to corrode, and for estimating the range of the corrosion potential.

Following the pioneering work of Pourbaix and his coworkers, further refinements of stability diagrams were made to extend their range of applicability. These refinements were made possible by the progress of the thermodynamics of electrolyte solutions and alloys. Specifically, further developments focused on

1. generation of diagrams at elevated temperatures,
2. taking into account the active solution species other than protons and water molecules,
3. introduction of solution nonideality, which influences the stability of species through realistically modeled activity coefficients,
4. introduction of alloying effects by accounting for the formation of mixed oxides and the nonideality of alloy components in the solid phase, and
5. flexible selection of independent variables, other than E and pH, for the generation of diagrams.

2.38.2.3.1 Diagrams at elevated temperatures

The key to the construction of stability diagrams at elevated temperatures is the calculation of the standard chemical potentials μ_i^0 of all individual species as a function of temperature. In earlier studies, the entropy correspondence principle of Criss and Cobble[21] was used for this purpose. Macdonald and Cragnolino[63] reviewed the development of *E*–pH diagrams at elevated temperatures until the late 1980s.

The HKF equation of state[22,23] formed a comprehensive basis for the development of *E*–pH diagrams at temperatures up to 300 °C for iron, zinc, chromium, nickel, copper, and other metals (Beverskog and Puigdomenech,[64–69] Anderko *et al.*[62]) An example of a high temperature *E*–pH diagram is shown for Fe at 300 °C in the lower diagram of **Figure 1**. Comparison of the Fe diagrams at room temperature and 300 °C shows a shift in the predominance domains of cations to lower pH values and an expansion of the domains of metal oxyanions at higher pH values. Such effects are relatively common for metal–water systems.

2.38.2.3.2 Effect of multiple active species

The concept of stability diagrams can be easily extended to solutions that contain multiple chemically active species other than H^+ and H_2O. In such a general case, the simple reaction eqn [15] needs to be extended as

$$v_X X + \sum_{i=1}^{k} v_i A_i = Y + v_e e^- \qquad [20]$$

where the species X and Y contain at least one common element and $A_i\,(i = 1, \ldots, k)$ are the basis species that are necessary to define equilibrium equations between all species containing a given element. Reaction eqn [20] is normalized so that the stoichiometric coefficient for the right-hand side species (Y) is equal to 1. Such an extension results in generalized

expressions for the boundaries between predominance regions (eqns [16] and [17]). The equilibrium expression for the chemical reactions (eqn [20] with $v_e = 0$) then becomes

$$\ln K = \left(\ln a_Y - v_X \ln a_X - \sum_{i=1}^{k} v_i \ln a_{A_i} \right) = \frac{1}{RT} \left(\mu_Y^0 - \mu_X^0 - \sum_{i=1}^{k} v_i \mu_{A_i}^0 \right) \quad [21]$$

and the expression for an electrochemical reaction (eqn [20] with $v_e \neq 0$) takes the form:

$$E_0 = E_0^0 + \frac{RT}{Fv_e} \left(\ln a_Y - v_X \ln a_X - \sum_{i=1}^{k} v_i \ln a_{A_i} \right) \quad [22]$$

Thus, the expression for the boundary lines become more complicated but the algorithm for generating the diagrams remains the same, that is, the predominance areas can still be determined semianalytically.

The strongest effect of solution species on the stability diagrams of metals is observed in the case of complex-forming ligands and species that form stable, sparingly soluble solid phases other than oxides or hydroxides (e.g., sulfides or carbonates). Several authors focused on stability diagrams for metals such as iron, nickel, or copper in systems containing sulfur species (Biernat and Robbins,[70] Froning *et al.*,[71] Macdonald and Syrett,[72] Macdonald *et al.*,[73,74] Chen and Aral,[58] Chen *et al.*,[59] Anderko and Shuler[75]). This is due to the importance of iron sulfide phases, which have a strong tendency to form in aqueous environments even at very low concentrations of dissolved hydrogen sulfide. The stability domains of various iron sulfides can be clearly rationalized using *E*–pH diagrams. Diagrams have also been developed for metals in brines (Pourbaix,[76] Macdonald and Syrett,[72] Macdonald *et al.*,[73,74] Kesavan *et al.*,[77] Muñoz-Portero *et al.*[78]). The presence of halide ions manifests itself in the stability of various metal–halide complexes. Typically, the effect of halides on the thermodynamic stability is less pronounced than the effect of sulfides. However, concentrated brines can substantially shrink the stability regions of metal oxides and promote the active behavior of metals. An example of such effects is provided by the diagrams for copper in concentrated bromide brines (Muñoz-Portero *et al.*[78]) Effects of formation of various carbonate and sulfate phases have also been reported (Bianchi and Longhi,[79] Pourbaix[76]).

The formation of complexes of metal ions with organic ligands (e.g., chelants) frequently leads to a significant decrease in the stability of metal oxides, which may, under some conditions, indicate an increased dissolution tendency in the passive state (Silverman,[80,81] Kubal and Panacek,[82] Silverman and Silverman[83]). An important example of the importance of complexation is provided by the behavior of copper and other metals in ammoniated environments. **Figure 2** (upper diagram) illustrates an *E*–pH diagram for copper in a 0.2 m NH_3 solution. As shown in the figure, the copper oxide stability field is bisected by the stability area of an aqueous complex. The formation of a stable dissolved complex indicates that the passivity of copper and copper-base alloys may be adversely affected in weakly alkaline NH_3-containing environments. In reality, ammonia attack on copper-base alloys is observed in steam cycle environments.[63] Stability diagrams are a useful tool for the qualitative evaluation of the tendency of metals to corrode in such environments.

2.38.2.3.3 Diagrams for nonideal solutions

As long as the solution is assumed to be ideal (i.e., $\gamma_i = 1$ in eqn [2]), the chemical and electrochemical equilibrium expressions can be written in an analytical form, and the *E*–pH diagrams can be generated semianalytically. For nonideal solutions, the analytical character of the equilibrium lines can be preserved if fixed activity coefficients are assumed for each species (see Bianchi and Longhi[79]). However, such an approach does not have a general character because the activities of species change as a function of pH. This is due to the fact that, in real systems, pH changes result from varying concentrations of acids and bases, which influence the activity coefficients of all solution species. In a nonideal solution, the activities of all species are inextricably linked to each other because they all are obtained by differentiating the solution's excess Gibbs energy with respect to the number of moles of the individual species.[2] Therefore, in a general case, the equilibrium expressions (eqns [21] and [22]) can no longer be expressed by analytical expressions. This necessitates a modification of the algorithm for generating stability diagrams. A general methodology for constructing stability diagrams of nonideal solutions has been developed by Anderko *et al.*[62]

In general, the nonideality of aqueous solutions may shift the location of the equilibrium lines because the activity coefficients may vary by one or even two orders of magnitude. Such effects become pronounced in concentrated electrolyte solutions and in mixed-solvent solutions, in which water is not necessarily the predominant solvent.

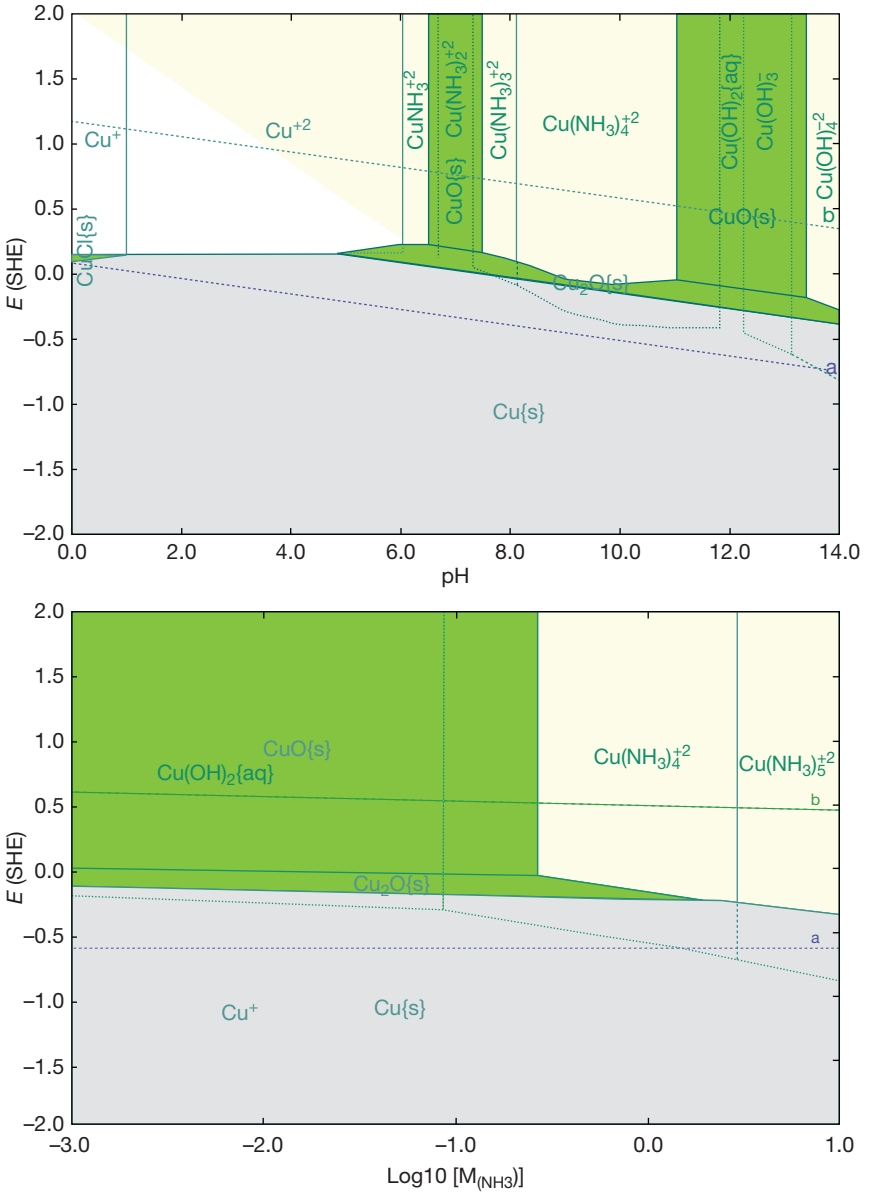


Figure 2 Use of thermodynamic stability diagrams to analyze the effect of ammonia on the corrosion of copper. The upper plot is an *E*–pH diagram for Cu in a 0.2 m NH_3 solution. The lower plot is a potential–ammonia molality diagram.

2.38.2.3.4 Diagrams for alloys

The vast majority of the published stability diagrams have been developed for pure metals. However, several studies have been devoted to generating stability diagrams for alloys, particularly those from the Fe–Ni–Cr–Mo family (Cubicciotti,[84,85] Beverskog and Puigdomenech,[69] Yang *et al.*,[86] Anderko *et al.*[87]). In general, a stability diagram for an alloy is a superposition of partial diagrams for the individual components of the alloy.[69] However, the partial diagrams are not independent because the alloy components form a solid solution and, therefore, their properties are linked. The superposition of partial diagrams makes it possible to analyze the tendency of alloy components for preferential dissolution. This may be the case when a diagram indicates that one alloy component has a tendency to form a passivating oxide, whereas another component has a tendency to form ions. Also, stability diagrams indicate in a simple way which alloy component is more anodic.

There are two key effects of alloying on thermodynamic equilibrium, that is, the formation of mixed solid corrosion products and the nonideality of alloy components in the solid phase. The formation of mixed solid corrosion products may be particularly significant for passive systems. Fe–Ni–Cr alloys may form mixed oxides, including $NiFe_2O_4$, $FeCr_2O_4$, and $NiCr_2O_4$. Such oxides may be more stable than the single oxides of chromium, nickel, and iron, which appear in stability diagrams for pure metals. *E*–pH diagrams that include mixed oxide phases have been reported by Cubicciotti,[84,85] Beverskog and Puigdomenech,[69] and Yang *et al.*[86] for Fe–Ni–Cr alloys in ideal aqueous solutions.

Another fundamental effect of alloying on the thermodynamic behavior results from the nonideality of the solid solution phase. In contrast to pure metals, the activity of a metal in an alloy is no longer equal to 1. This affects the value of the equilibrium potential for metal dissolution, which determines the upper boundary of the stability field of an alloy component on an *E*–pH diagram. As with the liquid phase, the activity of solid solution components can be obtained from a comprehensive excess Gibbs energy model of the solid phase. Thermodynamic modeling of alloys is a very wide area of research, which is beyond the scope of this chapter (see Lupis[88] and Saunders and Miodownik[89] for reviews). Detailed models have been developed for the thermodynamic behavior of alloys on the basis of high temperature data that are relevant to metallurgical processes. These models can be, in principle, extrapolated to lower temperatures that are of interest in aqueous corrosion. Despite the inherent uncertainties associated with extrapolation, estimates of activities of alloy components can be obtained in this way and incorporated into the calculation of stability diagrams. Such an approach was used by Anderko *et al.*[87] in a study in which solid solution models were coupled with an algorithm for generating stability diagrams for Fe–Ni–Cr–Mo–C and Cu–Ni alloys. In that work, the solid solution models of Hertzman,[90] Hertzman and Jarl,[91] and Anderson and Lange[92] were implemented to estimate the activities of alloy components.

It should be noted that the effect of varying activities of alloy components on *E*–pH diagrams is relatively limited. This is due to the fact that alloy phase nonideality, while crucial for modeling alloy metallurgy, is only a secondary contribution to the energetics of reactions between metals and aqueous species. The effect of alloy solution nonideality manifests itself in a limited shift of the upper boundary of the metal stability area in *E*–pH diagrams. The effect of mixed oxide phases is usually much more pronounced on stability diagrams.

2.38.2.3.5 Potential–concentration diagrams

Once the thermodynamic treatment of solution chemistry is extended to allow for the effect of species other than H^+ and H_2O (see eqns [21] and [22]), other independent variables become possible in addition to E and pH. The chemical and electrochemical boundaries can then be calculated as a function of concentration variables other than pH. In particular, the stability diagrams can be generated as a function of the concentration of active solution species that may react with metals through precipitation, complexation, or other reactions. Thus, an extension of *E*–pH diagrams to *E*–species concentration diagrams is relatively straightforward. The algorithm developed by Anderko *et al.*[62] for nonideal solutions is equally applicable to *E*–pH and *E*–concentration diagrams. An example of such a diagram is provided in the lower plot of **Figure 2**. This plot is an *E*–NH_3 molality diagram and illustrates the thermodynamic behavior of copper when increasing amounts of ammonia are added to an aqueous environment. In this simulation, pH varies with ammonia concentration. However, pH is less important in this case than the concentration of NH_3 because the stability of copper oxide is controlled by the formation of copper–ammonia complexes. The diagram indicates at which concentration of ammonia the copper oxide film becomes thermodynamically unstable, thus increasing the tendency for metal dissolution.

2.38.2.4 Chemical Equilibrium Computations

While the *E*–concentration diagrams are merely an extension of Pourbaix's original concept of stability diagrams, a completely different kind of thermodynamic analysis can also be performed using electrolyte thermodynamics. This kind of analysis relies on solving the chemical equilibrium expressions without using the potential as an independent variable. Assuming that temperature, pressure, and starting amounts of components (both metals and solution species) are known, the equilibrium state of the system of interest can be found by simultaneously solving the following set of equations:

1. chemical equilibrium equations (eqn [6]) for a linearly independent set of reactions that link all the species that exist in the system; this set of

equations also includes solid–liquid equilibrium reactions (or precipitation equilibria);
2. if applicable, vapor–liquid and, possibly, liquid–liquid equilibrium equations (i.e., the equality of chemical potentials of species in coexisting phases),
3. material balance equations for each element in the system, and
4. electroneutrality balance condition.

Once the solution is found, the potential can be calculated using the Nernst equation for any redox pair in the equilibrated system. It should be noted that procedures for solving these equations are very involved and require sophisticated numerical techniques. Methods for solving such electrolyte equilibrium problems were reviewed by Zemaitis *et al.*[4] and Rafal *et al.*,[7] and will not be discussed here.

The solution of the chemical equilibrium conditions provides detailed information about the thermodynamically stable form(s) of the metal in a particular system. This is in contrast with the information presented on the *E*–pH diagrams, which essentially indicate whether a given oxidation reaction (e.g., dissolution of a metal) can occur simultaneously with a given reduction reaction (e.g., reduction of water). In the chemical equilibrium algorithm outlined above, all possible reduction and oxidation equilibria are solved simultaneously. This provides very detailed information about the chemical identity of all species and their concentrations. On the other hand, this kind of information is less conducive to gaining qualitative insight into the key anodic and cathodic reactions that occur in the system.

To visualize the information obtained from detailed chemical equilibrium computations, stability diagrams can be constructed by selecting key independent variables. Such diagrams can be referred to as 'chemical' as opposed to 'electrochemical' because they do not involve the potential as an independent variable. An example of such a diagram is shown in **Figure 3**, which shows the stability areas of solid corrosion products of iron as a function of the amount of dissolved iron and hydrogen sulfide in 1 kg of water. Diagrams of this type can be generated by repeatedly solving the chemical equilibrium expressions outlined above and tracing the stability boundaries of various species (Lencka *et al.*,[94] Sridhar *et al.*[93]). Such diagrams are useful for illustrating the transition between various corrosion products (e.g., iron sulfide and magnetite in **Figure 3**) as a function of environmental conditions. Diagrams of this kind have been generated by Sridhar *et al.*[95] to identify the conditions under which $FeCO_3$, Fe_3O_4, and Fe_2O_3 coexist as corrosion products of carbon steel. This approach was used to elucidate conditions that are conducive to intergranular stress corrosion cracking in weakly alkaline carbonate systems, which has been associated with the transition between iron carbonate, iron (II), and iron (III) corrosion products.

It should be noted that simplified 'chemical' diagrams can be generated using activities rather than

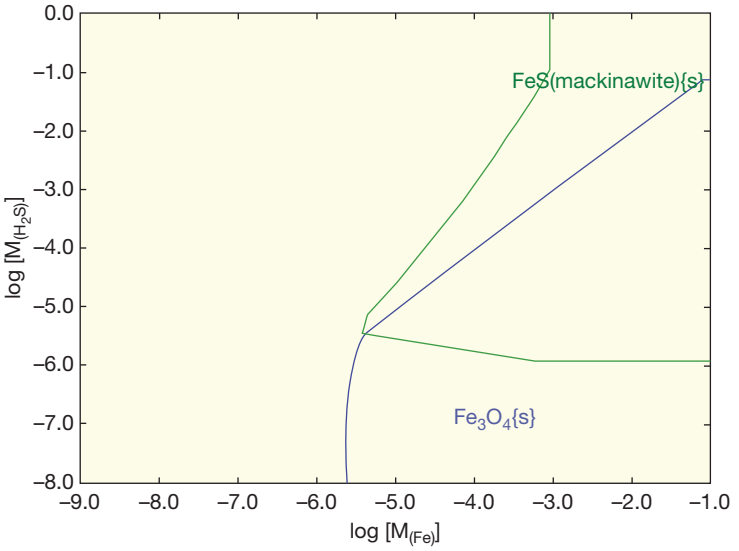


Figure 3 Stability diagram for solid iron corrosion products as a function of the molality of dissolved Fe and the total number of moles of H_2S per 1 kg of H_2O. The diagram has been constructed by computing the equilibrium states of solid–liquid–vapor reactions.[93]

concentrations of species as independent variables. Such an approach makes it possible to obtain the 'chemical' diagrams in a semianalytical way, as with the classical Pourbaix diagrams. An example of such diagrams is provided by Mohr and McNeil[96] for the Cu–H–O–Cl system. However, such simplified 'chemical' diagrams suffer from the fundamental disadvantage that species activities are not directly measurable and are, therefore, much less suitable as independent variables than concentrations. Comprehensive 'chemical' diagrams can be obtained only by simultaneously solving the chemical and phase equilibrium expressions for each set of conditions, preferably with the help of a realistic model for liquid-phase activity coefficients.

2.38.2.5 Problems of Metastability

The main strength of the stability diagrams lies in their purely thermodynamic nature. Accordingly, they can be generated using equilibrium thermochemical properties that are obtained from a variety of classical experimental sources (e.g., solubility, vapor pressure, calorimetric or electromotive force measurements for bulk systems) without recourse to any electrochemical kinetic studies of surfaces. At the same time, the purely thermodynamic nature of the diagrams is also their main weakness. An important limitation of the stability diagrams is the fact that they predict the equilibrium phases, whereas the actual corrosion behavior may be controlled by metastable phases. The same limitation applies to the detailed thermodynamic computations outlined earlier.

To illustrate the problems of metastability, it is instructive to examine how *E*–pH diagrams can be used to rationalize the Faraday paradox of iron corrosion in nitric acid, which historically played a great role in establishing the concept of passivity (Macdonald[97]). Faraday's key observation was that iron easily corroded in dilute nitric acid with the evolution of hydrogen. However, it did not corrode with an appreciable rate in concentrated HNO_3 solutions despite their greater acidity. When scratched in the solution, an iron sample would corrode for a short time and, then, rapidly passivate. A stability diagram for this system is shown in **Figure 4**. This figure presents a superposition of an *E*–pH diagram for iron species and another one for nitrogen species. Dilute nitric acid is a weak oxidizing agent and, therefore, the main cathodic reaction in this case is

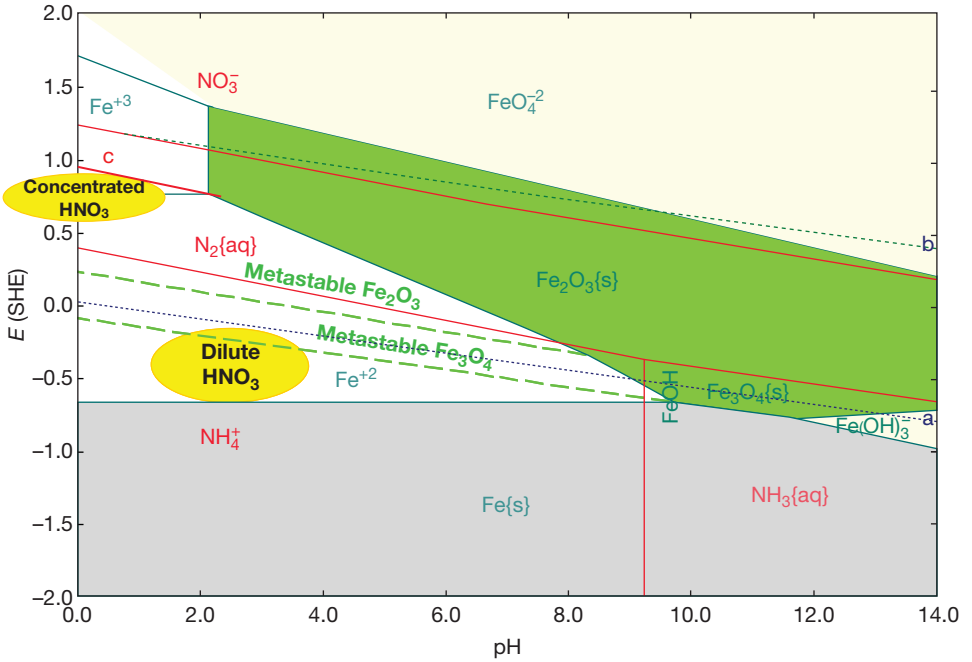


Figure 4 Metastability on the *E*–pH diagram and the interpretation of the Faraday paradox of iron corrosion in dilute and concentrated nitric acid. The stability fields of the oxides (Fe_3O_4 and Fe_2O_3) are extended into the metastable region (long dashed lines). An *E*–pH diagram for nitrogen is superimposed on the diagram for Fe. The line 'c' is a metastable equilibrium line for the reduction of HNO_3 to HNO_2. The ovals marked as 'concentrated HNO_3' and 'dilute HNO_3' indicate the mixed potential ranges that are expected for iron in concentrated and dilute HNO_3, respectively.

the common reaction of reduction of H^+ ions. The equilibrium potential for this reaction is given by the dotted line marked as 'a' in the diagram. The anodic process is the oxidation of Fe, for which the equilibrium potential is given by the equilibrium line between Fe(s) and Fe^{2+} in the diagram. Thus, the corrosion potential will establish itself between the 'a' line and the upper limit of the stability field of elemental iron. This corrosion potential range is approximately shown by the lower ellipsoid in **Figure 4**. In this region, the stable iron species is Fe^{2+} and, therefore, the stability diagram predicts the dissolution of iron with the formation of Fe^{2+} ions. Unlike dilute nitric acid, concentrated HNO_3 is a strong oxidizing agent. The main reduction reaction in this case is

$$NO_3^- + 2H^+ + 2e^- = NO_2^- + H_2O \qquad [23]$$

The equilibrium potential for this reaction is shown by a line marked as 'c' in **Figure 4**. It should be noted that the nitrite ions (NO_2^-) are metastable and, therefore, a stability field of nitrites does not appear on an *E*–pH diagram of nitrogen. Nevertheless, the equilibrium potential for reaction eqn [21] can be easily calculated as described above. It lies within the stability field of elemental nitrogen N_2(aq). The corrosion potential will then establish itself at a much higher potential, relatively close to the dominant cathodic line 'c.' The likely location of the corrosion potential is outlined in **Figure 4** by an ellipsoid marked 'concentrated HNO_3.' However, the stable species in this region are the Fe^{2+} and Fe^{3+} ions. Thus, the stability diagram does not explain the Faraday paradox as long as only stable species are considered. As indicated by Macdonald,[97] the *E*–pH diagram becomes consistent with experimental observation when the existence of metastable phases is allowed for. The dashed lines in **Figure 4** show the metastable extensions of the Fe^{2+}/Fe_3O_4 and Fe_3O_4/Fe_2O_3 equilibrium lines into the acidic range. Thus, it is clear that the corrosion potential of iron in concentrated HNO_3 is likely to establish itself in a potential range in which metastable iron oxides are stable. In contrast, the metastable solids are not expected to form at the low potentials that correspond to dilute HNO_3 environments (i.e., below the Fe^{2+}/metastable Fe_3O_4 boundary). If the surface is scratched, the large supply of H^+ ions near the scratch will create conditions under which the potential will be below the 'a' line and short-term hydrogen evolution will follow. However, this will lead to a rapid depletion of the H^+ ions, and the potential will shift to higher values, at which metastable phases can lead to passivation.

The formation of metastable solid phases is fairly common. The metastability of chromium oxide phases is important for the passivity of Fe–Ni–Cr alloys in acidic solutions. Although the *E*–pH diagrams indicate that chromium oxides/hydroxides cease to be stable in relatively weakly acidic solutions, the practical passivity range extends to a more acidic range for numerous alloys. Also, the metastability of iron sulfide phases plays an important role in the behavior of corrosion products in H_2S containing environments. Anderko and Shuler[75] used stability diagrams to evaluate the natural sequence of formation of various iron sulfide phases (e.g., amorphous FeS, mackinawite, pyrrhotite, greigite, marcasite, or pyrite). The sequence of formation of such phases could be predicted on the basis of the simple but reasonably accurate rule that the order of formation of solids is the inverse of the order of their thermodynamic stability. Stability diagrams are then used to predict the conditions under which various metastable phases are likely to form.

Thermodynamics provides a convenient starting point in the simulation of aqueous corrosion. Although it is inherently incapable of predicting the rates of interfacial phenomena, it is useful for predicting the final state toward which the system should evolve if it is to reach equilibrium. Furthermore, the computation of the final, thermodynamic equilibrium state can be refined by taking into account the metastable phases. In addition to determining the equilibrium state in metal–environment systems, electrolyte thermodynamics provides information on the properties of the environment. Such information, including speciation in the liquid phase, concentrations and activities of individual species, and phase equilibria, is necessary for constructing kinetic models of corrosion.

2.38.3 Modeling the Kinetics of Aqueous Corrosion

Aqueous corrosion is intrinsically an electrochemical process that involves charge transfer at a metal–solution interface. Because aqueous corrosion is a heterogeneous process, it involves the following fundamental steps:

1. reactions in the bulk aqueous environment,
2. mass transport of reactants to the surface,
3. charge transfer reactions at the metal surface,
4. mass transport of reaction products from the surface, and
5. reactions of the products in the bulk environment.

Reactions in the bulk environment (1 and 5) can be considered, with some important exceptions, to be equilibrium. As long as they are treated as equilibrium phenomena, they remain within the domain of electrolyte thermodynamics. On the other hand, the charge transfer reactions at the surface (3), and their coupling with mass transport (2 and 4) require the tools of electrochemical kinetics. In this section, we review the fundamentals of modeling approaches to electrochemical kinetics.

The objective of modeling the kinetics of aqueous corrosion is to relate the rate of electrochemical corrosion to external conditions (e.g., environment composition, temperature, and pressure), flow conditions, and the chemistry and metallurgical characteristics of the corroding interface. The two main quantities that can be obtained from electrochemical modeling are the corrosion rate (v_{corr}, often expressed as the corrosion current density i_{corr}) and corrosion potential (E_{corr}). For practical applications, the calculation of the corrosion rate is of primary interest for simulating general corrosion and the rate of dissolution in occluded environments such as pits or crevices. The value of the corrosion potential (E_{corr}) is also of interest because there is often a relationship between the value of the corrosion potential and the type of corrosion damage that occurs. In general, if the corrosion potential is above a certain critical potential (E_{crit}), a specific form of corrosion that is associated with E_{crit} can occur, typically at a rate that is determined by the difference $E_{corr} - E_{crit}$. This general observation applies to localized corrosion including pitting, crevice corrosion, intergranular stress corrosion cracking, and so on. An internally consistent model for general corrosion should simultaneously provide reasonable values of the corrosion rate and corrosion potential. Thus, the computation of the corrosion potential is of interest not so much for modeling general corrosion but for predicting other forms of corrosion.

2.38.3.1 Kinetics of Charge-Transfer Reactions

The theory of charge-transfer reactions is well developed and has been reviewed in detail by a number of authors including Vetter,[98] Bockris and Reddy,[99] Kaesche,[100] Bockris and Khan,[101] and Gileadi.[102] Here, we summarize the key relationships that form the basis of modeling.

Considering a simple reaction of transfer of n electrons between two species, a reduced form 'Red' and an oxidized form 'Ox':

$$\mathrm{Red} \underset{r_c}{\overset{r_a}{\rightleftharpoons}} \mathrm{Ox} + ne \qquad [24]$$

the current density associated with this reaction is, according to Faraday's law, equal to the difference between the anodic rate v_a and the cathodic rate v_c, multiplied by nF:

$$i = nF(v_a - v_c) \qquad [25]$$

According to the theory of electrochemical kinetics,[98] the rates of the anodic and cathodic reactions are related to the potential and the concentrations of the reacting species at the phase boundary, that is

$$i_a = nFv_a = nFk_a c_{r,s}^{x,r} \exp\left(\frac{\alpha_a nFE}{RT}\right) \qquad [26]$$

$$i_c = -nFv_c = -nFk_c c_{o,s}^{x,o} \exp\left(-\frac{\alpha_c nFE}{RT}\right) \qquad [27]$$

where k_a and k_c are the anodic and cathodic rate constants, α_a and α_c are the anodic and cathodic electrochemical transfer coefficients, $c_{r,s}$ and $c_{o,s}$ are the concentrations of the reduced (r) and oxidized (o) forms at the surface, respectively, and x,r and x,o are the reaction orders with respect to the reduced and oxidized species. For a given individual redox process, the anodic and cathodic electrochemical transfer coefficients are interrelated as $\alpha_c = 1 - \alpha_a$. The total current density for reaction eqn [24] is then

$$i = nFk_a c_{r,s}^{x,r} \exp\left(\frac{\alpha_a nFE}{RT}\right) - nFk_c c_{o,s}^{x,o} \exp\left(-\frac{\alpha_c nFE}{RT}\right) \qquad [28]$$

At the equilibrium (reversible) potential E_0, the current density i is equal to zero. In the absence of a net current, the concentrations of the species at the surface are equal to their bulk concentrations (i.e., $c_{r,s} = c_{r,b}$ and $c_{o,s} = c_{o,b}$). Then, the current density of the anodic process is equal to that of the anodic process and is defined as the exchange current density i_0, that is,

$$i_0 = nFk_a c_{r,b}^{x,r} \exp\left(\frac{\alpha_a nFE_0}{RT}\right) = nFk_c c_{o,b}^{x,o} \exp\left(-\frac{\alpha_c nFE_0}{RT}\right) \qquad [29]$$

Using eqn [29], the equation for the current density [28] can be expressed in terms of the exchange current density and the overvoltage $\eta = E - E_0$:

$$i = i_0 \left(\frac{c_{r,s}}{c_{r,b}}\right)^{x,r} \exp\left(\frac{\alpha_a nF(E - E_0)}{RT}\right) - i_0 \left(\frac{c_{o,s}}{c_{o,b}}\right)^{x,o} \exp\left(-\frac{\alpha_c nF(E - E_0)}{RT}\right) \qquad [30]$$

The ratios $c_{r,s}/c_{r,b}$ and $c_{o,s}/c_{o,b}$ depend on the transport of reactants and products to and from the corroding interface. If the mass transport is slow compared to charge transfer, the surface concentrations become different from those in the bulk. Conversely, if charge transfer is slow relative to mass transfer, the reaction is under charge transfer control and the ratios are equal to one not only at the equilibrium potential. In such a case, eqn [30] takes a particularly simple form and is usually referred to as the Butler–Volmer equation for charge-transfer reactions:

$$i = i_0 \exp\left(\frac{\alpha_a nF(E-E_0)}{RT}\right) - i_0 \exp\left(-\frac{\alpha_c nF(E-E_0)}{RT}\right) \quad [31]$$

The electrochemical transfer coefficient α depends on the mechanism of the charge transfer reaction. For some reactions, its value can be deduced from mechanistic considerations. However, in many cases, it needs to be determined empirically. It can be determined in the form of empirical Tafel coefficients defined as the slope of a plot of potential against the logarithm of current density, that is

$$\beta_a = \frac{dE}{d\ln i_a}; \quad \beta_c = \frac{dE}{d\ln i_c} \quad [32]$$

which yields

$$\beta_a = \frac{RT}{\alpha_a nF}; \quad \beta_c = \frac{RT}{\alpha_c nF} \quad [33]$$

or, in a more traditional decimal logarithm form:

$$b_a = \frac{2.303RT}{\alpha_a nF}; \; b_c = \frac{2.303RT}{\alpha_c nF} \quad [34]$$

Then, the Tafel coefficients b_a or b_c can be used in eqns [28] or [30] instead of the electrochemical transfer coefficient.

The above formalism includes both the cathodic and anodic process for a particular redox couple. However, in practical corrosion modeling, it is usually entirely sufficient to include only either a cathodic or an anodic partial current for a given redox process. Specifically, the cathodic partial process can be neglected for metal-ion reactions because the deposition of metal ions (i.e., the reverse of metal dissolution) is typically not of practical significance in corrosion. Similarly, the anodic partial process can be usually neglected for oxidizing agents because it is only their reduction that is of interest for corrosion. There are some exceptions to this rule, for example, in the case of relatively noble metals whose ions can be reduced under realistic conditions. Nevertheless, in the remainder of this review, we will separately consider partial anodic and cathodic processes for corrosion-related reactions.

As mentioned above, the concentrations of reactants and products at the surface depend on the mass transport to and from the corroding interface. In general, three mechanisms contribute to mass transport, that is, diffusion, migration, and convection. In many practical applications, migration can be neglected. This is the case for the transport of neutral molecules in any environment and, also, for the transport of charged species in environments that contain appreciable amounts of background electrolytic components (e.g., as supporting electrolyte). Migration becomes important in ionic systems in which there is no supporting electrolyte. We will return to the treatment of migration later in this review. If migration is neglected, the treatment of mass transfer by diffusion and convection can be simplified by using the concept of the Nernst diffusion layer. According to this concept, the environment near the corroding surface can be divided into two regions. In the inner region, called the Nernst diffusion layer, convection is negligible and diffusion is the only mechanism of transport. In the outer region, concentrations are considered to be uniform and equal to those in the bulk solution. Thus, the concentration changes linearly from the surface concentration to the bulk concentration over a distance δ, which is the thickness of the diffusion layer. In such a model, the flux of a species i in the vicinity of a corroding interface is given by Fick's law

$$\mathcal{J}_i = -D_i \left(\frac{\partial c_i}{\partial z}\right)_{z=0} \quad [35]$$

where D_i is the diffusion coefficient of species i and z in the direction perpendicular to the surface. Integration of eqn [35] over the thickness of the diffusion layer gives:

$$\mathcal{J}_i = -D_i \frac{c_{i,b} - c_{i,s}}{\delta_i} \quad [36]$$

It should be noted that the diffusion layer thickness δ is not a general physical property of the system. Rather, it is a convenient mathematical construct that makes it possible to separate the effects of diffusion and convection. It depends on the flow conditions, properties of the environment, and the diffusion coefficient of individual species. Thus, it may be different for various species. Methods for calculating δ will be outlined later in this review. Equation [36] can be applied to both the reactants that enter into electrochemical reactions at the interface and to corrosion products that leave the interface. Then, it can be combined with Faraday's law to obtain the current

density. For an oxidant o, eqn [36] yields an expression for a cathodic partial current density:

$$i_c = nF\mathcal{J}_o = -nFD_o\frac{c_{o,b} - c_{o,s}}{\delta_o} \quad [37]$$

According to eqn [37], the current density reaches a maximum, limiting value when the surface concentration $c_{o,s}$ decreases to zero. This condition defines the limiting current density, that is

$$i_{c,L} = -\frac{nFD_o c_{o,b}}{\delta_o} \quad [38]$$

For a corrosion product (e.g., Me ions), an analogous equation can be written for an anodic current density

$$i_a = nF\mathcal{J}_{Me} = -nFD_{Me}\frac{c_{Me,b} - c_{Me,s}}{\delta_{Me}} \quad [39]$$

In eqn [39], the surface concentration is typically limited by the solubility of corrosion products. Thus, a limiting anodic current density can be reached when the surface concentration of metal ions corresponds to the metal solubility, that is

$$i_{a,L} = -nFD_{Me}\frac{c_{Me,b} - c_{Me,sat}}{\delta_{Me}} \quad [40]$$

Combination of eqns [26]–[28] for charge-transfer processes and the simplified mass-transport equations (eqns [37] and [39]) yields a general formalism for electrochemical processes that are influenced by both charge transfer and mass transport. For example, for a cathodic process, the current density obtained from eqn [37] is equal to that obtained from eqn [27]. Thus, the surface concentration of the diffusing species can be obtained from eqn [37] and substituted into eqn [27]. The resulting equation can be solved analytically for i_c for some values of the reaction order x,o. If the reaction order is equal to one (i.e., x, o = 1 in eqn [27], a particularly simple relationship is obtained for i_c, that is

$$\frac{1}{i_c} = \frac{1}{i_{c,ct}} + \frac{1}{i_{c,L}} \quad [41]$$

where $i_{c,L}$ is the limiting current density (eqn [38]) and $i_{c,ct}$ is the charge-transfer contribution to the current density. The latter quantity is given by eqn [27] with the bulk concentration $c_{o,b}$ replacing the surface concentration, that is

$$i_c = -nFk_c c_{o,b}^{x,o}\exp\left(-\frac{\alpha_c nFE}{RT}\right) \quad \text{with } x,o = 1 \quad [42]$$

An analytical formula can also be obtained when x,o = 0.5.[103] For an arbitrary value of the reaction order, the current density can be computed numerically by solving a single equation.

It should be noted that eqns [26] and [27] are particularly simple forms for reactions of the type [24], in which no species other that *Red* and *Ox* participate. In general, the preexponential part of eqns [26] and [27] depends on the mechanism of a particular electrochemical reaction. In general, the preexponential terms of eqns [26] and [27] can be generalized using the surface coverage factors, θ_i, for reactive species that participate in electrochemical processes, that is,

$$i_a = nFk_a\theta_1^{x1}\theta_2^{x2}\ldots\theta_m^{xm}\exp\left(\frac{\alpha_a nFE}{RT}\right) \quad [43]$$

$$i_c = -nFk_c\theta_1^{x1}\theta_2^{x2}\ldots\theta_m^{xm}\exp\left(-\frac{\alpha_c nFE}{RT}\right) \quad [44]$$

The surface coverage factors, θ_i, are further related to the concentrations (or, more precisely, activities) of individual species at the metal surface through appropriate adsorption isotherms. In general, analysis of reaction mechanisms on the basis of experimental data leads to a substantial simplification of eqns [43] and [44]. In many cases, activities of species can be directly used in the kinetic expressions rather than the surface coverage fractions.

The above formalism makes it possible to set up a model of electrochemical kinetics on a corroding metal surface by considering the following steps:

1. determining all possible partial cathodic and anodic processes that may occur in a given metal-environment combination;
2. writing equations for the partial cathodic or anodic current densities associated with charge transfer reactions (eqns [43] and [44] or simplifications thereof); and
3. writing equations for the mass transport of the species that participate in the charge-transfer reactions (eqns [37] and [39]). In some simple, but realistic cases a combination of the charge-transfer and mass-transport equations results in analytical formulas such as eqn [41] for partial electrochemical processes.

Once the equations for the partial anodic and cathodic processes are established, the behavior of a corroding surface can be modeled on the basis of the Wagner–Traud[104] theory of metallic corrosion, often referred to as the mixed potential theory. According to the mixed potential theory, the sum of all partial anodic currents is equal to the sum of all cathodic

currents. Further, it is assumed that the electrical potential of the metal at an anodic site is equal to that at a cathodic site. These assumptions follow from the requirement that charge accumulation within a metal cannot occur and, therefore, the electrons produced as a result of oxidation processes must be consumed in the reduction processes. Therefore, in a freely corroding system, we have

$$\sum_j A_a i_{a,j} + \sum_j A_c i_{c,j} = 0 \text{ at } E = E_{corr} \qquad [45]$$

where A_a and A_c are the areas over which the anodic and cathodic reactions occur, respectively. The first sum in eqn [45] enumerates all anodic partial reactions and the second sum pertains to all cathodic reactions. Equation [45] is written using the sign conventions introduced earlier, in which the cathodic processes are written with a negative sign.

Equation [45] can be solved to obtain the corrosion potential, E_{corr}. Then, the corrosion current density and, equivalently, the corrosion rate, can be computed from the anodic current density for metal dissolution at the corrosion potential, that is

$$i_{corr} = i_{a,Me}(E_{corr}) \qquad [46]$$

At potentials that deviate from the corrosion potential, the left-hand side of eqn [45] represents the predicted current. Such a computed current versus potential relationship can be compared with experimentally determined polarization behavior.

In the case of general corrosion, $A_a = A_c$ and the solution of eqns [45] and [46] yields the corrosion potential and general corrosion rate. For localized corrosion, there is usually a great disparity between the areas on which the anodic and cathodic processes operate.

2.38.3.2 Modeling Adsorption Phenomena

Before discussing partial electrochemical reactions, it is necessary to outline the treatment of adsorption because the presence of adsorbed species is frequently assumed to derive expressions for electrochemical processes.

Adsorption of neutral molecules and ions on metals has been reviewed in detail by Gileadi,[105] Damaskin *et al.*,[106] and Habib and Bockris.[107] In general, adsorption leads to the reduction of the surface area that is accessible to electrochemical reactions. In such cases, adsorption results in a reduction in the rate of both anodic and cathodic processes. Thus, the rates of electrochemical reactions become modified by the factor $(1-\Sigma\theta_j)$, in which θ_j is a coverage fraction by species j. At the same time, adsorption may result in the formation of surface complexes that have different dissolution characteristics. An example of such a dual effect is provided by halide ions on Fe-group metal surfaces corroding in the active state. At relatively low or moderate halide concentrations, adsorption of halides leads to a reduction in electrochemical reaction rates. However, at higher halide concentrations, the adsorbed halide ions interfere with the mechanism of anodic dissolution of iron, which may lead to an increase in the corrosion rate.

A general formalism for modeling the effect of adsorption on electrochemical reactions is provided by the Frumkin isotherm. The Frumkin formalism takes into account the interactions between the species adsorbed on the surface. It results from the requirement that the rate of adsorption is equal to the rate of desorption in the stationary state,[105] that is

$$v_{ads,i} = v_{des,i} \qquad [47]$$

where the subscript i denotes any adsorbable species. The rate of adsorption is given by

$$v_{ads,i} = k_{ads,i}\left(1 - \sum_j \theta_j\right) a_i \exp\left(-\beta \sum_j A_{ij}\theta_j\right) \qquad [48]$$

where $k_{ads,i}$ is an adsorption rate constant, θ_i is a fraction of the surface covered by species i, a_i is the activity of species i in the solution, β is a transfer coefficient, and A_{ij} is a surface interaction coefficient between species i and j. The first term in parentheses on the right-hand side of eqn [48] represents the available surface, and the second term represents the effect of pairwise interactions between adsorbed species. The rate of desorption is given by

$$v_{des,i} = k_{des,i}\theta_i \exp\left((1-\beta)\sum_j A_{ij}\theta_j\right) \qquad [49]$$

where $k_{des,i}$ is the desorption rate constant. Combination of eqns [48] and [49] yields the Frumkin isotherm, that is

$$K_{ads,i}a_i = \frac{k_{ads,i}}{k_{des,i}}a_i = \frac{\theta_i}{1-\sum_j \theta_j}\exp\left(\sum_j A_{ij}\theta_j\right) \qquad [50]$$

where $K_{ads,i}$ is an adsorption equilibrium constant. Equation [50] can be simplified if it is assumed that the species are independently adsorbed. Then, the interactions between the species become zero, and eqn [50] takes the form of the well-known Langmuir isotherm:

$$K_{\mathrm{ads},i}a_i = \frac{\theta_i}{1-\sum_j \theta_j} \qquad [51]$$

The Langmuir isotherm is extensively used in electrochemical kinetics. The Frumkin isotherm has been used in some studies when more accurate modeling of adsorption is warranted by experimental data, for example, in the case of corrosion in very concentrated brines (Anderko and Young[108]).

It should be noted that detailed modeling of adsorption requires taking into account the effect of potential on adsorption. Equations [50] and [51] are strictly valid only when adsorption is not significantly influenced by metal dissolution. An approach to include the effect of dissolution and, hence, potential on adsorption has been developed by Heusler and Cartledge[109] who proposed an additional process in which a metal atom from an uncovered area $(1-\Sigma\theta_j)$ reacts with an adsorbed ion from the covered area θ_i to dissolve as ferrous ion. The adsorbed ion is then postulated to leave the surface during the reaction, thus contributing to the desorption process. Accordingly, eqn [49] is rewritten by adding an additional term, that is

$$v_{\mathrm{des},i} = k_{\mathrm{d}i}\theta_i \exp\left((1-\beta)\sum_j A_{ij}\theta_j\right) + i_{\mathrm{des},i} \qquad [52]$$

where the desorption current $i_{\mathrm{des},i}$ is given by

$$i_{\mathrm{des},i} = k_{ri}\theta_i\left(1-\sum_j \theta_j\right)a_X \exp\left(\frac{\beta FE}{RT}\right) \qquad [53]$$

where a_X is the activity of possible additional species (e.g., OH^-) that participate in the dissolution. In eqn [53], the desorption current is potential-dependent because it involves the dissolution of the metal. Equation [53] can be combined with eqns [48] and [52] to form a system of n equations for a solution with n adsorbable species. This system can be solved numerically for the coverage fractions θ_i of each adsorbed species. Because of the potential dependence, the model predicts that the adsorption coverage rapidly decreases above a certain potential range.

It should be noted that a detailed treatment of adsorption is not always necessary for modeling aqueous corrosion. In particular, the potential dependence of adsorption can be often neglected. In most cases, simplified approaches are warranted. Specifically, for low surface coverage, the fraction θ_i can be assumed to be proportional to the activity of the species i as shown by eqn [51]. Thus, eqns [43] and [44] can often be simplified by using activities of species rather than their surface coverages.

2.38.3.3 Partial Electrochemical Reactions

The behavior of a corroding system results from the interplay of at least two and, frequently, many partial electrochemical reactions. Such reactions include:

1. anodic dissolution of pure metals and alloys in both the active and passive state;
2. reduction of protons, which is usually the primary cathodic reaction in acid corrosion;
3. reduction of water molecules, which is frequently the main cathodic reaction in deaerated neutral and alkaline solutions;
4. reduction of dissolved species that can act as proton donors such as undissociated carboxylic acids, carbonic acid, hydrogen sulfide, and numerous ions that contain protons (e.g., bicarbonates, bisulfides, etc.);
5. reduction of oxygen, which is a common cathodic process in aerated solutions;
6. reduction of metal ions at high oxidation states such as Fe(III) or Cu(II), which can be reduced to a lower oxidation state;
7. reduction of oxyanions such as nitrites, nitrates, or hypochlorites in which a nonmetallic element is reduced to a lower oxidation state;
8. oxidation of water to oxygen, which occurs at high potentials and, therefore, is rarely important in freely corroding systems; and
9. oxidation of metals to higher oxidation states, for example, Cr(III) to Cr(VI), which may occur in the transpassive dissolution region of stainless steels and nickel base alloys.

In this section, we present illustrative examples of how these reactions can be modeled in practice.

2.38.3.3.1 Anodic reactions

The dissolution of several pure metals such as iron, copper, or nickel has been extensively investigated. Thus, it is possible to construct practical equations for the partial anodic dissolution processes on the basis of mechanistic information. For most alloys, detailed mechanistic information is not available and, therefore, it is necessary to establish kinetic expressions on a more empirical basis.

For iron dissolution, various multistep reaction mechanisms have been proposed. They have been reviewed in detail by Lorenz and Heusler,[110] Drazic,[111] and Keddam.[112] From the point of view of

modeling, particularly important parameters are the electrochemical transfer coefficient and reaction orders with respect to the ions that participate in anodic dissolution. Although there are substantial differences between the various proposed mechanisms, the dependence of the iron dissolution rate in acidic solutions on the activity of hydroxide ions is generally accepted. The mechanism proposed by Bockris *et al.*,[113] that is

$$\mathrm{Fe} + \mathrm{OH}^- \rightarrow \mathrm{FeOH}_{ads} + e^- \quad [54]$$

$$\mathrm{FeOH}_{ads} \rightarrow \mathrm{FeOH}^+_{ads} + \mathrm{e}^- \quad \text{(rate-determining step)} \quad [55]$$

$$\mathrm{FeOH}^+_{ads} \rightarrow \mathrm{Fe}^{2+} + \mathrm{OH}^- \quad [56]$$

predicts that the reaction order with respect to the OH^- ion is one because of the intermediate step of OH^- adsorption. The validity of this prediction has been verified for acidic solutions. Other mechanisms yield reaction orders between one and two. Additionally, the current density for iron dissolution has been found to depend on the activity of water (Smart and Bockris[114]). The mechanism of Bockris *et al.*[113] also predicts that the anodic transfer coefficient is $\alpha_{Fe} = 1.5$, which is consistent with experimentally observed Tafel slopes of 30–40 mV. Thus, the current density for Fe dissolution in acidic solutions can be expressed as

$$i_{Fe,OH} = i^*_{Fe,OH} a_{OH} a^c_{H_2O} \exp\left(\frac{\alpha_{Fe} F (E - E_{0,Fe})}{RT}\right) \quad [57]$$

where $i^*_{Fe,OH}$ is a temperature-dependent coefficient, the subscript Fe,OH indicates that the dissolution reaction is mediated by OH^- ions, and c is an empirically determined reaction order with respect to the activity of water. According to Smart and Bockris[114] $c = 1.6$. The effect of the activity of water on the current density becomes significant for concentrated solutions, for which the activity of water is usually significantly lower than 1.

Although the reaction order with respect to the OH^- ions is valid for acidic solutions, it has been found that iron dissolution proceeds with little influence of pH for solutions with pH above ~4. Bockris *et al.*[113] explained this phenomenon by assuming a certain nonzero reaction order with respect to Fe^{2+} and by considering the hydrolysis of the Fe^{2+} ions that result from the dissolution. Alternatively, the change in the reaction order with respect to OH^- ions can be reproduced by assuming that the exchange current density is proportional to the surface coverage by OH^- ions. This assumption is consistent with the reaction mechanism (see eqns [54]–[56]). Thus, eqn [57] can be generalized as[108]:

$$i_{Fe,OH} = i^*_{Fe,OH} \theta_{OH} a^c_{H_2O} \exp\left(\frac{\alpha_{Fe} F (E - E_{0,Fe})}{RT}\right) \quad [58]$$

Assuming that θ_{OH} follows the Langmuir adsorption model, eqn [58] can be rewritten as

$$i_{Fe,OH} = i^*_{Fe,OH} \frac{a_{OH}}{1 + K_{OH} a_{OH}} a^c_{H_2O} \exp\left(\frac{\alpha_{Fe} F (E - E_{0,Fe})}{RT}\right) \quad [59]$$

Equation [59] reduces to eqn [57] for low activities of OH^-, that is, for acidic solutions. For higher concentrations of hydroxide ions, the reaction order with respect to OH^- becomes zero. This is consistent with the lack of a dependence of the Fe oxidation reaction on pH in CO_2 corrosion of mild steel, which occurs at pH values above ~4 (Nešić *et al.*,[115] Nordsveen *et al.*[116]).

The effect of halide ions on the dissolution of iron and carbon steel is of particular interest. Adsorbed halide ions may accelerate the anodic dissolution, especially in concentrated halide solutions. A number of reaction mechanisms have been proposed to explain this phenomenon. In particular, Chin and Nobe[117] and Kuo and Nobe[118] developed a mechanism that postulates a reaction route that is parallel to eqns [52]–[54]. An essentially identical mechanism has also been proposed by Drazic and Drazic.[119] According to this mechanism, a halide-containing surface complex is responsible for the dissolution. Thus, eqn [54] is followed by the following parallel route:

$$\mathrm{FeOH}_{ads} + \mathrm{X}^- \rightarrow \mathrm{FeOHX}^-_{ads} \quad [60]$$

$$\mathrm{FeOHX}^-_{ads} \rightarrow \mathrm{FeOHX}_{ads} + \mathrm{e}^- \quad \text{(rate-determining step)} \quad [61]$$

$$\mathrm{FeOHX}_{ads} + \mathrm{H}^+ \rightarrow \mathrm{Fe}^{2+} + \mathrm{X}^- + \mathrm{H_2O} \quad [62]$$

The mechanism eqns [60]–[62] results in a dissolution current density that depends on the activities of both halide and hydroxide ions. In acidic solutions, an equation analogous to eqn [57] can be written as

$$i_{Fe,X} = i^*_{Fe,X} a^s_{X^-} a^t_{OH^-} \exp\left(\frac{\alpha_{Fe} F (E - E_{0,Fe})}{RT}\right) \quad [63]$$

where the subscript X indicates the halide ions that mediate the reaction. For chloride systems, $s = 0.4$ and $t = 0.6$ when concentrations are used instead of

activities.[118] Since the mechanism described by eqns [60]–[62] is assumed to be parallel to the mechanism under halide-free conditions, the total current density of anodic dissolution can be assumed to be a sum of the contributions of two mechanisms. Also, eqn [63] can be generalized to neutral solutions in analogy with eqn [59]. Additionally, the desorption current density (eqn [53]) contributes to the total current, although it becomes important only at relatively high potentials, and its numerical significance is usually limited. Thus, the expression for the total active Fe dissolution current in halide solutions becomes[108]

$$i_{Fe} = i_{Fe,OH} + i_{Fe,X} + i_{des,i} \qquad [64]$$

As with iron, anodic dissolution of copper has also been extensively investigated. Kear *et al.*[120] reviewed the mechanisms and associated expressions for the current density of anodic dissolution of copper in the active state in chloride environments. Copper dissolution is generally thought to proceed through the formation of cuprous chloride complexes and to be under mixed, charge-transfer and transport, control close to the corrosion potential. Several authors (Lee and Nobe,[121] Deslouis *et al.*,[122] King *et al.*[123]) assumed the following mechanism:

$$Cu + Cl^- \underset{k_{-1}}{\overset{k_1}{\longleftrightarrow}} CuCl + e^- \qquad [65]$$

$$Cu + Cl^- \underset{k_{-2}}{\overset{k_2}{\longleftrightarrow}} CuCl_2^- \qquad [66]$$

The expression for the anodic current density derived from eqns [65] and [66] is

$$\frac{i_{Cu}}{nF} = \frac{k_1 k_2}{k_{-1}} a_{Cl}^2 \exp\left(\frac{F(E - E_{0,Cu})}{RT}\right) - k_{-2} a_{CuCl_2^-} \qquad [67]$$

Since the reaction rate depends on the activity of the reaction products ($CuCl_2^-$) at the surface (see the second term on the right-hand side of eqn [67]), the reaction is partially controlled by the mass transport of the $CuCl_2^-$ ions, and the anodic current density is simultaneously equal to

$$\frac{i_{Cu}}{nF} = D_{CuCl_2^-} \frac{a_{CuCl_2^-}}{\delta_{CuCl_2^-}} \qquad [68]$$

Equation [68] is a special case of eqn [39] when the bulk concentration of $CuCl_2^-$ is negligible.

Much less mechanistic information is available for the anodic dissolution of alloys in the active state. In the case of stainless steels and nickel-base alloys, this is due to the fact that dissolution of these metals in the passive state is more important than in the active state. For these alloys, active dissolution is of importance only in acidic solutions. In this case, expressions for anodic dissolution need to be established on an empirical basis. For stainless steels and nickel-base alloys, a positive reaction order between one and two with respect to hydroxide ions is observed. While such values are similar to those observed for Fe, the exchange current densities are very different and need to be determined separately for individual alloys.

2.38.3.3.2 Cathodic reactions

Among the numerous possible partial cathodic processes, the reduction of protons, water molecules, and dissolved oxygen is ubiquitous in aqueous corrosion. Reduction of protons is an important cathodic process in acidic solutions. The overall reaction is given by

$$H^+ + e^- \rightarrow 0.5H_2 \qquad [69]$$

The mechanisms of this reaction have been reviewed by Vetter[98] and Kaesche.[100] Proton reduction proceeds in two steps according to two alternative mechanisms. The Volmer–Heyrovsky mechanism applies to most metals, whereas the Volmer–Tafel mechanism may be observed on certain noble metals. The Volmer–Heyrovsky mechanism can be represented as a sequence of two elementary reactions, that is

$$H^+ + e^- \rightarrow H_{ads} \qquad [70]$$

$$H^+ + H_{ads} \rightarrow H_2 \qquad [71]$$

It is generally accepted that the H^+ reduction reaction may proceed under activation or mass transfer control. The cathodic process of H^+ reduction can be modeled assuming that the reaction order with respect to the protons is equal to one. Then, eqns [41] and [42] can be directly used for modeling. In addition to its dependence on the activity of protons, there is empirical evidence that the H^+ reduction depends on the activity of water. According to Smart *et al.*,[124] the reaction order with respect to water activity is 2.2 on iron. The electrochemical transfer coefficient can be assumed to be equal to ~0.5 for carbon steels and many corrosion-resistant alloys, which corresponds to a Tafel slope of 118 mV at 25 °C.

As the pH of a solution increases, the importance of the proton reduction reaction rapidly decreases. In neutral and alkaline solutions, the reduction of water molecules becomes predominant unless stronger oxidizing agents (e.g., oxygen) are present in the system. The water reduction is given by

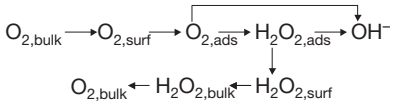
$$H_2O + e^- \rightarrow 0.5H_2 + OH^- \quad [72]$$

and is thermodynamically equivalent to the reduction of protons. However, its kinetic characteristics are different. Unlike the reduction of protons, the water reduction reaction typically does not exhibit a limiting current density because there are no diffusion limitations for the transport of H_2O molecules to the surface. This remains true as long as the system is predominantly aqueous. The water reduction process can be modeled by assuming the same reaction order with respect to H_2O as that for proton reduction. Also, practically the same value of the electrochemical transfer coefficient can be assumed.

Reduction of oxygen, that is

$$O_2 + 4H^+ + 4e^- \rightarrow 2H_2O \quad [73]$$

is the predominant cathodic reaction in aerated aqueous solutions unless the solution contains stronger oxidizing agents such as ferric, cupric, or hypochlorite ions. The mechanism of oxygen reduction is substantially more complex than the mechanisms of H^+ or H_2O reduction. Oxygen reduction on iron and carbon steel has been reviewed by Jovancicevic and Bockris,[125] Zecevic *et al.*,[126] and Jovancicevic.[127] On stainless steels, it has been analyzed by Le Bozec *et al.*,[128] Kapusta,[129] and in papers cited therein. On copper, it has been studied by King *et al.*[130] In general, it has been established that the reaction may proceed either through a four-electron pathway, which leads to the reduction of O_2 to H_2O, or through a two-electron pathway, which leads to the formation of H_2O_2 as an intermediate. An overall reaction scheme may be represented as



[74]

where the subscripts 'surf' and 'ads' denote the oxygen in the diffusion layer close to the surface and oxygen adsorbed on the surface, respectively. The absorbed intermediate H_2O_2 can be either further reduced to OH^- or desorbed and dissolved in the solution or converted back to oxygen through decomposition or reoxidation. The actual reaction pathway is influenced by many factors such as the surface treatment of the electrode.[128] However, the four-electron reduction path from O_2 to OH^- seems to predominate.[129] Oxygen reduction may be under charge transfer or mass transfer control, due to the diffusion of dissolved oxygen molecules. For passive metals, the process is usually under charge transfer control because the limiting current density for oxygen reduction is usually greater than the passive current density at typical dissolved oxygen concentrations.

For modeling purposes, the key parameters are the electrochemical transfer coefficient and the reaction orders with respect to dissolved oxygen and protons. These parameters determine the dependence of the reduction reaction on dissolved oxygen concentration (or, equivalently, the partial pressure of oxygen) and on pH. Once these parameters are known, the oxygen reduction process can be modeled on a semi-empirical basis. The current density for oxygen reduction can be written as:

$$i_{O_2} = i^*_{O_2} a^q_{O_2,s} a^r_{H^+,s} \exp\left[\frac{-\alpha_{O_2} F\left(E - E_{0,O_2}\right)}{RT}\right] \quad [75]$$

Equation [75] needs to be coupled with eqn [37] with $n = 4$ for mass-transfer limitations. The reaction orders q and r in eqn [75] are, in general, specific to the metal surface although they are expected to be similar within families of alloys. For stainless steels, there seems to be a consensus that the reaction order with respect to dissolved oxygen is 0.5 (Kapusta,[129] Sridhar *et al.*[131]), whereas the order with respect to protons ranges from 0.5 to 1 (or, equivalently, the order with respect to hydroxide ions varies from −0.5 to −1). For passive iron or carbon steel, the reaction order with respect to O_2 has been reported as 0.5 (Calvo and Schiffrin[132]) or 1 (Jovancicevic and Bockris,[125] Jovancicevic[127]). For copper, a value of 1 has been reported (King *et al.*[130]).

Another important cathodic reaction is the reduction of transition metal ions such as Fe^{3+} and Cu^{2+} to lower oxidation states, for example

$$Fe^{3+} + e^- \rightarrow Fe^{2+} \quad [76]$$

This process can be modeled as a first-order reaction with respect to the activity of ferric ions by taking into account the mass transport limitations (eqn [37]).

All the cathodic reactions discussed above may proceed under mass transfer limitations due to the diffusion of reactants to the corroding surface. However, cathodic limiting current densities may also arise because of limitations due to homogeneous reactions in the solution. A prominent example of such a reaction is the reduction of carbonic acid, which is the key cathodic process in CO_2 corrosion of carbon steel. This reaction accounts for the substantially higher corrosivity of CO_2 solutions than mineral acid solutions at the same pH. Carbonic acid results from the hydration of dissolved CO_2, that is

$$CO_2 + H_2O = H_2CO_3 \quad [77]$$

Reaction eqn [77] is followed by the reduction of H_2CO_3 on the surface, that is

$$H_2CO_3 + e^- \rightarrow 0.5H_2 + OH^- \quad [78]$$

which is thermodynamically equivalent to the reduction of protons, but is characterized by different kinetics. The H_2CO_3 reduction is under activation or chemical reaction control, and can be modeled using eqn [41]. The charge transfer current is expressed as (Nešić *et al.*[115]):

$$i_{H_2CO_3} = i^*_{H_2CO_3}\, a_{H_2CO_3}\, a_{H^+}^{-0.5} \exp\left[\frac{-\alpha_{H_2CO_3} F\left(E - E_{0,H}\right)}{RT}\right] \quad [79]$$

where the transfer coefficient can be assumed to be equal to that for H_2O reduction. The limiting current density can be calculated from an equation developed by Nešić *et al.*[115] on the basis of a formula derived by Vetter[98] for processes with a rate-determining homogeneous reaction in the solution. Here, the rate-determining reaction is the hydration of CO_2 and the limiting current density is:

$$i_{H_2CO_3,L} = Fa_{H_2CO_3}\left(D_{H_2CO_3} K_{H_2CO_3} k^f_{H_2CO_3}\right)^{1/2} \quad [80]$$

where $D_{H_2CO_3}$, $K_{H_2CO_3}$, and $k^f_{H_2CO_3}$ are the diffusion coefficient of H_2CO_3, equilibrium constant for the hydration of CO_2, and forward reaction constant for the hydration reaction, respectively.

2.38.3.3.3 Temperature dependence

The rates of the majority of partial anodic and cathodic processes are strongly dependent on temperature. This temperature dependence can be modeled by assuming that the concentration-independent part of the exchange current density (here denoted by i^*) is expressed as

$$i^*(T) = i^*(T_{ref}) \exp\left[-\frac{\Delta H^{\neq}}{R}\left(\frac{1}{T} - \frac{1}{T_{ref}}\right)\right] \quad [81]$$

Equation [81] is equivalent to assuming a constant enthalpy of activation $\Delta H^{\neq}$ for each partial process.

2.38.3.4 Modeling Mass Transport Using Mass Transfer Coefficients

To calculate the mass-transport effects on electrochemical kinetics according to eqns [37]–[40], it is necessary to predict the diffusion layer thickness δ_i or, equivalently, the limiting current density. Theoretical formulas for these quantities cannot be obtained for arbitrary flow conditions and, therefore, empirical approaches are necessary for most practical applications.

In the case of a rotating disk electrode, a theoretical solution has been derived by Levich.[133] It is noteworthy that Levich's solution preceded experimental results. The thickness of the diffusion layer on a rotating disk electrode is

$$\delta_i = 1.61 D_i^{1/3} \nu^{1/6} \omega^{-1/2} \quad [82]$$

where D_i is the diffusion coefficient of the reacting species, ω is the rotation rate, and ν is the kinematic viscosity, which is the ratio of the dynamic viscosity and density, that is

$$\nu = \eta/\rho \quad [83]$$

In view of the relationships between the thickness of the diffusion layer and the limiting current density (eqns [38] and [40]), a physically equivalent predictive expression can be written for the limiting current density. For example, the limiting current density of a cathodic reaction (eqn [38]) then becomes:

$$i_{c,L} = -0.6205 nF c_{o,b} D_o^{2/3} \nu^{-1/6} \omega^{1/2} \quad [84]$$

For many other flow geometries, mass transport can be calculated using empirical correlations expressed in terms of the mass transfer coefficient k_m. In general, the mass transfer coefficient is defined as

$$k_m = \frac{\text{Reaction rate}}{\text{Concentration driving force}} \quad [85]$$

For an electrochemical reaction, the reaction rate is expressed using the current density, and eqn [36] for a mass transport-limited reaction can be rewritten in terms of the mass transfer coefficient k_m as

$$\mathcal{J}_i = \frac{i_i}{n_i F} = -D_i \frac{c_{i,b} - c_{i,s}}{\delta_i} = -k_{m,i}(c_{i,b} - c_{i,s}) \quad [86]$$

This indicates a relationship between δ_i and $k_{m,i}$, that is

$$k_{m,i} = \frac{D_i}{\delta_i} \quad [87]$$

Mass transport rates can be expressed using dimensionless groups, for which empirical correlations can be developed for a number of flow patterns. The mass transfer coefficient k_m enters into the Sherwood number Sh, which is defined as

$$Sh = \frac{k_m d}{D} \quad [88]$$

where d is a characteristic dimension (e.g., a pipe or rotating disk diameter); D, a diffusion coefficient; and the subscript i has been dropped for convenience as it is understood that eqn [88] is written for individual reacting species. The Sherwood number can be correlated with the Reynolds (Re) and Schmidt (Sc) numbers, which are defined as

$$Re = \frac{Vd}{\nu} \quad [89]$$

$$Sc = \frac{\nu}{D} \quad [90]$$

where V is the linear velocity. It can be shown by dimensional analysis that Sh is a function of Re and Sc. This function typically has the form[134,135]

$$Sh = \text{Const} \times Re^{x} \times Sc^{y} \quad [91]$$

where x is usually between 0.3 and 1 and y is about 1/3. For example, the theoretically derived results for the rotating disk can be recast in terms of the mass transfer coefficient as

$$Sh = 0.6205\ Re^{0.5}\, Sc^{0.33} \quad [92]$$

Empirical expressions for other flow geometries have been reviewed by Poulson[134,135] for single-phase flow conditions. Equations of the type eqn [91] exist for the rotating cylinder, impinging jet, nozzle or orifice, and pipe flow. For the rotating cylinder, the correlation of Eisenberg *et al.*[136] is widely used

$$Sh = 0.0791 Re^{0.70}\, Sc^{0.356} \quad [93]$$

For single-phase flow in a straight pipe, several correlations have been developed. Among these equations, Berger and Hau's[137] correlation has found use in a number of corrosion modeling studies:

$$Sh = 0.0165 Re^{0.86}\, Sc^{0.33} \quad [94]$$

The earlier pipe flow formulas have been reviewed by Poulson,[134] and the use of more recent equations has been discussed by Lin *et al.*[138] Equations of this kind are not as well developed for multiphase flow. Correlations are available for stratified flow (Wang and Nešić[139]), but a comprehensive treatment is not available for various regimes of multiphase flow. Therefore, a convenient alternative is to base the computation of mass transfer coefficients on the well-known analogy between heat and mass transfer.

The analogies between the transport of mass, momentum, and heat can be understood by considering the similarity between their respective mathematical formulations, namely, Fick's law of diffusion, Newton's law of viscosity, and Fourier's law of heat conduction. Thus, once a relationship has been established for a given phenomenon in terms of dimensionless numbers, it can further serve for the calculation of another phenomenon that takes place under the same geometric and physical conditions but with different velocities, dimensions, and physical properties of the system. In particular, the correlations established for heat transfer can be used for mass transfer calculations. The analogy between the heat, mass, and momentum transfer has been stated in a dimensionless form by Chilton and Colburn[140] as

$$\frac{Sh}{Re\, Sc^{1/3}} = \frac{Nu}{Re\, Pr^{1/3}} = \frac{f}{2} \quad [95]$$

where the Nusselt number, Nu, and the Prandtl number, Pr, are the heat transfer equivalents of the Sherwood and Schmidt numbers in mass transfer, respectively, and f is a friction factor. Correlations for the friction factor are available as a function of pipe roughness, its diameter, and the Reynolds number (Frank[141]). The exponent 1/3 in eqn [95] can be replaced with a generalized exponent n. The Nusselt and Prandtl numbers are defined as

$$Nu = \frac{hd}{\lambda} \quad [96]$$

$$Pr = \frac{\nu}{\alpha} \quad [97]$$

where h is the convection heat transfer coefficient; λ, the thermal conductivity; and α, the thermal diffusivity ($\alpha = \lambda/\rho C_p$).

This relationship makes it possible to determine the mass transfer coefficient in two-phase flow systems for which experimental heat transfer correlations are available. Heat transfer correlations take the form

$$Nu = \text{Const} \times Re^{x} Pr^{y} \quad [98]$$

They have been reviewed by Kim *et al.*[142] and Adsani *et al.*[143,144] for annular, slug, and bubbly flow in horizontal and vertical tubes. Also, Adsani *et al.*[143,144] developed a correlation for calculating the two-phase Nusselt number, which is a generalization of the Chilton-Colburn[140] heat transfer expression (i.e., the second equality of eqn [95]):

$$Nu_{\text{two-phase}} = C_1 f_{\text{L}}^{C_2} \left(\frac{V_{\text{L}} d}{\nu}\right)^{C_3} Pr_{\text{L}}^{1/3} \quad [99]$$

where the liquid-phase friction factor f_{L} and velocity V_{L} are calculated using flow models for annular, slug, and bubbly flow, and C_1, C_2, and C_3 are fitting constants. Subsequently, the Sherwood number and the

mass transfer coefficient can be obtained from the first equality in eqn [95].

As shown in the above equations, the computation of the mass transfer coefficient requires the diffusion coefficients, viscosity, and density. Density is a thermodynamic property and, as such, can be calculated from any comprehensive thermodynamic model for electrolyte systems. For example, Wang *et al.*[54] describe how to calculate densities in a way that is consistent with other thermodynamic properties. The computation of viscosities and diffusion coefficients requires separate models, which are beyond the scope of this chapter. Viscosity and diffusivity models have been reviewed and critically evaluated by Corti *et al.*[145] with particular emphasis on systems at elevated temperatures.

2.38.3.4.1 Example of electrochemical modeling of general corrosion

To illustrate the application of the principles described earlier, **Figure 5** shows the computation of the corrosion rate and potential of type 316 stainless steel in aqueous solutions of HF. These calculations have been made using the model of Anderko *et al.*[108,146] as implemented in the Corrosion Analyzer software.[61] The upper diagram shows the partial cathodic and anodic processes in a 2 m HF solution. Three cathodic processes are taken into account in this system: reduction of protons (H^+), reduction of undissociated HF molecules, and reduction of water molecules. These partial processes are marked in **Figure 5(a)** as (1), (2), and (3), respectively. The H^+ reduction reaction is modeled using eqns [41] and [42] as described earlier. The HF reduction process is calculated using the same equations, but with a different exchange current density. Both the H^+ and HF reduction processes show partial current densities because of mass transport limitations for the transport of H^+ and HF to the surface. Because of the low degree of dissociation of HF, the reduction of HF (line two in **Figure 5(a)**) plays a much more important role than the reduction of H^+ ions (line 1). The partial anodic curve for the oxidation of 316 SS is labeled as (4). It is assumed that the alloy components dissolve congruently, and the dependence of the partial anodic current density on the acidity of the solution is analogous to that observed for Fe (see eqn [59]), but with a different exchange current density. The superposition of the partial cathodic and anodic processes yields a predicted polarization curve, which is shown by a thick line in **Figure 5(a)**. The location of the mixed potential is calculated according to eqn [45] with $A_a = A_c$ and marked with a triangle. The complete model reproduces the experimental corrosion rates (**Figure 5(b)**) and corrosion potentials (**Figure 5(c)**) as a function of HF concentration and temperature.

2.38.3.5 Detailed Modeling of Mass Transport

The treatment of mass transport by the use of mass transfer coefficients is computationally efficient and is capable of reproducing steady-state corrosion behavior with good accuracy. However, it is subject to some limitations including:

1. it is suitable for calculating only steady-state behavior and, therefore, it is not appropriate for modeling time-dependent corrosion,
2. it neglects migration which introduces errors especially for dilute systems without a background electrolyte, and
3. it is not convenient for modeling transport in systems with geometrical constraints, especially when the cathodic and anodic areas are spatially separated; thus, it is not well suited for modeling the propagation of localized corrosion.

These limitations can be eliminated by a more comprehensive (but much more computationally demanding) treatment of transport phenomena. This treatment is based on the conservation laws for each species in the solution (Newman[150]):

$$\frac{\partial c_k}{\partial t} = -\nabla \mathcal{J}_k + R_k \quad k = 1, \ldots, K \qquad [100]$$

where C_k is the concentration of species k, t is the time, $\mathcal{J}_k$ is the flux of species k, ∇ is the vector differential operator (which reduces to $\partial/\partial x$ in a one-dimensional case), and R_k is the rate of production (source) or depletion (sink) of this species as a result of chemical reactions. In the vast majority of practical applications, the dilute solution theory is used to calculate the flux of the species, that is

$$\mathcal{J}_k = -D_k \nabla c_k - z_k F u_k c_k \nabla \varphi + c_k v \qquad [101]$$

where u_k is the mobility of species k, φ is the electrostatic potential in the solution, v is the fluid velocity, and the other symbols were defined previously. In eqn [101], the first term on the right-hand side represents the contributions of diffusion, the second term describes migration, and the third term is a contribution of convection. In the migration term, the mobility can be calculated from the diffusivity using the Nernst–Einstein equation:

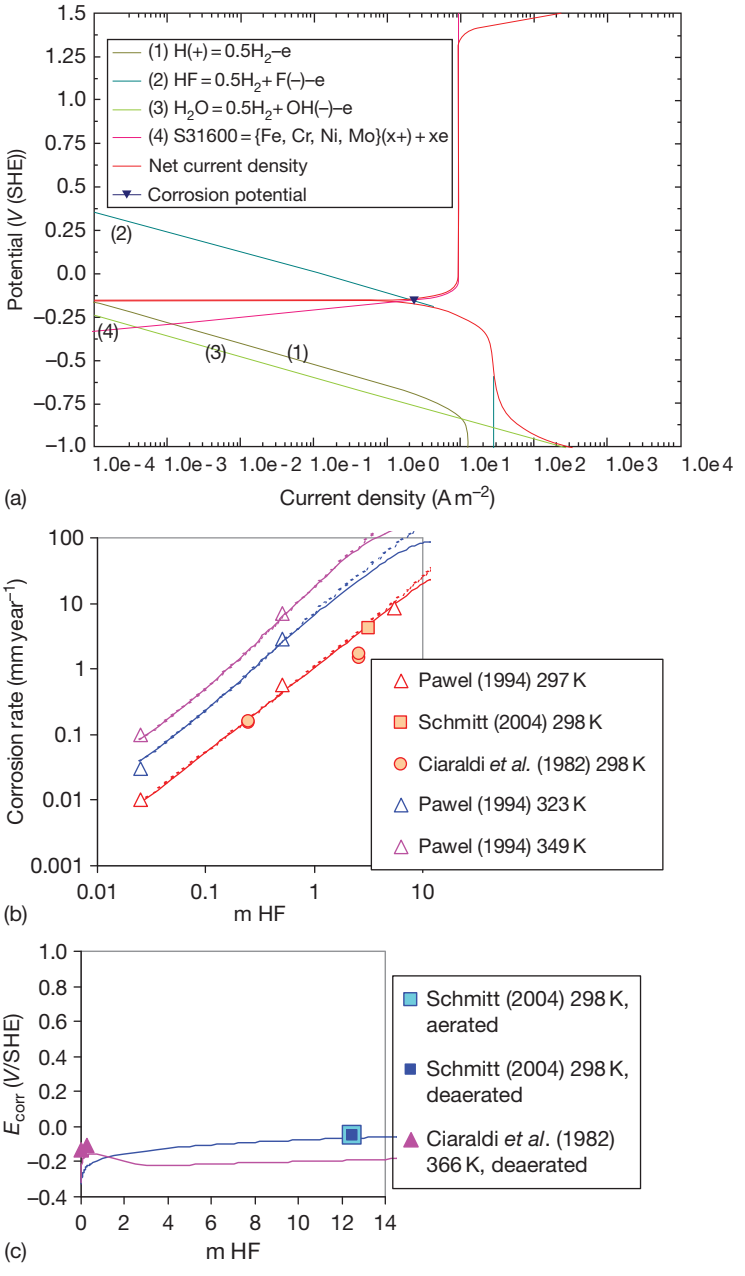


Figure 5 Application of an electrochemical model of general corrosion[108,146] to type 316L stainless steel in aqueous HF solutions. The upper diagram (a) shows the partial cathodic and anodic processes in a 2 m HF solution. The middle (b) and lower (c) diagrams compare the calculated and experimental[147–149] corrosion rates and potentials, respectively.

$$u_k = \frac{D_k}{RT} \qquad [102]$$

which is exact for species at infinite dilution and provides a good approximation at finite concentrations. An additional condition for determining the potential in eqn [101] is given by the Poisson equation:

$$\nabla^2 \varphi = -\frac{F}{\varepsilon} \sum_k z_k c_k \qquad [103]$$

where ε is the dielectric permittivity of the solution. It can be shown[150] that, due to the large value of the ratio F/ε, even a very small separation of charges (i.e., $\sum z_k c_k \neq 0$) results in a large potential gradient,

which in turn prevents any appreciable separation of charge. Therefore, eqn [103] is very often replaced in practice by the simple electroneutrality condition, that is

$$\sum z_k c_k = 0 \qquad [104]$$

which is one of the basic equations for computing equilibrium properties of electrolytes.

It should be noted that eqn [101] is rigorous only for dilute solutions. For concentrated solutions, its more general counterpart is[150]

$$\mathcal{J}_k = c_k v_k \qquad [105]$$

where the velocities v_k of species k are determined by the multicomponent diffusion equations:

$$c_k \nabla \mu_k = RT \sum_j \frac{c_k c_j}{c_{\text{tot}} D_{kj}} (v_j - v_k) \qquad [106]$$

where D_{kj} are the mutual diffusion coefficients, and c_{tot} is the total concentration of all components. The application of eqns [105] and [106] is very difficult due to the lack of a general methodology for computing D_{kj} and computational complexity. Therefore, eqns [105] and [106] have found few applications in practical models. However, a practical simplified form can be obtained for moderately dilute solutions for which the concentrations of solute species are smaller than the concentration of the solvent. Then, eqns [105] and [106] simplify to

$$\mathcal{J}_k = -\frac{D_k}{RT} c_k \nabla \mu_k + c_k v \qquad [107]$$

Considering that $\mu_k = \mu_k^0 + RT \ln c_k \gamma_k + z_k F \varphi$, the flux equation becomes[150]

$$\mathcal{J}_k = -D_k \nabla c_k - z_k F u_k c_k \nabla \varphi + c_k v - D_k c_k \nabla \ln \gamma_k \qquad [108]$$

which is only moderately more complex than the dilute-solution eqn [101], but benefits from the information on solution nonideality that is embedded in the activity coefficient and can be calculated from an electrolyte thermodynamic model.

In the convective term of eqn [101] or [108], the instantaneous fluid velocity (v) can be calculated, in principle, by the methods of computational fluid dynamics. However, such calculations involve a large computational effort and are, in practice, limited with respect to flow geometries and conditions. For turbulent flow, a practical approach relies on introducing turbulent diffusion. Accordingly, instantaneous velocity is divided into steady and turbulent components. The steady component is parallel to the surface and does not contribute to transport to and from the surface. Then, the convection term in eqn [101] or [108], $c_k v$, is approximated by a turbulent diffusivity term, $-D_t \nabla c_k$, which can be lumped with the molecular diffusion term thus defining an effective diffusion coefficient (Davis,[151] Nordsveen *et al.*[116])

$$D_k^{\text{eff}} = D_k + D_t \qquad [109]$$

where D_t can be obtained from empirical correlations with fluid properties.[151,152] For example, Davis's[151] correlation has been used in the CO_2 corrosion model of Nordsveen *et al.*[116]:

$$D_t = 0.18 \left(\frac{z}{\delta}\right)^3 \frac{\eta}{\rho} \qquad [110]$$

where z represents the distance from the surface (either a metal surface or a surface covered with corrosion products), η is the viscosity, ρ is the density, and δ is the thickness of the laminar boundary layer, which can be calculated for a pipe with a diameter d as

$$\delta = 25 Re^{-7/8} d \qquad [111]$$

It is noteworthy that this formalism can be shown to be physically equivalent to the treatment of turbulent flow through mass transfer coefficients as described in the previous section. Specifically, Wang and Nešić[139] showed a relationship between the mass transfer coefficient and D_t:

$$\frac{1}{k_{m,i}} = \int_0^{\delta} \frac{\mathrm{d}z}{D_i + D_t} \qquad [112]$$

The computation of the rates of production or depletion R_k is necessary in order to apply eqn [100]. A general matrix formalism for calculating the R_k terms in a system with multiple reactions has been developed by Nordsveen *et al.*[116] The main limitation here is the fact that rate data are available only for a very limited number of reactions such as precipitation of common scales ($CaCO_3$, $FeCO_3$) and selected homogeneous reactions (e.g., hydration of H_2CO_3). For the vast majority of reactions, only equilibrium equations are available and, in fact, there is no physical need for kinetic expressions for most homogeneous reactions because they are fast relative to mass transport. Therefore, arbitrary rate expressions may be assumed as long as they are constrained by the equilibrium constant (i.e., the ratio of the forward and reverse rate constants is equal to the equilibrium constant), give appropriately fast reaction rates and change direction as the equilibrium point is crossed. A convenient expression for the production or depletion rate for species k can be defined in terms of the departure of the ionic product from equilibrium (Walton[153]):

$$R_k = \sum_{m=1}^{M} \left[-r_m v_{km} \ln\left(\frac{\prod c_k^{v_{km}}}{K_m} \right) \right] \qquad [113]$$

where r_m is an adjustable numerical rate parameter for reaction m, K_m is the equilibrium constant for reaction m, and v_{km} is the stoichiometric coefficient for species k in the mth reaction. Alternatively, the transport equations can be first solved separately from the chemical effects and, then, at the end of each sufficiently small time step, thermodynamic equilibrium calculations can be performed in each elementary volume.

To solve the system of transport eqns [100] and [101] or [108], boundary conditions are required. In the bulk solution, the equilibrium concentrations are the natural boundary conditions. For each species involved in electrochemical reactions, the flux at the metal surface is determined from

$$\mathcal{J}_k = -\frac{i_k}{n_k F} \qquad [114]$$

where the current density i_k is calculated from appropriate expressions for cathodic and anodic partial processes as a function of concentrations at the metal surface. This provides a link to the mechanistic or empirical electrochemical expressions described above. For the species that are not involved in the reactions, the flux at the interface is zero. It should be noted that the application of this formalism of mass transport can become quite computationally involved, especially for systems with numerous species, because it requires solving a system of differential equations (eqns [100] and [101] or [108]) with constraints (eqns [103] or [104] and chemical terms).

2.38.3.5.1 Effect of the presence of porous media

The above transport equations need to be modified when mass transport occurs through porous media such as corrosion products, calcareous deposits, soil or sand, and various man-made environments, including concrete and ceramics. Such a generalization can be formulated in terms of two characteristic quantities, porosity and tortuosity. Porosity (ε) is the volumetric void fraction of the medium, whereas tortuosity (τ) is defined as the ratio of the distance that an ion or molecule travels around solid particles to the direct path. For practical applications, tortuosity can be correlated with porosity thus leaving porosity as the only parameter to affect transport equations. The generalization of transport equations to porous media has been discussed by Newman[150] and Bear.[154] A detailed corrosion model that includes the transport in porous corrosion products has been developed for CO_2 corrosion by Nordsveen *et al.*[116] and Nešić and Lee.[155]

2.38.3.6 Active–Passive Transition and Dissolution in the Passive State

The expressions for anodic partial current densities discussed above (eqns [54]–[68]) are limited to the dissolution in the active state. However, dissolution in the passive state and the transition between the active and passive state are equally important for modeling aqueous corrosion. In fact, passivity is the key to our metal-based civilization (Macdonald[97]) and has been extensively investigated since the pioneering work of Faraday and Schönbein in the 1830s. Theories of passivity have been reviewed by many investigators (Frankenthal and Kruger,[156] Froment,[157] Marcus and Oudar,[158] Natishan *et al.*,[159] Macdonald[97,160]) and are beyond the scope of this chapter. In this section, we focus solely on practical models for calculating the anodic current density in the passive and active–passive transition regions as a function of solution chemistry.

Passivity manifests itself by a sharp drop in the anodic current density at a certain critical potential as the metal is polarized in a negative-to-positive potential direction. For calculation purposes, empirically determined anodic polarization curves can be reproduced using a suitable fitting function. For example, such a function has been developed by Macdonald.[161] Then, empirical fitting functions can be used within the framework of the mixed potential theory as described above.

A convenient way to introduce the active–passive transition into a computational model is to consider a current that leads to the formation of a passive layer in addition to the current that leads to active dissolution (Ebersbach *et al.*,[162] Anderko and Young[108]). For this purpose, a certain fraction of the surface θ_P can be assumed to be covered by a passive layer. The change of the passive layer coverage fraction with time can be expressed as

$$\left(\frac{\partial \theta_P}{\partial t} \right)_{E,a_i} = c i_{MeO}(1 - \theta_P) - K\theta_P \qquad [115]$$

where i_{MeO} is the current density that contributes to the formation of a passive layer. The second term on the right-hand side of eqn [115] represents the rate of dissolution of the passive layer, which is

proportional to the coverage fraction. Solution of this equation in the steady-state limit yields an expression for the anodic dissolution current:

$$i_{\mathrm{Me,TOT}} = \frac{i_{\mathrm{Me}} + i_{\mathrm{MeO}}}{1 + (ci_{\mathrm{MeO}}/K)} = \frac{i_{\mathrm{Me}} + i_{\mathrm{MeO}}}{1 + (i_{\mathrm{MeO}}/i_{\mathrm{P}})} \qquad [116]$$

where i_{Me} is the dissolution current density in the active state and the ratio $i_{\mathrm{P}} = c/K$ constitutes the passive current density. The current i_{Me} is calculated using the active dissolution models described above. The current i_{MeO} is expressed using the usual expression for process under activation control, that is

$$i_2 = i_2^0 \exp\left(\frac{\alpha_2 F(E - E_{\mathrm{F}})}{RT}\right) \qquad [117]$$

in which the parameters can be adjusted to reproduce the observable characteristics of the active–passive transition including the critical current density (i_{crit}) and Flade potential (E_{F}).[146] Equation [116] can be then used for the anodic process of metal dissolution within the framework of the mixed-potential theory (eqns [45] and [46].

An example of mixed-potential calculations for a passive metal is shown in **Figure 6**. This figure illustrates the computation of the corrosion potential of alloy 600 in a dilute LiOH solution as a function of dissolved oxygen concentration. As in **Figure 5**, the upper and middle diagrams of **Figure 6** show the predicted partial E versus i curves for the anodic and cathodic processes. The upper diagram (**Figure 6(a)**) shows the predictions for a very low O_2 concentration (0.013 ppm), whereas the concentration in the middle diagram is somewhat higher (0.096 ppm). In a weakly alkaline solution, the alloy is passive as indicated by the vertical portion of the anodic curve (line labeled as (3)). Two main cathodic processes are taken into account in this system, that is, the reduction of H_2O (line (1)) and the reduction of O_2 (line (2)). At the lower O_2 concentration, the limiting current density is lower than the passive current density, and the main cathodic process is the reduction of H_2O (i.e., the mixed potential lies at the intersection of the lines (1) and (3) in **Figure 6(a)**). As the O_2 concentration increases, the O_2 reduction reaction becomes predominant and determines the mixed potential, which then lies at the intersection of lines (2) and (3). This behavior explains the experimentally determined s-shaped dependence of E_{corr} on O_2 concentration as shown in **Figure 6(c)**. The s-shape is due to the transition from H_2O reduction to O_2 reduction as the dominant cathodic process. The transition depends on flow conditions because the O_2 reduction is partially under mass transport control.

Passive dissolution and active–passive transition strongly depend on solution chemistry. In the absence of specific active ions, the dissolution of oxide films depends primarily on the pH of the solution. Appropriate kinetic expressions can be constructed by considering dissolution reactions between the passive oxide/hydroxide surface layers and solution species (Anderko *et al.*,[146] Sridhar *et al.*[131,164]) In acidic solutions, the key reaction involves the protons from the solution:

$$\equiv \mathrm{MeO}_a(\mathrm{OH})_b + (2a + b - x)\mathrm{H}^+ = \mathrm{Me(OH)}_x^{(2a+b-x)+} + (a + b - x)\mathrm{H_2O} \qquad [118]$$

where the symbol '≡' denotes surface species. The corresponding kinetic equation is

$$i_{\mathrm{p,H^+}} = k_{\mathrm{H^+}} a_{\mathrm{H^+,s}}^q \qquad [119]$$

where $a_{\mathrm{H^+,s}}$ denotes the surface concentration of hydrogen ions and q is a reaction order, which is not necessarily related to the stoichiometric coefficient in the dissolution reaction. In neutral solutions, the predominant dissolution reaction can be written as

$$\equiv \mathrm{MeO}_a(\mathrm{OH})_b + a\mathrm{H_2O} = \mathrm{Me(OH)}_{(2a+b),\mathrm{aq}}^0 \qquad [120]$$

where the predominant species on the right-hand side of eqn [120] is a neutral complex as indicated by the superscript 0. The corresponding kinetic equation is

$$i_{\mathrm{p,H_2O}} = k_{\mathrm{H_2O}} a_{\mathrm{H_2O,s}}^r \qquad [121]$$

where the reaction order with respect to water indicates that dissolution may be affected by water activity. Similarly, the predominant reaction in alkaline solutions is

$$\equiv \mathrm{MeO}_a(\mathrm{OH})_b + (x - 2a - b)\mathrm{OH^-} + a\mathrm{H_2O} = \mathrm{Me(OH)}_x^{(x-2a-b)-} \qquad [122]$$

with a corresponding kinetic equation given by

$$i_{\mathrm{p,OH^-}} = k_{\mathrm{OH^-}} a_{\mathrm{OH^-,s}}^s \qquad [123]$$

The total passive current density as a function of pH is given by

$$i_{\mathrm{p}} = i_{\mathrm{p,H^+}} + i_{\mathrm{p,H_2O}} + i_{\mathrm{p,OH^-}} \qquad [124]$$

It should be noted that the passive dissolution may be influenced by mass transport. For example, aluminum dissolution in alkaline solutions is known to be partly under mass transport control due to the transport of

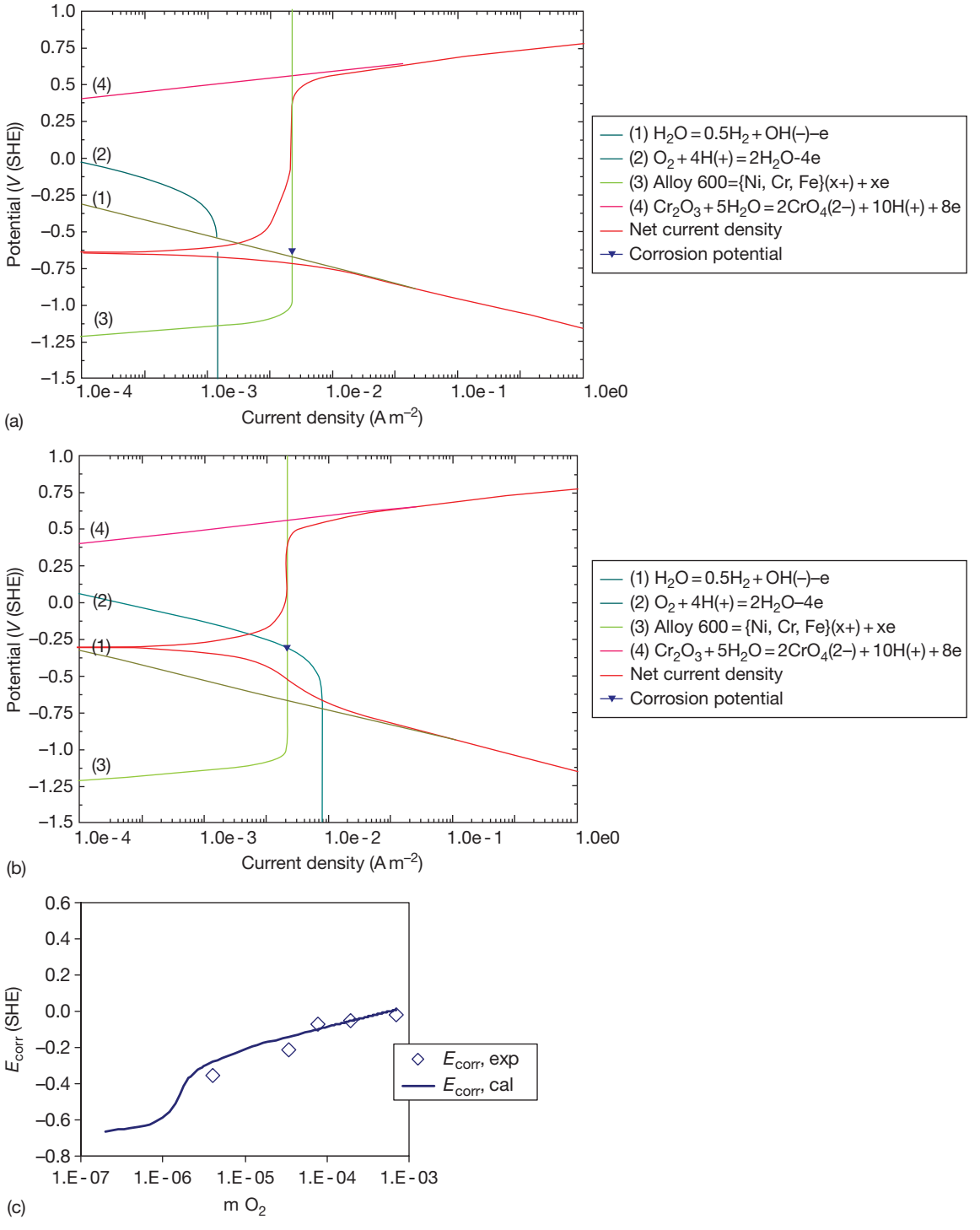


Figure 6 Modeling of the corrosion potential of alloy 600 in a 0.1 M LiOH solution at 200 °C as a function of dissolved oxygen concentration. The upper and middle diagrams (a and b) show the calculated partial electrochemical reactions and predicted polarization curve for solutions containing 4e-7 m (0.013 ppm) and 3×10^{-6} m (0.096 ppm) O_2, respectively. The lower diagram (c) compares the calculated corrosion potential with experimental data[163] at 200 °C as a function of dissolved oxygen molality.

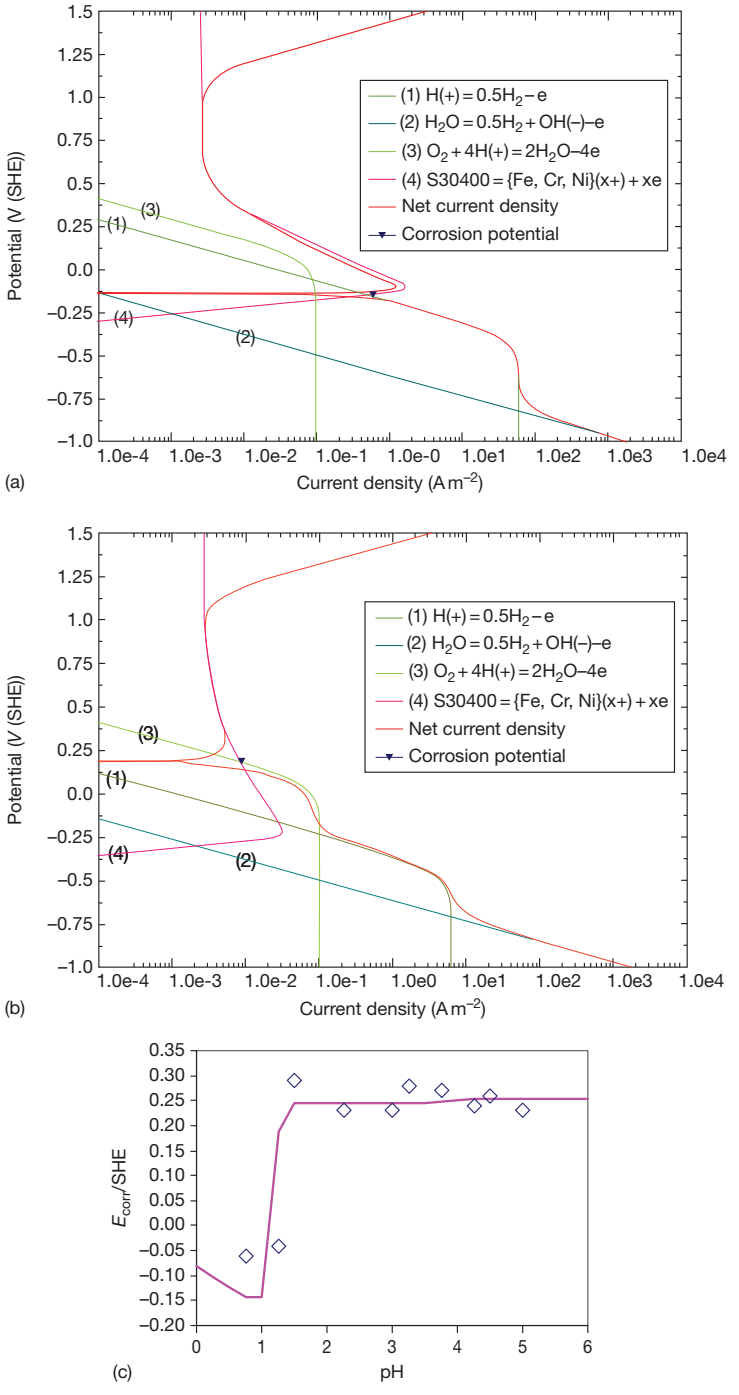


Figure 7 Electrochemical modeling of the depassivation pH and corrosion potential of type 304 stainless steel in aerated 0.1 M $Na_2SO_4 + H_2SO_4$ solutions. The upper (a) and middle (b) diagrams show the calculated partial electrochemical processes and predicted polarization curve for pH = 0.8 and 1.8, respectively. The lower diagram (c) compares the calculated corrosion potentials with experimental data.[165]

OH^- ions from the bulk to the interface. Then, the contributions to the passive current density (eqn [124]) should be coupled with mass-transfer equations such as eqn [36].[131]

Figure 7 illustrates the electrochemical modeling of the pH dependence of the active–passive transition of type 304 stainless steel. The upper and middle diagrams in Figure 7 show the partial electrochemical

processes in Na_2SO_4 solutions with pH = 0.8 and 1.8, respectively. The anodic curve (line 4) was modeled using a model based on eqns [116] and [117]. The main cathodic process is the reduction of O_2. Because of the pH effect on the active–passive transition, the mixed potential moves from the active dissolution region for pH = 0.8 to the passive region for pH = 1.8. This explains the dependence of the experimentally determined corrosion potential on pH (**Figure 7(c)**). The pH value at which an abrupt change of E_{corr} occurs can be identified with the depassivation pH.

In addition to pH effects, some active ions may influence the magnitude of the passive current density. The effect of active species on the dissolution in the passive state can be modeled by considering surface reactions between the metal oxide film and solution species (Blesa *et al.*,[166] Anderko *et al.*[146]):

$$\equiv \mathrm{MeO}_a(\mathrm{OH})_b + c_i X_i = \equiv \mathrm{MeO}_a(\mathrm{OH})_b X_{c_i} + e_i \mathrm{OH}^- \qquad [125]$$

where X_i is the ith reactive species in the solution, and the subscripts a, b, c_i, and e_i represent the reaction stoichiometry. In general, eqn [125] may be written for any active, aggressive, or inhibitive species i in the solution ($i = 1, \ldots, n$). It is reasonable to assume that eqn [125] is in quasi-equilibrium. The surface species that forms as a result of reaction eqn [125] may undergo irreversible dissolution reactions such as:

$$\equiv \mathrm{MeO}_a(\mathrm{OH})_b X_{c_i} + a\mathrm{H}_2\mathrm{O} \rightarrow \mathrm{Me(OH)}^0_{2a+b,\mathrm{aq}} + c_i X_i \qquad [126]$$

in which dissolved metal species are formed in analogy to those described by eqns [118], [120], and [122]. Mathematical analysis of reactions eqns [125] and [126][108,146] yields a relationship between the passive current density and activities of reactive species:

$$i_\mathrm{p} = i^0_\mathrm{p}(\mathrm{pH}) \frac{1 + \sum_i l_i (a^{c_i}_{X_i} / a^{e_i}_{\mathrm{OH}^-})}{1 + \sum_i K_i (a^{c_i}_{X_i} / a^{e_i}_{\mathrm{OH}^-})} \qquad [127]$$

where $i^0_\mathrm{p}(\mathrm{pH})$ is given by eqn [124], l_i is the forward rate of reaction eqn [126], and K_i is the equilibrium constant of reaction eqn [125].

Figure 8 illustrates the effect of active ions on the rate of general corrosion using alloy 22 in mixed $HNO_3 + HF$ solutions as an example. The upper diagram (**Figure 8(a)**) shows the predicted partial electrochemical processes in a 20% HNO_3 solution and the middle diagram (**Figure 8(b)**) shows how these processes change when a moderate amount of HF (1.57%) is added. A 20 wt.% HNO_3 solution is an oxidizing medium, and, therefore, reduction of NO_3^- ions in an acidic environment is the main cathodic process. This results in a high corrosion potential as shown in **Figure 8(a)**. The corrosion rate is controlled by the dissolution rate of the oxide film. When HF is added, the dissolution rate of the oxide substantially increases even though a moderate amount of HF has practically no effect on the acidity of the system. This effect is reproduced by eqn [127] and manifests itself by the increased passive current density in **Figure 8(b)**. The predicted effects can be compared with the observed corrosion rates in 20% HNO_3 solution as a function of HF concentration (**Figure 8(c)**).

2.38.3.7 Scaling Effects

In addition to passive dissolution and active–passive transition, modeling of surface scale formation is of great practical importance. Scales form as a result of deposition of corrosion products (e.g., iron carbonate or sulfide) or other solids that reach supersaturation near metal interfaces (e.g., calcareous deposits). Scales can be distinguished from passive films in that they do not give rise to the classical active–passive transition such as that shown in **Figure 7**. Rather, they reduce the rate of dissolution by providing a barrier to the diffusion of species to and from the surface and by partially blocking the interface, thus reducing the overall rate of electrochemical reactions. In general, there may be multiple mechanisms of scale formation depending on the chemistry of the precipitating solids.

One mechanism of scale formation can be quantified in terms of the competition between the rate of scale formation, which results in the precipitation of a corrosion product, and the rate of corrosion under the scale, which leads to the 'undermining' of the scale. When the rate of precipitation exceeds the rate of corrosion, dense protective films are formed. Conversely, when the corrosion rate is greater than the precipitation rate, the scale still forms, but the precipitation rate is not fast enough to fill the growing voids. Then, the scale becomes unprotective even though it may be thick. Nešić and Lee[155] developed a model to represent this phenomenon for $FeCO_3$ scale formation. In Nešić and Lee's[155] model, the local change in the volumetric concentration of the scale-forming solid is given by a redefined eqn [100]:

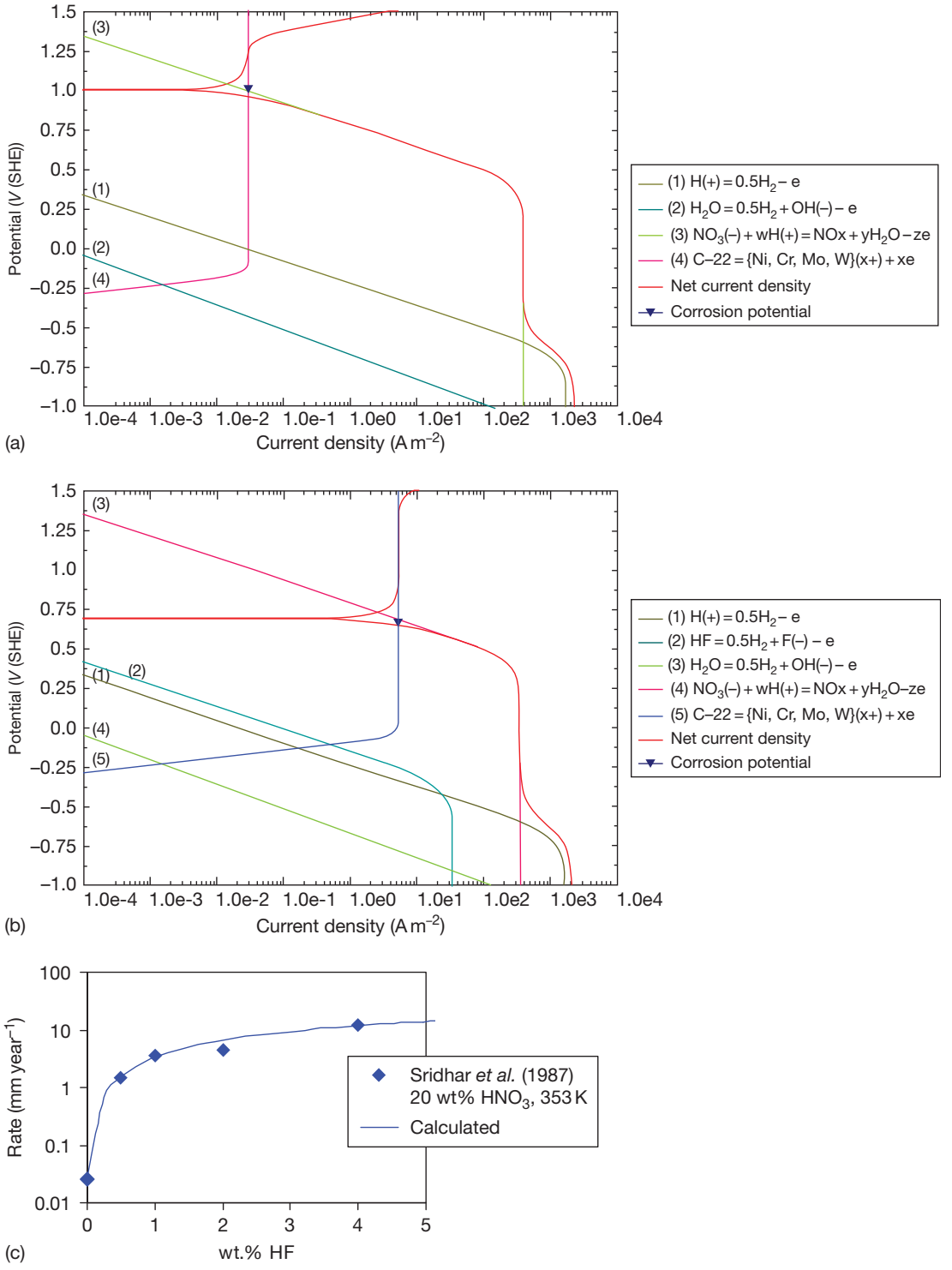


Figure 8 Electrochemical modeling of the effect of HF concentration on the corrosion rate of alloy 22 in $HNO_3 + HF$ solutions. The upper (a) and middle (b) diagrams show the partial electrochemical processes in 20% HNO_3 solutions without HF and with 1.57% HF, respectively. The lower diagram compares the calculated results with experimental data[167] as a function of HF concentration.

$$\frac{\partial c_{\text{solid}}}{\partial t} = R_{\text{solid}} - \text{CR}\frac{\partial c_{\text{solid}}}{\partial z} \quad [128]$$

where the first term on the right-hand side represents the rate of the scale formation and the second term is the scale undermining rate. In eqn [128], CR is the corrosion rate and z is a direction perpendicular to the surface. The rate of formation of the scale is, in general, a product of the scale particles' surface area-to-volume ratio A/V, a function of temperature, the thermodynamic solubility product K_{sp}, and an empirical function of supersaturation S:

$$R_{\text{solid}} = \frac{A}{V} f(T) K_{\text{sp}} f(S) \quad [129]$$

where supersaturation is defined as

$$S = \frac{\prod a_i^{\nu_i}}{K_{\text{sp}}} = \frac{\prod c_i^{\nu_i}}{K_{\text{sp}}} \quad [130]$$

While the $f(T)$ and $f(S)$ functions can be, in principle, derived from precipitation kinetics data that are independent of corrosion, Sun and Nešić[168] have determined that much more reliable precipitation rates can be obtained from corrosion weight loss and gain measurements than from kinetic measurements that start from dissolved metal ions. The A/V ratio depends on the porosity of the scale on the metal surface. Nešić and Lee[155] developed an empirical function of porosity that is consistent with the experimental data for $FeCO_3$ scale formation.

A different model is necessary for scales whose formation does not follow the kinetics of precipitation processes. For example, FeS scales form very fast in highly undersaturated solutions, in which they would be thermodynamically unstable in the bulk, and their formation appears not to be influenced by solution supersaturation. Thus, the effect of FeS scales can be modeled by assuming a solid-state reaction at the metal surface that is mediated by the adsorption of H_2S (Anderko and Young,[169] Nešić *et al.*[170]) The formation of FeS scales is further complicated by the existence of an outer layer that results from the growth, cracking, and delamination of the FeS film. A model that accounts for these phenomena was developed by Sun and Nešić.[171]

2.38.3.8 Modeling Threshold Conditions for Localized Corrosion

Modeling of the evolution of localized corrosion has been the subject of extensive research during the past three decades, and a number of important models have been developed for the initiation, stabilization, propagation, and stifling of individual pits, crevices, and cracks and for the statistical behavior of their ensembles in corroding structures. However, this topic is outside the scope of this chapter and will be reviewed in the chapter '**Predictive Modeling of Corrosion**' in this volume.

In this chapter, we focus solely on models that predict the conditions for the occurrence of localized corrosion without going into the treatment of the evolution of localized corrosion events in time and space. Such models are designed to find the threshold criteria for localized corrosion. In general, localized corrosion occurs when the corrosion potential of an alloy in a given environment exceeds a critical potential. The meaning, experimental determination, and interpretation of the key potentials that characterize localized corrosion have been reviewed by Szklarska-Smialowska.[172] While this general concept is well accepted, what constitutes a critical potential continues to be debated. The selection of the critical potential depends on the particular phenomenon that is to be modeled.

The applicability of the critical potential concept to modeling localized corrosion is qualitatively illustrated in **Figure 9**. In this figure, the arrows indicate the conditions at which localized corrosion is expected. For a given alloy, the critical potential decreases with an increase in the concentration of aggressive species (e.g., halide ions) as shown in **Figure 9(a)**. The shape of the E_{crit} curve corresponds to that of the repassivation potential curve, but the qualitative pattern is more general. Unlike the critical potential, the corrosion potential is usually not a strong function of aggressive ion concentration unless significant localized corrosion occurs. The critical aggressive species concentration for localized corrosion is observed when E_{corr} exceeds E_{crit}. Similarly, for a given aggressive chemical environment, a critical temperature exists (see **Figure 9(b)**). The critical potential is also strongly affected by the presence of inhibitors. As shown in **Figure 9(c)**, this gives rise to a critical inhibitor concentration. In many environments, the presence of oxidants may increase E_{corr} so that localized corrosion may occur beyond a critical concentration of redox species (**Figure 9(d)**). The actual conditions in a system may be a combination of the four idealized cases shown in **Figure 9**. Thus, the key is to predict both the corrosion potential and the repassivation potential. The corrosion potential can be obtained from a general-corrosion, mixed-potential model for passive metals as described above. For the critical potential, separate models are necessary.

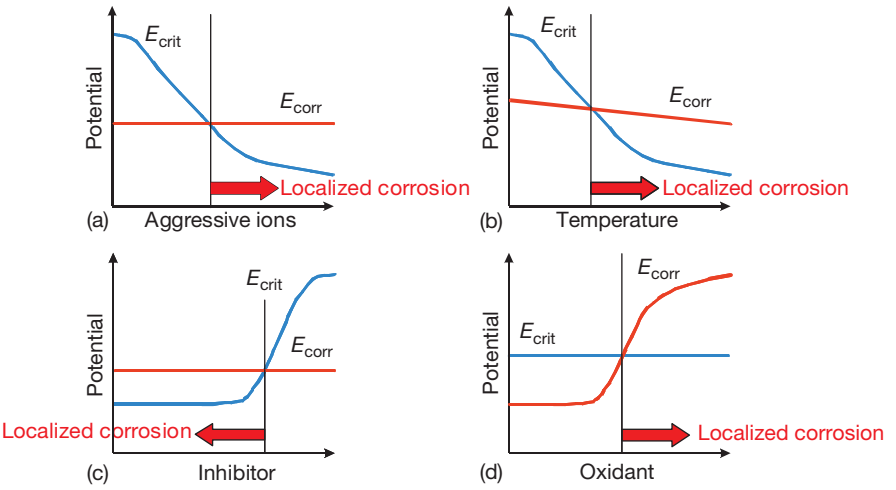


Figure 9 A general conceptual scheme of the use of the corrosion potential (E_{corr}) and critical potential (E_{crit}) to predict the effect of aggressive ions, temperature, inhibitors, and oxidizing redox species on localized corrosion. The arrows marked 'localized corrosion' denote the potential ranges in which localized corrosion can be expected.

In this chapter, we briefly review the computation and applicability of the passivity breakdown potential and the repassivation potential.

2.38.3.8.1 Breakdown of passivity

To predict the initiation of localized corrosion, it is necessary to calculate the critical passivity breakdown potential. Several theories have been developed to relate the breakdown potential to the concentration of aggressive species in the solution (Heusler and Fischer,[173] Strehblow and Titze,[174] Lin *et al.*,[175] Okada,[176] McCafferty,[177] Haruna and Macdonald,[178] Macdonald,[97] Yang and Macdonald,[179] and papers cited therein). A common theoretical result, confirmed by experimental data, is the linear dependence of the passivity breakdown potential on the logarithm of the concentration of aggressive ions. While this observation is generally accepted, its generalization to systems with multiple aggressive and inhibitive ions is not immediately obvious.

A particularly comprehensive treatment of passivity breakdown is provided by the point defect model of Macdonald and coworkers (Lin *et al.*,[175] Haruna and Macdonald,[178] Macdonald,[97] Yang and Macdonald[179]). According to the point defect model, passivity breakdown results from the condensation of cation or metal vacancies at the interface between the metal and the passive barrier layer. The vacancies are envisaged to be generated at the barrier layer–solution interface in an autocatalytic, anion-induced process. For systems containing only aggressive (halide) ions X, the critical breakdown potential is expressed as

$$E_b = \frac{4.606RT}{\chi\alpha F}\log\left[\frac{\mathcal{J}_m}{\hat{a}Du^{-\chi/2}}\right] - \frac{2.303RT}{\alpha F}\log a_X \quad [131]$$

where $\mathcal{J}_m$ is the rate of annihilation of cation vacancies at the metal/barrier layer interface, $\hat{a}$ and u are thermodynamic parameters related to the absorption of an aggressive ion into an oxygen vacancy, D is the cation vacancy diffusivity, and α is the polarizability of the film–solution interface. Yang and Macdonald[179] extended eqn [131] to systems containing both aggressive ions X^- and inhibitive ions Y^{z-}:

$$E_b = E_b^0 - \frac{\beta}{\alpha}\text{pH} - \frac{2.303}{\alpha\alpha_0 F}\log\frac{a_{X^-}}{a_{Y^{z-}}} \quad [132]$$

where the constant E_b^0 is a function of adsorption and elementary reaction rate parameters that is derived from a competitive adsorption model for the X^- and Y^{z-} species, β is the dependence of the potential drop across the barrier layer–solution interface on pH, and α_0 is a transfer coefficient. The predictions of the point defect model have been found to be in agreement with experimental phenomena including the linear dependence of the breakdown potential on the concentrations of aggressive and inhibitive ions, the dependence of the induction time on potential and chloride concentration, dependence of the breakdown potential on the scan rate and the inhibition of pitting by Mo and W in the alloy. It should be noted that the breakdown potential is a distributed quantity that can be described with a normal distribution function. The distribution in E_b has been reproduced by assuming that the cation diffusivity is normally distributed.[97]

While the logarithmic dependence of the breakdown potential on the aggressive species concentration is predicted by most passivity breakdown models, the induction time provides a more stringent criterion for testing alternative models. Accordingly, Milošev *et al.*[180] tested the validity of the point defect model, the two-dimensional nucleation model of Heusler and Fischer,[173] and the halide nuclei model of Okada[176] for the pitting of copper. The point defect model was found to yield the best agreement with experimental data.

2.38.3.8.2 Repassivation potential and its use to predict localized corrosion

While the breakdown potential is the critical parameter for the initiation of pitting, the repassivation potential (E_{rp}) has been used for predicting the long-term occurrence of pitting and crevice corrosion. The repassivation potential (also called protection potential) is the potential at which a stably growing pit or crevice corrosion will cease to grow. Thus, localized corrosion cannot occur at potentials below E_{rp}. The use of E_{rp} for engineering predictions can be justified by the fact that only the fate of stable pits or crevice corrosion is important for predicting the possibility of failure, and metastable pits do not adversely affect the performance of engineering structures. It has been shown by Dunn *et al.*[181,182] that E_{rp} is practically independent of the amount of charge passed in a localized corrosion process as long as it is above a certain minimum amount of charge. As a result, the repassivation potential is relatively insensitive to prior pit depth and surface finish. As a corollary, it has been shown that the repassivation potential for pitting (i.e., measured on an open sample) and the repassivation potential for crevice corrosion (i.e., measured on a creviced sample) coincide at high pit depths. This has demonstrated the utility of the repassivation potential for engineering design as it provides a reproducible and inherently conservative threshold for the occurrence of localized corrosion. Thus, the prediction of long-term occurrence of localized corrosion can be separated into two independent parts, that is, the calculation of the repassivation and the corrosion potentials. The separation of localized corrosion modeling into these two steps is valid as long as the initial stages of stable localized corrosion are considered because the corrosion potential is not affected at this stage by the progress of the localized corrosion process and the interaction between pits can be ignored. The separation remains valid as long as significant pit or crevice corrosion growth does not occur and the area of an actively corroding pit does not become significant compared to the overall area.

A model for calculating the repassivation potential has been developed by Anderko *et al.*[183] by considering the electrochemistry of a metal M that undergoes dissolution underneath a layer of concentrated metal halide solution MX. The concentrated solution may or may not be saturated with respect to a hydrous solid metal halide. In the process of repassivation, a thin layer of oxide forms at the interface between the metal and the hydrous metal halide. The model assumes that, at a given instant, the oxide layer covers a certain fraction of the metal surface. This fraction increases as repassivation is approached. Further, the model includes the effects of multiple aggressive and nonaggressive or inhibitive species, which are taken into account through a competitive adsorption scheme. The aggressive species form metal complexes, which dissolve in the active state. On the other hand, the inhibitive species and water contribute to the formation of oxides, which induce passivity. The model assumes that the measurable potential drop across the interface can be expressed as a sum of four contributions, that is

$$E = \Delta\Phi_{M/MX} + \Delta\Phi_{MX} + \Delta\Phi_{MX/S} + \Delta\Phi_{S} \qquad [133]$$

where $\Delta\Phi_{M/MX}$ is the potential difference at the interface between the metal and metal halide, which may be influenced by the partial coverage by the metal oxide, $\Delta\Phi_{MX}$ is the potential drop across the hydrous halide layer, $\Delta\Phi_{MX/S}$ is the potential difference across the metal halide–solution interface, and $\Delta\Phi_{S}$ is the potential drop across the boundary layer within the solution. Expressions for the potential drops can be derived using the methods of nonequilibrium thermodynamics.[184] In general, these expressions are complex and can be solved only numerically. However, a closed-form equation has been found in the limit of repassivation, that is, when the current density reaches a predetermined low value i_{rp} (typically $i_{rp} = 10^{-2}$ A m^{-2}) and the fluxes of metal ions become small and comparable to those for passive dissolution. Then, eqn [133] can be used to arrive at a closed-form expression for the repassivation potential. This closed-form expression, which can be solved numerically to calculate E_{rp}, is given by:

$$1 + \sum_k \left[\left(\frac{i_{rp}}{i_p} - 1\right)\frac{l_k''}{i_{rp}}\right]\theta_k^{n_k} \exp\left(\frac{\xi_k F E_{rp}}{RT}\right) = \sum_j \frac{k_j''}{i_{rp}}\theta_j^{n_j}\exp\left(\frac{\alpha_j F E_{rp}}{RT}\right) \qquad [134]$$

where i_p is the passive current density, T is the temperature, R is the gas constant, and F is the Faraday constant. The partial coverage fraction of a species j is related to the activity of this species in the bulk solution by

$$\theta_j = \frac{K_{\text{ads},j} a_j}{1 + \sum_k K_{\text{ads},k} a_k} \quad [135]$$

where

$$K_{\text{ads},j} = \exp\left(-\frac{\Delta G_{\text{ads},j}}{RT}\right) \quad [136]$$

and $\Delta G_{\text{ads},j}$ is the Gibbs energy of adsorption. The parameters k''_j and l''_k in eqn [134] are rate constants for surface reactions mediated by the adsorption of aggressive and inhibitive species, respectively. The inhibitive species include water, as it is necessary for oxide formation. The parameter n_j is the reaction order with respect to species j, and α_j and ξ_k are the electrochemical transfer coefficients for reactions mediated by aggressive and inhibitive species, respectively. Some parameters ($\Delta G_{\text{ads},i}$, α_j, and n_k) can be assigned default values. The remaining parameters need to be regressed from a limited amount of experimental E_{rp} measurements. Since E_{rp} data are most abundant for chloride solutions, the rate constant for the chloride ions (k''_{Cl}), reaction order with respect to chlorides (n_{Cl}), rate constant for water ($l''_{\text{H}_2\text{O}}$), and electrochemical transfer coefficient for water (ξ_{H2O}) are determined based on the data for chloride solutions. The determination of parameters is greatly simplified by the fact that the parameters for Fe–Ni–Cr–Mo–W–N alloys can be correlated with alloy composition,[185,186] thus enhancing the predictive value of the model. The k''_j and, if necessary, n_j parameters are determined for other aggressive species j (e.g., bromide ions) using E_{rp} data for either pure or mixed solutions containing such ions. Finally, the l''_k parameters for inhibitive ions k are determined on the basis of data for mixed solutions containing chlorides and inhibitors. Data for mixed systems are necessary because E_{rp} is undefined in solutions containing only inhibitors.

The repassivation potential model has a limiting character, that is, it accurately represents the state of the system in the repassivation potential limit. In addition to the value of the repassivation potential, the model predicts the correct slope of the current density versus potential relationship as the potential deviates from E_{rp}.[183] The current density predicted by the model as a function of potential is given by

$$i = \frac{\sum_j k''_j \theta_j^{n_j} \exp\left(\frac{\alpha_j FE}{RT}\right) + \sum_j l''_j \theta_j^{n_j} \exp\left(\frac{\xi_j FE}{RT}\right)}{1 + \frac{1}{i_p}\sum_j l''_j \theta_j^{n_j} \exp\left(\frac{\xi_j FE}{RT}\right)} \quad [137]$$

Equation [137] reduces to eqn [134] for $E = E_{\text{rp}}$ and $i = i_{\text{rp}}$. Since eqn [137] is a limiting law, its accuracy gradually deteriorates as the potential increasingly deviates from E_{rp}. Equation [137] cannot be regarded as a model for the propagation rate of an actively growing pit or crevice because it does not take into account the factors such as the ohmic potential drop, transport limitations, and so on. However, the current density predicted using eqn [137] for $E > E_{\text{rp}}$ is useful because it provides an estimate of the maximum propagation rate of an isolated pit as a function of potential. Such an upper estimate is convenient because it relies only on parameters that are calibrated using repassivation potential data.

Figure 10 shows the application of the repassivation potential model to alloy CuNi 7030 in chloride

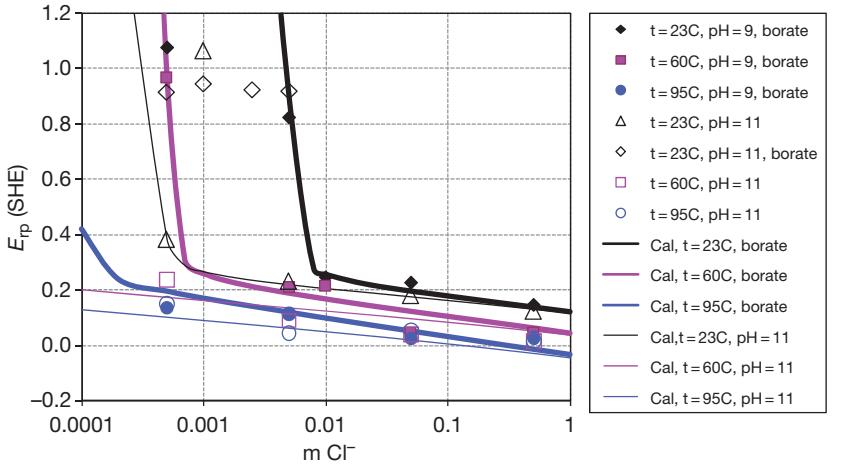


Figure 10 Modeling the repassivation potential of alloy CuNi 7030 (UNS C71500).[217] as a function of chloride activity at various temperatures.

solutions at three temperatures. As shown in the figure, the slope of the repassivation potential changes as a function of chloride activity. A steeper slope is observed at low chloride concentrations. This is a general phenomenon for alloys and becomes more pronounced for more corrosion-resistant alloys.[183] The transition between the low-slope and high-slope segments of the curves strongly depends on the alloy and temperature.

A particularly useful application of the repassivation potential model is for investigating the competing effects of aggressive and inhibitive species. For example, **Figure 11** shows the inhibitive effect of nitrate ions on localized corrosion of alloy 22 in concentrated chloride solutions. The E_{rp} versus NO_3^- concentration curves have a characteristic shape with two distinct slopes. As the concentration of the NO_3^- ions is increased, the slope of the E_{rp} versus NO_3^- concentration curve initially slowly increases with a low slope. At a certain concentration of NO_3^-, the slope of the E_{rp} curve rapidly increases and the repassivation potential attains a high value. At NO_3^- concentrations that lie beyond the high-slope portion of the E_{rp} versus NO_3^- curve, localized corrosion becomes impossible even in systems with a high corrosion potential. Thus, there is a fairly narrow range of inhibitor concentrations over which the E_{rp} curve transitions from a low-slope region (in which localized corrosion is possible depending on the value of the corrosion potential) to a high-slope region that constitutes the upper limit of inhibitor concentrations for localized corrosion. The exact location of the transition region depends on the temperature and chloride concentration and can be accurately reproduced using the repassivation potential model.

Figure 12 illustrates an application of the corrosion potential and repassivation potential models to predict the critical crevice temperature. At temperatures below critical crevice temperature (CCT), the calculated corrosion potential (E_{corr}) should lie below the repassivation potential, whereas it should exceed E_{rp} above CCT. Thus, the intersection of the E_{corr} and E_{rp} curves versus temperature provides an estimate of CCT. **Figure 12** shows the results of such calculations for alloy C-276 in 6% $FeCl_3$ solutions.[188] The repassivation potential shows an initially steep decrease followed by a moderate decrease at higher temperatures. On the other hand, the corrosion potential shows a much weaker temperature dependence. The intersection points of the E_{corr} and E_{rp} curves can be compared with experimental critical crevice temperatures (Hibner[189]).

It should be noted that while the approach based on computing E_{corr} and E_{rp} can predict the long-term occurrence and maximum propagation rate of localized corrosion, it gives no spatial or temporal information. For predicting the spatial and temporal evolution of localized corrosion, models are required that include a detailed treatment of mass transport and take into account the geometric constraints of crevices, pits, and so on. Such models are outside the scope of this chapter.

2.38.3.9 Selected Practical Applications of Aqueous Corrosion Modeling

In this section, we briefly outline selected models that have been developed for practical applications on the basis of the principles discussed above.

Extensive efforts have been devoted to the modeling of aqueous corrosion in oil and gas environments.

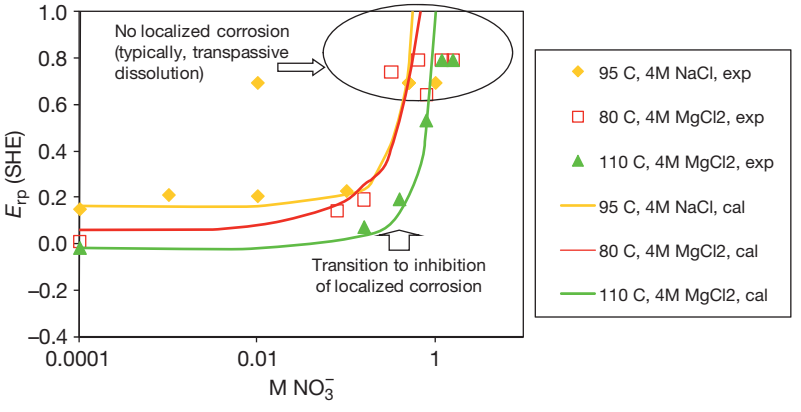


Figure 11 Modeling the effect of nitrate ions on the repassivation potential of alloy 22 in concentrated chloride solutions. The experimental data are from Dunn *et al*.[187]

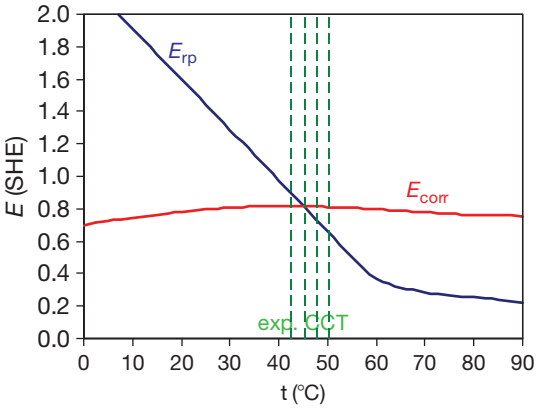


Figure 12 Prediction of the critical crevice temperature (CCT) for various alloys in a 6% $FeCl_3$ solution using corrosion potential and repassivation potential models.[188] The vertical lines show the location of the experimental CCT values.[189] The intersection of the calculated corrosion potential and repassivation potential lines shows the predicted critical crevice temperature.

This is due to the great practical importance of corrosion in oil and gas production and transmission, and to the fact that the number of key corrosive components in such environments is relatively limited (primarily to CO_2, H_2S, acetic acid, and O_2), thus making the modeling task manageable despite the inherent complexity of corrosion mechanisms and their dependence of flow conditions. Corrosion modeling in this area has been reviewed by Nešić *et al.*,[190] Nyborg,[191] and Papavisanam *et al.*[192–194] The CO_2/H_2S models range from expert systems based on laboratory and field data to mechanistic models that recognize the partial electrochemical processes in an explicit way. Among the electrochemical models, the treatment of flow effects using mass transfer coefficients has found wide applicability for both single-phase and two-phase flow (Nešić *et al.*,[115] Dayalan *et al.*,[195] Anderko and Young,[169] Pots and Kapusta,[196] Deng *et al.*[197]). While the methodology for modeling CO_2 corrosion is well established, the modeling of H_2S effects is still in a state of flux and is a subject of significant research efforts (Anderko and Young,[169] Nešić *et al.*,[170] Sun and Nešić[171]). A CO_2 corrosion model based on a detailed treatment of transport (eqns [100]–[103]) has been developed by Nordsveen *et al.*,[116] Nešić *et al.*,[198] and Nešić and Lee.[166] This model has been further extended to CO_2/H_2S corrosion by considering mechanistic aspects of FeS scale formation (Nešić *et al.*,[170] Sun and Nešić[171]). The model has been integrated with a multiphase flow model, which made it possible to predict the effect of water entrainment and water wetting in oil–water systems (Nešić *et al.*[170]). Also, a stochastic algorithm for predicting localized corrosion due to partially protective $FeCO_3$ scales has been integrated with this model (Nešić *et al.*[199]).

Another area in which electrochemical models have found wide applicability is the prediction of the corrosion potential in dilute aqueous solutions that exist in power generation industries. In particular, mixed-potential models have been developed for stainless steels in water as a function of oxygen and hydrogen content (Macdonald,[161] Lin *et al.*,[138] Kim[103]). Such models are useful for calculating E_{corr} in boiling water reactors and related environments. E_{corr} is then further used for localized corrosion modeling. In such systems, modeling of the initiation and evolution of crevice corrosion, stress corrosion cracking, and corrosion fatigue is of primary importance. For this purpose, several electrochemical models that are suitable for alloys in high temperature and low-conductivity water have been developed (see Betts and Boulton,[200] Turnbull,[201–203] Engelhardt *et al.*[204,205] and references therein).

Corrosion in halide-containing natural and industrial environments such as seawater, deliquescing liquids, or production brines is another important application for modeling. For example, King *et al.*[123] developed a model for calculating the corrosion potential of copper in aerated chloride solutions. Anderko and Young[108] modeled general corrosion of steel in concentrated bromide brines used in absorption cooling. Sridhar *et al.*[131] developed a model that predicts both the general corrosion and the occurrence of localized corrosion for stainless steels and aluminum in seawater. A mixed-potential model has been developed to predict the behavior of nuclear fuel in steel containers (Shoesmith *et al.*[206]). Models for the initiation, stabilization, and propagation of pitting, crevice corrosion, and cracking in halide systems have been developed by a number of authors (Turnbull and Ferris,[207] Turnbull,[203,208] Betts and Boulton,[200] Engelhardt *et al.*,[209,210] Cui *et al.*,[211] and references therein).

Modeling corrosion in the process industries is a potentially fruitful area but is subject to great difficulties because of the complex and variable nature of the chemical environments. In principle, acid systems are amenable to modeling because of their well-defined chemistry. In particular, corrosion in acids has been modeled by Sridhar and Anderko[212] in the moderate concentration range. Rahmani and Strutt[213] developed a model for very concentrated sulfuric acid solutions, in which corrosion can be assumed to

be exclusively under mass transport control. Veawab and Arronwilas[214] developed a model for the general corrosion in amine–CO_2 systems. Models are also available for the corrosivity in wet porous media (Huet *et al.*[215]). Anderko *et al.*[216] applied the localized corrosion model based on calculating the repassivation and corrosion potentials to predict the occurrence of pitting and estimate the worst-case propagation rates of localized corrosion in a process environment.

2.38.4 Concluding Remarks

Over the past three decades, tremendous progress has been achieved in the development of computational models of aqueous corrosion. A number of practically important models have been developed for applications as diverse as oil and gas production and transmission, nuclear and fossil power generation, seawater service, and various chemical processes.

Thermodynamic models of electrolyte systems have reached a level of sophistication that extended their applicability range from dilute aqueous solutions to multiphase, multicomponent systems ranging from infinite dilution to solid saturation or pure solute limits. Although they were originally developed mostly for applications other than corrosion (especially chemical processing and geology), they are increasingly used to predict the solution chemistry of corrosive environments and to understand the effect of phase behavior on corrosion. Electrochemical models of corroding interfaces have been developed to predict the kinetics of anodic and cathodic reactions that are responsible for corrosion and to relate them to bulk solution chemistry and flow conditions. Semiempirical and mechanistic models of passivity have been developed to predict the behavior of passive metals and the breakdown of passivity. Also, models are available to predict the threshold conditions for localized corrosion.

The main focus of the models reviewed in this chapter is on relating the chemistry of the environment to electrochemical corrosion phenomena on metal surfaces. However, this is often only the first stage of corrosion modeling. Beyond this stage, the models discussed here serve as a basis for simulating the spatial and temporal evolution of localized and general corrosion damage in various engineering structures subject to localized and general corrosion. This level of modeling will be discussed in other chapters of this volume.

Acknowledgments

The work reported in this paper has been supported by the Department of Energy (award number DE-FC36-04GO14043) and cosponsored by Chevron-Texaco, DuPont, Haynes International, Mitsubishi Chemical, Shell, and Toyo Engineering.

References

1. Pourbaix, M. *Atlas of Electrochemical Equilibria in Aqueous Solutions*; Pergamon Press: New York, NY, 1966.
2. Pitzer, K. S. *Thermodynamics*, 3rd ed.; McGraw-Hill: New York, NY, 1995.
3. Anderko, A.; Wang, P.; Rafal, M. *Fluid Phase Equil.* **2002**, *194–197*, 123–142.
4. Zemaitis, J. F., Jr.; Clark, D. M.; Rafal, M.; Scrivner, N. C. *Handbook of Aqueous Electrolyte Thermodynamics*; DIPPR, AIChE: New York, NY, 1986.
5. Renon, H. *Fluid Phase Equil.* **1986**, *30*, 181–195.
6. Pitzer, K. S. Ed. *Activity Coefficients in Electrolyte Solutions*, 2nd ed.; CRC Press: Boca Raton, FL, 1991.
7. Rafal, M.; Berthold, J. W.; Scrivner, N. C.; Grise, S. L. In *Models for Thermodynamic and Phase Equilibria Calculations*; Sandler, S. I., Ed.; Marcel Dekker: New York, NY, 1994; pp 601–670.
8. Loehe, J. R.; Donohue, M. D. *AIChE J.* **1997**, *43*, 180–195.
9. Chase, M. W.; Davies, C. A.; Downey, J. R.; Frurip, D. J.; McDonald, D. A.; Syverud, A. N. *JANAF Thermochemical Tables,* 3rd ed., supplement no. 1, *J. Phys. Chem. Ref. Data* **1985**, *14*, 1–1856.
10. Barin, I.; Platzki, G. *Thermochemical Data of Pure Substances*, 3rd ed.; Wiley-VCH: Verlag, 1997.
11. Cox, J. D.; Wagman, D. D.; Medvedev, V. A. *CODATA Key Values for Thermodynamics*; Hemisphere Publishing: New York, 1989.
12. Glushko, V. P.; Medvedev, V. A.; Bergman, G. A.; Vasil'ev, B. P.; Kolesov, V. P.; Gurvich, L. V.; Yungman, V. S.; Khodakovskii, I. L.; Resnitskii, L. A.; , Smirnova, N. L.; *et al. Thermal Constants of Compounds;* Academy of Sciences: Moscow, USSR, 1965–1981.
13. Gurvich, L. V.; Veyts, I. V.; Alcock, C. B.; Iorish, V. S. *Thermodynamic Properties of Individual Substances;* Hemisphere Publishing: New York, 1988–1998; Vols 1–5.
14. Kelly, K. K. Contribution to the Data on Theoretical Metallurgy XIII. High Temperature Heat Content, Heat Capacity and Entropy Data for the Elements and Inorganic Components, Bull. 584, Bureau of Mines, US Dept. of Interior, 1960.
15. Robie, R. A.; Hemingway, B. S.; Fisher, J. R. *US Geol. Survey Bull* **1978**, 1452.
16. Shock, E. L.; Helgeson, H. C. *Geochim. Cosmochim. Acta* **1988**, *52*, 2009–2036; *Geochim. Cosmochim. Acta* **1990**, *54*, 915–943.
17. Shock, E. L.; Helgeson, H. C.; Sverjensky, D. A. *Geochim. Cosmochim. Acta* **1989**, *53*, 2157–2183.
18. Shock, E. L.; Sassani, D. C.; Willis, M.; Sverjensky, D. A. *Geochim. Cosmochim. Acta* **1997**, *61*, 907–950.
19. Stull, D. R.; Westrum, E. F., Jr.; Sinke, G. L. *The Chemical Thermodynamics of Organic Compounds*; R. F. Krieger: Malabara, FL, 1987.

20. Wagman, D. D.; Evans, W. H.; Parker, V. B.; Shum, R. H.; Halow, I.; Bailey, S. M.; Churney, K. L.; Nuttal, R. L. *J. Phys. Chem. Ref. Data* **1982**, *11*(2), 1–392.
21. Criss, C.; Cobble, J. *J. Am. Chem. Soc.* **1964**, *86*, 5385–5390; *J. Am. Chem. Soc*. **1964**, *86*, 5390–5393.
22. Helgeson, H. C.; Kirkham, D. H.; Flowers, G. C. *Am. J. Sci.* **1974**, *274*, 1089–1198; *Am. J. Sci.* **1974**, *274*, 1199–1261; *Am. J. Sci.* **1976**, *276*, 97–240; *Am. J. Sci.* **1981**, *281*, 1241–1516.
23. Tanger, J. C., IV; Helgeson, H. C. *Am. J. Sci.* **1988**, *288*, 19–98.
24. Sverjensky, D. A.; Shock, E. L.; Helgeson, H. C. *Geochim. Cosmochim. Acta* **1997**, *61*, 1359–1412.
25. Johnson, J. W.; Oelkers, E. H.; Helgeson, H. C. *Comput. Geosci.* **1992**, *18*, 899–947.
26. Sedlbauer, J.; O'Connell, J. P.; Wood, R. H. *Chem. Geol.* **2000**, *163*, 43–63.
27. Sedlbauer, J.; Majer, V. *Eur. J. Mineral.* **2000**, *12*, 1109–1122.
28. Kolker, A.; de Pablo, J. J. *Ind. Eng. Chem. Res.* **1996**, *35*, 228–233; *Ind. Eng. Chem. Res*. **1996**, *35*, 234–240.
29. Pitzer, K. S. *J. Phys. Chem.* **1973**, *77*, 268–277.
30. Debye, P.; Hückel, E. *Physik. Z.* **1923**, *24*, 185–206; *Physik. Z*. **1923**, *24*, 305–325; *Physik. Z*. **1924**, *25*, 97–107.
31. Wu, R.-S.; Lee, L. L. *Fluid Phase Equil.* **1992**, *78*, 1–24.
32. Papaiconomou, N.; Simonin, J.-P.; Bernard, O.; Kunz, W. P. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4435–4443.
33. Guggenheim, E. A. *Philos. Mag* **1935**, *19*, 588–643.
34. Guggenheim, E. A.; Turgeon, J. C. *Trans. Faraday Soc.* **1955**, *51*, 747–761.
35. Helgeson, H. C. *Am. J. Sci.* **1969**, *267*, 729.
36. Bromley, L. A. *AIChE J.* **1973**, *19*, 313–320.
37. Zemaitis, J. F., Jr. *ACS Symp. Series* **1980**, *133*, 227–246.
38. Meissner, H. P. *ACS Symp. Series* **1980**, *133*, 495–511.
39. Pitzer, K. S.; Simonson, J. M. *J. Phys. Chem.* **1986**, *90*, 3005–3009.
40. Clegg, S. L.; Pitzer, K. S. *J. Phys. Chem.* **1992**, *96*, 3513–3520.
41. Chen, C.-C.; Britt, H. I.; Boston, J. F.; Evans, L. B. *AIChE J.* **1982**, *28*, 588–596.
42. Liu, Y.; Watanasiri, S. *Fluid Phase Equil.* **1996**, *116*, 193–200.
43. Abovsky, V.; Liu, Y.; Watanasiri, S. *Fluid Phase Equil.* **1998**, *150*, 277–286.
44. Chen, C.-C.; Mathias, P. M.; Orbey, H. *AIChE J.* **1999**, *45*, 1576–1586.
45. Chen, C.-C.; Song, Y. *AIChE J.* **2004**, *50*, 1928–1941.
46. Sander, B.; Fredenslund, A.; Rasmussen, P. *Chem. Eng. Sci.* **1986**, *41*, 1171–1183.
47. Macedo, E. A.; Skovborg, P.; Rasmussen, P. *Chem. Eng. Sci.* **1990**, *45*, 875–882.
48. Kikic, I.; Fermeglia, M.; Rasmussen, P. *Chem. Eng. Sci.* **1991**, *46*, 2775–2780.
49. Dahl, S.; Macedo, E. A. *Ind. Eng. Chem. Res.* **1992**, *31*, 1195–1201.
50. Iliuta, M. C.; Thomsen, K.; Rasmussen, P. *Chem. Eng. Sci.* **2000**, *55*, 2673–2686.
51. Li, J.; Polka, H.-M.; Gmehling, J. *Fluid Phase Equil.* **1994**, *94*, 89–114.
52. Yan, W.; Topphoff, M.; Rose, C.; Gmehling, J. *Fluid Phase Equil.* **1999**, *162*, 97–113.
53. Zerres, H.; Prausnitz, J. M. *AIChE J.* **1994**, *40*, 676–691.
54. Wang, P.; Anderko, A.; Young, R. D. *Fluid Phase Equil.* **2002**, *203*, 141–176.
55. Wang, P.; Anderko, A.; Springer, R. D.; Young, R. D. *J. Molec. Liq.* **2006**, *125*, 37–44.
56. Prausnitz, J. M.; Lichtenthaler, R. N.; Gomes de Azevedo, E. *Molecular Thermodynamics of Fluid Phase Equilibria*; Prentice-Hall: Englewood Cliffs, NJ, 1998.
57. Malanowski, S.; Anderko, A. *Modeling Phase Equilibria: Thermodynamic Background and Practical Tools*; Wiley: New York, 1992.
58. Chen, C. M.; Aral, K. *Corrosion* **1982**, *38*, 183–190.
59. Chen, C. M.; Aral, K.; Theus, G. J. *Computer Calculated Potential–pH Diagrams to 300 °C,* EPRI NP 3137; Electric Power Research Institute: Palo Alto, CA, 1983; Vols. 1–2.
60. Partridge, E.; Hall, R. *Trans. Am. Soc. Mech. Eng* **1939**, *61*, 597.
61. OLI Systems Inc. Corrosion Analyzer software version 2.0, www.olisystems.com, 2007.
62. Anderko, A.; Sanders, S. J.; Young, R. D. *Corrosion* **1997**, *53*, 43–53.
63. Macdonald, D. D.; Cragnolino, G. A. In *The ASME Handbook on Water Technology for Thermal Power Systems*; Cohen, P., Ed.; American Society of Mechanical Engineers: New York, 1989; Chapter 9.
64. Beverskog, B.; Puigdomenech, I. *Corros. Sci.* **1996**, *38*, 2121–2135.
65. Beverskog, B.; Puigdomenech, I. *J. Electrochem. Soc.* **1997**, *144*, 3476–3483.
66. Beverskog, B.; Puigdomenech, I. *Corros. Sci.* **1997**, *39*, 43–57.
67. Beverskog, B.; Puigdomenech, I. *Corros. Sci.* **1997**, *39*, 969–980.
68. Beverskog, B.; Puigdomenech, I. *Corros. Sci.* **1997**, *39*, 107–114.
69. Beverskog, B.; Puigdomenech, I. *Corrosion* **1999**, *55*, 1077–1087.
70. Biernat, R. J.; Robbins, R. G. *Electrochim. Acta* **1972**, *17*, 1261–1283.
71. Froning, M. H.; Shanley, M. E.; Verink, E. D., Jr. *Corros. Sci.* **1976**, *16*, 371–376.
72. Macdonald, D. D.; Syrett, B. C. *Corrosion* **1979**, *35*, 471–474.
73. Macdonald, D. D.; Syrett, B. C.; Wing, S. S. *Corrosion* **1979**, *35*, 1–11.
74. Macdonald, D. D.; Syrett, B. C.; Wing, S. S. *Corrosion* **1979**, *35*, 367–378.
75. Anderko, A.; Shuler, P. J. *Comput. Geosci.* **1997**, *23*, 647–658.
76. Pourbaix, M. In *Localized Corrosion*; Staehle, R. W., Brown, B. F., Kruger, J., Agarwal, A., Eds.; NACE: Houston, TX, 1974; pp 12–33.
77. Kesavan, S.; Mozhi, T. A.; Wilde, B. E. *Corrosion* **1989**, *45*, 213–214.
78. Muñoz-Portero, M. J.; Garcia-Anton, J.; Guiñon, J. L.; Perez-Herranz, V. *Corrosion* **2004**, *60*, 749–756.
79. Bianchi, G.; Longhi, P. *Corros. Sci.* **1973**, *13*, 853–864.
80. Silverman, D. C. *Corrosion* **1982**, *38*, 541–549.
81. Silverman, D. C. *Corrosion* **1988**, *44*, 606–610.
82. Kubal, M.; Panacek, F. *Br. Corros. J.* **1995**, *30*, 309–311.
83. Silverman, D. C.; Silverman, A. L. *Corrosion/2007*; NACE International: Houston, TX, 2007; Paper no. 07616.
84. Cubicciotti, D. *J. Nucl. Mat.* **1989**, *167*, 241–248.
85. Cubicciotti, D. *J. Nuclear Mat.* **1993**, *201*, 176–183.
86. Yang, L. X.; Yang, X. Z.; Pourbaix, A. *Corrosion/2001*; NACE International: Houston, TX, 2001; Paper no. 01084.
87. Anderko, A.; Wang, P.; Young, R. D.; Riemer, D. P.; McKenzie, P.; Lencka, M. M.; Babu, S. S.; Angelini, P. Prediction of Corrosion of Alloys in Mixed-Solvent Environments, Final report, DE-FC07–00CH11019, DOE OIT, 2003, www.osti.gov/servlets/purl/811533-wUOQwz/native/.
88. Lupis, C. H. *Chemical Thermodynamics of Materials*; North Holland: New York, 1983.

89. Saunders, N.; Miodownik, A. P. *CALPHAD – Calculation of Phase Diagrams – A Comprehensive Guide*; Pergamon/Elsevier Science: New York, 1998.
90. Hertzman, S. *Metall. Trans. A* **1987**, *18A*, 1753–1766; *Metall. Trans. A*, **1987**, *18A*, 1767–1778.
91. Hertzman, S.; Jarl, M. *Metall. Trans. A* **1987**, *18A*, 1745–1752.
92. Andersson, J.-O.; Lange, N. *Metall. Trans. A* **1988**, *19A*, 1385–1394.
93. Sridhar, N.; Dunn, D. S.; Anderko, A.; Lencka, M. M.; Schutt, H. U. *Corrosion* **2001**, *57*, 221–235.
94. Lencka, M. M.; Nielsen, E.; Anderko, A.; Riman, R. E. *Chem. Mater.* **1997**, *9*, 1116–1125.
95. Sridhar, N.; Dunn, D. S.; Anderko, A. In *Environmentally Assisted Cracking: Predictive Methods for Risk Assessment and Evaluation of Materials, Equipment and Structures*; ASTMP STP 1401; Kane, R. D., Ed.; American Society for Testing and Materials: West Conshohocken, PA, 2000; p 241.
96. Mohr, D. W.; McNeil, M. B. *J. Nucl. Mater.* **1992**, *190*, 329–342.
97. Macdonald, D. D. *Pure Appl. Chem.* **1999**, *71*, 951–978.
98. Vetter, K. J. *Electrochemical Kinetics*; Academic Press: New York, London, 1967.
99. Bockris, J. O'M.; Reddy, A. K. N. *Modern Electrochemistry*; Plenum Press: New York, 1970.
100. Kaesche, H. *Metallic Corrosion*; NACE International: Houston, TX, 1985.
101. Bockris, J. O'M.; Khan, S. U. M. *Surface Electrochemistry: A Molecular Level Approach*; Plenum Press: New York, London, 1993.
102. Gileadi, E. *Electrode Kinetics*; VCH Publishers: New York, 1993.
103. Kim, Y. J. *Corrosion* **2002**, *58*, 208–215.
104. Wagner, C.; Traud, W. *Z. Elektrochem* **1938**, *44*, 391–402.
105. Gileadi, E. Ed. *Electrosorption*; Plenum Press: New York, 1967.
106. Damaskin, B. B.; Petrii, O. A.; Batrakov, V. V. *Adsorption of Organic Compounds on Electrodes*; Plenum Press: New York, 1971.
107. Habib, M. A.; Bockris, J. O'M. In *Comprehensive Treatise of Electrochemistry*; Bockris, J. O'M., Conway, B. E., Yeager, E., Eds.; Plenum Press: New York, 1980; Vol. 1.
108. Anderko, A.; Young, R. D. *Corrosion* **2000**, *56*, 543–555.
109. Heusler, K. E.; Cartledge, G. E. *J. Electrochem. Soc.* **1961**, *108*, 732–740.
110. Lorenz, W.; Heusler, K. In *Corrosion Mechanisms*; Mansfeld, F., Ed.; Marcel Dekker: New York, 1987.
111. Drazic, D. M. In *Modern Aspects of Electrochemistry*; Conway, B. E., Bockris, J. O'M., White, R. E., Eds.; Plenum Press: New York, 1989; Vol. 19, pp 69–192.
112. Keddam, M. In *Corrosion Mechanisms in Theory and Practice*; Marcus, P., Oudar, J., Eds.; Marcel Dekker: New York, 1995.
113. Bockris, J. O'M.; Drazic, D.; Despic, A. R. *Electrochim. Acta* **1961**, *4*, 325–361.
114. Smart, N. G.; Bockris, J. O'M. *Corrosion* **1992**, *48*, 277–280.
115. Nešić, S.; Postlethwaite, J.; Olsen, S. *Corrosion* **1996**, *52*, 280–294.
116. Nordsveen, M.; Nešić, S.; Nyborg, R.; Stangeland, A. *Corrosion* **2003**, *59*, 443–456.
117. Chin, R. J.; Nobe, K. *J. Electrochem. Soc.* **1972**, *119*, 1457–1461.
118. Kuo, H. C.; Nobe, K. *J. Electrochem. Soc.* **1978**, *125*, 853–860.
119. Drazic, V. J.; Drazic, D. M. *J. Serb. Chem. Soc.* **1992**, *57*, 917–926.
120. Kear, G.; Barker, B. D.; Walsh, F. C. *Corros. Sci.* **2004**, *46*, 109–135.
121. Lee, H. P.; Nobe, K. *J. Electrochem. Soc* **1986**, *133*, 2035–2043.
122. Deslouis, C.; Tribollet, B.; Mengoli, G.; Musiani, M. M. *J. Appl. Electrochem* **1988**, *18*, 374–383.
123. King, F.; Litke, C. D.; Quinn, M. J.; LeNeveu, D. M. *Corros. Sci* **1995**, *37*, 833–851.
124. Smart, N. G.; Gamboa-Aldeco, M.; Bockris, J. O'M. *Corros. Sci.* **1993**, *34*, 759–777.
125. Jovancicevic, V.; Bockris, J. O'M. *J. Electrochem. Soc.* **1986**, *133*, 1797–1807.
126. Zecevic, S.; Drazic, D.; Gojkovic, S. *Electrochim. Acta* **1991**, *36*, 5–14.
127. Jovancicevic, V. In *Electrochemistry in Transition*; Murphy, O., Ed.; Plenum Press: New York, 1992; p 127.
128. Le Bozec, N.; Compere, C.; L'Her, M.; Laouenan, A.; Costa, D.; Marcus, P. *Corros. Sci* **2001**, *43*, 765–786.
129. Kapusta, S. D. *Corrosion/2004*; NACE International: Houston, TX, 2004; Paper no. 04655.
130. King, F.; Quinn, M. J.; Litke, C. D. *J. Electroanal. Chem* **1995**, *385*, 45–55.
131. Sridhar, N.; Brossia, C. S.; Dunn, D. S.; Anderko, A. *Corrosion* **2004**, *60*, 915–936.
132. Calvo, E. J.; Schiffrin, D. J. *J. Electroanal. Chem* **1988**, *243*, 171–185.
133. Levich, V. G. *Physicochemical Hydrodynamics*; Prentice-Hall: Englewood Cliffs, NJ, 1962; p 700.
134. Poulson, B. *Corros. Sci* **1983**, *23*, 391–430.
135. Poulson, B. *Corros. Sci.* **1993**, *35*, 655–665.
136. Eisenberg, M.; Tobias, C. W.; Wilke, C. R. *J. Electrochem. Soc.* **1954**, *101*, 306–319.
137. Berger, F. P.; Hau, K.-F. F. L. *Int. J. Heat Mass Trans.* **1977**, *20*, 1185–1194.
138. Lin, C. C.; Kim, Y.-J.; Niedrach, L. W.; Ramp, K. S. *Corrosion* **1996**, *52*, 618–625.
139. Wang, S.; Nešić, S. *Corrosion/2003*; NACE International: Houston, TX, 2003; Paper no. 03631.
140. Chilton, T. H.; Colburn, E. I. *Ind. Eng. Chem* **1934**, *26*, 1183–1187.
141. Frank, M. W. *Viscous Fluid Flow*, 2nd ed.; McGraw-Hill: New York, 1991.
142. Kim, D.; Sofya, Y.; Ghajar, A. J.; Dougherty, R. L. ASME Proceedings, National Heat Transfer Conference **1997**; Vol. 34, pp 119–130.
143. Adsani, E.; Shirazi, S. A.; Shadley, J. R.; Rybicki, E. R. *Corrosion/2002*; NACE International: Houston, TX, 2002; Paper no. 02492.
144. Adsani, E.; Shirazi, S. A.; Shadley, J. R.; Rybicki, E. R. *Corrosion/2006*; NACE International: Houston, TX, 2006; Paper no. 06573.
145. Corti, H.; Trevani, L.; Anderko, A. In *Aqueous Systems at Elevated Temperatures and Pressures; Physical Chemistry in Water, Steam and Hydrothermal Solutions*; Palmer, D. A., Fernandez-Prini, R., Harvey, A. H., Eds.; Academic Press, 2004; Chapter 10, p 752.
146. Anderko, A.; McKenzie, P.; Young, R. D. *Corrosion* **2001**, *57*, 202–213.
147. Ciaraldi, S. W.; Berry, M. R.; Johnson, J. M. *Corrosion/82*; NACE International: Houston, TX, 1982; Paper no. 98.
148. Pawel, S. J. *Corrosion* **1994**, *50*, 963–971.
149. Schmitt, G.; Losacker, S.; Renner, M. H. W.; Horn, E. M. *Corrosion/2004*; NACE International: Houston, TX, 2004; Paper no. 04229.
150. Newman, J.; Thomas-Alyea, K. E. *Electrochemical Systems*, 3rd ed.; Prentice Hall: Englewood Cliffs, NJ, 2004.
151. Davis, J. T. *Turbulence Phenomena*; Academic Press: London, 1972.

152. Lin, C. S.; Moulton, R. W.; Putnam, G. L. *Ind. Eng. Chem* **1953**, *45*, 636–640.
153. Walton, J. C. *Corros. Sci.* **1990**, *30*, 915–928.
154. Bear, J. *Dynamics of Fluids in Porous Media*; Dover Publications: New York, NY, 1972.
155. Nešić, S.; Lee, K.-L. J. *Corrosion* **2003**, *59*, 616–628.
156. Frankenthal, R. P.; Kruger, J. Eds. *Passivity of Metals*; The Electrochemical Society: Princeton, NJ, 1978.
157. Froment, M. Ed. *Passivity of Metals and Semiconductors*; Elsevier: Amsterdam, 1983.
158. Marcus, P.; Oudar, J. Eds. *Corrosion Mechanisms in Theory and Practice*; Marcel Dekker: New York, 1995.
159. Natishan, P. M. Ed. *Passivity and its Breakdown*; The Electrochemical Society: Princeton, NJ, 1998; Vol. PV97-26.
160. Macdonald, D. D. *J. Electrochem. Soc* **2006**, *15*, B213–B224.
161. Macdonald, D. D. *Corrosion* **1992**, *48*, 194–205.
162. Ebersbach, U.; Schwabe, K.; Ritter, K. *Electrochim. Acta* **1967**, *12*, 927–938.
163. Park, H. Y.; Tsuruta, T.; Macdonald, D. D. EPRI Project RP 1166–1 report Electric Power Research Institute: Palo Alto, CA, 1980.
164. Sridhar, N.; Brossia, C. S.; Dunn, D. S.; Buckingham, J. P.; Anderko, A. *Corrosion/2002*; NACE International: Houston, TX, 2002; Paper no. 2204.
165. Leckie, H. P. *Corrosion* **1968**, *24*, 70–74.
166. Blesa, M. A.; Morando, P. J.; Regazzoni, A. E. *Chemical Dissolution of Metal Oxides*; CRC Press: Boca Raton, FL, 1994.
167. Sridhar, N.; Wu, J. B. C.; Corey, S. M. *Mater. Perform.* **1987**, *26*(10), 17–23.
168. Sun, W.; Nešić, S. *Corrosion/2006*; NACE International: Houston, TX, 2006; Paper no. 06365.
169. Anderko, A.; Young, R. D. *Corrosion/99*; NACE International: Houston, TX, 1999; Paper no. 31.
170. Nešić, S.; Cai, J.; Lee, K.-L. J. *Corrosion/2005*; NACE International: Houston, TX, 2005; Paper no. 05556.
171. Sun, W.; Nešić, S. *Corrosion/2007*; NACE International: Houston, TX, 2007; Paper no. 07655.
172. Szklarska-Smialowska, Z. *Pitting and Crevice Corrosion*; NACE International: Houston, TX, 2005.
173. Heusler, K. E.; Fischer, L. *Werkst. Korros* **1976**, *27*, 551–556; *Werkst. Korros. 27*, 788–791.
174. Strehblow, H. H.; Titze, B. *Corros. Sci.* **1977**, *17*, 461–472.
175. Lin, L. F.; Chao, C. Y.; Macdonald, D. D. *J. Electrochem. Soc* **1981**, *128*, 1194–1198.
176. Okada, T. J. *Electrochem. Soc* **1984**, *131*, 241–247.
177. McCafferty, E. *J. Electrochem. Soc* **1990**, *137*, 3731–3737.
178. Haruna, T.; Macdonald, D. D. *J. Electrochem. Soc* **1997**, *144*, 1574–1581.
179. Yang, S.; Macdonald, D. D. *Electrochim. Acta* **2007**, *52*, 1871–1879.
180. Milošev, I.; Metikoš-Huković; Drogowska, M.; Ménard, H.; Brossard, L. *J. Electrochem. Soc.* **1992**, *139*, 2409–2418.
181. Dunn, D. S.; Sridhar, N.; Cragnolino, G. A. *Corrosion* **1996**, *52*, 1583–1612.
182. Dunn, D. S.; Cragnolino, G. A.; Sridhar, N. *Corrosion* **1996**, *56*, 90–104.
183. Anderko, A.; Sridhar, N.; Dunn, D. S. *Corros. Sci.* **2004**, *46*, 1583–1612.
184. Okada, T. *J. Electrochem. Soc.* **1984**, *131*, 1026–1032.
185. Anderko, A.; Sridhar, N.; Brossia, C. S.; Dunn, D. S. *Corrosion/2004*; NACE International: Houston, TX, 2004; Paper no. 04061.
186. Anderko, A.; Sridhar, N.; Tormoen, G.; Brossia, C. S. *Corrosion/2006*; NACE International: Houston, TX, 2006; Paper no. 06215.
187. Dunn, D. S.; Pensado, O.; Pan, Y.-M.; Pabalan, R. T.; Yang, L.; He, X.; Chiang, K. T. Report CNWRA 2005-02 Southwest Research Institute: San Antonio, TX, 2005.
188. Anderko, A.; Sridhar, N.; Brossia, C. S. *Corrosion/2005*; NACE International: Houston, TX, 2005; Paper no. 05053.
189. Hibner, E. L. *Corrosion/86*; NACE International: Houston, TX, 1986; Paper no. 181.
190. Nešić, S.; Postlethwaite, J.; Vrhovac, M. *Corros. Rev* **1997**, *15*, 211–240.
191. Nyborg, R. *Corrosion/2002*; NACE International: Houston, TX, 2002; Paper no. 02233.
192. Papavinasam, S.; Revie, R. W.; Doiron, A. *Corrosion/2005*; NACE International: Houston, TX, 2005; Paper no. 05643.
193. Papavinasam, S.; Revie, R. W.; Doiron, A. *Corrosion/2005*; NACE International: Houston, TX, 2005; Paper no. 05644.
194. Papavinasam, S.; Revie, R. W.; Friesen, W. I.; Doiron, A.; Paneerselvam, T. *Corros. Rev.* **2006**, *24*, 173–230.
195. Dayalan, E.; de Moraes, F. D.; Shadley, J. R.; Shirazi, S. A.; Rybicki, E. F. *Corrosion/98*; NACE International: Houston, TX, 1998; Paper no. 51.
196. Pots, B. F. M.; Kapusta, S. D. *Corrosion/2005*; NACE International: Houston, TX, 2005; Paper no. 05550.
197. Deng, C.; Sand, K.; Teevens, P. J. *Corrosion/2006*; NACE International: Houston, TX, 2006; Paper no. 06565.
198. Nešić, S.; Nordsveen, M.; Nyborg, R.; Stangeland, A. *Corrosion* **2003**, *59*, 489–497.
199. Nešić, S.; Xiao, Y.; Pots, B. F. M. *Corrosion/2004*; NACE International: Houston, TX, 2004; Paper no. 04628.
200. Betts, A. J.; Boulton, L. H. *Br. Corros. J.* **1993**, *28*, 2121–2135.
201. Turnbull, A. *Corros. Sci.* **1993**, *34*, 921–960.
202. Turnbull, A. *Corros. Sci.* **1997**, *39*, 789–805.
203. Turnbull, A. *Corrosion* **2001**, *57*, 175–189.
204. Engelhardt, G.; Macdonald, D. D.; Millett, P. J. *Corros. Sci.* **1999**, *41*, 2165–2190.
205. Engelhardt, G.; Macdonald, D. D.; Urquidi-Macdonald, M. *Corros. Sci.* **1999**, *41*, 2267–2302.
206. Shoesmith, D. W.; Kolar, M.; King, F. *Corrosion* **2003**, *59*, 802–816.
207. Turnbull, A.; Ferris, D. H. *Corros. Sci.* **1987**, *27*, 1323–1350.
208. Turnbull, A. *Br. Corros. J.* **1993**, *28*, 297–308.
209. Engelhardt, G.; Urquidi-Macdonald, M.; Macdonald, D. D. *Corros. Sci.* **1997**, *39*, 419–441.
210. Engelhardt, G.; Urquidi-Macdonald, M.; Macdonald, D. D. *Corros. Sci.* **2004**, *46*, 1159–1187.
211. Cui, F.; Presuel-Moreno, F.; Kelly, R. G. *Corros. Sci.* **2005**, *47*, 2987–3005.
212. Sridhar, N.; Anderko, A. *Corrosion/2001*; NACE International: Houston, TX, 2001; Paper no. 1348.
213. Rahmani, M.; Strutt, J. E. *Hydrodynamic Modeling of Corrosion of Carbon Steels and Cast Irons in Sulfuric Acid, Materials Technology Institute*; NACE International: Houston, TX, 1992.
214. Veawab, A.; Aroonwilas, A. *Corros. Sci.* **2002**, *44*, 967–987.
215. Huet, B.; L'hostis, V.; Santarini, G.; Feron, D.; Idrissi, H. *Corros. Sci.* **2007**, *49*, 1918–1932.
216. Anderko, A.; Sridhar, N.; Yang, L. T.; Grise, S. L.; Saldanha, B. J.; Dorsey, M. H. *Corros. Eng. Sci. Technol* **2005**, *40*, 33–42.
217. Jakab, M. A.; Sridhar, N.; Anderko, A. *Corrosion/2007*; NACE International: Houston, TX, 2007; Paper no. 07243.

2.40 Neural Network Methods for Corrosion Data Reduction*

R. A. Cottis
Corrosion and Protection Centre, School of Materials, University of Manchester, Manchester M60 1QD, UK

This article is a revision of an article that appeared in *Material and Design* by R. A. Cottis, Li Qing, G. Owen, S. J. Garland and I. A. Helliwell, volume 20(4), pp 169–178, © 2010 Elsevier B.V.

2.40.1 Introduction

The reliable prediction of corrosion behavior is a fundamental requirement for the effective control of corrosion. At first sight, this is relatively easy. There is an extensive corrosion literature, and we can simply look up the relevant information. Unfortunately, real world corrosion never seems to involve quite the same conditions as have been tested before, and there is also the difficult question of the inherent variability of the corrosion process. If we are concerned with only one variable such as the effect of temperature on the corrosion of a particular grade of stainless steel in a specific concentration of sulfuric acid, we can plot all the data that we can find and estimate the behavior for the temperature that we are interested in. However, this approach is limited. If we are interested in 22% H_2SO_4, we shall find very few data points, such that we shall probably not be able to plot a reasonable curve. We can extend the number of points available by considering a spread of values, but even then we shall be using only a very small fraction of the relevant data available.

However, in the particular case where we do have a large collection of information, we need techniques to make efficient use of that information. In this chapter, we are concerned with techniques that rely only on the available data. It is important to appreciate that such techniques are necessarily limited to conditions that lie within the region for which we have data. In order to extrapolate outside the range of the available data, we need to use techniques that take advantage of the knowledge of the controlling processes.

2.40.2 Interpolation Techniques

In its simplest form, the problem is one of interpolation. Given the data for points around the point that we are interested in, how do we estimate data for our point? For one or two variables, it is relatively easy to plot graphs, fit lines or surfaces to the graph, estimate standard error and so forth. However, in corrosion, we generally need to consider more than two controlling variables. Even for the very simple case of austenitic stainless steels corroding in sulfuric acid, our data set has a minimum of six controlling variables (Cr, Ni, and Mo content of the steel, temperature and concentration of the acid, and flow conditions). We could easily and justifiably increase the number of variables, for example, by considering minor alloying elements, such as sulfur, phosphorus, and carbon. In the absence of a good model of the behavior, it is difficult to know which variables are significant. If we ignore the effects of significant variables, we introduce uncertainty into the prediction; if we reject data that do not have the 'correct'

*This article is based on a paper originally published in *Materials and Design*.[1]

values of parameters, we ignore much of the information contained within our data set.

We therefore need interpolation methods that can accommodate large numbers of input variables. The most obvious approach is to assume that the corrosion rate (y) depends linearly on the n controlling variables (x_1, x_2, ..., x_n) and that the effects of the variables are independent.

Then, we can write

$$y = a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

where a_0 to a_n are coefficients that we must estimate from the available data.

This approach is quite widely used, and provided we have more data points than controlling variables, it is relatively easy to find the values of a_0 to a_n that minimize the error, using techniques of regression analysis. Unfortunately, the underlying assumptions are rather questionable. Corrosion rate often has a rather nonlinear dependence on the controlling variables, and the effects of the variables are rarely independent. In theory, we can reduce this limitation by fitting higher-order functions, for example, the second-order function

$$y = a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n + a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n + \cdots + a_{nn}x_n^2$$

Note, however, that the second-order function has a large number of coefficients (for $n = 6$, there are 28 coefficients) and the second-order function is not particularly flexible (in one dimension, it is a quadratic). We can get more flexibility by using higher-order functions, but the number of coefficients involved increases markedly.

A better approach is to use rather more flexible functions that can adapt to the data, and thereby fit it better with fewer coefficients. In effect, this is the approach used by artificial neural networks (so-called because they are based on simplified models of animal nervous systems). However, we tend not to describe the artificial neural network by the function that is used to model the data, because the overall function used is very complicated. Rather, we describe the neural network as an assembly of connected components, where each component implements a somewhat simpler function. The components are known as nodes or neurons, and each node has one or more inputs and one output. Several transfer functions can be used for the nodes, but a common one is the sigmoidal function (see **Figure 1**):

$$y = \frac{1}{1 + (1/\exp\ x')}$$

where $x' = a_0 + a_1x_1 + \cdots + a_nx_n$. Then a neural network is constructed by assembling nodes in a network. In principle, the network can be arbitrarily complex, but in practice we generally use fairly simple layered structures, such as that shown in **Figure 2**. Each node of the network takes a number of inputs; each input value (x_1 to x_n) is multiplied by a weight

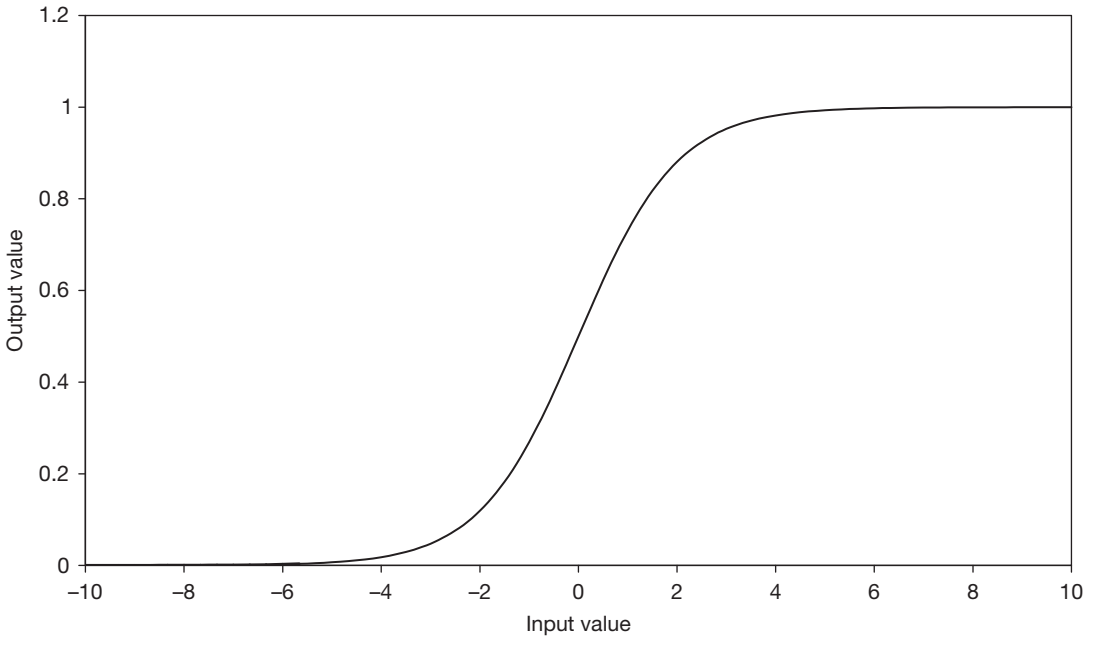


Figure 1 Sigmoidal transfer function typically used by neural network nodes or neurons.

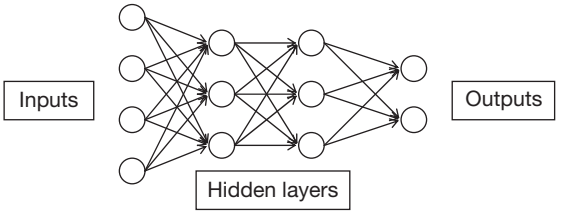


Figure 2 Layered structure of typical artificial neural network.

(a_1 to a_n) and the sum of all the weighted inputs plus a bias (a_0) is used as the input (x') to the sigmoidal function.

As with simpler linear and polynomial functions, the problem is then to find the values of the coefficients that give the 'best' fit between the output of the neural network and the known data. Owing to the complexity of the function implemented by the neural network, we cannot define an analytical method to find the coefficients, and a range of heuristic methods are used to search for the values of the coefficients that give the best fit, usually starting from randomized values. The process of seeking the coefficients is known as 'training' the network (by analogy with the training of biological neural networks).

It is important to appreciate the limitations of neural network methods:

- Neural networks cannot produce reliable predictions for input conditions that are a long way from the data used to train them. The precise meaning of 'a long way' in this context is rather difficult to define, but it is clear that extrapolation beyond the bounds of the training data is unlikely to be reliable (except by good luck).
- If there are insufficient training data for the complexity of the network, the network may 'learn' the individual data points very precisely, but the behavior between the training points will be unpredictable. It has been suggested that there should always be more training points than weights in the network in order to avoid this problem. While this seems plausible, the rule has often been broken, apparently without problems. However, it is difficult to inspect the multidimensional function implemented by the network, and it may therefore be the case that problems with the fit produced by the network have simply not been detected.
- Neural networks are purely phenomenological and do not inherently produce a mechanistic understanding of the process being modeled. However, the generalization of the behavior that is achieved with a well-fitted neural network may well contribute to the development of mechanistic understanding.
- Large training data sets are required, especially when the data are noisy (it might be considered that this should advise against their use to model corrosion processes).

The training of a neural network involves the iterative refinement of the weights for all of the nodes. Typically, this involves applying inputs to the network for which the output values are known. Then, the error (the difference between the value predicted by the network and the known value) is passed back through the network and used to modify the weights so as to reduce the error. In order to help in deciding when to stop training the network, a number of data points (called the validation or selection set) are normally held back, and the network is deemed to be trained as well as possible when the error for the validation set starts to increase. A further set may also be reserved (the test set) to test the error performance of the network on data that have not been used in any way in the training.

2.40.3 Application of Neural Networks to Corrosion

An early published attempt to apply a neural network to a corrosion problem was that of Smets and Bogaerts.[2] They developed a series of neural networks to predict the SCC of type 304 stainless steel in near-neutral solutions as a function of chloride content, oxygen content and temperature. They found that the neural network approach outperformed traditional regression techniques (this should always be the case with a well-fitted network, as it is always possible for the network to model a linear fit). The network used had almost as many weights (35) as the number of examples in the data set used to train it (39). Thus, there may have been a tendency for the network to fit the 'noise' in the data as well as the mean behavior.

Urquidi-Macdonald *et al.*[3] developed a neural network model for predicting the number and depth of pits in heat exchangers. No information was given about the network size (other than that it had two hidden layers) or the number of training points, although there appear to be rather few. Hence, there must be some uncertainty about the reliability of the predictions obtained. On the other hand, the predicted evolution of pit depth and number do appear plausible.

Ben-Hain and Macdonald[4] describe the use of neural network models to predict the influence of various parameters on the acidity of simulated geological brines. The solutions were based on NaCl plus $MgCl_2$; the network inputs were the Na^+ and Mg^{2+} concentration and the temperature, and the output was the predicted pH value. A relatively large training set was available (101 points, of which 90 were used for training, with the remaining 11 retained as a test set), and a very simple network was used (just two hidden nodes in a single layer), and it is perhaps not surprising that the network achieved a good result. The prediction error was of the same order as the experimental uncertainty.

Silverman and Rosen[5] combined artificial neural networks with an expert system in order to predict the type of corrosion from a polarization curve. Inputs to the networks included the passive current density, the pitting potential and the repassivation potential, while outputs were the risks of crevice, pitting and general corrosion. Two approaches were used: independent networks for each type of corrosion, and a single combined network producing all three outputs. The expert system was used to interpret the outputs produced by the two approaches. This combination of neural network with an expert system 'supervisor' offers interesting possibilities, although in this particular case it seems probable that the available training set (87 examples) was not really adequate to describe the dependence of the type of corrosion on the seven input variables. Hence, a larger training set could have led to neural networks that were sufficient by themselves.

Silverman[6] also applied neural network methods to the prediction of service behavior of polymeric linings from laboratory test data. Inputs to the network were parameters obtained from a short-term solvent-uptake test, while the output was a pass–fail description of service behavior. Taking into account the inherent difficulty of this problem, the results obtained seem promising. For example, the neural network was able to produce a correct prediction for three out of four examples of a lining material that was not included in the training data. Silverman's neural network models have since been adapted for web access.[7]

Ramamurthy *et al.*[8] used a neural network to analyze impedance data for automobile paint finishes subjected to stone impacts. The paper contains very little detail on the data set or the network, and it is difficult to assess the performance of the network. However, the data presented suggest that the network is too large, and fits the noise in the training data rather than the mean behavior.

Trasatti and Mazza[9] developed a neural network for the prediction of crevice corrosion behavior of stainless steels. The network was trained from long-term laboratory and field tests. Seventeen input variables were used with one hidden layer of five nodes. Six hundred training examples were available; 450 of these were used for training and the remaining 150 as a test set. The performance of the network was reasonably good, but the very large number of input variables might be expected to present difficulties in training with a relatively small data set. A 17-dimensional hypercube has 2^{17} (about 130 000) 'corners,' so the data space is inevitably very sparsely populated.

Nesic and Vrhovac[10] have studied the prediction of corrosion of steel in CO_2-containing solutions. The prediction ability was found to be significantly better than conventional models, although there are some questions about the number of coefficients used in the network.

Neural network methods have been used to fit the corrosion rate of austenitic stainless steels in sulfuric acid using the data set described earlier.[11]

Since the paper on which this chapter is based was published in 1999, there have been about 100 articles published on the application of neural network methods to corrosion. The subject areas that have been covered include:

- atmospheric corrosion[12–23]
- corrosion and degradation of steel-reinforced concrete[24–35]
- marine corrosion[36–45]
- erosion–corrosion[46]
- organic coatings[47–51]
- corrosion fatigue[52–56]
- pipeline and oilfield corrosion[37,39,57–63]
- alloy optimization[64]
- damage assessment[65]
- pitting corrosion[66–68]
- general corrosion[69–73]
- interpretation or simulation of polarization curves[74–77]
- corrosion in aircraft[78–85]
- refinery corrosion[86–88]
- corrosion monitoring and inspection[89–91]
- stray current corrosion[92]
- metallic coatings[93]
- CO_2 corrosion [94–96]
- cathodic protection anode performance[97]
- oxidation in high temperature water and steam[98–100]

- stress corrosion cracking[63,101]
- identification of corrosion type[102–104]
- high temperature oxidation and corrosion[105,106]
- hydrogen effects.[107]

One of the difficulties of working with prediction in multiple dimensions is that of representing the results. All that can be done on two-dimensional paper is to plot representative sections through the data, either as simple two-dimensional plots, or as pseudo-three-dimensional plots. The only complete way of reporting the results of a neural network fit is to present the coefficients of the individual nodes, together with a description of the network. Clearly, this is not a very helpful representation for the casual reader. The alternative representation would be a computer program that implements the neural network, and thereby provides means of interrogating the network. However, even this is not very effective in representing the essence of the neural network model in a way that provides a general view of the relationships involved. We do not, as yet, have a good method of achieving this.

2.40.4 Variability of Corrosion Data

While it is useful to represent the expected corrosion rate as a function of the controlling parameters, this is not a complete representation of the corrosion behavior. Corrosion is inherently variable; as well as knowing the average corrosion rate, it is also important to know the variability of the behavior, as it is the extreme behavior that will typically cause a corrosion failure, rather than the average. Methods are currently being developed to derive the variability of the training data. A simple approach illustrated in **Figure 3** is to assume that the deviation of the training data points from the fitted neural network is an indication of the inherent variance of the training data. Then, a second network can be fitted to this variance information to derive a model of the variance of the corrosion rate as a function of the controlling variables. While this is a relatively transparent way of obtaining an estimate of the variance, there may be better ways of achieving this and this is the subject of current research. One thing that is clear is that a larger data set is required to obtain a good prediction of the variance as a function of the input variables, since the training data for variance are inherently noisy. In the absence of sufficient data points to model the variance as a function of the input variables, the best that can be done to overcome this problem is to estimate the average behavior as the mean-square residual, although it may be better to use the residual as a proportion of the mean at each point (i.e., to assume that the coefficient of variation is constant over the problem domain).

2.40.5 Confidence in Fitted Function

In general, one limitation of fitting techniques is the way in which they isolate the user from the original

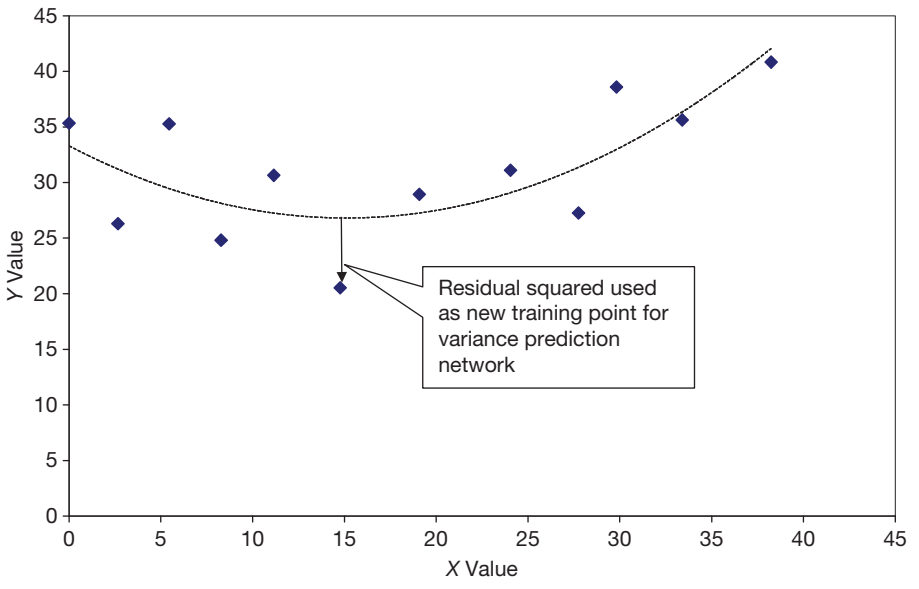


Figure 3 Estimation of variance.

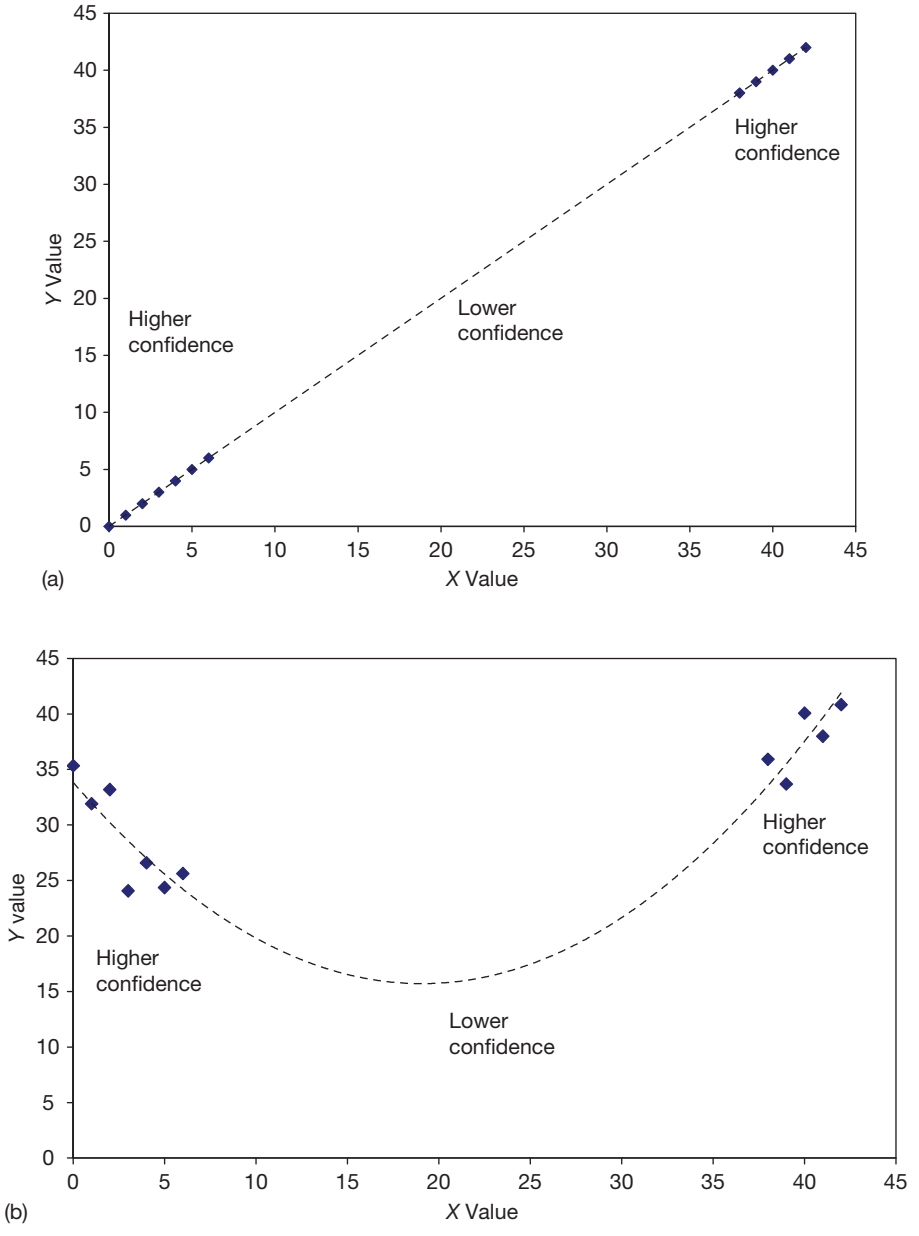


Figure 4 Regions of low confidence in a neural network prediction: (a) More reliable interpolation? and (b) Less reliable interpolation?

data. It is apparent from Figure 4 that any attempt to predict the corrosion rate in the central region is unlikely to be reliable. However, when the corrosion rate is produced as an output from a neural network or any other fitting function, this limitation is lost, and a naive user may assume that the prediction is as reliable as a prediction from regions with a high density of training data. Indeed, the imagined infallibility of computer programs may lead users to place greater faith in the prediction of the neural network. Consequently, it is important to provide some indication of the confidence that can be placed in the prediction of the network. If the original training data are available, the 'distance' from the nearest training points can be determined, and this can be used as an indicator of confidence. Alternatively, an additional network can be trained on the local density of points. This may not be an entirely accurate estimate. Figure 4 shows cases with equivalent distance between the training points, but with a somewhat

different perceived confidence in the interpolation or extrapolation. The continuous, consistent slope in **Figure 4(a)** gives a strong expectation that the line will continue between the two data regions. On the other hand, the inconsistent slopes in **Figure 4(b)** lead to uncertainty about the intervening region and hence a much lower confidence than in **Figure 4(a)**. These perceptions may, of course, be misleading, but this just underscores the difficulty of estimating the confidence that can be placed in the prediction of a neural network (or any other interpolation method). With a simple two-dimensional system such as that shown in **Figure 4**, it is relatively easy to see the implications of the interpolation process. However, in six-dimensional space it is much more difficult to visualize the relationship between the training data points and the interpolated result, and consequently it is important to develop a quantitative estimate of the confidence that can be placed in the prediction. Again, this is the basis of ongoing research.

2.40.6 Inconsistent Data Sets

As several groups have shown, neural network techniques can be very effective for modeling corrosion behavior. However, these methods require consistent data sets. To look at another example, work on pitting corrosion of austenitic stainless steels has used a number of measurement techniques, including the measurement of pitting potential, the ferric chloride exposure test (ASTM G 48) for critical pitting temperature (CPT) or various electrochemical methods for the measurement of CPT. If we are interested in the effect of steel composition, we can use neural network methods to fit the individual types of measurement (this has been done for the particular groups mentioned above[108]). This provides useful information, but fails to take full advantage of the 'relatedness' of the three data sets. In this case, it is apparent that we should be able to get better estimates if we could take account of all the data in the related data sets. We do not, as yet, have ways of doing this.

A similar problem is concerned with the relatively poor quality of many corrosion results. In this context, poor quality may indicate one or more of many limitations in the recorded data:

- Errors in the data arising from poor experiment design, faulty equipment or miscalculation. It is hoped such errors are rare.
- Failure to measure, control or report significant variables. For any work concerned with the effect of alloy composition on corrosion behavior, the reporting of nominal, rather than actual, compositions introduces a significant uncertainty, and degrades the value of the result. Similarly, failure to report or control environmental variables such as flow rate, oxygen concentration or temperature restrict the value of the data recorded.
- The summarization of data, for example, by plotting lines without the data points on which the line is based, seriously limits the use of such data for further analysis, and, in particular, tends to lose information about the variability and reliability of the data. It is less of a problem to report mean and standard deviation (e.g., as 'error bars') for measured data, as examples of the original data can be recreated by Monte Carlo methods. However, it is still better to know the exact data.

2.40.7 Training Data Requirements

In the introduction to prior work, we have placed significant emphasis on the number of training points, the number of input variables and the number of weights in the network. The significance of these values has often been misunderstood in the application of neural methods to corrosion problems. Even the neural network literature is somewhat unclear about what is acceptable for a good performance from a network. It may therefore be useful to examine this in a little more detail, although it should be appreciated that there is only a limited theoretical basis for this discussion.

We can consider the basic objective of the training of the neural network to be the accurate modeling of the expected value of the outer parameters (in the statistical sense of most likely value) over the relevant regions of the problem domain. By the problem domain we mean the multidimensional space defined by all the input parameters, with the 'relevant region' being the range of input parameters for which we want the output to be valid. Note that in many problems there may be strong correlation between input parameters, such that only a limited region of the problem domain is significant. In effect, this means that the effective dimensionality of the problem is less than the number of input variables would imply. This has important consequences for the number of training points that will be required, but the general analysis of this problem requires further work. We shall therefore consider only the case

where there is little correlation between the input variables, other than through possible clustering.

In general, an M-dimensional hypercube has 2^M 'corners.' A reasonable objective of a training data set would be to have a point on each corner of the hypercube. In experimental terms, this corresponds to taking two values of each variable and performing experiments for every combination of the two values. Thus, for one input variable this corresponds to taking one measurement at each of the maximum and minimum values of the input parameter and drawing a straight line between them. On this basis, two points might be rather low, and 3, 4, or more points might be better. However, the application of this approach over multiple dimensions leads to rapidly escalating requirements for the number of training points. If we cover all the possible combinations of N values for each M variable, we require N^M measurements. For example, for the relatively modest case of six input parameters we need 64 729 and 4096 measurements to provide 2, 3, and 4 points along each axis.

Unfortunately, corrosion data are generally expensive to produce, and large corrosion data sets with sufficient consistency to be suitable for training neural networks are rare. It is therefore generally the case that neural networks for corrosion data reduction must be trained on fewer points than is ideal. This leads to questions about acceptable network architectures, dimensions and training methods. This is a complex subject, and one that has not been fully defined.

From a simple information theory viewpoint, each weight used in the network provides one degree of freedom for the function produced by the network, and there should be at least one training point for each degree of freedom. If the training data are noisy, more training data (or fewer weights) will be required, otherwise the network will model the noise rather than the average behavior. On this basis, a good rule of thumb would appear to be that there should be at least twice as many training points as weights in the network. However, this rule is frequently broken, apparently without ill-effects, and it appears that over-large networks may produce acceptable behavior if the training method used is not too aggressive. Furthermore, methods have been proposed that allow the less important connections to be 'pruned' (their weights set at or near zero), thereby permitting a more complex network to be used while still retaining a good generalization ability. This is another area that merits further study. Until these questions have been clarified, we suggest that the guideline of two training points for every weight[109] should be taken as the ideal.

Recently, Liu *et al.*[52] have used Newton interpolation to create artificial data points between existing points in order to increase the number of training points and the density of points in the problem domain. However, at first sight, this is effectively training the network to perform Newton interpolation, and it is questionable whether it offers real advantages.

2.40.8 A Particular Example

To conclude this introduction to the use of neural networks for the handling of corrosion data, we shall show the results obtained for a 'purpose-built' experimental data set that examined the pitting potential for type 304 austenitic stainless steel as a function of solution composition and temperature. The input variables were the concentrations of Cl^-, I^-, Br^-, F^-, SO_4^{2-}, CO_3^{2-}, OH^-, NO_3^-, and $S_2O_3^{2-}$ and the temperature of the solution, and the output was the pitting potential (the latter term being interpreted rather imprecisely as that potential at which the anodic current started to increase rapidly – this may therefore correspond to transpassive corrosion or oxygen evolution).

The data set consisted of 200 measurements distributed at random through the 10-dimensional problem domain. Various network architectures and training methods were examined, but for the results reported here, 2/3 of the data were used for training, with the remaining 1/3 being used as a validation set to determine when to terminate training. A very simple network was used with only four nodes in a single hidden layer. This produced a network with a total of 75 weights, just over half the number of training points.

As noted previously, it is very difficult to represent or test the properties of a 10-dimensional function. The best that we have been able to do so far is to compare the predictions for simple situations, in which only one variable is changed with the corresponding experimental data. The results of this exercise are shown in **Figures 5–10**, in which the network predictions are compared with the experimental results of Man and Gabe[110] and with data measured using the same material and test method as were used to produce the training data (these results were not included in the training or validation data). In general, it can be seen that the network predictions are reasonable, but not

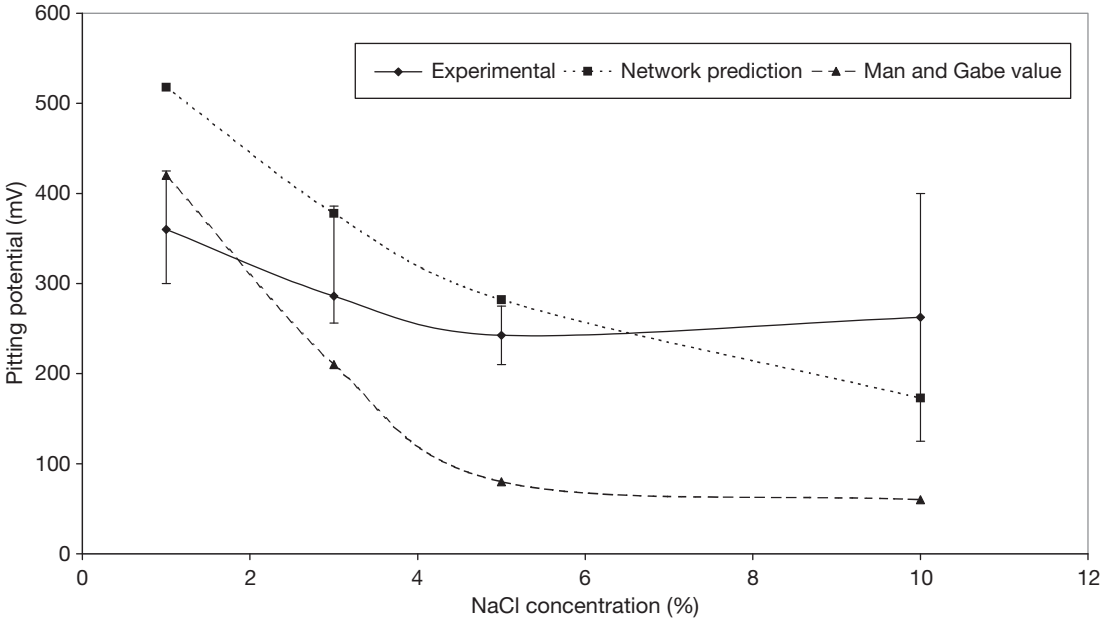


Figure 5 Effect of chloride concentration on pitting potential.

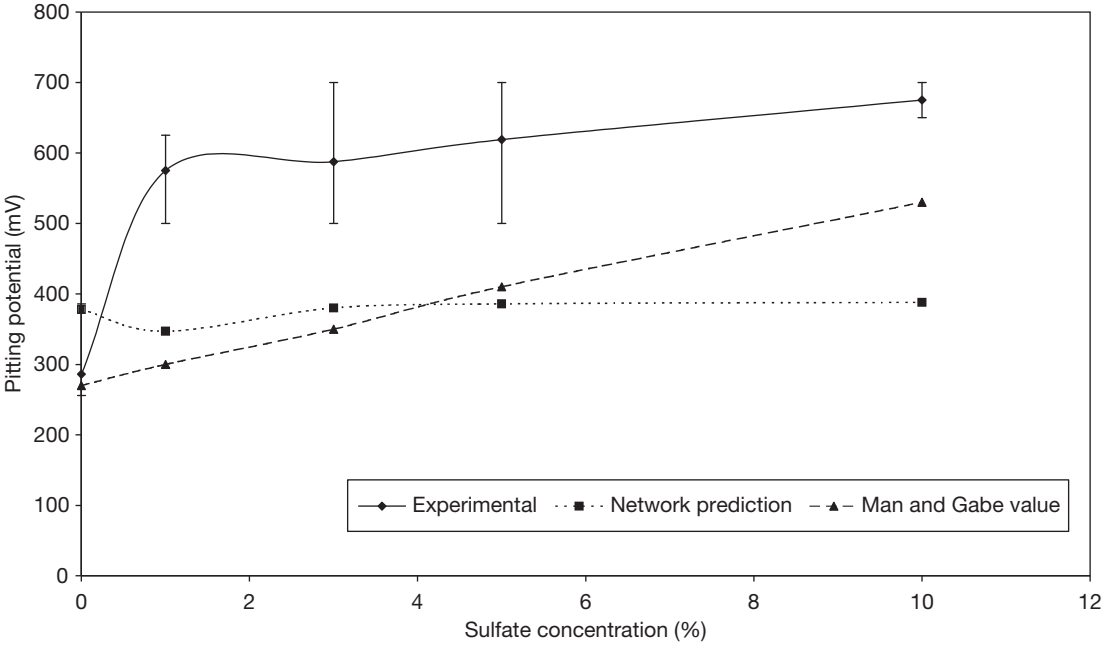


Figure 6 Effect of sulfate on pitting potential in 3% NaCl solution.

outstanding. The failure to provide completely accurate predictions is understandable, and demonstrates the fundamental problems of this approach to corrosion prediction. The training data set is small by comparison to the size of the problem (the 10-dimensional hypercube has 1024 corners, compared to our training data set of only 133 points). This, then, requires the use of a small network in order to avoid the possibility of memorizing the data, rather than generalizing. The conditions used for plotting, with 7 or 8 of the 9

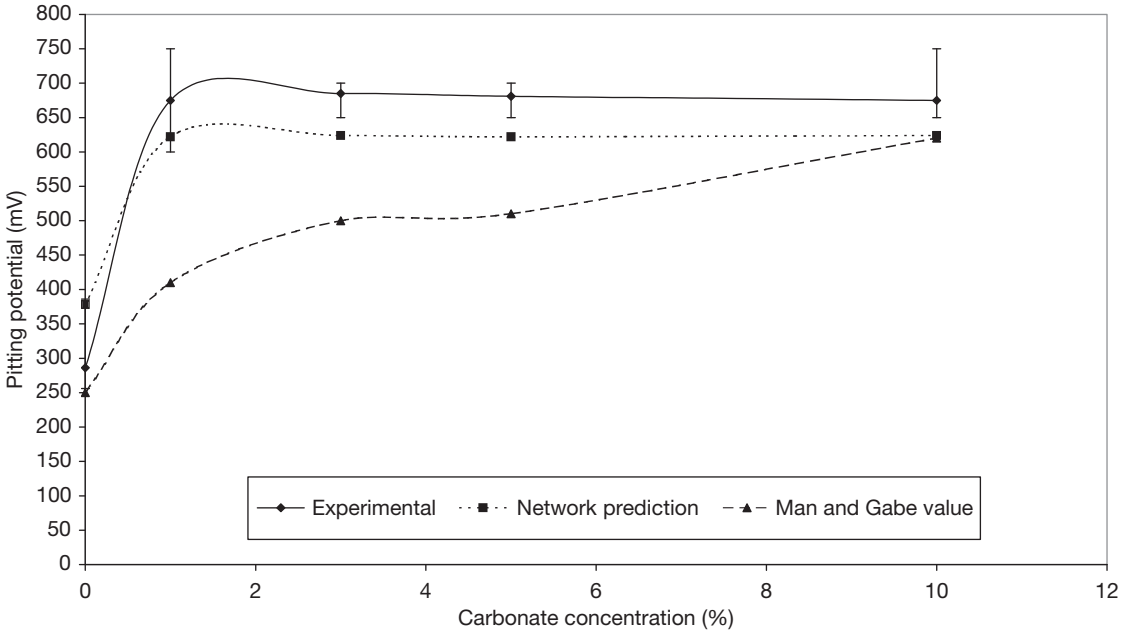


Figure 7 Effect of carbonate concentration on pitting potential in 3% NaCl solution.

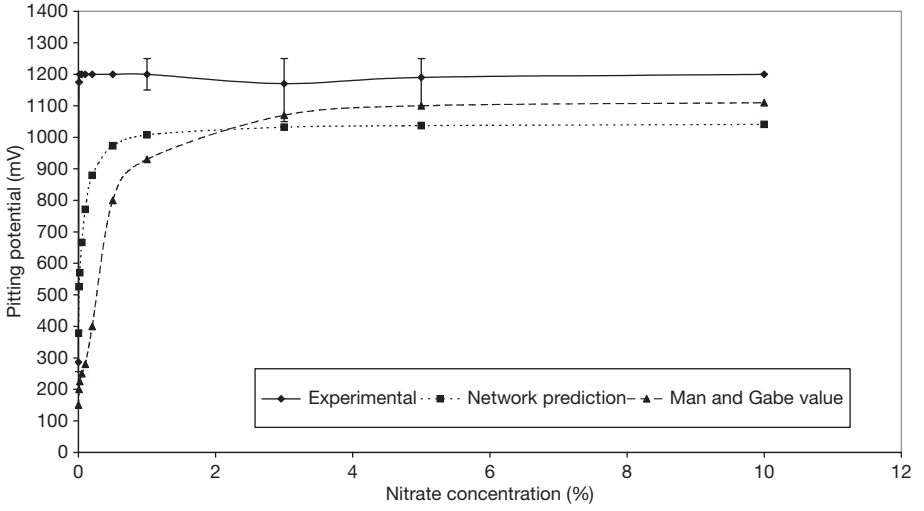


Figure 8 Effect of nitrate concentration on pitting potential in 3% NaCl solution.

concentrations set to 0 (modeled as 10^{-6} M to accommodate the log scaling used for concentration) lie right at the 'edges' of the problem domain, implying that an element of extrapolation is inevitably occurring. When these difficulties are taken into account, it seems reasonable to suggest that the neural network model performs remarkably well.

Owing to the rather small training set relative to the size of the problem, it is not realistic, at least with our current methods, to try to extract the variance information from the network. However, the residual error (that part of the training data that is not accounted for by the network) is of the same order (around 100 mV) as the standard deviation found when repeating experiments for the same conditions, which suggests that the network is modeling the mean behavior reasonably well, and that the residual errors are largely a result of real variability in the

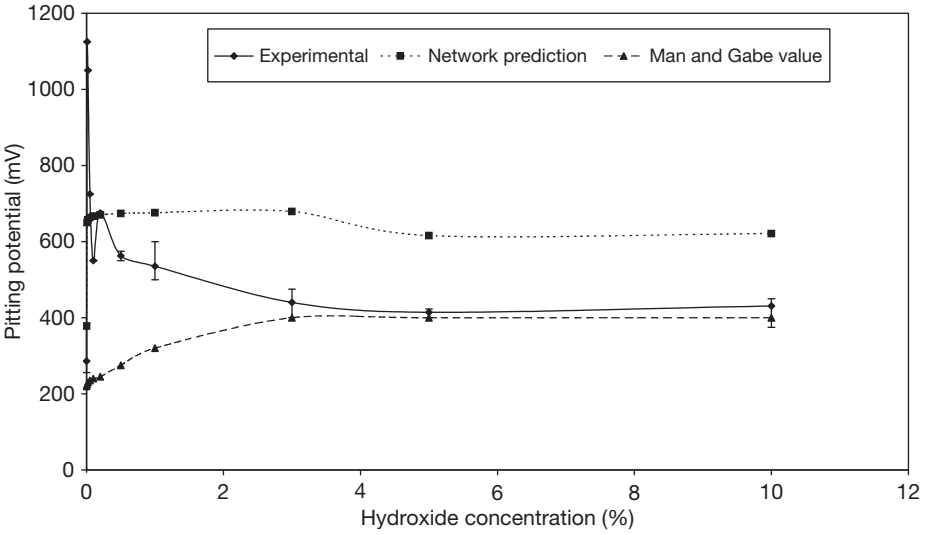


Figure 9 Effect of hydroxide concentration on pitting potential in 3% NaCl solution.

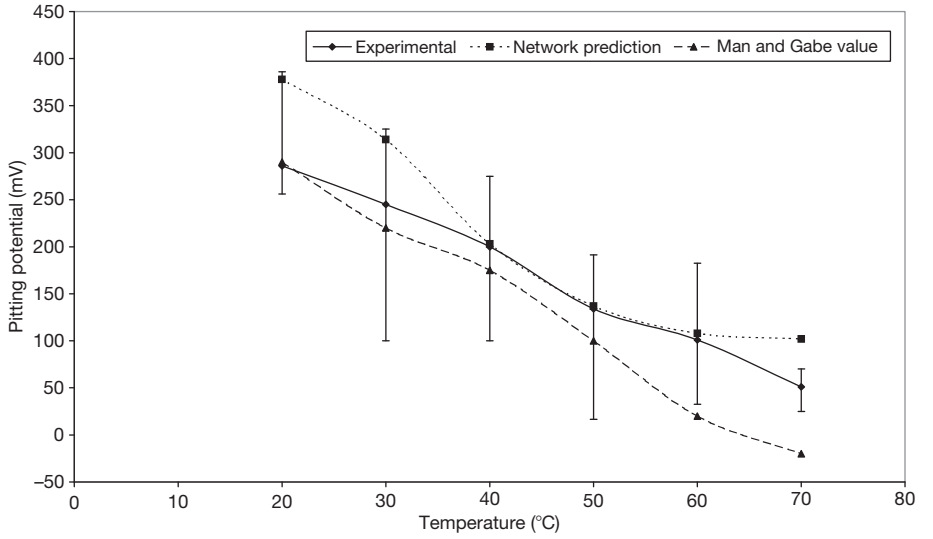


Figure 10 Effect of temperature on pitting potential in 3% NaCl solution.

measured value. As the training data have been deliberately distributed throughout the problem domain in a relatively uniform way, there is no point in assessing the confidence that can be attributed to the prediction of the network.

2.40.9 Conclusions

- Artificial neural networks have generally been claimed to be successful in modeling various types of corrosion behavior. However, it is difficult to evaluate the quality of the model that is provided by a neural network with many input variables, and many studies have only validated the performance of the neural network to a limited extent.
- There is a need to model the variability in the corrosion behavior as well as the mean behavior. This can be done, but requires rather larger training data sets than the modeling of the mean.
- There is a need to provide an indication of the confidence that can be placed in the prediction of a neural network. This requires further work towards developing the necessary techniques.

References

1. Cottis, R. A.; Li Qing; Owen, G.; Gartland, S. J.; Helliwell, I. A.; Turega, M. *Mater. Des.* **1999**, *20*(4), 169–178.
2. Smets, H. M. G.; Bogaerts, W. F. L. *Mater. Perform.* **1992**, *31*(a), 64–67.
3. Urquidi-Macdonald, M.; Eiden, M. N.; Macdonald, D. D. In *Modifications of Passive Films, Paris*, 1993, pp 336–343.
4. Ben-Hain, M.; Macdonald, D. D. *Corros. Sci.* **1994**, *36*(2), 385–393.
5. Silverman, D. C.; Rosen, E. M. *Corrosion* **1992**, *48*, 734–745.
6. Silverman, D. C. *Corrosion* **1994**, *50*(6), 411–418.
7. Silverman, A. L.; Silverman, D. C. In *Corrosion/2005*; Houston, TX, 3–7 April 2005; p 12.
8. Ramamurthy, A. C.; Lorenzen, W. I.; Urquidi-Macdonald, M. *Electrochim. Acta.* **1993**, *14*(14), 2083–2091.
9. Trasatti, S. P.; Mazza, F. *Br. Corros. J.* **1996**, *31*(2), 105–112.
10. Nesic, S.; Vrhovac, M. *J. Corros. Sci. Eng.* **1998**, *1* Paper 6.
11. Cottis, R. A.; Helliwell, I.; Turega, M. In *Corrosion/96*; NACE, 1996; Paper 379.
12. Fang, S. F.; Wang, M. P.; Qi, W. H.; Zheng, F. *Comput. Mater. Sci.* **2008**, *44*(2), 647–655.
13. Diaz, V.; Veronica; Lopez, C. *Corros. Sci.* **2007**, *49*(3), 949–962.
14. Wang, H.-T.; Han, E.-H.; Ke, W. *J. Chin. Soc. Corros. Protect.* **2006**, *26*(5), 272–274281.
15. Wang, H.-T.; Han, E.-H.; Ke, W. *Corros. Sci. Protect. Technol.* **2006**, *18*(2), 144–147.
16. Wang, S.; Song, S. *J. Chin. Soc. Corros. Protect.* **2005**, *25*(5), 257–261.
17. Abdul-Wahab, S. A.; Bakheit, C. S.; Siddiqui, R. A.; Al-Alawi, S. M. *J. Corros. Sci. Eng.* **2004**, *Vol. 5*, 1–17.
18. Song, S.; Wang, G.; Wang, S. *J. Chin. Soc. Corros. Protect.* **2003**, *23*(1), 56–64.
19. Diaz, V.; Lopez, C.; Rivero, S. *Revista de Metalurgia.* no. Extraordinario. **2003**, *Vol. 39*, 188–193.
20. Ma, X.; Qu, Z.; Li, C. *Corros. Sci. Protect. Technol. (China)* **2002**, *14*(1), 52–54.
21. Ma, X.; Qu, Z.; Li, C. *J. Univ. Sci. Technol. Beijing (China)* **2001**, *23*(2), 123–126.
22. Pintos, S.; Queipo, N. V.; Troconis de Rincon, O.; Rincon, A.; Morcillo, M. *Corros. Sci.* **2000**, *42*(1), 35–52.
23. Cai, J.; Cottis, R. A.; Lyon, S. B. *Corros. Sci.* **1999**, *41*(10), 2001–2030.
24. Ukrainczyk, N.; Ukrainczyk, V. *Mag. Concr. Res.* **2008**, *60*(7), 475–486.
25. Shu, Y.; Du, Y.-P.; Wang, X.-M.; Yan, S.-W.; Yan, C.-H. *Gongcheng Lixue (Eng. Mech.)* **2008**, *25*(suppl. 1), 102–106.
26. Fan, L.; Niu, D.-T.; Dong, Z.-P.; Qin, H.-F. *Hunningtu (Concrete)* **2008**, *3*, 27–29.
27. Ukrainczyk, N.; Pecur, I. B.; Bolf, N. *Civil Eng. Environ. Syst.* **2007**, *24*(1), 15–32.
28. Sun, L.; Wang, T. *Jianzhu Jiegou Xuebao/J. Build. Struct.* **2006**, *27*(suppl. 1), 713–716.
29. Sun, L.-Z.; Wang, T.-C.; Zhang, H.-B. *Tianjin Daxue Xuebao (J. Tianjin Univ. Sci. Technol.)* **2006**, *39*(3), 284–288.
30. Parthiban, T.; Ravi, R.; Parthiban, G. T.; Srinivasan, S.; Ramakrishnan, K. R.; Raghavan, M. *Corros. Sci.* **2005**, *47*(7), 1625–1642.
31. Morcous, G.; Lounis, Z. *Comput. Aided Civil Infrastruct. Eng.* **2005**, *20*(2), 108–117.
32. Hou, H.-B.; Zhang, G.-Z. *J. Wuhan Univ. Technol. (Mater. Sci. Ed.)* **2004**, *19*(4), 6–8.
33. Otsuki, N.; Yodsudjai, W.; Nishida, T.; Yamane, H. *ACI Mater. J.* **2004**, *101*(2), 146–453.
34. Fan, Y.-F.; Zhou, J.; Feng, X. *Key Eng. Mater.* **2004**, *274–276*(Part 2), 667–672.
35. Iitsuka, Y.; Iyoda, T.; Kato, Y.; Uomoto, T. *Seisan-Kenkyu (J. Inst. Ind. Sci., Univ. Tokyo) (Japan)* **2000**, *52*(2), 34–37.
36. Liu, W.; Zhao, X.; Deng, C.; Li, W. *J. Chin. Soc. Corros. Protect.* **2008**, *28*(4), 201–204.
37. Kong, T.; Wang, J.; Zhong, L. *Corros. Sci. Protect. Technol.* **2008**, *20*(1), 58–61.
38. Du, C.-W.; Zhao, Y.-Y.; Lu, L.; Gao, J.; Li, X.-G. *Zhuangbei Huanjing Gongcheng (Equip. Environ. Eng.)* **2007**, *4*(3), 85–87.
39. Deng, C.-L.; Sun, M.-X.; Li, W.-J.; Lin, Z.-J.; Chen, G.-Z.; Guo, W.-M.; Liu, W.; Mu, Z.-J. *Zhuangbei Huanjing Gongcheng (Equip. Environ. Eng.)* **2006**, *3*(3), 58–62, 88.
40. Deng, C.-L.; Li, W.-J.; Shun, M.-X. *Corros. Sci. Protect. Technol.* **2006**, *18*(1), 54–57.
41. Liu, X.; Tang, X.; Wang, J. *J. Chin. Soc. Corros. Protect.* **2005**, *25*(1), 11–14.
42. Wang, S.; Song, S. *J. Chin. Soc. Corros. Protect.* **2004**, *24*(2), 108–111.
43. Kong, D.; Wang, S.; Song, S. *J. Chin. Soc. Corros. Protect. (China)* **2001**, *21*(6), 352–356.
44. Owen, G.; Cottis, R. A.; Turega, M. In *Corrosion 2000*; NACE, 2000; Paper 489.
45. Kong, D.; Hou, G.; Song, S. *Corros. Sci. Protect. Technol. (China)* **2000**, *12*(1), 16–19.
46. Pai, P. S.; Mathew, M. T.; Stack, M. M.; Rocha, L. A. *Tribol. Int.* **2008**, *41*(7), 672–681.
47. Zhao, X.; Wang, J.; Kong, T.; Zhang, W.; Wang, Y.-H.; Pan, G.-Y. *Corros. Sci. Protect. Technol.* **2008**, *20*(4), 275–278.
48. Yao, D.; Gong, L.-H.; Wang, H.-L.; Wang, H.-X. *Jiefangjun Ligong Daxue Xuebao (J. PLA Univ. Sci. Technol. (Nat. Sci. Ed.) (China)* **2004**, *5*(5), 67–69.
49. Kumar, G.; Buchheit, R. G. In *Corrosion/2004*, New Orleans, LA, 28 March–1 April 2004; p 24.
50. Gao, Z.; Wang, S.; Song, S. *Corros. Sci. Protect. Technol. (China)* **2001**, *13*(suppl), 464–466.
51. Castaneda, H.; Urquidi-Macdonald, M. In *Corrosion 2000*, Orlando, FL, 26–31 March 2000; pp 00768.1–00768.13.
52. Liu, Z.-G.; Mu, Z.-T.; Jin, P. *Zhuangbei Huanjing Gongcheng (Equip. Environ. Eng.)* **2008**, *5*(3), 24–27, 57.
53. Fan, C.; He, Y.; Zhang, H.; Li, H.; Li, F. *Key Eng. Mater.* **2007**, *353–358*(Part 2), 1029–1032.
54. Ji, D.-M.; Zhou, C.-Y.; Hu, Y.-R. *Chuanbo Lixue (J. Ship Mech.)* **2005**, *9*(6), 103–112.
55. Haque, M. E.; Sudhakar, K. V. *Int. J. Fatigue* **2001**, *23*(1), 1–4.
56. Pleune, T. T.; Chopra, O. K. *Nucl. Eng. Des. (Switzerland)* **2000**, *197*(1–2), 1–12.
57. Xu, S.-H.; Ma, K.; Liu, X.-D. *Jisuanji Yingyong Yanjiu/ Applic. Res. Computers* **2008**, *25*(6), 1673–1675.
58. Zhang, Z.; Liu, X.; Li, Z. *J. Chin. Soc. Corros. Protect.* **2008**, *28*(3), 173–176.
59. Huang, Y.; Ji, D.-W. ISOPE-2007: Seventeenth (2007) International Offshore and Offshore and Polar Engineering Conference Proceedings, 2007.
60. Castaneda, H.; Urquidi-Macdonald, M. *Corrosion* **2004**, *60*(6), 538–547.
61. Gao, Z.; Song, S.; Wang, S.; Chen, S.; Lai, G. *J. Chin. Soc. Corros. Protect.* **2004**, *24*(2), 100–104.

62. Sinha, S. K.; Pandey, M. D. *Comput. Aided Civil Infrastruct. Eng.* **2002**, *17*(5), 320–329.
63. Cassidy, J. M.; McCoy, T. H. In *Corrosion/2002*, Denver, CO, 7–11 April 2002; p 12.
64. Yan, G.-F.; Zhang, Y.; Wu, J.-J. *Hebei Gongye Daxue Xuebao/J. Hebei Univ. Technol.* **2007**, *36*(6), 18–22.
65. Pidaparti, R. M. *Structural Health Monitor.* **2007**, *6*(3), 245–259.
66. Liang, C.; Zhang, W. *J. Wuhan Univ. Technol. (Mater. Sci. Ed.)* **2007**, *22*(3), 389–393.
67. Zhang, W.; Liang, C.-H. *J. Iron Steel Res. Int.* **2005**, *12*(6), 59–62.
68. Leifer, J.; Mickalonis, J. I. *Corrosion (USA)* **2000**, *56*(6), 563–571.
69. Shang, J.; Hu, Z.-L. *Corros. Sci. Protect. Technol.* **2007**, *19*(3), 225–228.
70. Mahjani, M. G.; Jalili, S.; Jafarian, M.; Jaberi, A. In *EUROCORR 2004: Long Term Prediction and Modelling of Corrosion*; Nice: France, 12–16 September 2004.
71. Tan, X.; Chen, Y.; Mu, Z.; Wei, Z. *J. Chin. Soc. Corros. Protect.* **2004**, *24*(4), 218–221.
72. Wang, H.-T.; Han, E.-H.; Ke, W. *Corros. Sci. Protect. Technol.* **2004**, *16*(3), 147–150.
73. Mao, X.-H.; Gan, F.-X.; Wang, D.-H. *Mater. Protect. (China)* **2001**, *34*(9), 55–56.
74. Aleksanyan, A. Y.; Reformatskaya, I. I.; Podobaev, A. N. *Protect. Met. (Russia)* **2007**, *43*(2), 125–128.
75. Kiselev, V. D.; Ukhlovtsev, S. M.; Podobaev, A. N.; Reformatskaya, I. I. *Protect. Met. (Russia)* **2006**, *42*(5), 452–458.
76. Nor, A. M.; Cottis, R. A. In *Corrosion/2004*, New Orleans, LA, 28 March–1 April 2004; p 8.
77. Sundara Raj, A.; Ravi, R.; Parthiban, T.; Radhakrishnan, G. *Bull. Electrochem. (India)* **1999**, *15*(12), 552–555.
78. Pidaparti, R. M.; Neblett, E. J. *Comput. Mater. Continua.* **2007**, *5*(1), 1–10.
79. Yu, D.; Chen, Y.; Duan, C. *J. Chin. Soc. Corros. Protect.* **2006**, *26*(1), 19–21.
80. Kumar, M.; Ramuhalli, P.; Liu, Z. *Rev. Progr. Quant. Nondestruct. Eval.* **2005**, *24A*, 720–727.
81. Chen, Y.; Lu, G.; Duan, C. *Xibei Gongye Daxue Xuebao* **2002**, *20*(3), 368–372.
82. Pidaparti, R. M.; Jayanti, S.; Sowers, C. A.; Palakal, M. J. *J. Aircraft* **2002**, *39*(3), 486–492.
83. Pidaparti, R. M.; Jayanti, S.; Palakal, M. J. *J. Aircraft* **2002**, *39*(1), 175–180.
84. Palakal, M. J.; Pidaparti, R. M. V.; Rebbapragada, S.; Jones, C. R. *AIAA J.* **2001**, *39*(10), 1936–1943.
85. Bailey, R. A.; Pidaparti, R. M.; Jayanti, S.; Palakal, M. J. In *AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference and Exhibit*, 41st, Atlanta, GA, 3–6 April 2000.
86. Trasatti, S. P.; Gabetta, G. *Corros. Eng. Sci. Technol.* **2006**, *41*(3), 200–211.
87. Fu, D.; Li, X.; Dong, C.; Li, M. *J. Univ. Sci. Technol. Beijing (China)* **2001**, *23*(2), 189–192.
88. Li, X.; Fu, D.; Dong, C.; Jin, Y. *Corros. Sci. Protect. Technol. (China)* **2001**, *13*(1), 56–59.
89. Prateepasen, A.; Kaewtrakulpong, P.; Jirarungsatean, C. *Key Eng. Mater.* **2006**, *Vol. 321–323*(Part 1), 549–552.
90. Coates, C. W.; Olanubi, O. O.; Sokoloski, R. D.; Singleton, J. In *45th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Palm Springs, CA, 19–22 April 2004.
91. Spanner, J.; Udpa, L.; Polikar, R.; Ramuhalli, P. *J. Nondestruct. Testing Ultrasonics (Germany)*; **2000**, *5*(7).
92. Li, W.; Wang, Y.-Q. *Corros. Sci. Protect. Technol.* **2005**, *17*(6), 438–441.
93. Xu, J.; Liu, W.; Xu, Z. *Surf. Rev. Lett.* **2005**, *12*(4), 569–572.
94. Hernandez, S.; Nesic, S.; Weckman, G.; Ghai, V. In *Corrosion/2005*, Houston, TX, 3–7, April 2005; p 33.
95. Wu, J.; Li, M.; Li, X. *J. Chin. Soc. Corros. Protect.* **2003**, *23*(1), 26–29.
96. Nesic, S.; Nordsveen, M.; Maxwell, N.; Vrhovac, M. *Corros. Sci. (USA)* **2001**, *43*(7), 1373–1392.
97. Lai, Y.-Q.; Chen, X.-T.; Qin, Q.-W.; Li, J.; Liu, Y.-X. *Zhongnan Daxue Xuebao (Ziran Kexue Ban)/J. Central South Univ. (Sci. Technol.)* **2004**, *35*(6), 896–901.
98. Osgerby, S.; Fry, A. In *Fourth International Conference on Advances in Materials Technology for Fossil Power Plants*; Hilton Head Island, SC, 25–28 October 2004; pp 388–401.
99. Guo, H.; Lu, Z.; Feng, G.; Cai, X.; Yang, W. *J. Chin. Soc. Corros. Protect.* **2004**, *24*(1), 25–28.
100. Zhou, J.-M.; Liu, D.-Z.; Bai, Z.-Q. *Corros. Sci. Protect. Technol.* **2003**, *15*(6), 342–344351.
101. Ye, H.; Xiong, J.-P.; Zhao, J.-M.; Zuo, Y. *Corros. Sci. Protect. Technol.* **2003**, *15*(6), 365–368.
102. Song, G.-X.; He, S.-F.; Cao, H.; Zhang, Z.; Zhong, Q.-P. *Trans. Mater. Heat Treatment (China)* **2003**, *24*(1), 82–8489.
103. Verma, A.; Ibragimov, A.; Ramachandran, S. *Neural Parallel Sci. Computat.* **2003**, *11*(1–2), 19–40.
104. Howell, P. A.; Winfree, W. P. *AIP Conf. Proc.* **2003**, *657A* (22), 568–574.
105. Simms, N. J.; Oakey, J. E.; Saunders, S. R. J.; Osgerby, S. *Mater. High Temp.* **2003**, *20*(2), 137–151.
106. Makkonen, P. In *BALTICA V, Porvoo*, Finland, 6–8 June 2001; pp 185–195.
107. Jin, Y.; Dong, C.; Fu, D.; Li, X. *J. Chin. Soc. Corros. Protect. (China)* **2001**, *21*(6), 368–373.
108. Gartland, S. J. Ph.D. Thesis, UMIST 1998.
109. Tarrasenko, L. *Guide to Neural Computing Applications*; Wiley, 1998.
110. Man, H.; Gabe, D. *Corros. Sci.* **1981**, *21*(9), 713.

2.39 Predictive Modeling of Corrosion

D. D. Macdonald
Center for Electrochemical Science and Technology, Department of Materials Science and Engineering, Pennsylvania State University, University Park, PA 16802, USA
G. R. Engelhardt
OLI Systems, Inc., 108 The American Road, Morris Plains, NJ 07950, USA

Abbreviations

BWR Boiling water reactor
CDF Cumulative distribution function
CEFM Coupled environment fracture model
CEM Coupled environment model
CEP Corrosion evolutionary path
CF Corrosion fatigue
CGR Crack growth rate
CT Compact tension
DAH Differential aeration hypothesis
DFA Damage function analysis
ECL Electrochemical crack length
ECP Electrochemical corrosion potential
EVD Extreme value distribution
HWX Hydrogen water chemistry
IGSCC Stress corrosion cracking
LPST Low-pressure steam turbines
MCL Mechanical crack length
MPM Mixed potential model
NWC Normal water chemistry
PDM Point defect model
SCC Stress corrosion cracking

Symbols

a Depth of a corrosion event
a_{max} Deepest corrosion event
a_{cr} Critical depth
b Bulk
b Inverse Tafel constant
C_k Concentration of species k
d Thickness of the wall
$\bar{D}$ Average diffusivity of breakdown sites
D_k Diffusion coefficient of species k
D_t Turbulent diffusion coefficient
E Potential
E_{corr} Corrosion potential

E_{crit} Critical potential for localized corrosion
E_{rp} Repassivation potential
E_0 Equilibrium potential
F Faraday's constant
F_k Integral damage function for corrosion event, k
f_k Differential damage function for corrosion event, k
G^{ex} Excess Gibbs energy
I Current density
I_{ss} Steady-state passive current density
i_a Anodic current density
i_c Cathodic current density
i_{corr} Corrosion current density
i_p Passive current density
i_0 Exchange current density
i^0_{corr} Current density on the bare surface
i^* Current density calculated in the absence of a potential drop in the cell
J_{ca} Flux of cation vacancies
J_m Annihilation flux of metal vacancies
j_k Flux of species k
K Constant in eqn [70]
k_i Reaction rate for ith reaction
K Equilibrium constant
K_I Stress intensity factor
K_{ISCC} Critical stress intensity factor
k_m Rate of chemical reaction m
K_v Electrochemical equivalent volume
L Thickness of barrier layer
L_{ss} State thickness of barrier layer
m Constant in eqn [70]
N Total number of nucleated stable pits
n_k Rate of nucleation of defect k
Q Quasipotential
P_f Probability of failure
R Gas constant
R_k Rate of production or depletion of species k
Re Reynolds number
s Surface
S Surface area
Sc Schmidt number
T Temperature
t Time
t_s Service time
t_{in} Incubation time
t_{pr} Propagation time
u Central parameter
u_i Mobility of species i
U_0 Open circuit potential
v Hydrodynamic velocity
V_k Rate of propagation of corrosion defect k
V_0 Initial pit propagation rate
$\overline{V}_0$ Mean initial pit propagation rate
V_{app} Applied voltage
V_c Critical voltage
V_m Electrode potential
X_i Internal independent variable
X_m Average depth of the largest pit
x Depth of penetration
x_0 Characteristic depth
Y_i External independent variable
z Direction perpendicular to the surface
z_k Charge of species k
α (1) Anodic transfer coefficient, (2) scale parameter, (3) polarizability of the barrier layer/solution (outer layer) interface
β (1) Tafel coefficient using natural logarithms, (2) dispersion in Laplace's distribution
ΔG^0_S Standard Gibbs energy change for the chloride absorption reaction
ΔG^0_S Change of Gibbs energy for the Schottky-pair reaction
ΔK Stress intensity factor range
ε (1) Dielectric permittivity, (2) electric field strength within the barrier layer
$\dot{\varepsilon}_{ct}$ Crack tip strain rate
ε_f Fracture strain
γ Delay repassivation constant
γ_k Activity coefficient of species k
Γ Ratio of the bare surface of the crack tip to the total geometric surface
η Dynamic viscosity
ΔK_{th} Threshold stress intensity factor range
κ Conductivity
λ (1) Constant in eqn [83], (2) function defined by eqn [97]
μ_k Chemical potential of species k
ν Kinematic viscosity
ν_k Stoichiometric coefficient of species k
ρ Density
ξ Critical areal concentration of vacancies
ζ Survival probability
σ Standard deviation
τ Dissolution time
Φ Cumulative distribution function
φ Electrical potential
Ψ Extreme value distribution
ψ Laplace's distribution function
Ω Mole volume of the barrier layer per cation

2.39.1 Introduction

As our industrial and infrastructure systems (refineries, power plants, pipelines, etc.) age, there is a considerable economic incentive to avoid unscheduled outages and to extend operation beyond the design lifetime. The avoidance of unscheduled outages is of particular interest, because the failure of even a minor component can result in the complete shutdown of a facility. For example, the unscheduled shutdown of a 1000 MWe nuclear power plant due to the failure of a valve may cost the operator between $1 million and $3 million per day, depending upon the cost of replacement power and other factors. However, if component failures could be accurately predicted, maintenance could be performed during scheduled outages, the cost of which has already been built into the price of the product, thereby minimizing the economic impact of the failure. With regard to life extension, the successful extension of operation beyond the design life translates into enhanced profits and the avoidance of costly licensing and environmental impact assessments associated with the development and construction of a new facility. However, in this case as well, the key to successful operation is the ability to avoid unscheduled downtime and hence maintain continuity of production. However, eventually, the frequency and severity of unscheduled outages render continued operation uneconomic and, at that point, replacement of the facility becomes necessary.

Corrosion is a major cause of component failure, and hence the occurrence of unscheduled downtime, in complex industrial systems. In particular, the various forms of localized corrosion, including pitting corrosion, crevice corrosion, stress corrosion cracking (SCC), and corrosion fatigue (to name the common forms), are particularly deleterious, because they frequently occur without any outward sign of accumulating damage and because they often result in sudden and catastrophic failures. Thus, the development of effective general and especially localized corrosion damage prediction technologies is essential for the successful avoidance of unscheduled downtime and for the successful implementation of life extension strategies.

There are two main approaches for predicting corrosion damage – empiricism and determinism. ('Determinism' is used here in the physics sense to describe a model whose predictions are constrained to the 'physically viable' realm by the natural laws. The term 'determinism' is often used in engineering disciplines to indicate a model that provides a definite output in response to a definite input.) Empiricism is the philosophy that everything we can ever know must have been experienced. On the other hand, determinism is the philosophy that we may predict the future from the past via the natural laws. Within these two classes, there exist numerous subclasses. For example, within the empirical class, there are functional models, in which (discrete) data are represented by continuous mathematical functions; statistical models; and artificial neural networks, to name a few. Within the broad class of 'deterministic' models, there can exist 'definite' models that yield a single output for a given set of input values; and probabilistic models, in which the inputs are distributed resulting in a distributed output from which the probability of an event occurring can be estimated.

It must be noted that, up to now, the prediction of corrosion damage has been largely based on the application of empirical models and only in the past decade, or so, have deterministic models been developed.

It is important to note that there are particular difficulties in using purely empirical models for predicting corrosion damage in real industrial systems, caused by the following factors: (i) Empirical models are generally expensive, because of the need for large databases covering many independent variables for calibration. Complex industrial systems are unique, even when they are of the same design, because they endure different operating conditions and histories. Thus, for example, an airplane that has been operating in a hot, humid environment might require more inspection and maintenance than an identical aircraft based in a desert environment. (ii) Failures are rare events. Accordingly, it is generally impossible to develop an effective database for model calibration based on the failure statistic of any given system. (iii) Empirical models also fail to capture the mechanism of failure.

Of course, deterministic models also pose many challenges. The most important of these are as follows: (i) Corrosion is an extremely complex phenomenon that depends on a multitude of factors, including the chemistry of the environment, metallurgy, and thermomechanical history of the corroding metal, hydrodynamics of multiphase flow, geometry, stress, temperature, pressure, and so on; (ii) the lack of information on kinetic parameters of the corroding system; and (iii) the need to define the corrosion evolutionary path (CEP). Because any model that is developed to describe the damaging process, and on which the deterministic prediction of damage is based, is only a figment of our imagination, and is based upon inputs from imperfect senses that are interpreted through an imperfect intellect, it is clear

that a working deterministic model can yield only an approximate description of a real system. Accordingly, pure 'determinism' is an ideal concept that is probably never achieved in reality, and the complementary, use of deterministic and empirical model provides the most effective method for predicting corrosion damage. Nevertheless, through the application of the 'scientific method,' in which a model is continually revised as it is cyclically tested against new observations, many deterministic models evolve toward providing accurate descriptions of real systems.

At this point, it is necessary to contrast the philosophical bases of 'mechanistic models' and 'deterministic models,' because great confusion exists in the literature on this issue. A mechanistic model is based upon a realistic mechanism for the process that is underpinned by a valid theory that, ideally, accounts for all of the observed properties. There is no requirement that the output must be constrained by the natural laws, as is the case for a deterministic model. Frequently, it is found that the values of unknown parameters are simply determined by calibration, whereas, in a deterministic model those same parameter values would be determined by the constraints. Given sufficient complexity and enough unknown parameters, a mechanistic model can be made to fit any data set imaginable and then yield predictions that are frequently at odds with reality. This cannot happen in the case of a deterministic model because the constraints, which are expressions of scientific generality (i.e., the 'natural laws'), limit the predictions to a realm that is consistent with the constraints themselves (i.e., to physical 'reality').

The need to define the CEP has been explained above. The CEP is defined as the path taken by the system in terms of those independent variables that have a significant impact on the rate of damage accumulation in transitioning from the present state to the future state. This path must be continuous and is required for both deterministic and empirical damage prediction models. Examples of defining the CEP are given later in this chapter, when we discuss the prediction of damage in practical systems.

In this article, we review recent advances in the development of some deterministic models for predicting corrosion damage (the review of some empirical models, including statistical models and artificial neural networks can be found in other articles in this volume). This subject is much too broad for justice to be done in a single chapter of the current length, but, by limiting the scope, the authors wish to illustrate the essential elements of theory and deterministic model-building and to demonstrate how the models can make useful predictions. The cases discussed have been chosen to address practical problems in current science and engineering, and, where possible, applications of the models to real engineering problems have been selected.

2.39.2 Definition of Corrosion Damage

As is well known, corrosion damage can be classified into two categories: uniform (general) corrosion and localized corrosion, and the quantitative description of these two cases are quite different. In this review, principal attention will be devoted to the case of localized corrosion. However, it must be noted that, in the general case, it is impossible to describe the propagation of localized corrosion damage without having reliable deterministic models for general corrosion, because the latter models yield the most important value for predicting the rate of localized corrosion; namely, the corrosion potential, E_{corr}.

In the case of general corrosion, it is natural to define corrosion damage at a given point on a metal surface as being the thickness of the metal layer that has corroded, a. This definition means that we can predict general corrosion damage if we can calculate a as a function of time and of the independent variables controlling the damaging process, that is, a is predicted in the form:

$$a = a(t, X_i, Y_i) \qquad [1]$$

Here, X_i and Y_i are internal and external independent variables, respectively, that determine the damage propagation rate. Examples of internal variables are grain size and orientation, texture, electrochemical kinetic parameters, and other microstructural properties. The external variables include loading and environmental conditions.

Even the term 'uniform corrosion' shows that, usually, the corroding layer thickness, a, depends slightly on the coordinates on the metal surface, if internal and external variables can be considered to be approximately constant for different parts of the system. In other words, it is assumed that, for a given set of conditions, uniform corrosion damage can be characterized for a given time and for given environmental conditions by a single value – average thickness of the corroded layer. Accordingly, for the case of general corrosion, the service life of the system, t_s, can be defined as the duration before the corroding

layer thickness achieves some average critical thickness, a_{cr}, (or reaches some minimum allowed value). Usually, the uniform corrosion data are analyzed by using the normal distribution. It is also assumed that the standard deviation of the thickness of the corroded layer, σ, can be estimated by using eqn [1], if the standard deviations of values X_i and Y_i are known by using, for example, the standard method for calculating the propagation of error.

In the case of localized corrosion, the situation becomes more complicated. Generally speaking, the accumulation of localized corrosion damage in a system is completely defined if we know how many pits or other corrosion events (per square centimeter) have depths between x and $x + dx$ for a given observation time, t, at a given location on the metal surface. We will denote this quantity as $f_k(x, t)dx$ where $f_k(x, t)$ is the so-called differential damage function.[1] Here, the index k denotes different types of localized defects, such as active and passive pits, cracks, crevices, and so on.

However, in the overwhelming majority of practical cases, such complete information is not required to effectively predict the failure time due to the penetration of the deepest event. Thus, very often, it is sufficient to obtain information about only the deepest corrosion event (pit, crack), because failure in the system commonly occurs when the depth of the deepest corrosion event a_{max} exceeds some critical value, a_{cr}. Usually, a_{cr} is the thickness of the wall of a pipe, for example, or the depth of a pit transitioning into an unstable crack. With regard to the first case, in order to describe the damage, we can use an equation of the previous form, that is,

$$a_{max} = a_{max}(t, X_i, Y_i) \quad [2]$$

Accordingly, the service life of the system, t_s (that is sometimes denoted as being the time to failure, t_f) can be expressed as the sum of incubation period, t_{in}, and propagation period, t_{pr}, in order for the defect to attain a critical depth:

$$t_s = t_{in} + t_{pr} \quad [3]$$

The period of propagation, t_{pr}, can be divided into periods corresponding to different forms of propagating of the corrosion defect; for example, in the form of pit, t_{pit}, or in the form of crack, t_{cr}, if the crack nucleates from the pit. In turn, for example, t_{cr} can be subdivided into the period $t_{cr,c}$ (the time required for the surface crack to grow into a through crack) and $t_{cr,g}$ (the time for a through crack to grow to a prescribed critical length), and so on.

It is important to understand that even if relation [2] can be obtained, its applicability would be questionable. The problem is that eqn [2] yields a single number for a_{max} for a given value for t and for other parameters of the system. However, in the general case, it is impossible to describe available experimental data by a single number. Thus, **Table 1** shows typical data for the depths of the deepest pits experimentally measured[2] on 20 sampled areas of 300×300 mm on the bottom outer surface near the periphery of a circular tank made of SS41 carbon steel, storing heavy petroleum, after 7 years of service; all of the pit depths not smaller than 0.5 mm were recorded, and the maximum pit depth in each sampled area was noted. The experimental data were taken from Harlow and Wei.[3–7]

We see that the difference between the observed values is greater than 280%. It is clear that, in the general case, prediction of corrosion damage can be done only in probabilistic terms. Apparently, the best form of predicting localized corrosion damage would be prediction of the probability of failure, P_f:

$$P_f = P_f(a_{cr}, t, X_i, Y_i, S) \quad [3]$$

that is, the probability that at least one corrosion event of any form (pit, stress corrosion crack, fatigue crack) reaches some critical depth, a_{cr}, at a given observation time, t, for any given set of environmental conditions. It is essential that P_f must be a function of the total area of the system, because the larger the area of the system, the larger will be the number of corrosion events and hence the greater the probability that a deeper event will exist. Accordingly, P_f must increase with area.

Table 1 Maximum pit depth distribution, $a_{max,i}$ (in ascending order) for oil storage tank

	i								
	12	***13***	***14***	***15***	***16***	***17***	***18***	***19***	***20***
$a_{max,i}$ (mm)	0.65	0.71	0.75	0.84	0.90	1.07	1.18	1.25	1.82

Index i numerates samples. Samples with $i < 12$ have maximum pit depth of <0.5 mm.

Let us consider possible ways for theoretically estimating this probability. The most intensive work that has been reported on predicting corrosion (in aluminum alloys) is that of Harlow and Wei,[3–7] whose important achievement has been to insert into prediction technology the realization that an analytical description (not necessarily a deterministic description) of the accumulation of damage is possible. Their approach is as follows: It is assumed that a pit is nucleated at an initial moment, $t = 0$, with an initial depth of a_0, and that the depth increases with increasing observation time, t. At some critical depth, a_{tr}, the pit transitions into a crack. By using mechanically-based models for calculating pit or crack propagation rates, an analytical expression for the depth of the corrosion event, a (in the form of a pit or crack), as a function of time is found in analytical form, as described by eqn [1]. All variables X_i and Y_i (where only temperature has been included in the environmental variables) are subdivided into two parts: deterministic parameters (in the engineering sense, i.e., parameters with fixed values) and random parameters, with the latter obeying, for example, the Weibull distribution. Subsequently, eqn [1], together with the change-of-variable theorem, or standard Monte Carlo technique, may be used to estimate the probability of failure, $Pr(a > a_{cr},t)$, that is, the probability that the crack length, a, exceeds the prescribed critical length of the crack, a_{cr}, at given observation time, t. The obvious problem with this approach is that it is completely devoid of environmental effects, such as those of potential, pH, $[Cl^-]$, and solution conductivity on pit and crack nucleation and growth, even though these effects have been established experimentally and are consistent with field observations. Accordingly, the underlying model for damage propagation fails to explain all of the experimental observations and hence is not viable. The second problem is that, by selecting a single pit, the significant literature demonstrating that pitting, which leads to cracking, is a progressive phenomenon in which new pits nucleate as existing pits grow and die (repassivate), is ignored. Thus, by selecting a particular pit initially, there is no way of knowing, *a priori*, whether that pit will survive to grow to a critical depth and hence, nucleate a crack. Indeed, the approach taken by Harlow and Wei is, in form, very similar to that employed by Liu and Macdonald[8,9] in the early 1990s to describe the failure of low pressure steam turbines, except that the latter study incorporated environmental effects, and the predictions of the component models were constrained by the appropriate natural laws (including the conservation of charge, recognizing the electrochemical nature of the damaging processes).

The question then arises whether the models of Harlow and Wei, for example, can be generalized by including the influence of environmental parameters, other than temperature, in explicit form. Even if they were able to do so, which, in itself, would require a massive reformulation of the models to incorporate environmental effects, in our opinion, the applicability of their approach would remain highly questionable, because environmental effects cannot be included in an *ad hoc* fashion. In addition, as noted above, the current Harlow and Wei[3–7] models consider only a single event, whereas it is well known that corrosion damage due to pitting, stress corrosion cracking, and corrosion fatigue accumulates progressively (that is, new pits/cracks nucleate while existing pits/cracks grow and die). In other words, it is impossible, in the general case, to attribute corrosion damage to a single corrosion event on the surface in isolation from all other events. Instead, the development of damage must be described in terms of an ensemble of localized corrosion events.

This thesis is confirmed by the following: The probability that any particular pit will repassivate (die) during the service life of a system is generally very high. Thus, let us consider the classical measurements of pit depth distribution versus time performed by Aziz[10] on Alcan aluminum alloy 2S-O in Kingston tap water (**Figure 1**). This study showed that by the end of 2 months, the bulk of the pits represented by the bell-shaped curve had ceased to grow (i.e., they had 'died') and only the deeper pits continued to propagate. In other words, by the end of two months from the beginning of the corrosive attack (in this particular case), the overwhelming majority of the pits repassivate (die). Accordingly, there is no way of knowing whether the pit selected had the necessary characteristics to grow to the critical length. In the general case, the probability that the depth of the deepest pit will exceed some critical value depends on the number of nucleated, stable pits and on the probability of repassivation. In some extreme cases, the probability that the depth of the deepest pit increases beyond some critical value is vanishingly small,[11] with the result that cracks will not nucleate. Obviously, this is a scenario that can only be described in terms of a large ensemble of events, rather than in terms of a single event.

In addition, the interaction between growing pits must be taken into account. The presence of existing

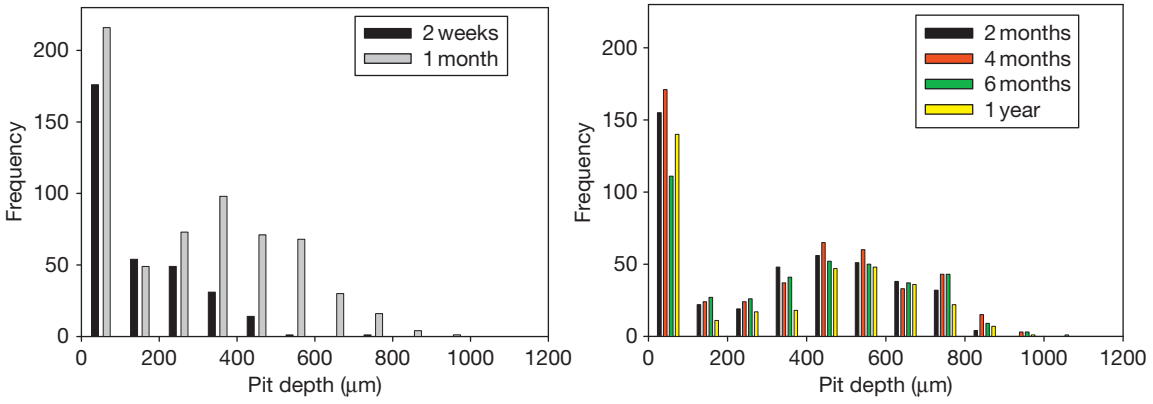


Figure 1 Pit depth distributions for Alcan 2S-O aluminum alloy immersed in Kingston water for different periods of time. The increment in pit depth is 100 μm. Reproduced from Aziz, P. M. *Corrosion* **1956**, *12*, 35–46.

pits impacts the probability of nucleation of other stable pits within the neighboring region. This phenomenon has been experimentally reported by Macdonald[12] and is explained by the competition of the pits for the available resources on the external surfaces (oxygen reduction) or because the existing pit cathodically protects the neighboring surface and hence inhibits the nucleation of the second pit. The interaction between the 'hemispheres of influence' (subtending the external area over which the cathodic reaction occurs) of growing, stable pits can also reduce the propagation rate of each of the pits, simply by reducing the metal potential at the pit bottom. Finally, we also have to take into account the overlapping between growing pits that also influences the shape and propagation rate. All of this shows that any method for estimating the probability of failure on the basis of the propagation of a single corrosion event (pit or crack) cannot be general.

2.39.3 Damage Function Analysis

One of the approaches that considers the accumulation of corrosion damage in terms of the evolution of an ensemble of pits and cracks is damage function analysis (DFA).[1,13,14] (The other possible method, that based on the Monte Carlo simulation of movement of ensemble of corrosion events will also be described below). As mentioned above, the differential damage function, $f_k(x,t)$ yields the complete description of corrosion damage. It is easy to obtain the differential equation for this function. The function f_k has a dimension of $\#/(\mathrm{cm}^2\ \mathrm{cm}) = \#/\mathrm{cm}^3$, analogous to the concentration of particles. Accordingly, it is very convenient to regard each defect as a 'particle' that moves in the x direction (perpendicular to the surface, with $x = 0$ being at the metal surface). The coordinate of this particle, x, coincides with the depth of penetration into the surface. Accordingly, f_k must obey the law of mass conservation

$$\frac{\partial f_k}{\partial t} + \frac{\partial j_k}{\partial x} = R_k, \quad k = 1, 2, \ldots, K \qquad [4]$$

where j_k and R_k are the flux density and the bulk source (sink) of the 'particle' k respectively. Thus, the subscript k enumerates the corrosion defect (e.g., pit or crack) and K is the total number of different corrosion defects in the system. By definition, $R_k(x,t)\mathrm{d}x\mathrm{d}t$ yields the number of defects k (per square centimeter) with depths between x and $x + \mathrm{d}x$ that arise (or disappear) during the period of time between t and $t + \mathrm{d}t$, due to transformation (repassivation, in the case of pits).

The system of eqns [4] can be solved with the appropriate boundary and initial conditions.

$$j_k = n_k(t) \quad \text{at } x = 0, t > 0 \qquad [5]$$

and

$$f_k = f_{k0}(x) \quad \text{at } x \geq 0, t = 0 \qquad [6]$$

where $f_{k0}(x)$ is the initial distribution of defect k (usually we can assume that $f_{k0}(x) = 0$, i.e., no damage exists at zero time) and $n_k(t)$ is the nucleation rate of the same defect (i.e., $n_k(t)\ \mathrm{d}t$ is the number of stable defects (per square centimeter) that nucleate in the induction time interval between t and $t + \mathrm{d}t$).

Thus, because the defect propagation flux, j_k, must be nonnegative (the depth of a corrosion event can only increase with time), the following, simplest numerical upwind finite difference scheme can be used for numerically solving eqns [4]–[6].

$$f_{k,m}^{n+1} = f_{k,m}^{n} - \frac{\Delta t}{\Delta x}\left(j_{k,m}^{n} - j_{k,m-1}^{n}\right) + R_k^n \Delta t \qquad [7]$$

Here, we use the straightforward approach of choosing equally spaced points along both the t- and x-axes: $x_j = x_0 + m\Delta x$, $m = 0, 1, \ldots, \mathcal{J}$ and $t_j = t_0 + n\Delta x$, $n = 0, 1, \ldots, N$, and we denote $f_{k,m}^{n} = f_k(t_n, x_m)$. The values $f_{k,0}^{n}$ and $f_{k,m}^{0}$ are calculated from the boundary and initial conditions [4] and [5]. Of course, it is assumed that we know (i.e., can calculate) fluxes, j_k, and sources/sinks, R_k, as functions of the spatial coordinates and time, and, in nonlinear cases, as functions of the unknown values of f_k.

It is important to note that, sometimes, the equation of continuity is presented in the simplified form[15,16]

$$\frac{\partial f_k}{\partial t} + \frac{\partial [V_k(x) f_k]}{\partial x} = R_k, k = 1, 2, \ldots, K \qquad [8]$$

This expression implies that the rate of kth defect, V_k, depends only on the depth of penetration, x, and accordingly, for the flux density, j_k, we have

$$j_k(x, t) = f_k(x, t) V_k(x) \qquad [9]$$

In the simplest cases, it is even possible to obtain analytical solutions for the damage functions. As an example, let us consider the case of pitting corrosion under constant external conditions. In this instance, we have two kinds of defects ($K = 2$): active pits with the damage function, f_a, and passivated pits (i.e., those that have 'died' through delayed repassivation) with the damage function, f_p. Let us assume that the flux density of active pits is described by eqn [8] with $V_a(x) = V(x)$. By definition, the flux of passivated pits is zero (i.e., these pits are 'dead'). It is evident that functions R_a and R_p must obey the relation $R_p = -R_a$ (a pit must be either alive or dead). If, in addition, we assume that the pit repassivation process obeys a first order decay law, the function R_a has the form

$$R_a(x, t) = -\gamma f_a(x, t) \qquad [10]$$

where γ is the delayed repassivation ('death') constant (i.e., the rate constant for repassivation of stable pits). In the general case, γ depends on the depth of the pit, x, and on time, t, when the external conditions depend on time. However, in this example, we assume that γ is a constant, that is, we suppose that pits repassivate accidentally and that the probability of repassivation does not depend on pit depth. (This is clearly a gross oversimplification, since the probability of delayed repassivation is expected to increase with pit age, but this does not change the logic of the argument). Accordingly, the system of equations for calculating the DFs has the form

$$\frac{\partial f_a}{\partial t} + \frac{\partial [V(x) f_a]}{\partial x} = -\gamma f_a \qquad [11]$$

and

$$\frac{\partial f_p}{\partial t} = \gamma f_a \qquad [12]$$

The boundary and initial conditions are given as

$$V f_a = n(t) \text{at } x = 0, t > 0 \text{ and } f_a = f_p = 0 \quad \text{at } t = 0 \qquad [13]$$

where $n(t)$ is the nucleation rate of pits on the surface. Analytical solutions to this system of linear partial first-order differential equations can be obtained by using the characteristic method[17] and have the following form:

$$f_a = \frac{\exp[-\gamma\theta(x)] n[t - \theta(x)]}{V(x)} \text{ and}$$
$$f_p = \frac{\gamma \exp[-\gamma\theta(x)] N[t - \theta(x)]}{V(x)} \qquad [14]$$

where

$$\theta_{pit}(x) = \int_0^x \frac{dx'}{V(x')} \qquad [15]$$

is the age of a pit with depth x and

$$N(t) = \int_0^t n(t')dt' \qquad [16]$$

is the number of stable pits (per square centimeter) that nucleate in the time interval between 0 and t.

In some cases, it is convenient to use the so-called integral damage function

$$F_k(x, t) = \int_x^\infty f_k(x', t)dx' \qquad [17]$$

which yields the number (per square centimeter) of corrosion events with depths larger than x for a given observation time, t. It is important to note that, experimentally, only the sum of the damage functions for active and passive pits $f = f_a + f_p$ is determined and in many instances the integral damage function only is measured. Accordingly, it is important for practical reasons to obtain the equation for the integral damage function, $F = F_a + F_p$, for the sum of the active and passive pits. Note that, for the considered case, the integral damage function corresponds to the

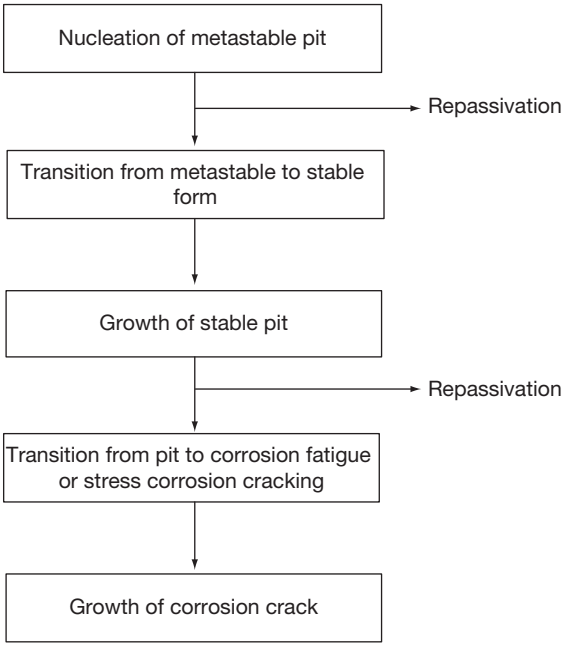


Figure 2 Schematic history of the nucleation and propagation of corrosion damage.

number of remaining pits on a surface, as the surface is removed layer by layer in preestablished increments. From eqns [14] and [17], we therefore have

$$F(x,t) = \exp[-\gamma\theta(x)]N[t-\theta(x)] \qquad [18]$$

According to the theory outlined above, calculation of the damage functions requires the determination of three independent functions for each kind of corrosion defect, k: (1) the rate of defect nucleation, n_k; (2) the flux density (growth rate) of the defect, j_k; and (3) the rate of transition of one kind of defect into another, R_k (e.g., the transition of an active pit into a passivated pit or the transition of a pit into a crack). In other words, we need to have quantitative models for describing each stage of corrosion damage as indicated in Figure 2. Below, we will discuss the feasibility of calculating each of these three functions. However, as noted above, the detailed deterministic description of any phase of corrosion propagation is impossible without reliable information about the corrosion potential of the system, E_{corr}. Accordingly, in the first step, the methods for calculating E_{corr} will be considered.

2.39.4 Mixed Potential Model

In 1937, Wagner and Traud[18] formulated their mixed potential theory in electrochemistry, in which the potential adopted by an electrode in contact with an aqueous solution containing both oxidizing and reducing species is determined by a balance of the cathodic (reduction) and anodic (oxidation) partial processes occurring at the surface. If the electrode is inert, the resulting potential is known as the redox potential, E_{redox}. On the other hand, if the electrode is electroactive and undergoes electrooxidation (corrosion), thereby contributing to the total partial anodic current, the potential is known as the electrochemical corrosion potential, ECP (or E_{corr}), of the substrate. Thus, if the partial current densities do not depend on the coordinate on the metal surface, the charge conservation condition for the interface may be written as

$$\sum_k i_{a,k} + \sum_m i_{c,m} = 0 \text{ at } E = E_{corr} \qquad [19]$$

where $i_{a,k}$ and $i_{c,m}$ are partial anodic and cathodic current densities, corresponding to the kth anodic and mth cathodic reaction correspondingly. This simple theory has proved to have a profound impact on how we interpret the corrosion of metals and alloys in a wide variety of systems.

As an example, let us consider one of the first comprehensive applications of the mixed potential theory in corrosion science to an industrial system, namely the calculation of ECP in the coolant circuits of water-cooled nuclear reactors, particularly boiling water reactors (BWRs), using a mixed potential model (MPM).[19,20] In this complex chemical system (the reactor coolant circuit), radiolysis of the coolant water by ionizing radiation (γ-photons and neutrons) produces a myriad of electroactive species, some of which are oxidizing species (e.g., O_2, H_2O_2, OH) and others that are reducing species (H_2, H). These species all contribute to the current flow across the interface, but the mixed potential theory predicts that the contribution that any given species can make to the potential is roughly proportional to its concentration. Thus, in BWR primary coolant circuits, the only species of practical importance are O_2, H_2O_2, and H_2, since the concentrations of these species are orders of magnitude greater than those of the other radiolytic species. The reactions describing each of the electroactive species are described as redox reactions, which can be written in the general form as

$$R \leftrightarrow O + ne^- \qquad [20]$$

where R and O are reduced and oxidized species, respectively, with the kinetics of the reaction being

described in terms of the generalized Butler–Volmer equation

$$i_{R/O,j} = \frac{e^{b_{a,j}\eta_j} - e^{-b_{c,j}\eta_j}}{\frac{1}{i_{0,j}} + \frac{e^{b_{a,j}\eta_j}}{i_{l,f,j}} - \frac{e^{-b_{c,j}\eta_j}}{i_{l,r,j}}} \quad [21]$$

where b_a and b_c are the anodic (oxidation) and cathodic (reduction) inverse Tafel constants, i_0 is the exchange current density, $i_{l,f}$ and $i_{l,r}$ are the mass transport limiting current densities, and η is the overpotential that is defined as the difference between the potential and the equilibrium potential for reaction j. The parameters b_a and b_c are normally measured in separate experiments, but in principle can be calculated *ab initio*. The exchange current density, i_0, is almost always measured directly, as theory is not sufficiently well developed to calculate this quantity from first principles. In addition to the redox partial reactions (oxidation and reduction processes), the anodic oxidation of the substrate also contributes to the total current density. In this case, the point defect model (PDM)[21] (the PDM will be considered in detail in the next section) provides the functional form of the anodic oxidation current density as

$$i_a = A + Be^{CE} \quad [22]$$

where A, B, and C are constants that depend upon the values of fundamental parameters in the model and upon the properties of the system (e.g., pH). For metals and alloys that form n-type passive films, in which cation interstitials and/or oxygen vacancies are the dominant defects, $B = 0$, and the anodic oxidation current density is independent of potential, unless there is a change in the oxidation state of the metal ion (cation or interstitial) being ejected from the barrier oxide layer at the barrier layer–solution interface. Systems of this type include iron and carbon steel, the stainless steels, and the chromium-containing, nickel-based alloys, such as Alloys 600 and 22. On the other hand, for metals that form p-type passive films (e.g., nickel), $A = 0$, and B and C are greater than zero. In this case, the passive current density is described by the Tafel equation.

The conservation of charge requires that the sum of the partial current densities at the interface be zero, with this condition being expressed as

$$i_a(E) + \sum_{j=1}^{\mathcal{J}} \frac{e^{b_{a,j}\eta_j} - e^{-b_{c,j}\eta_j}}{\frac{1}{i_{0,j}} + \frac{e^{b_{a,j}\eta_j}}{i_{l,f,j}} - \frac{e^{-b_{c,j}\eta_j}}{i_{l,r,j}}} = 0 \quad [23]$$

where $\mathcal{J}$ is the total number of redox reactions in the system. Note that in eqn [19], the summation takes place over the anodic and cathodic reactions separately, whereas in eqn [23], index j enumerates the redox reactions, each comprising conjugate partial anodic and cathodic partial reactions. Solution of eqn [23] yields the ECP or the corrosion potential, $E_{corr.}$ Note that $\eta_j = E - E_j^e$, where E_j^e is the equilibrium potential for the jth redox reaction. The limiting currents can be written in terms of the mass transfer correlations for the flow geometry and flow regime of interest.[22]

$$i_{l,j} = \pm n_j F A' Re^{\lambda} Sc^{\gamma} C_{O/R}^{b} \quad [24]$$

where n_j is the number of electrons involved in the reaction, A' is a constant, $Re = dV/\nu$ is Reynolds number, d is the hydrodynamic diameter of the channel, V is the flow velocity, $\nu = \hat{\eta}/\rho$ is the kinematic viscosity, $\hat{\eta}$ and ρ are the dynamic viscosity and the density, respectively, of the medium, and $C_{O/R}^{b}$ is the concentration of the reactive species in the bulk environment. The Schmidt number is defined as $Sc = \nu/D$, where D is the diffusivity of the reacting species. The sign convention is '+' for the forward direction (left to right) of the reaction [20], as written, and '−' for the reverse direction. Mass transfer correlations of the type expressed by eqn [24] are available in the literature for a wide variety of flow geometries and regimes (Selman and Tobias,[22] and citations therein), with the latter corresponding to specific ranges in Re and Sc. Once the corrosion potential is known (by solving eqn [23]), the corrosion current density is readily calculated using eqn [22].

As noted above, the MPM has been used extensively to calculate the corrosion potentials of stainless steel components in the coolant circuits of BWRs, and we will use this case to illustrate the power of this model in defining the response of metals and alloys to the properties of the environment in a complex industrial system. The typical measured and calculated ECP data for a cell attached to the recirculation piping of the Liebstadt BWR in Switzerland as a function of the amount of hydrogen added to the feedwater are plotted in **Figure 3**. This represents a 'double blind' comparison in that those who performed the calculations[19,23] did not have access to the data measured on the reactor, and those who measured the ECP on the reactor (Sierra Nuclear) did not have access to the calculated values. As shown, the measured and predicted corrosion potential data are in good agreement, except for the initial point (zero added hydrogen). However, in this case, it was found that the level of agreement could be greatly improved by changing the mass transport parameters (flow velocity, hydrodynamic diameter, etc.).

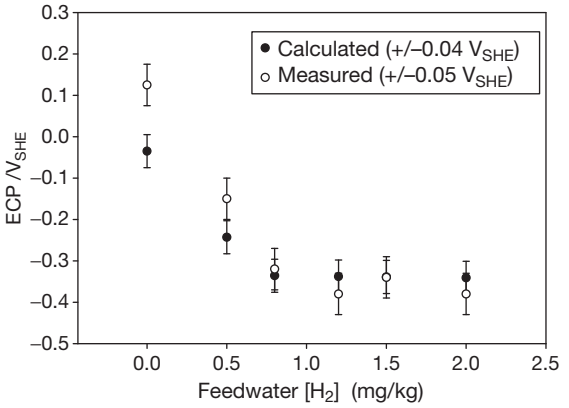


Figure 3 Comparison of calculated and measured ECP for the Leibstadt BWR in Switzerland. The 'measured' data were acquired in a test cell attached to the recirculation piping. The data employed in the calculation are summarized in Macdonald and Urquidi-Macdonald[19] and Macdonald *et al.*[23] The estimated accuracies of the calculated and measured data are indicated in the box in the figure.

Since these parameters are poorly characterized for an autoclave at very low flow rates, the disagreement in the absence of added hydrogen is of little consequence.

The accuracy of any MPM model for predicting E_{corr} and the rate of uniform corrosion depends on the accuracy of modeling of the rates of particular electrochemical reactions along with the completeness of the set of reactions included in the summation in eqn [19]. A detailed review of different models for the kinetics of corrosion can be found in the corresponding chapter of this book **Chapter 2.38, Modeling of Aqueous Corrosion**. Here we would only like to mention that already existing commercial software yields reliable prediction of the corrosion potential and, correspondingly, the rate of uniform corrosion under relatively complicated conditions, including the influence of adsorption, active–passive transition effects, scaling, transport in porous media, and so on.

Thus, the models of Anderko *et al.*[24–26] and those of others[8,9,19–21,23] have been developed for simulating the rates of general corrosion of selected metals (including carbon steels, stainless steels, aluminum, and nickel-based alloys) in aqueous solutions. The model consists of thermophysical and electrochemical modules. The thermophysical module is used to calculate the speciation of aqueous solutions and to obtain concentrations, activities, and the transport properties of individual species. The electrochemical module simulates partial oxidation and reduction processes on the surface of the metal. It is capable of reproducing the active–passive transition and the effect of solution species on passivity. The model has been implemented in a program that can be used to simulate the effects of various conditions, such as temperature, pressure, pH, component concentrations, and flow velocity on the corrosion potential and corrosion rate. The model clearly demonstrates the influence of inhibitors on corrosion rate. Good agreement with experimental data has been obtained.

Other comprehensive models have been developed for describing general corrosion under specific conditions. Thus, Nordsveen *et al.*[27] developed a mechanistic model of uniform carbon dioxide (CO_2) corrosion that takes into account such phenomena as diffusion of species between the metal surface and the bulk including diffusion through porous surface films, migration due to the establishment of potential gradients, and the existence of homogenous chemical reactions, including the precipitation of surface films. Nešić *et al.*[28] developed a comprehensive model for calculating internal uniform corrosion rates under multiphase flow conditions in mild steel pipelines.

However, it is important to note that, in the general case, for calculating ECP and uniform corrosion rate, where extensive localized corrosion may exist on a surface, thereby imparting significant nonuniformity in the current densities, we have to use, instead of eqn [19], the following equation.

$$\int_S \sum_m i_{a,m}\, dS + \int_S \sum_k i_{c,k}\, dS = 0 \quad \text{at } E = E_{corr} \quad [25]$$

It is evident that eqn [25] reduces to eqn [19] when we can assume that the partial current densities do not depend on the coordinates (location) on the metal surface, or if we can neglect any such dependencies. Thus, let us consider the case of pitting corrosion. Of course, the anodic current density inside pits can exceed the passive corrosion current density by several orders of magnitude. However, if the area of the active dissolution (area of pit surfaces) is much smaller than the total area of the metal surface (active corrosion current is much smaller than the passive corrosion current), it may be possible to use eqn [19] instead of eqn [25]; otherwise, such a simplification would be incorrect.

Experiments show that, often, both possibilities can be realized in an industrial system.[29,30] Thus, in the case of the corrosion of Alloy 20Kh13 (the Russian analog of Type 403 SS) in NaCl solutions, the corrosion potential is observed to decrease with time.[30] This reduction occurs due to the intensive growth of corrosion pits, because the total cathodic partial current has

to increase in order to compensate the increasing total partial anodic current, to conserve charge in the system. At some point, this growth stops due to the transport limitation for the cathodic reactant (oxygen), that is, the system is now limited by the lack of cathodic resources to support the ever-increasing corrosion current. On the other hand, such reduction is not observed in cases where the total area of observed pits is very small and hence, the nonuniformity in the partial current densities is minor and can be ignored.

These examples clearly show that, in the general case, eqn [25] must be used for predicting corrosion potential, at least in the case of steels that are not highly resistant to localized corrosion in the prevailing environment. It is evident that because of the statistical character of pit distribution in size and position, the imposition of charge conservation can be done correctly only by considering propagation of pitting damage as the evolution of an ensemble of pits that initiate, propagate, and repassivate on the metal surface.

2.39.5 Rate of Pit Nucleation

2.39.5.1 Empirical Models

It would be natural to assume, and experiment confirms this at least for the case of stainless steels, that the rate of nucleation of stable pits, n, is proportional to the rate of nucleation of metastable pits, n_{mp},[31–33] that is,

$$n = \varsigma n_{mp} \qquad [26]$$

where coefficient, ς, can be considered as the probability of nucleation of a stable pit from a metastable pit and is termed the 'survival probability.' This parameter can be measured experimentally. Thus, for example, for Type 304L stainless steel in chloride-containing solution, the experimentally measured survival probability has a value of the order of 10^{-2} to 10^{-4},[31,32] depending upon the potential and solution conditions. Accordingly, the most probable number of stable pits (per square centimeter), N_0, will be $N_0 = \varsigma N_{mp,0}$, where $N_{mp,0}$, is the most probable number of available sites (per square centimeter) for metastable pits.

In many practical cases, it is possible to assume that all stable pits on a given surface nucleate during an initial period of time that is much smaller than the observation time, t, or the service life of the system, t_s. In this case, the process is termed 'instantaneous nucleation.' For example, for the case of the pitting corrosion of aluminum in tap water, as described by Aziz,[10] practically all the pits were found to nucleate within the first 2 weeks (see **Figure 1**). Under these conditions, the total number of nucleated stable pits (per square centimeter) that nucleate in the time interval between 0 and t can be simply represented as

$$N(t) = N_0 U_+(t) \qquad [27]$$

where $U_+(t)$ is the asymmetrical unit function ($U_+ = 0$ at $t \leq 0$ and $U_+ = 1$ at $t > 0$).

If pit nucleation cannot be regarded as being 'instantaneous,' the simplest assumption concerning the pit nucleation rate of metastable sites, $n_{mp}(t) = dN_{mp}/dt$, is that $n_{mp}(t)$ is proportional to the number of available metastable sites, that is, $dN_{ms}/dt = [N_{mps,0} - N_{mp}(t)]/t_0$,[32] which yields

$$N(t) = N_0[1 - \exp(-t/t_0)] \qquad [28]$$

where t_0 is some characteristic time. It is evident that t_0 must depend on the corrosion potential, temperature, and electrolyte composition, and some experimental data indicating such dependencies can be found in the literature.[31–34] However, no theoretical (deterministic) model has been reported for estimating t_0 as a function of the environmental conditions and the kinetic parameters of the system.

2.39.5.2 Point Defect Model

A comprehensive model, in the form of the PDM[35–39] has been developed for estimating the nucleation rate of metastable pits which, when combined with eqn [26], yields the sought-after nucleation rate of stable pits. (Note that only the growth of stable pits gives rise to pitting damage on a surface.) The PDM was originally developed in the early 1980s to provide an atomic scale description of the growth of passive films on a metal surface, but was subsequently expanded to describe metastable passivity breakdown. Thus, it is clearly evident that any deterministic model for describing pit nucleation rate must simultaneously describe the properties of passive films existing on the metal surface and clearly specify the criteria for passivity breakdown itself. The conditions under which passive films exist on metal surfaces are a matter of great theoretical and practical interest, because the phenomenon of passivity is the enabler for our current, metals-based civilization.[21] Thus, our industrial systems and machines are fabricated primarily from the reactive metals and their alloys, including iron, nickel, chromium, aluminum, titanium, copper, zinc, zirconium, stainless steels, nickel-base alloys, and

aluminum alloys, to name a few. Although the phenomenon of 'passivity' has been known for about 170 years,[40] and the conditions under which metals and alloys become passive have been systematically explored over the past 70 years, until recently no reasonably unifying theoretical treatment of the limits of passivity has emerged. While many theories and models for the passive state have been developed,[21] most of the presently available models describe an already existing passive film and do not address the conditions under which the film may form or disappear. One of the few attempts to address this issue is the one by Engell,[41] who postulated that passive films can be thermodynamically stable or metastable, with film formation being governed by equilibrium thermodynamics in the first case and by the relative rates of formation and dissolution in the second. While Engell's work[41] made a valuable contribution to the theory of passivity, it did not resolve the theoretical issues with sufficient precision to allow specification of the exact conditions under which passivation/depassivation might occur (see below).

A comprehensive review of the conditions under which passivity may occur and be lost has been explored within the framework of the PDM by using phase–space analysis (PSA) and can be found in Macdonald.[42] It has been shown that the PDM provides a comprehensive basis for describing the formation and destruction of passive films and hence allows specification of the conditions for the use of reactive metals in our metals-based civilization. A brief description of some results is given below.

The PDM postulates that passive films that form on metal and alloy surfaces in contact with oxidizing environments are bilayer structures comprising a highly (point) defective barrier layer, which grows into the metal and an outer layer that forms via the hydrolysis of cations transmitted through the barrier layer and the subsequent precipitation of a hydroxide, oxyhydroxide, or oxide, depending upon the formation conditions, or by transformation of the outer surface of the barrier layer itself (an 'Ostwald ripening' process). In many systems (e.g., Ni and Cr), the barrier layer appears to be substantially responsible for the phenomenon of passivity. In other systems, such as the valve metals and their alloys (Al, Ta, Ti, Nb, Zr), and iron (particularly at elevated temperatures), for example, the outer layer may form a highly resistive coating that effectively separates the reactive metal and the barrier layer from the corrosive environment. The 'sealing' of anodized aluminum is an example of how the outer layer may be manipulated to achieve high corrosion resistance. In the present analysis, only the barrier layer is considered, because the passivity of chromium-containing alloys appears to be due to a thin barrier layer of defective Cr_2O_3 that forms on the surface in contact with the alloy. Thus, in these cases, the barrier layer is clearly 'the last line of defense.'

The PDM further postulates that the point defects present in a barrier layer are, in general, cation vacancies ($V_M^{\chi'}$), oxygen vacancies ($V_O^{\bullet\bullet}$), and cation interstitials ($M_i^{\chi+}$), as designated by the Kroger–Vink notation (**Figure 4**). Cation vacancies are electron acceptors, which result in doping the barrier layer

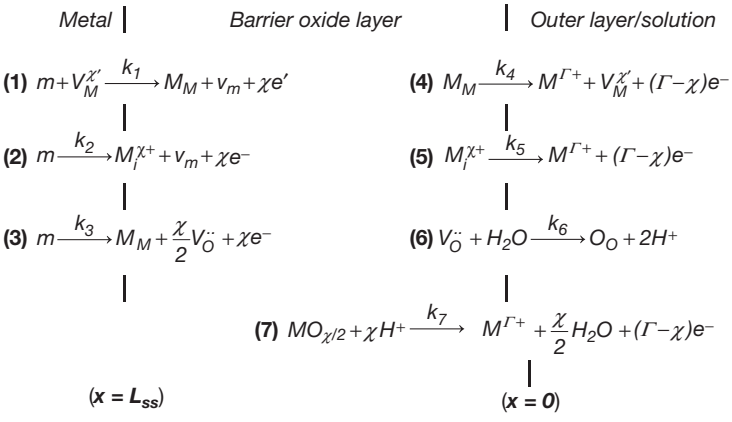


Figure 4 Interfacial defect generation/annihilation reactions that are postulated to occur in the growth of anodic barrier oxide films according to the point defect model.[35] M, metal atom; $V_M^{\chi'}$, cation vacancy on the metal sublattice of the barrier layer; $M_i^{\chi+}$, interstitial cation; M_M, metal cation on the metal sublattice of the barrier layer; $V_O^{\bullet\bullet}$, oxygen vacancy on the oxygen sublattice of the barrier layer; O_O, oxygen anion on the oxygen sublattice of the barrier layer, $M^{\Gamma+}$, metal cation in solution.

p-type, whereas oxygen vacancies and metal interstitials are electron donors, resulting in n-type doping. Thus, on both pure metals and alloys, the barrier layer is essentially a highly doped, point defect semiconductor, as demonstrated by Mott–Schottky analysis[21] for example. Not unexpectedly, the situation with regard to alloys is somewhat more complicated than that for the pure metals. Thus, while the barrier layers on pure chromium and on Fe–Cr–Ni alloys (including the stainless steels) are commonly described as being 'defective Cr_2O_3,' those on pure chromium are normally p-type in electronic character[43] and those on the stainless steels[44] are n-type. The latter can be described as $Cr_{2+x}O_{3-y}$, recognizing that the barrier layer may be metal rich (via metal interstitials) or oxygen-deficient (via oxygen vacancies), or both, whereas that on pure chromium apparently is metal- and oxygen-deficient, or $Cr_{2-x}O_{3-y}$, with the cation vacancy being the dominant defect in the system. It is not known whether the apparent differences in the barrier layers on pure chromium and on chromium-containing alloys are due to doping of the barrier layer by other alloying elements, or the inhibition of cation vacancy generation relative to the generation of oxygen vacancies and metal interstitials, in the barrier layer on the alloys compared with that on pure chromium.

The defect structure of the barrier layer can be understood in terms of the set of defect generation and annihilation reactions occurring at the metal–barrier layer interface and at the barrier layer–outer layer (solution) interface, as depicted in **Figure 4**.[42] Regardless of the electronic type, that is, irrespective of the identity of the dominant defect in the system, reactions [3] and [7] (**Figure 4**) are responsible for the growth and destruction of the barrier layer and any analysis of the stability of the layer must focus on these two reactions. That the barrier layer always contains oxygen vacancies is self-evident, since the rate of dissolution at the barrier layer–solution interface is always finite.

As noted elsewhere,[21,42] the rate of change of the barrier layer thickness for a barrier layer that forms irreversibly on a metal or alloy surface can be expressed as

$$\frac{dL}{dt} = \Omega k_3^0 e^{a_3 V} e^{b_3 L} e^{c_3 \mathrm{pH}} - \Omega k_7^0 (C_{H^+}/C_{H^+}^0)^n e^{a_7 V} e^{c_7 \mathrm{pH}} \quad [29]$$

where $a_3 = \alpha_3(1-\alpha)\chi\gamma, a_7 = \alpha_7\alpha(\Gamma-\chi)\gamma, b_3 = -\alpha_3\chi\varepsilon\gamma$, $c_3 = -\alpha_3\chi\beta\gamma$, and $c_7 = \alpha_7\beta(\Gamma-\chi)\gamma$. In these expressions, Ω is the mole volume of the barrier layer per cation, ε is the electric field strength within the barrier layer (postulated to be a constant and independent of the applied voltage in the steady state, because of the buffering action of Esaki tunneling),[21] k_i^0 and α_i are the standard rate constant and transfer coefficient, respectively, for the appropriate reactions depicted in **Figure 4** (i.e., reactions [3] and [7]), α is the polarizability of the barrier layer–solution (outer layer) interface (i.e., the dependence of the voltage drop across the interface, $\phi_{f/s}$, on the applied voltage, V), β is the dependence of $\phi_{f/s}$ on pH (assumed to be linear), $\gamma = F/RT, \chi$ is the oxidation state of the cation in the barrier layer, Γ is the corresponding quantity for the cation in solution, C_{H^+} is the concentration of hydrogen ion, $C_{H^+}^0$ is the standard state hydrogen ion concentration, and n is the kinetic order of the barrier layer dissolution reaction with respect to H^+.

By setting the left side of eqn [29] equal to zero, the steady state thickness of the barrier layer, L_{ss}, is readily derived as

$$L_{ss} = \left[\frac{1-\alpha}{\varepsilon} - \frac{\alpha\alpha_7}{\alpha_3\varepsilon}\left(\frac{\Gamma}{\chi}-1\right)\right]V + \left[\frac{2.303n}{\alpha_3\varepsilon\chi\gamma} - \frac{\alpha_7\beta}{\alpha_3\varepsilon}\left(\frac{\Gamma}{\chi}-1\right) - \frac{\beta}{\varepsilon}\right]\mathrm{pH} + \frac{1}{\alpha_3\varepsilon\chi\gamma}\ln\left(\frac{k_3^0}{k_7^0}\right) \quad [30]$$

Note that in deriving these expressions, the convention has been adopted that, for the rate of barrier layer dissolution, C_{H^+} and $C_{H^+}^0$ have units of mol cm^{-3}, but when used for defining pH, the units are the conventional mol 1^{-1}. Thus, the standard states for the dissolution reaction (second term on the right side of eqn [39]) and for the pH are 1.0 mol cm^{-3} and 1.0 mol 1^{-1}, respectively. The introduction of a standard state into the dissolution rate renders the units of k_7^0 independent of the kinetic order, n, without altering the numerical value of the rate.

The steady state passive current density is readily derived[21] as

$$I_{ss} = \Gamma F\left[k_2^0 e^{a_2 V} e^{b_2 L_{ss}} e^{c_2 \mathrm{pH}} + k_4^0 e^{a_4 V} e^{c_4 \mathrm{pH}} + k_7^0 e^{a_7 V} e^{c_7 \mathrm{pH}} \cdot (C_{H^+}/C_{H^+}^0)^n\right] \quad [31]$$

where the first, second, and third terms arise from the generation and transport of cation interstitials, cation vacancies, and oxygen vacancies, respectively, with the term due to the latter being expressed in terms of the rate of dissolution of the barrier layer.[21] This expression is derived, in part, by noting that the fluxes of a given defect at the two interfaces under steady-state conditions are equal; in this way, the

expression for the current can be formulated so as to avoid the defect concentrations at the interfaces.

The passive state is not perfectly protective and, for a variety of reasons, passivity breakdown occurs, resulting in enhanced corrosion rates. Of particular concern is localized passivity breakdown, which results in the nucleation and growth of pits and subsequently, the nucleation and growth of cracks if, the requisite tensile stress is present in the system. A review of the literature reveals the following generalizations of the experimental data for passivity breakdown of metals and alloys in a wide variety of environments[21]:

1. Localized corrosion is initiated by passivity breakdown and occurs on a wide variety of passive metals and alloys in a wide variety of environments.
2. Certain species (e.g., Cl^- and Br^-) induce passivity breakdown by interacting with the barrier layer. These aggressive species apparently do not penetrate through the barrier layer but may be incorporated into the precipitated outer layer.
3. Passivity breakdown occurs at a wide variety of sites on metal and alloy surfaces.
4. Passivity breakdown is a dynamic deterministic process, being predetermined and (in principle) predictable on the basis of known physical laws.
5. The transition of a metastable event to a stable event is a rare event.
6. Two fundamentally different repassivation phenomena may be identified: (i) 'Prompt' repassivation and (ii) 'delayed' repassivation (sometimes referred to as 'stifling').
7. A single passivity breakdown site is characterized by a critical voltage (V_c) and induction time (t_{ind}). V_c is found to be near-normally distributed while t_{ind} displays a left acute distribution. The parameters V_c (and $\bar{V}_c$) and t_{ind} exhibit highly characteristic dependencies on the activities of the breakdown-inducing aggressive species (a_x) and on the applied voltage (t_{ind} only) for a wide variety of systems, suggesting commonality in mechanism.
8. V_c (and $\bar{V}_c$) is found to depend on the identity of the aggressive ion within a homologous series. Thus, in the case of iron and nickel, the propensities of the halide ions for inducing passivity breakdown lie as $F^- < Cl^- > B^- < I^-$, whereas, in the case of titanium, bromide ion is the most aggressive. These trends are readily explained by the absorption of the halide into oxygen vacancies in the surface of the barrier layer, with the extent of absorption being determined by the competitive needs to dehydrate the ion and expand the vacancy.[21]
9. The mean breakdown voltage, $\bar{V}_c$ is found to decrease linearly with $\log(a_X)$ with a slope that exceeds $2.303RT/F$, which is attributed to the value of the polarizability of the barrier layer–solution interface lying between 0 and 1, for essentially all systems (metal–solution) that have been investigated. Likewise, the induction time for essentially all systems investigated display a common form of the dependencies of $\log(t_{ind})$ on potential and $[X^-]$. These relationships strongly suggest commonality in breakdown mechanism.
10. Certain oxyanions (e.g., nitrate, borate, and nitrite) strongly inhibit passivity breakdown, with the effect being accounted for by competitive absorption with the aggressive anion into surface oxygen vacancies in the barrier layer.[21]
11. In many systems (e.g., Al, Ga, Zr, stainless steel), blister formation is observed to be the precursor to passivity breakdown.
12. Certain alloying elements (e.g., Mo in Ni) cause a positive shift in V_c (and $\bar{V}_c$) and in a lengthening of the induction time. The effect is greater for a larger difference in the oxidation states between the solute and host.
13. Incident electromagnetic radiation, with a photon energy that is greater than the bandgap of the barrier layer oxide also results in a positive shift in V_c (and $\bar{V}_c$) and in a lengthening of the induction time. The defect (electronic and crystallographic) structures of the barrier layer are modified by irradiation.

It has been shown that PDM is able to explain all the experimental data for passivity breakdown of metals and alloys mentioned above and accordingly, can be considered as providing a suitable theoretical basis for a deterministic model treatment of passivity breakdown and the nucleation of metastable pits in a metal surface.[21]

The PDM, as it relates to passivity breakdown, postulates that certain aggressive anions, for example, F^-, Cl^-, Br^-, and I^- absorb into oxygen vacancies in the surface of the barrier layer, resulting in the generation of cation vacancies and hence, to an enhanced flux of the same species across the barrier layer toward the metal–barrier layer interface, as depicted in **Figure 5**. The PDM postulates that if the cation vacancies arriving at the metal–barrier layer interface

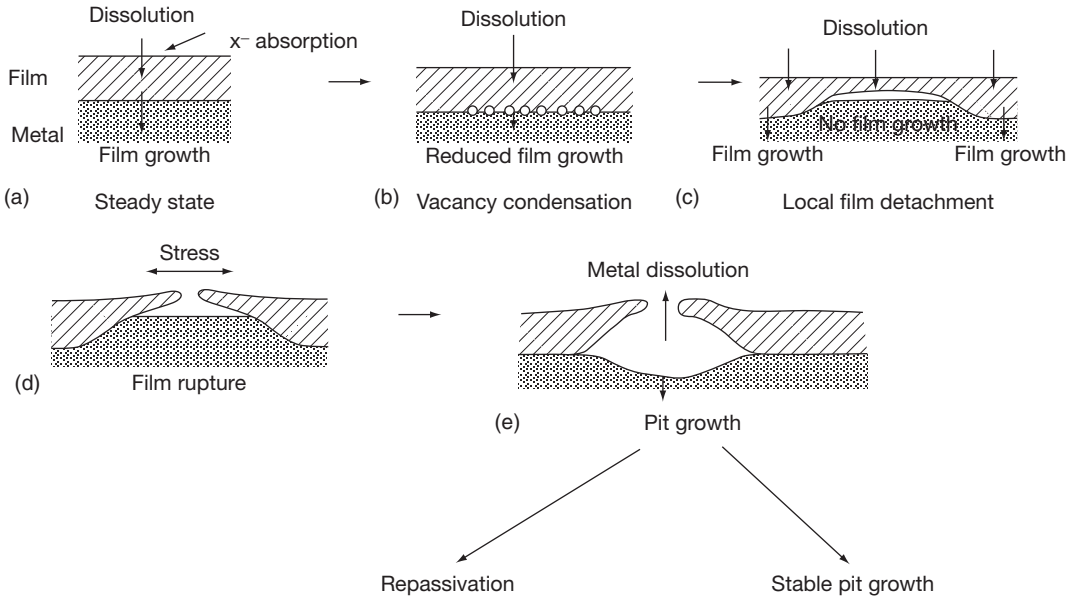


Figure 5 Sequence of events in passivity breakdown, according to the point defect model.[21] Note that the initial event is the absorption of an aggressive anion into an oxygen vacancy at the film–solution interface (a), resulting in an enhanced flux of cation vacancies across the film and eventually in condensation of cation vacancies at the metal–film interface (b). The film stops growing into the metal beneath the cation vacancy condensate while it continues to dissolve at the film–solution interface (c), eventually resulting in rupture (d) and repassivation or the formation of a stable pit (e).

cannot be annihilated at a sufficiently high rate via reaction [1] (**Figure 5**), the excess vacancies will condense locally and hence cause local separation of the barrier layer from the substrate metal. Once separation has occurred, reaction [3] (**Figure 5**) can no longer occur so that the barrier layer at that locale is prevented from growing into the metal. However, the barrier layer continues to grow into the metal at the periphery of the cation vacancy condensate and also continues to dissolve at the barrier layer–solution interface. This results in local thinning of the 'cap' over the cation vacancy condensate with the cap eventually rupturing because of the growth stresses in the film and in the near-surface substrate. The 'weak points' on the surface where passivity breakdown is predicted to occur correspond to regions of high cation vacancy flux. These regions are assumed to be regions of high local discontinuity in the barrier layer, such as the points of intersection of the barrier layer with precipitates, inclusions (e.g., MnS), and other 'second phase' particles. Support for this mechanism stems from the almost general observation of the formation of blisters ('cation vacancy condensate') as precursors to passivity breakdown on a wide variety of metals and alloys, the passivity breakdown on liquid versus solid gallium, and the potential sweep rate dependence of the apparent breakdown voltage.[21] The latter evidence is particularly convincing because the test involves no adjustable parameters and yields a quantity (the concentration of condensed vacancies) that can be compared with the same quantity calculated from fundamental principles (crystal structure of the barrier layer).

The description of possible mechanisms for the generation of cation vacancies at the barrier layer–solution interface upon the absorption of the aggressive anion into a positively charged oxygen vacancy in the surface of the barrier layer can be found in Macdonald.[21]

Mathematically, the condition for the initial formation of the cation vacancy condensate can be expressed as[35]

$$(\mathcal{J}_{ca} - \mathcal{J}_m)(t - \tau) \geq \xi \qquad [32]$$

where $\mathcal{J}_{ca}$ is the flux of cation vacancies across the barrier layer at the breakdown site, $\mathcal{J}_m$ is the annihilation flux (i.e., the rate of reaction [1], **Figure 5**), t is the time, τ is the dissolution time (see below), and ξ is the critical areal concentration of vacancies ($\#/cm^2$). Noting that $\mathcal{J}_{ca}$ is voltage dependent, the critical breakdown voltage corresponds to that at which breakdown takes an infinite time to occur that is, when $\mathcal{J}_{ca} = \mathcal{J}_m$. This condition, in turn, leads to the following expressions for the breakdown voltage and

the induction time for passivity breakdown at a single site on the metal surface as[35]

$$V_C = \frac{4.606RT}{\chi\alpha F}\log\left[\frac{\mathcal{J}_m}{\mathcal{J}^O u^{-\chi/2}}\right] - \frac{2.303RT}{\alpha F}\log a_X \quad [33]$$

$$t_{ind} = \xi'\left[\exp\left(\frac{\chi\alpha F\Delta V}{2RT}\right) - 1\right]^{-1} + \tau \quad [34]$$

where

$$\mathcal{J}^O = \chi KD[N_v/\Omega]^{1+\chi/2}\exp[-\Delta G_S^0/RT] \quad [35]$$

and

$$u = (N_A/\Omega)\exp[\Delta G_A^0/RT]\exp[-F(\beta pH + \varphi_{f/s}^0)/RT] \quad [36]$$

where N_A is the Avogadro number, ΔG_S^0 is the Gibbs energy change for the Schottky-pair reaction, β is the dependence of the potential drop across film–solution interface on pH, $\varphi_{f/s}^0$ is a constant (potential drop at the film–solution interface for $V_{app} = 0$, and pH = 0) and ΔG_A^0 is the standard Gibbs energy change for the chloride–oxygen vacancy absorption reaction.

Additionally, the 'relaxation' time, τ, which is now identified with the time taken for the cap over the vacancy condensate to thin sufficiently for rupture to occur[21] from the point of initial cation vacancy condensation, can be expressed as

$$\tau \le L_{ss}/(dL/dt)_{dissolution} = L_{ss}/\Omega k_s\left(C_{H^+}/C_{H^+}^0\right)^n \quad [37]$$

In these expressions, L_{ss} is the steady-state thickness of the barrier layer at the voltage at which cation vacancy condensation just begins. This value is given by eqn [30]. The other parameters are as defined above.

Typical plots of breakdown voltage versus chloride activity, plotted in accordance with eqn [33] are presented in **Figure 6**. The data labeled 'PDM' were calculated from eqn [33] using parameter values estimated by different experiments (primarily from electrochemical impedance spectroscopy, EIS). It should be noted that the PDM has been subjected to numerous experimental tests and, to the authors' knowledge, no substantial discrepancies have been noted.[21] The model has also been extended to account for transients in barrier layer thickness and passive current in response to potential step and linear potential sweep perturbations. The model has also been extended to account for electrochemical impedance data and, indeed, optimization of the model on impedance data as a function of frequency and voltage has proven to be a very effective way of determining values for various model parameters. Finally, the PDM has also been used to describe the cathodic formation of metal hydride films on metals, such as Li and Zr, and, again, optimization has proven to be an effective means of determining values for model parameters. Discussion of these topics is beyond the

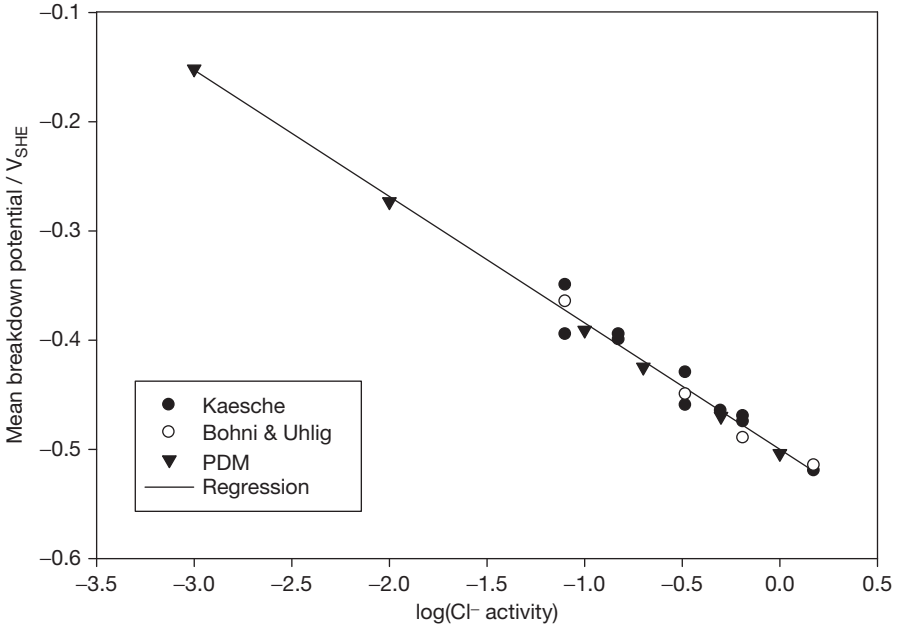


Figure 6 Plot of the mean critical breakdown potential vs. log(Cl^- activity) for aluminum in sodium chloride solutions at 25 °C. Experimental data are taken from Kaesche[45] and Böhni and Uhlig.[46]

scope of the present chapter, and the reader is referred to the literature for additional information.[21]

On any real surface, a large number of potential breakdown sites exist, corresponding to a distribution in the properties of the 'weak spots.' Thus, examination of data reported by Shibata[47] and Fratesi,[48] among others, suggests that the breakdown voltage is nearly normally distributed. The PDM assumes that the breakdown sites, with respect to the diffusivity of cation vacancies, are described approximately by a normal distribution function,[39] that is,

$$\frac{dN(D)}{dD} = -A \exp\left[-\frac{(D-\bar{D})^2}{2\sigma_D^2}\right] \quad [38]$$

Here, $N(D)$ is the number of breakdown sites (per square centimeter) that have diffusivities larger than D, and $\bar{D}$ and σ_D are the average value and the standard deviation, respectively, of the diffusivity for the population of the breakdown sites. The negative sign in eqn [38] means that $N(D)$ decreases with increasing D. Parameter A does not depend on D, so that normalization of the diffusivity distribution using the condition $N(0) = N_{mp,0}$, where $N_{mp,0}$ is the total number of breakdown sites (per square centimeter), yields

$$N(D) = N_{mp,0}\ \mathrm{erfc}\left(\frac{D-\bar{D}}{\sqrt{2}\sigma_D}\right)/\mathrm{erfc}\left(-\frac{\bar{D}}{\sqrt{2}\sigma_D}\right) \quad [39]$$

Because the transport of cation vacancies across the barrier layer from the barrier layer–solution interface to the metal–barrier layer interface occurs primarily by electro-migration, the cation vacancy flux density, $\mathcal{J}_{ca}$, is proportional to the diffusion coefficient of the vacancies, D, that is,

$$\mathcal{J}_{ca} = DB \quad [40]$$

where the function B depends on the external conditions (applied voltage, V_{app}, temperature, T, chloride activity, a_x, etc.) and on the electric field strength within the film. Thus, for the case of passivity breakdown in a solution containing an aggressive anion, X^- (e.g., chloride ion), the PDM yields

$$B = \hat{a}u^{-\chi/2}\exp\left(\frac{\chi F\alpha V_{app}}{2RT}\right)a_x^{\chi/2} \quad [41]$$

where α is the polarizability of the film–solution interface (i.e., dependence of the potential drop across the barrier layer–solution interface on the applied potential), χ is the cation oxidation state in the barrier layer, R is the gas constant, and F is Faraday's constant. In turn, parameters $\hat{a}$ and u are defined as[35]

$$\hat{a} = \chi(F\varepsilon/RT)[N_A/\Omega]^{1+\chi/2}\exp\left(-\Delta G_S^0/RT\right) \quad [42]$$

where parameter u is defined by eqn [36].

Substituting eqns [40] into eqn [32] yields the following criterion for metastable pit nucleation as

$$D \geq D_{cr} = \frac{\xi + \mathcal{J}_m(t-\tau)}{B(t-\tau)} \quad [43]$$

Criterion [43] states that the nucleation of metastable pits occurs within the observation time, t, on those and only those sites that have $D \geq D_{cr}$. From Equations [26], [39], and [43], we have

$$N(t) = N_0\ \mathrm{erfc}\left(\frac{a}{t-\tau}+b\right)/\mathrm{erfc}(b) \quad [44]$$

where $a = \xi/\left(B\sqrt{2}\sigma_D\right)$, and $b = (\mathcal{J}_m/B - \bar{D})/\left(\sqrt{2}\sigma_D\right)$. Accordingly, for the rate of pit nucleation, we obtain the following expression,

$$n(t) = \frac{dN}{dt} = \frac{N_0 2a}{\mathrm{erfc}(b)\sqrt{\pi}}\frac{\exp\left[-\left(\frac{a}{t-\tau}+b\right)^2\right]}{(t-\tau)^2} \quad [45]$$

It is important to note that, in accordance with eqn [28], the maximum pit nucleation rate must be observed at the beginning of corrosion attack, that is, at $t = 0$. On the other hand, in accordance with PDM, as follows from eqn [45], the maximum pit nucleation rate must be observed at

$$t_{max} = a\left(b+\sqrt{b^2+4}\right)/2 \quad [46]$$

Accordingly, if observation time, t (service life t_s), satisfies the condition

$$t \gg t_{max} \quad [47]$$

practically all pits nucleate during a very short period of time at the beginning of the observation time, that is, the PDM predicts the case of instantaneous nucleation. Calculation shows[13] that, in some cases (especially at high concentrations of Cl^-) criterion [47] holds very well and the nucleation of pits on a metal surface may be regarded as an 'instantaneous nucleation' phenomenon. However, t_{max}, increases very sharply with decreasing chloride concentration, leading to the case of progressive nucleation, in which new pits nucleate on the surface as existing pits grow and die (repassivate). The fact that the surface concentration of pits grows at a maximum rate after the appearance of the first pit has been experimentally observed, for example,

the case of the pitting corrosion of passive iron in borate buffer solution containing chloride ion.[49]

2.39.6 Rate of Pit (Cavity) Propagation

The quantitative description of pit (or cavity) growth remains one of the key problems in predicting corrosion damage in many practical systems. This follows from the fact that the calculated corrosion damage that is based only on this (growth) stage can be compared with experiments in many limiting cases. For example, in the case of pitting corrosion, when all pits nucleate 'instantaneously,' or when the induction time for pit nucleation is much smaller than the observation time, it is possible to ignore the initial stage of pit nucleation when estimating the damage. In addition, if the probability of survival of a corrosion defect is sufficiently high, we must take into account the possibility that a stable corrosion defect (pit or crack) nucleates immediately after the start of operation and propagates without repassivation and hence, the same defect that nucleated in the beginning is still active at the end of the observation time. In any case, calculations based only on the growth stage yield the most conservative estimate of the service life, $t_{s,min}$, of the system. We can be sure that if calculation of the service life is based on growth alone, the real service life, t_s, will at least be not less than $t_{s,min}$.

Moreover, it is also natural to assume that the rate of propagation of an individual pit (crack) without neighbors will be greater than that for the same pit (crack) with neighbors (i.e., within an ensemble of pits (cracks)), because the neighboring pits reduce the potential at the surface and multiple, neighboring cracks are expected to reduce the stress intensity factor experienced by a single crack for a given loading stress. Accordingly, the proper modeling of propagation of individual corrosion defects can yield the possible minimum survival time for the system as a whole (e.g., a machine). If this time extends beyond the projected service life of the system, we can be sure that the system will survive under real conditions.

2.39.6.1 General Approach for Calculating Propagation Rates

Apparently, one of the main theoretical problems in describing localized corrosion damage is the prediction of the shape and dimensions of corrosion cavities at any time as a function of the parameters controlling the process (potential of the metal, species concentrations, such as velocity of the electrolyte, thickness of the passive film, diameter of the pit mouth, etc.). It seems obvious that mathematical models describing the pit (cavity) growth should start with the assumption that the pit form and size are not known *a priori*, but should be found during the solution of the appropriate equations.

From the mathematical point of view, the problem of determining the shape and size of a developing pit belongs to the class of Stefan problems.[50] These problems are reduced to the solution of the system differential equations of parabolic or elliptic type with unknown boundary conditions because the velocity of this boundary is connected to the sought after solution by some differential relationship.

The implicit equation $\Phi(t, x_1, x_2, x_3) = 0$ shall describe the pit surface with time, relative to the Cartesian coordinate system x_1, x_2, x_3. This expression fulfills the relation $d\Phi = (\partial\Phi/\partial t + \nabla\Phi \cdot V_s)dt = 0$, where $V_s(t, dx_1/dt, dx_2/dt, dx_3/dt)$ is the velocity of the dissolving metal surface. In this formulation, subscript refers to the electrode surface and subscript ∞ to the bulk of the solution. Let $\vec{n} = \nabla\Phi/|\Phi|$ be a unit vector pointing normally from the solution to the metal surface. According to Faraday's law, $V_s = K_V i_{corr}\vec{n}$, where i_{corr} is the current density for metal dissolution, and K_v is the electrochemical equivalent volume, that is, the volume of dissolving metal when one Faraday of charge (96 487 int. coulombs) passes through the interface.

Introducing the correlation for V_s into the expression for the differential $d\Phi$ yields an equation describing the change of the electrode surface

$$\frac{\partial\Phi}{\partial t} + K_V i_{corr}|\nabla\Phi| = 0 \qquad [48]$$

Equation [48] is frequently used to describe some important processes in electrochemical technology, for example, electrochemical machining of metals, electrochemical shape formation, and so on.[51] Of course, we can also expect that, in the general case, some additional component, not connected with metal dissolution (e.g., due to mechanical spallation at very high fluid velocities or hydrogen embrittlement) should be incorporated into the expression for the V_s, but this issue is not germane to the present discussion and will not be considered further here.

Besides the information about the boundary position in the initial period, $\Phi(0, x_1, x_2, x_3)$, the integration of eqn [48] requires information about the

dissolution current density, i_{corr}, that is, eqn [48] must be solved with a set of nonstationary equations for mass balance for each component in the solution, namely

$$\frac{\partial c_k}{\partial t} = -\nabla \cdot \vec{\jmath}_k + R_k; \quad k = 1, \ldots, K \qquad [49]$$

where C_k is the concentration of species k, $\vec{\jmath}_k$ is the flux density of species k, and R_k is the rate of production (source) or depletion (sink) of this species as a result of homogeneous chemical reactions. The current density, $\vec{i}$, in the electrolyte solution is defined as

$$\vec{i} = F \sum_k z_k \vec{\jmath}_k \qquad [50]$$

In the vast majority of practical applications, dilute solution theory is used to calculate the flux of the species, that is,

$$\vec{\jmath}_k = -D_k \nabla C_k - z_k F u_k C_k \nabla \phi + C_k \vec{v} \qquad [51]$$

where ϕ is the electrostatic potential in the solution, v is the fluid (hydrodynamic) velocity, and u_k is the mobility of species k, which can be estimated by using the Nernst–Einstein equation:

$$u_k = \frac{D_k}{RT} \qquad [52]$$

In eqn [51], the first term on the right-hand side represents the contributions of diffusion, the second term describes migration, and the third term is a contribution of convection. An additional condition for determining the electrostatic potential is the equation

$$\sum_k z_k C_k = 0 \qquad [53]$$

It must be noted that, generally speaking, the condition of electroneutrality is not a law of nature and can be considered as an approximation of Poison's equation:

$$\nabla^2 \phi = -\frac{F}{\varepsilon} \sum_k z_k C_k \qquad [54]$$

where ε is the dielectric permittivity (dielectric constant multiplied by the permittivity of free space) of the solution. However, due to the large value of the ratio F/ε, an appreciable separation of charge would require unrealistically large electric forces. Visible deviation from electroneutrality can be observed only in a very thin double layer near an electrode surface (which is of the order of 1–10 nm in thickness) or within a doped semiconductor junction that can be taken into account in the boundary conditions for the problem. Accordingly, the electroneutrality approximation is fulfilled very well, and is widely accepted by the electrochemical community.[52] In spite of this fact, some papers where Poison's equation is used directly can be found in the literature.[53] In accordance with our opinion, this approach introduces only unnecessary complications, due to excessive calculational time.

The hydrodynamic velocity, $\vec{v}$, within the framework of dilute solution theory, can be found apart from the solution of the mass transfer problem by using the corresponding mechanical equations (e.g., the Navier–Stokes equations, in the case of laminar flow[52,54]). In the case of turbulent flow, which often exists outside the corrosion cavity, it is convenient to use some effective diffusion coefficient instead of the usual diffusion coefficient, D_k,

$$D_k^{\text{eff}} = D_k + D_t \qquad [55]$$

where diffusion coefficient D_t depends on the distance from the wall, hydrodynamic conditions, and the physical properties of the liquid. Some empirical correlations for D_t can be found in the literature.[52,54,55]

Strictly speaking, eqn [51] for species flux densities is valid in the case of dilute solutions. In the case of concentrated solutions, eqn [51] should be replaced by

$$\vec{\jmath}_k = C_k \vec{v}_k \qquad [56]$$

where the velocity, $\vec{v}_k$, of species k can be found from the equation of multicomponent diffusion.[52]

$$C_k \nabla \mu_k = RT \sum_i \frac{C_k C_i}{C_T D_{ki}} (\vec{v}_i - \vec{v}_k) \qquad [57]$$

where μ_k is the electrochemical potential of species k, D_{ij} are diffusion coefficients, and C_T, is the sum of all concentrations of all components, including the solvent. However, in the literature, there are no significant applications of the theoretical principles describing transport in concentrated solutions for the case of corrosion, with the exception of the preliminary work of Popov *et al.*[56] The reasons for this state of affairs are the increased complexity of the theory and the lack of appropriate input data for practical application.

However, it is important to note that, in the overwhelming majority of corroding systems, the concentrations of all solute species, C_k, are much smaller than the concentration of the solvent, C_0. Under these

conditions, only one term on the right side of eqn [50] is important. Accordingly, eqn [57] can be rewritten in the following form for the theory of moderately dilute solutions[52]

$$\vec{J}_k = -D_k \nabla \mu_k + C_k \vec{v} \qquad [58]$$

If we present μ_k in the form

$$\mu_k = RT \ln(\gamma_k C_k) + z_k F \varphi \qquad [59]$$

where the activity coefficient, γ_k, depends on the concentration of all species in the solution, the equation of the flux becomes

$$\vec{J}_k = -D_k \nabla C_k - \frac{D_k z_k F}{RT} C_k \nabla \phi + C_k \vec{v} - D_k C_k \nabla \ln(\gamma_k) \qquad [60]$$

This equation for ion flux densities has been used, for example, in Walton *et al.*[57] for describing crevice corrosion. However, the activity coefficients in Walton *et al.*[57] are calculated within the framework of Debye–Hückel theory that, is strictly speaking, applicable only for the case of dilute solutions.

We would like to emphasize that, in many real corrosion systems, the concentration drops within a corroding cavity may be not very high, because corrosion is not a fast process. Accordingly, eqn [1] can be used with sufficient accuracy in relatively concentrated solutions, assuming that D_k is referred to the bulk electrolyte, but not to infinitely dilute solutions.

In accordance with the rules, the homogeneous terms, R_k, in the balance equations can be written in general form as

$$R_k = \sum_{m=1}^{M} \left[-k_m v_{km} \left\{ \prod_{v_{km}>0} C_k^{v_{km}} - K_m \prod_{v_{km}<0} C_k^{-v_{km}} \right\} \right] \qquad [61]$$

where k_m is the rate constant of reaction m, $\sum_k v_{km} M_k = 0$ (m=1,2,...,M), K_m is the equilibrium constant for reaction m, v_{km} is the stoichiometric coefficient for species k in mth chemical reaction, and M_k is the symbol for the chemical formula of species k.

The presence of the chemical terms in balance eqns [49] greatly complicates their solution. First of all, we have practically no information about the rates of homogeneous reactions (usually, we only have information about their equilibrium constants). However, we can reasonably assume that these reactions are fast and they are practically in equilibrium at any given point and time. The most widely applied method for dealing with these problems is to eliminate chemical terms by adding or subtracting balance eqns [49]. The new set of governing equations is then supplemented by equations of equilibrium.[58] The great disadvantage of this approach is as follows. Every time, we want to add or delete from consideration any chemical reaction, we have to completely change the computer code for numerical solution of balance equations because of their change in form. Moreover, the forms of these equations do not coincide with the standard form of the equation of mass balance, and this adds to the complexity of the process of numerical solution.

The alternative approach for modeling the equilibrium state is to keep the governing equation in rate format (i.e., in the form of eqn [49]), and hence to assume that reactions are very fast. As long as reaction rates are large, relative to the rates of mass transport, the reactions will remain at equilibrium, and the solution composition will be independent of the kinetics assumed. All kinetic expressions that are physically consistent (i.e., are stoichiometrically true) yield zero net reaction at equilibrium, change sign as the equilibrium point is crossed, and give sufficiently fast reaction rates that can be used.[59] The choice becomes a matter of numerical stability and convenience. Therefore, Walton[60] used the following expression for the rate of production or depletion of species k by the chemical reaction

$$R_k = \sum_{m=1}^{M} \left[-r_m v_{km} \ln \left\{ \prod_{k=1}^{K} C_k^{v_{km}} / K_m \right\} \right] \qquad [62]$$

which satisfies all of the conditions noted above.

Another alternative approach is the following.[57,61] It is assumed that characteristic times of chemical reactions in aqueous solution are much shorter than those of the mass transport or corrosion processes, which is equivalent to assuming high reaction rates. Accordingly, the set of transport equations are first solved separately from the chemical terms. After that, at the end of each sufficiently small time step, the resulting aqueous solution composition, within each elementary volume, is solved to equilibrium by calling an equilibrium solver. For example Walton *et al.*[57] determined the equilibrium composition of the solution by Gibbs free energy minimization.

The homogeneous terms, R_k, in the balance equations can be written in a general form as: It is assumed that at a point far away from the mouth of a pit or crevice, the concentrations and potentials have their bulk values, that is,

$$C_k = C_{k,\infty}; \phi = \phi_\infty = 0 \quad [63]$$

It is also assumed that the normal fluxes on the solid surface, $\mathcal{J}_{ks}$, can be expressed as a known function of surface temperature, T_s, surface concentrations, C_{ks}, and surface potential, E_{corr}–φ_s, (on the metal surface), that is,

$$-D_k\frac{\partial C_k}{\partial n} = \mathcal{J}_{ks}(E_{corr} - \phi, C_{1,s}, C_{2,s}, \ldots, C_{K,s}, T_s) \quad [64]$$

If some component does not participate in any heterogeneous reaction (chemical or electrochemical), its flux density will be equal to zero. Of course, on an insulator, all fluxes are equal to zero.

In formulating the initial conditions, it would be natural to assume that concentrations of all species and potentials coincide with the corresponding bulk values at $t = 0$. However, simple calculation shows that, under normal conditions, the velocity of the interface is so slow that the steady-state approximation can be used for describing the transport processes for the anodic dissolution of the metal.[50,51] The sense of this approximation is as follows: The movement of the metal surface is so slow that the concentration distribution in the solution is approximately that corresponding to steady flow at a given position of the boundary and for given boundary conditions. In this case, it is possible to omit the derivative over time in balance, eqn [49]. Accordingly, it is possible to separate the solution of the mass transfer problem from the movement of the boundary. After solving the steady-state problem in the region with the fixed boundaries, and after calculating the corrosion current density, the new position of the metal surface is found by using Faraday's law (see eqn [48]). After that, the solution of the steady problem for the new boundary is found, and the process is repeated until the desired time is achieved. Of course, it is not possible to omit the derivative over time in the balance equation in the case of corrosion fatigue, caused by the sharply changing hydrodynamic velocity, which is a result of periodic loading.

However, it must be noted that there are a very limited number of studies where the problem of pit propagation is regarded as a problem having a free boundary in multidimension space. Practically, (in multidimension space) consideration of a moving boundary is limited to the simplest cases, when the system of transport equations can be reduced to the solution of a single Laplace equation for the relative concentration of a single species[62] or electric potential.[63]

On the other hand, such problems are successfully solved in heat transfer studies, for example, in connection with the problem of metal ingot solidification.[64] The presence of migration is the fundamental difference between the problems of ionic transport and those in nonelectrolyte solutions or problems in heat transfer. This difference does not permit the direct use of a wide range of methods and computer programs that have been developed in connection with the problems of heat transfer for solving ion transport problems. As a result, to solve each ion transport problem, for example, using the so-called Newman method,[52] it is necessary to create new programs that require substantial expenditure of time and effort.

It is also possible to use iterative methods to reduce ion transport problems to a sequence of transport problems for nonelectrolyte solutions, which enables one to directly use the methods and even the existing computer programs available in heat engineering.[65–67]

The very useful method of quasipotential transformation for modeling transport processes in dilute electrochemical systems has been developed by Pillay and Newman.[68] It can be shown that, under steady-state quiescent conditions, with equilibrium homogeneous chemical reactions, and for a single electrochemical reaction, the electrostatic potential and all concentration distributions can be represented as a single-valued harmonic function of the quasipotential, q. This harmonic function vanishes at infinity and satisfies the following boundary conditions

$$\frac{\partial q}{\partial \boldsymbol{n}} = 0 \quad [65]$$

at an insulating surface, and

$$-\frac{\partial q}{\partial \boldsymbol{n}} = i \quad [66]$$

at an electrode (conducting) surface, where $\boldsymbol{n}$ is the unit normal vector at the electrode surface pointed towards the solution. Accordingly, the set of coupled second-order, nonlinear partial differential equations governing mass transfer by diffusion and migration in electrochemical systems is transformed into Laplace's equation for the quasipotential and hence, into a set of coupled, first-order nonlinear ordinary differential equations. This method has been applied to the case of the development of an active hemispherical pit.[69] Unfortunately, this method cannot be generalized, for example, for the nonstationary case, for the case of

multiple electrochemical reactions, or for mass transfer in flowing electrolytes.

It is also important to note that, at the present time, commercial software is available for solving transport equations with migration terms with fixed boundaries (see, e.g., COMSOL software),[70] although some algorithms also handle moving boundaries now. Accordingly, the most difficult problem that arises very often during the solution of mass transfer problems is not the solution of the differential equations themselves, but in defining the transport coefficients (e.g., diffusion coefficients) in multicomponent systems, rates of chemical reactions, and especially in establishing real kinetic relations for the species fluxes on the metal surfaces (particularly for the rate of metal dissolution), such as those contained in eqn [64]. The concrete dependencies of the rates of corrosion propagation as a function of potential, concentrations of species near the metal surface, and mechanical conditions for the cases of pitting corrosion, stress corrosion cracking, corrosion fatigue, crevice corrosion, etc. are discussed in the corresponding chapters of this book. Commercial software for calculating transport properties (diffusion coefficients, activities, viscosities, density rates of particular electrochemical reactions on particular metals and alloys) is also available (see, e.g., OLI Systems software).[71]

A detailed review of a great number of papers dealing with mathematical modeling of transport phenomenon in pitting and crevice corrosion is presented, for example, in Turnbull,[58] Sharland,[61] Papavinosan *et al.*[72] In the majority of these models, instead of boundary conditions [64], the boundary conditions of the Type II kind (i.e., prescription of current density on the metal surface) are employed. In the latter case, the principal aim of these models is not to predict corrosion damage (corrosion rates are assumed to be known in advance), but to estimate species concentrations, and the potential distribution in corrosion cavities, as a function of many parameters like cavity dimension, bulk solution composition, temperature, and so on. Very often such models serve to aid in the understanding of results of particular experimental systems. However, at present, the overwhelming majority of the models devoted to estimating the size of a corrosion cavity have adopted the one-dimensional approximation, that is, they are reduced to calculating the depth of the cavity, a, as a function of time, when eqn [48] is reduced to the simplest relation

$$\frac{da}{dt} = K_V i_{corr} \qquad [67]$$

Because of the mathematical complexity of the problem, analyses of the transfer processes occurring within corrosion cavities have generally employed the one-dimensional approximation. For example, it is often assumed, in the case of corrosion pits, that the cavity has a cylindrical shape with a depth a, which is much larger than the radius, r. Likewise, crevices and cracks are often viewed as being one-dimensional slots of length a, such that a is much greater than the opening displacement, w.[57,60,61,73–82] Very often, only metal dissolution at the bottom is assumed. However, two-dimensional analyses of corroding hemispherical pits have been described for the case of a well mixed electrolyte by Newman *et al.*,[83] and for the case of a quiescent system, in which concentration gradients exist.[63,69,84] In many cases, the results obtained by mathematical simulation or by experimental studies of 'ideal' cells (one-dimensional or hemispherical) have been used to describe corrosion processes under real conditions.

In the one-dimensional case, the balance equations are reduced to the form

$$\frac{\partial(wC_k)}{\partial t} = -\frac{\partial w\mathcal{J}_k}{\partial x} + wR_k + \mathcal{J}_{S1} + \mathcal{J}_{S2}, \quad k = 1, 2, \ldots, K \qquad [68]$$

where $\mathcal{J}_k$ is the flux density along the direction down the crevice, x, averaged over the width of the crevice, w; $\mathcal{J}_{S1}$ and $\mathcal{J}_{S2}$ are the flux densities at a metal–solution interface or on the side walls (usually it is assumed that $\mathcal{J}_{S1} = \mathcal{J}_{S2}$). In the case of the cylindrical pit of radius, r, we simply have

$$\frac{\partial C_k}{\partial t} = -\frac{\partial \mathcal{J}_k}{\partial x} + R_k + 2\mathcal{J}_S/a, \quad k = 1, 2, \ldots, K \qquad [69]$$

It must be emphasized that, in the overwhelming majority of the work devoted to describing localized corrosion in the one-dimensional approximation, it has been tacitly assumed that it is possible to neglect the potential drop in the external environment (outside the corrosion cavity), that is, to assume that boundary conditions [56] are fulfilled at the mouth of the crevice. The following section will show that such an approximation is incorrect from the physical point of view and in many cases can lead to significant errors in the estimation of corrosion damage.

2.39.6.2 Coupled Environment Corrosion Cavity Growth Models

Many localized corrosion processes, including pitting, stress corrosion cracking, corrosion fatigue, and crevice corrosion may be described within the framework of the differential aeration hypothesis (DAH), which was first postulated by U. R. Evans in the 1920s.[85] This postulate attributes localized corrosion to a spatial separation of the local anode and local cathode, with the local anode occurring in that region of the system that has the least access to the cathodic depolarizer (e.g., oxygen), while the local cathode occurs in that region that has the greatest access to the cathodic depolarizer. In the case of a pit or a crack, as depicted schematically in **Figure 7**, the local anode exists within the cavity, whereas the local cathode exists on the bold external surfaces.

The voltage difference generated between the cavity and the external surface causes a positive current to flow through the solution from the local anode to the local cathode. Negative electron current flows through the metal in the reverse direction and the two currents mutually annihilate at the external surface via a charge transfer reaction, in this case, the reduction of oxygen. The current is known as the 'coupling current' and is easily measured.[86] The coupling current has been shown to contain a wealth of information concerning the processes that occur within the cavity. The coupling current is generated within the cavity by anodic oxidation of the metal that may be unassisted by mechanical processes (pitting corrosion and crevice corrosion) or assisted by the presence of a constant stress (stress corrosion cracking, SCC) or a cyclic stress (corrosion fatigue, CF). In any event, the flow of positive current out of the cavity occurs because of the existence of a potential gradient in the solution from the cavity tip to the cavity mouth, such that $\phi_t > \phi_m > \phi_\infty$, where ϕ_t, ϕ_m and ϕ_∞ are the corresponding electrostatic potentials in the solution at the cavity tip, cavity mouth, and at a point on the external surface that is at an effectively infinite distance from the mouth (i.e., the 'throwing power,' which turns out to be about 10–20 crack mouth opening displacements, but which also depends upon the conductivity of the environment). The sign of this potential gradient is such that anions are transported into the crack, where they neutralize the positive charge that is injected into the cavity solution in the form of metal cations. The metal cations hydrolyze to produce protons that acidify the environment. The concentration factors for both Cl^- and H^+ from electromigration and hydrolysis, respectively, can exceed 10^4. Thus, for an environment containing 10 ppb of chloride ion and having a pH of 7, the chloride concentration and pH in the cavity may exceed 100 ppm and below 3, respectively. Accordingly, noting the propensity of chloride ions to induce passivity breakdown, the DAH accounts for the development of aggressive conditions within the cavity, even though the external conditions might be quite benign.

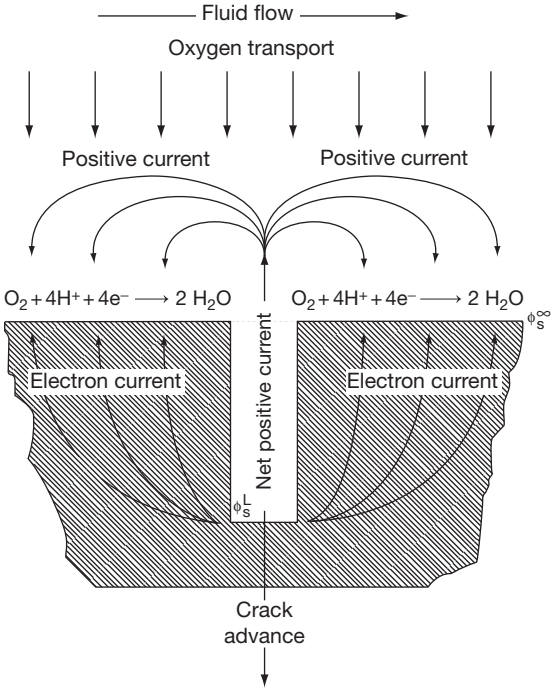


Figure 7 Coupling of crack internal and external environments. Note that in the steady state, the crack can grow only as fast as the positive current flowing from the crack can be consumed on the external surfaces by oxygen reduction. Reproduced from Macdonald, D. D.; Lu, P.-C.; Urquidi-Macdonald, M.; Yeh, T.-K. *Corrosion* **1996**, *52*, 768.

It is evident that, under free corrosion conditions, the conservation of charge requires that the charge passed by the cathodic reaction be matched by the charge passed by the partial anodic reaction (see eqn [25]). Recognizing that, at sufficiently large distances from the crack mouth on the external surface, the net current, $i_N = i_a + i_c$ (i_a and i_c are the local anodic and cathodic partial current densities) must be zero, corresponding to 'free,' general corrosion, the partial current density distributions are such that eqn [25] must be satisfied. Note that eqn [25] does not stipulate that cathodic partial reactions cannot occur within the cavity or that anodic partial reactions cannot occur on the external surfaces; only that, in the 'cathodic' and

'anodic' regions, the net current densities are negative and positive, respectively, and the total currents integrated over the respective areas are equal.

It is also important to glean an understanding of the critical nature of the properties that exist at the crack mouth. Thus, note that solving the entire problem involves solving two closely coupled sub problems, one for the external environment and the other for the crack enclave. These two problems are coupled because they share common boundary conditions in potential, ϕ_m, and in current, i_m. The objective in all coupled environment models (CEMs) developed to date has been to find the appropriate values for ϕ_m and i_m, such that eqn [25] is satisfied. It is the constraint imposed by eqn [25] that imparts determinism to the model.

Traditionally, models for localized corrosion cavity growth have assumed that the electrical potential in the solution at the crack mouth is the negative of the free corrosion potential or that the external environment presents no impedance to current flow (i.e., the current may flow to infinity). Neither postulate is correct. In the first case, no potential gradient would exist in the external environment and hence, the coupling current must be zero, contrary to experimental observation. Indeed, these models actually predict that SCC cannot occur. In the second case, if there exists no impedance for current flow to infinity, then there should be no dependence of the coupling current and crack growth rate on the catalytic properties of neighboring external surfaces, again contrary to experiment. The requirement that the underlying theory must account for all of the experimental observations is therefore not met by both classes of models. However, the authors note that the assumptions may be valuable approximations in solving the mass transport equations for the cavity, but, in doing so, it must be recognized that in ignoring the external environment, the physical description of the system is incomplete. Accordingly, and emphasizing again that a viable model must account for all experimental observations, such an approach would fail to recognize the important impact that catalysis or inhibition of oxygen reduction on the external surfaces has on crack growth rate in sensitized Type 304 SS in high temperature water, as noted above.[87,88]

To date, we[86,89–95] and others[96,97] have developed 'coupled environment' models for stress corrosion cracking (coupled environment fracture model, CEFM[89,90,92,93]), corrosion fatigue (coupled environment corrosion fatigue model, CECFM[94]), pitting (coupled environment pitting model, CEPM[91]), and crevice corrosion (coupled environment crevice model, CECM[95]). Details of the algorithms are not given here as they have been adequately described in the literature. In the interests of space, only a brief outline of some of the results from the CEFM is given. We note that alternate models for crack growth in sensitized stainless steels have been developed by others, most notably by Ford and coworkers,[98,99] Shoji,[100,101] and Vankeerberghen and Gavrilov.[102] Of these models, only the model of Vankeerberghen and Gavrilov[102] constrains the solution by the conservation of charge and hence, can be regarded as being deterministic. The models of Shoji[100,101,103] are essentially mechanical in form, with any electrochemistry being introduced 'inadvertently,' and, like Ford *et al.*'s models,[98,99] while they are 'mechanistic,' they fail to be constrained by the relevant natural law (conservation of charge), recognizing that IGSCC is primarily an electrochemical phenomenon.

The first of the CEMs to be developed was the coupled environment models (CEMs) in 1991.[89] Since then, the model has been extensively developed by the authors and their colleagues as a deterministic model for predicting stress corrosion crack growth rate in a variety of systems, including the coolant circuits of water-cooled nuclear reactors.[104] The models are based upon the following general experimental observations that apply strictly to the growth of intergranular cracks in sensitized Type 304 SS in high temperature water, but that are believed to be general correlations for SCC and other forms of localized corrosion in other systems:

- Localized corrosion generally follows the differential aeration hypothesis, first proposed by Evans in the 1920s.[85]
- A positive coupling current is observed to flow through the solution from the crack mouth to the external surfaces, while an equal but opposite electron current flows through the metal in the reverse direction.[86]
- The crack growth rate increases roughly exponentially with the potential of the metal if it is sufficiently high. At lower potentials, the CGR is potential-independent, corresponding to the mechanical creep (**Figure 8**).[90–101,103]
- The CGR depends upon the electrochemical crack length, which is defined as the shortest distance between the crack front and the exposed external surface. This length is generally different from the mechanical loading crack length.[90,102]

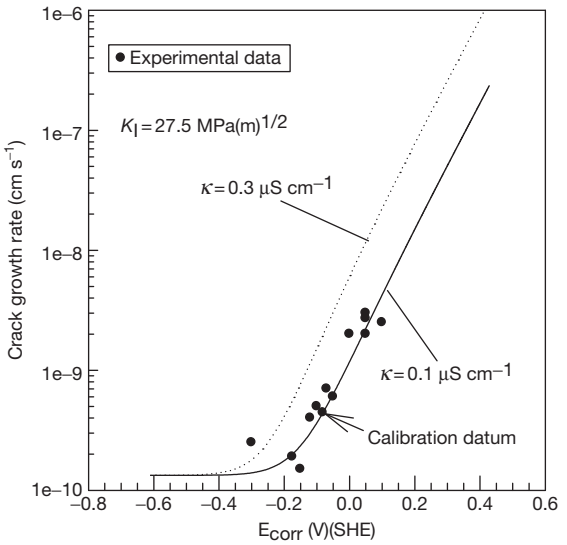


Figure 8 Measured and analytically calculated crack growth rate of sensitized Type 304 stainless steel as a function of corrosion potential. Calculations were done using the Congleton's correlation for the crack-tip strain rate. Experimental data were taken from Ford *et al.*[99]

- The environmentally-mediated CGR is proportional to the magnitude of the coupling current.[87,105]
- Coating the external surfaces with an insulator, and hence inhibiting the reduction of oxygen, causes the coupling current to decrease sharply and stops the crack from growing.[87]
- Catalyzing the reduction of oxygen on the external surface results in an increase in the coupling current and hence, an increase in the CGR.[86,88]
- The crack growth rate passes through a maximum with increasing temperature at about 175 °C.[106]

It is well known that, in the fracture of sensitized stainless steels and nickel alloys in high temperature (250–300 °C), oxidizing aqueous media, the CGR becomes independent of potential at sufficiently negative ECP values (see **Figure 8**) and that the fracture morphology changes from intergranular brittle fracture to ductile failure. The ductile fracture surfaces frequently yield evidence of microvoid coalescence, with ductile tearing of the matrix between the voids. These voids appear to nucleate at intergranular precipitates, such as carbides (e.g., $Cr_{23}C_7$) on the grain boundaries, but nucleation at intragranular precipitates is also observed. Thus, in the CEFM, which was developed originally to describe fracture in sensitized stainless steels in BWR primary heat transport circuits,[83,84,90] it seemed appropriate to describe crack growth at the 'creep' limit in terms of a cavitation model.

A modified version of the cavitation model developed by Wilkinson and Vitek[107] was used to estimate CGR at sufficiently negative potentials, where environmental effects are not evident. A detailed account of the creep model will not be given here and the reader is referred to the original paper by Wilkinson and Vitek[107] and papers by Macdonald *et al.*[90,93,108] describing the application of the model in the CEFM. It suffices to note that the Wilkinson–Vitek model accurately describes creep crack growth in stainless steels over the temperature range of interest (25–300 °C).

The numerical and analytical solutions employed in the CEFM yield very reasonable results for the environmentally assisted and creep fracture of sensitized Type 304 SS,[89] and have yet to yield a prediction of crack growth rate that is at odds (i.e., lies outside the experimental error range) with experiment. The calibration factor appears to take care of a number of simplifying assumptions in the numerical CEFM (e.g., inert crack walls and linear potential drop down the crack) and compensates for some less quantified effects (e.g., parameters associated with the crack tip process). Nevertheless, it is only necessary to calibrate the model with a single crack growth rate under specified conditions (see **Figure 8**) and to choose an appropriate activation energy for the crack tip strain rate (this is actually equivalent to calibrating the model with CGR data at two temperatures), in order to accurately describe crack growth rate in sensitized Type 304 SS over temperatures ranging from 25 to 288 °C.

One of the persistent issues that is raised in debates on the basis of the CEFM concerns the role of the external environment in determining the crack growth rate. If the differential aeration hypothesis (DAH) is accepted as the basis for localized corrosion, including SCC, then the importance of the external environment is unequivocal. That the external environment is intimately involved has been demonstrated experimentally by detecting and measuring the coupling current that flows from the crack to the external surfaces where it is consumed by oxygen reduction.[88,89] Furthermore, unless the cathodic processes are confined to the crack, in which case the measured coupling current would be zero, and hence at odds with experiment,[86] the conservation of charge *requires* consideration of the external environment in any deterministic

description of crack growth. Finally, catalysis[88] and inhibition[87] of the oxygen reduction reaction on the external surface is found to increase and decrease the crack growth rate, respectively. In the case of inhibition, which was affected by depositing Zirconia on the external surfaces, the extent of inhibition was in agreement with that calculated by the CEFM by reducing the exchange current density for the oxygen electrode reaction by the amount indicated by electrochemical impedance spectroscopy measurements of the specific interfacial impedance using a fast redox couple $[Fe(CN)_6^{3-/4-}]$.[87] These studies unequivocally demonstrate the importance, and possibly the dominance, of the processes occurring, on the external surfaces in determining the crack growth rate in sensitized Type 304 SS in high temperature aqueous systems.

The role of the external environment is further demonstrated by the calculated polarization data plotted in **Figure 9**.[93,108] Thus, it is seen that, at low temperatures (e.g., 50 °C), significant polarization is predicted to occur in the external environment (170 mV). As the temperature increases, the polarization in the external environment is predicted to decrease, but nevertheless remains significant at temperatures up to 250 °C. Even at higher temperatures, where the external polarization has been decreased to small values, because of increased conductivity of the external environment and increased rate of oxygen reduction on the external surfaces, the external environment must still be considered because of the need for charge conservation. The temptation has been to assume that, in this case, the potential at the crack mouth can be equated to $\phi_m = -E_{corr}$, and hence to eliminate the need to consider the external environment altogether. However, this would be to assume that no potential gradient exists in the external environment, in which case no coupling current should be detected, again at odds with experiment,[89] leading to the prediction that SCC cannot occur. Finally, it is interesting to note that the decrease in the polarization in the external environment is predicted to be at the expense of an increase in the potential drop down the crack.

As noted elsewhere,[90,104] the CEFM predicts that the crack growth rate for given values of stress intensity, ECP, conductivity, etc. depends upon the crack length. This prediction is consistent with the available experimental data.[19,88,105] In discussing this topic, it is necessary to differentiate between the mechanical crack length (MCL), which is traditionally referred to as the 'crack length' in fracture studies, and the 'electrochemical crack length' (ECL), as noted previously.[19,88,105] The electrochemical

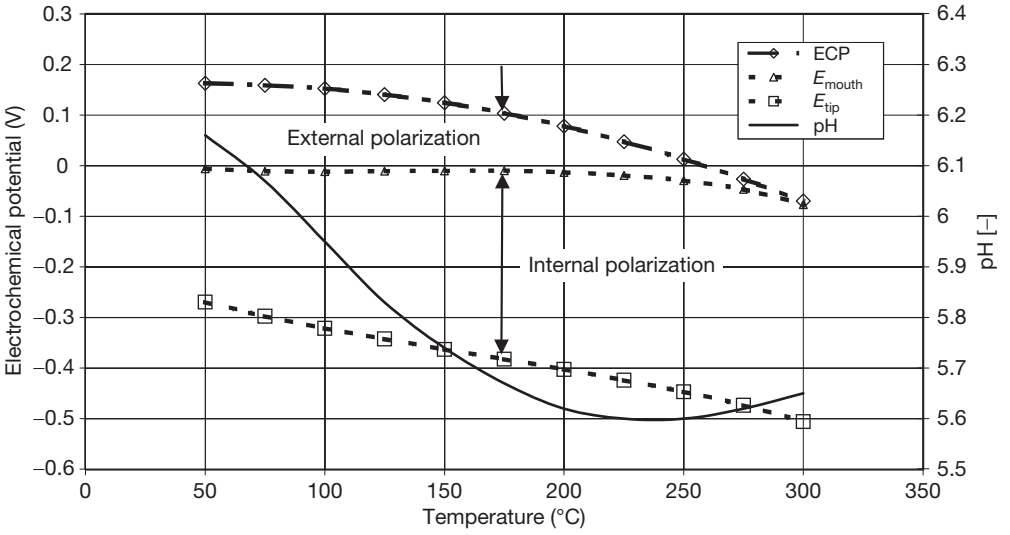


Figure 9 The calculated effect of temperature on the pH of the external environment, the electrochemical potential at the crack tip, E_{tip}, the potential at the crack mouth, E_{mouth}, and on the potential in the external environment (ECP), during crack growth in Type 304 stainless steel in dilute sulfuric acid solution having an ambient temperature (25 °C) conductivity of 0.27 μS cm^{-1} and a dissolved oxygen concentration of 200 ppb. The data were calculated using the CEFM after calibration at 288 °C and assuming a crack tip strain rate thermal activation energy of 100 kJ mol^{-1} (Congleton crack tip strain rate model).[93]

crack length is defined as the shortest path through the solution from the crack front, where the coupling current is generated, to the external surface, where the current is consumed. For a CT specimen, the MCL corresponds physically to the distance between the load line (for example) and the crack tip, and this distance increases as the crack grows through the specimen. On the other hand, because of the through thickness nature of a crack in a CT specimen, the ECL is constant and is essentially independent of the MCL. Furthermore, a CT geometry yields a distribution in ECL, depending upon where the current originates at the crack front. This dependence is such that the crack growth rate is highest at the crack edges (smallest ECL), but is lowest for the coupling current that originates from the crack front in the center of the specimen (largest ECL) by virtue of the dependence of crack growth rate on the electrochemical crack length.[90,102,104] This phenomenon is responsible for the generation of convex crack fronts, when the crack grows by SCC, in contrast to the concave crack front obtained for creep crack growth.[105]

Finally, we note that the CEFM is used extensively in various corrosion damage codes developed by Macdonald and coworkers (DAMAGE PREDICTOR, ALERT, REMAIN, FOCUS)[19,23] to predict stress corrosion cracking damage in BWR primary coolant circuits, as described below. To date, fourteen BWRs have been modeled and, where comparison is possible, the predicted damage is found to be in excellent agreement with that observed in the field.

2.39.6.3 Simplified Approach for Calculating Propagation Rates

It is well known from experiment that the rates of individual pit (crack) propagation as a function of time at constant environmental conditions can be approximated by a relatively simple function of time. Thus, for the case of pitting corrosion, the power function,

$$a = kt^m \qquad [70]$$

has been suggested, where k and m are empirical constants.[11,58,109] Here, a is the characteristic pit size (e.g., pit depth or radius of the pit mouth). Published values of m are very often approximately equal to $1/3$, $1/2$, $2/3$ or 1, but they can also vary over wider ranges.[11,109–111]

However, this dependence of a on t cannot be used directly in mathematical calculations for small times, because of the nonphysical limit

$$V_s = \frac{da}{dt} = kmt^{m-1} \to \infty, \text{at } t \to 0 \text{ for } m<1 \qquad [71]$$

This is why, instead of eqn [70], the following interpolation equation for pit propagation rate, V_s, has been suggested[91]

$$V_s = \frac{da}{dt} = V_0(1 + t/t_0)^n \qquad [72]$$

where $n = m - 1$ and t_0 are constants, and V_0 is the initial, finite rate of pit propagation. Equation [72] yields $V = V_0$ at $t/t_0 << 0$ and $V = V_0(t/t_0)^n$ at $t >> t_0$.

Integration of eqn [72] with the boundary condition $a = 0$ at $t = 0$ yields

$$a = x_0[(1 + t/t_0)^m - 1] \qquad [73]$$

where $x_0 = t_0V_0/m$. For small times, as follows from eqn [73], a can be presented as the linear function of t, $a = V_0t$ and for large times, a takes the form of eqn [70]. In many cases, the period of time over which the approximation

$$V(t)V_0 = \text{Constant} \qquad [74]$$

is valid can be comparable with the observation time (or even with the service life of the system). The reason is that corrosion is, generally speaking, a slow process and under real, practical conditions, values of the critical pit depth of the system, x_{cr}, and typical service life, t_s, impose significant restrictions on the values of the initial and average corrosion current densities and, thus, on the potential and concentration drops that might be observed in a corrosion cavity.[112]

For constant external conditions, the dependence of cavity propagation rate as a function of the cavity depth, x, can be written in the form

$$V_s = \frac{da}{dt} = V_0\varsigma(x) \qquad [74]$$

where function $\varsigma(x)$ satisfies the evident boundary condition $\varsigma(x) \to 1$ at $x \to 0$. For the particular case when eqn [72] holds, $\zeta(x)$ can be expressed in the form:

$$\varsigma(x) = \frac{1}{(1 + x/x_0)^{(1-m)/m}} \qquad [75]$$

where $x_0 = v_0\, t_0/m$.

The general approach based on the numerical solution of the system of mass transfer equations does not present the depth of pit in the form of eqn [70] or [73]. As noted in Turnbull,[58] the mechanistically based models are complex to set up and, at the

present stage of development, are not user-friendly. For us, it is most important that the solution of the full system of mass transfer equations for the ensemble of pits is practically unrealistic for usual systems, from the point of view of the required computation time. One way to deal with this problem is to perform numerous calculations of pit propagation in advance and approximate the obtained results by eqns [70] or [73]. The other way is to try, after some simplification that does not significantly reduce the accuracy of calculations, to develop a model that can yield a cavity propagation law of the form of eqns [70] or [73], and that may even yield analytical expressions for parameters k and m. The principal concept in this simplification is that, if the rate of corrosion reaction depends explicitly only on the local potential and some surface concentrations of particular species (e.g., Cl^- and H^+), the pit growth rate depends only on the concentrations of those species that determine the value of electrostatic potential near the corroding surface.

Below, we will consider the example of such an approach for the simplest (but perhaps the most important, from the practical point of view) case of the corrosion of steels in neutral solutions. Thus, it has been suggested that, in the mathematical simulation of the corrosion of Fe in NaCl solutions, at least six species in the solution must be taken into the account.[91]

$$S_1 \equiv Fe^{2+};\ S_2 \equiv Fe(OH)^+;\ S_3 \equiv Na^+;\ S_4 \equiv Cl^-;$$
$$S_5 \equiv H^+;\ \text{and}\ S_6 \equiv OH^-$$

These species include iron ions from the dissolution process, sodium and chloride ions (for example) that are commonly included to control the bulk conductivity, hydrogen and hydroxyl ions from the dissociation of water, and a metal hydrolysis product (for example). In the simplest case, the following homogeneous hydrolysis reactions are assumed to occur

$$Fe^{2+} + H_2O = Fe(OH)^+ + H^+ \quad [76]$$
$$Fe(OH)^+ + H_2O = Fe(OH)_2(s) + H^+ \quad [77]$$
$$H_2O = H^+ + OH^- \quad [78]$$

with $Fe(OH)_2(s)$ representing the precipitated hydrolysis product. Homogeneous reactions include hydrolysis processes and also precipitation processes. If the degree of supersaturation is sufficiently high, direct precipitation will occur in the solution itself; otherwise, precipitation will tend to occur on the walls preferentially. However, if an appropriate balance is obtained between the rate of production of the hydrolyzing species and the rate of precipitation, general saturation is expected to occur.[60]

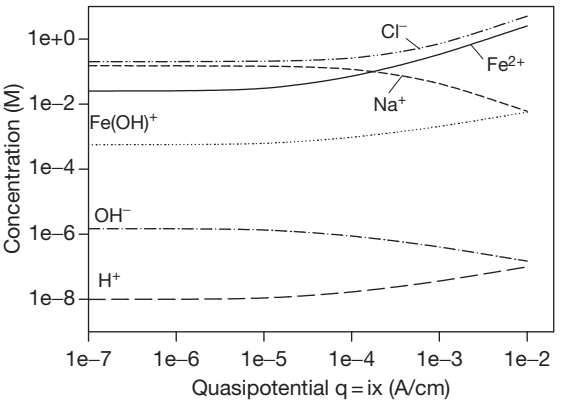


Figure 10 The concentration of species down a corroding cavity (crevice) in 0.1 M NaCl at 25 °C as functions of the quasipotential.

We will assume that a single electrochemical reaction

$$Fe \rightarrow Fe^{2+} + 2e \quad [79]$$

occurs on the walls of the cavity.

Numerical solution of the corresponding boundary problem has been performed[91] by using the values of kinetic parameters from Sharland *et al.*[78] **Figure 10** shows the concentrations profiles as a function of the quasipotential, q.[68] Please note that, for the case of a deep, one-dimensional cylindrical pit with passive walls or for a crevice with parallel passive sides, $q = ix$, where x is the distance down the cavity, and i is the current density in the crevice that is determined by the dissolution rate.[91,112] We can see that there are three species that dominate in the pit, Fe^{2+}, Cl^-, and Na^+, and only these species determine the electrostatic field in the cavity.

Let us now assume in addition, that, the electrodissolution reaction occurring in the cavity due to oxidation of the metal, can be described in terms of the Tafel equation as

$$i = i_0 \exp\left\{\frac{\alpha F(V_m - U_0 - \phi_s)}{RT}\right\} = i^* \exp\left\{-\frac{\alpha F \phi_s}{RT}\right\} \quad [80]$$

Here, i_0 is the exchange current density, α is the anodic transfer coefficient, V_m is the local electrode potential, U_0 is the open-circuit potential (measured on the external surface remote from the pit), ϕ_s is the electrostatic potential in the solution near the

electrode surface, all referred to a reference electrode at infinity, and $i^* = i_0 \exp\{\alpha F(V_m - U_0)/RT\}$ is the current density calculated in the absence of a potential drop in the cell ($\phi_s = 0$ corresponding to the maximum possible current density on the electrode surface at the given potential of the metal).

Hence, if we only want to calculate the potential distribution and rate of corrosion in this system, we can approximate the complete system involving six species (Fe^{2+}, $Fe(OH)^+$, Na^+, Cl^-, H^+ and OH^-) by a reduced system involving only three species (Fe^{2+}, Na^+ and Cl^-).

In this case, the solution of the mass transfer equations for this simplified system can be easily performed even in analytical form. In particular, if the polarization curve for reaction [79] can be described by Tafel's law [80] with the symmetry factor, α, it was shown that the depth of one dimensional pit with passive walls can be described by eqn [73] with

$$m = \frac{\alpha + 1}{2\alpha + 1} \qquad [81]$$

and

$$x_0 = 3FD_1 C^{b}_{Cl^-}/i^*,\ V_0 = K_V i^*,\ t_0 = mx_0/V_0 \qquad [82]$$

where D_1 is the diffusion coefficient of Fe^{2+}, and $C^{b}_{Cl^-}$ is the bulk concentration of chloride ions. Moreover, if $\alpha = 1$ (typical value for corrosion of steels), eqn [81] yields $m = 2/3$.

If, for example, $C^{b}_{Cl^-} = 0.6\,M$ (which corresponds approximately in the case of seawater), $i^* = 10^{-4}$ A cm^{-2} (which corresponds to an initial propagation rate, V_0, of ~1 mm $year^{-1}$) and $D_1 = 0.72\ cm^2/s$, eqn [82] yields $x_0 \approx 12$ cm. Accordingly, if the critical depth of the pit penetration x_{cr} is of the order of 1 cm, as follows from eqns [74] and [75], it is possible to conclude that the pit will propagate at a constant rate because, when x_0 is a large number $\varsigma(x)$ becomes equal to one. The physical reason for this uniform propagation is that the potential (and concentration) drops are very small in this system for the adopted values of dimensions, kinetic parameters, and environmental conditions. We can also state that the mere fact that hemispherical (or approximately hemispherical) pits exist can be explained by the insignificance of the potential drop in the corrosion cavity. Otherwise, the corrosion current density on the edge of the pit becomes higher than that on the bottom, and the pit will quickly assume a shallow ('saucer') form. This initial rate can be determined by using the experimentally measured polarization curves in the region of active dissolution, and can be calculated for certain metals and alloys for different environmental conditions by using available software.[71]

However, it is also possible that pits can propagate with a constant rate, in spite of the fact that significant potential drops may exist down the cavity. In many cases, the rate of corrosion depends on the concentration of aggressive anions (e.g., chloride ions) near the metal surface. This dependence is usually approximated by the following relation[11]:

$$i = k_a C^{\lambda}_{As} \exp\left\{\frac{\alpha F(V_m - U_0 - \phi_s)}{RT}\right\} \qquad [83]$$

where C_{As} is the surface concentration of the aggressive anion, k_a is the rate constant for the reaction proceeding in the anodic direction, and λ is the effective kinetic order of the metal dissolution reaction with respect to the anion concentration. The value of λ is usually restricted to $0 \le \lambda \le 1$.[11] It was shown that, in this case, instead of eqn [81], we can use the alternate equation[91,112]

$$m = \frac{\alpha_{eff} + 1}{2\alpha_{eff} + 1} \qquad [84]$$

where

$$\alpha_{eff} = \alpha - \lambda \qquad [85]$$

It is clear that if $\alpha \approx \lambda$, we have $m \approx 1$, and the pit propagates with an approximately constant rate.

Generally speaking, under real, practical conditions, values of the critical pit depth of the system, x_{cr}, and typical service life, t_s, impose significant restrictions on the corrosion current densities (averaged and initial) and thus on the potential and concentration drops that might be observed in the corrosion cavity. Thus, if x_{cr} does not exceed the order of 1–10 mm, and if the order of t_s is not less than 1 year, initial corrosion current densities in real, open pits cannot exceed values of 10^{-4} to 10^{-3}A cm^{-2}, with the understanding that the polarization curve (corrosion current density versus potential) and the surface concentrations of the species do not change as the pit propagates.[112] Simple analytical expressions for predicting the potential and concentration drops in open corrosion cavities and for predicting cavity propagation rate in systems containing uni- and bivalent anions in stagnant electrolytes under well mixed, external conditions have been obtained.[112]

After the transition of a pit into a crack or into a corrosion fatigue crevice, the total rate of corrosion cavity propagation, V_s, cannot be described simply by Faraday's law, but also has to contain mechanical

and, in the general case, some additional environmental component (e.g., hydrogen embrittlement component). It would be natural to assume that V_s is a known function of the environmental (concentration of species near metal surface, temperature) and mechanical parameters (stress), and so on. The explicit forms of these dependences are discussed in the corresponding chapters of this book. Here, only the most general relations will be discussed.

Figure 8 clearly shows that at low corrosion potential, the crack propagation rate does not depend on E_{corr}, that is, V_s reduces to the mechanical (creep) component. It is also clear that at high values of E_{corr}, crack propagation rate does not depend on creep rate and is described by an exponential dependence of the Tafel type. Accordingly, in the first approximation, we can present V_s as the sum of two contributions

$$V = V_{env} + V_{mech} \qquad [86]$$

where V_{env} is the environmental/electrochemical component, which is often determined by Faraday's law, and V_{mech} is the mechanical component (crack advance associated with mechanical fatigue/creep). Of course, in some cases, the synergistic interaction of mechanical and environmental attack can play a role and a simple summation, as expressed by eqn [86], may not be appropriate.[113]

A review of many models for describing environment-induced cracking based directly or indirectly on anodic reaction processes or on hydrogen embrittlement can be found.[114,115] Thus, in accordance with the slip-dissolution model, the average crack growth velocity will be given by

$$V_{env} = K_v i^0_{corr} \Gamma \qquad [87]$$

where i^0_{corr} is the current density on the bare surface, and Γ is the ratio of the bare surface of the crack tip to the total geometrical surface. It can be shown that

$$\Gamma = \frac{t_0^n}{(1-n)\varepsilon_f^n}(\dot{\varepsilon}_{ct})^n \qquad [88]$$

where ε_v is the fracture strain of the passive film at the crack apex, $\dot{\varepsilon}_{ct}$ is the crack tip strain rate, and t_0 and n are material constants. It is evident that $\Gamma \leq 1$. In turn, the expressions for $\dot{\varepsilon}_{ct}$ can be found as described in Engelhardt *et al.*,[92] Turnbull,[114] and Peng *et al.*[116]

In the case of corrosion fatigue, the crack growth rate is usually expressed in terms of the increment of crack growth during each fatigue cycle, $da/dN = V_s/f$ (where f is the frequency of the applied stress). Very often, this rate follows the well known power-law relationship (Paris' law)

$$\frac{da}{dN} = C(\Delta K)^n \qquad [89]$$

for both fatigue (no corrosion) and corrosion fatigue. Here, ΔK is the stress intensity factor range and C and n are empirical parameters that depend, in the general case, on environmental conditions. While the dependencies of the Paris equation parameters on environmental conditions is well recognized through numerous experimental studies, few of the models for corrosion fatigue developed to date incorporate environmental effects in an explicit, mechanistic manner. This is another example of where the underlying (mechanical) theory fails to account for the experimental data, and the resultant models fail to meet the requirements of determinism.

Until now, it has been tacitly assumed that the rate of pit (crack) propagation is unequivocally determined by its depth and by the external conditions, that is, there is no distribution in cavity propagation rate for pit cavities of equal depth. However, as noted above, a distribution in pit (crack) propagation rate might be observed in practical systems because of underlying distributions in system parameters that affect the growth rate. Usually, such problems are considered in the following way. In the case of pitting corrosion, it is assumed that the pit depth is described by eqn [70] with distributed parameters k and m.[117] However, in Turnbull *et al.*,[118] it is assumed for simplicity, that only parameter k is normally distributed; parameter m is assumed to be fixed, and its value is obtained by fitting the model to the experimental data. In contrast, in Harlow and Wei,[3] it is assumed that the pit maintained a hemispherical geometry and grew at a constant volumetric rate. From this fact, it immediately follows that $m = \text{const} = 1/3$. On the other hand, it was assumed that parameter k is described by a Weibull distribution.

Analogously, in Turnbull *et al.*,[118] for the case of SCC, the crack propagation rate was assumed to be given by

$$\frac{da}{dt} = C\sigma^p a^q \qquad [90]$$

where σ is the applied stress. It was assumed that parameters p and q are fixed and were determined by fitting eqn [90] to the experimental data, but parameter C was assumed to be normally distributed. Additionally, for example, in Harlow and Wei,[3] for the case of corrosion fatigue, it was assumed that crack propagation is described by eqn [89] with

fixed crack growth exponent, n, and distributed parameter C.

Of course, few of the models that are discussed in this section are 'deterministic' in nature, although they are certainly 'mechanistic' in form. Determinism is readily introduced by first identifying the relevant natural laws and then ensuring that the predictions are constrained by these laws. In many cases, this would be most effectively done by introducing the models, which for the most part, describe the 'anodic' part of the corrosion cell, into a 'coupled environment' framework to impose the conservation of charge. This again illustrates the important difference between 'mechanistic' and 'deterministic' models, with only the latter constraining the output ('predictions') to those that are 'physically real' (i.e., that are consistent with scientific knowledge).

2.39.7 Rate of Pit Repassivation and Transition of Pits into Cracks

As noted above, we assume that the repassivation process obeys a first-order decay law

$$R_a(x,t) = -\gamma f_a(x,t) \qquad [91]$$

where γ is the delayed repassivation ('death') constant (i.e., the rate constant for repassivation of stable pits), and f_a is the differential damage function for active pits. The repassivation constant, γ, is, in general, expected to be a function of the external conditions, including the corrosion potential, temperature, and electrolyte composition. Generally speaking, γ is also expected to be a function of the depth of the pit, x, because the local potential in the solution at the cavity surface depends on the IR potential drop in the cavity, that is, γ might be a function of both the spatial coordinates and time. Of course, if the potential and concentration drops inside the corrosion cavity are insignificant during pit propagation, it is possible to neglect changes in γ (see above). However, the value of this constant still depends on the external conditions, such as potential, pH, and concentration of aggressive species in the bulk electrolyte. Finally, active pits may no longer be viable if the potential, E, at the pit internal surface is less than the repassivation value, E_{rp}. Accordingly, if the value of E_{rp} is reached at some pit depth, x_{rp}, active pits passivate and cannot penetrate further into the metal. The value of repassivation potential E_{rp}, is a function of the metal potential and surface concentrations at the pit tip. It can be calculated, for example, by using methodology described in Anderko *et al.*[119]

Regarding the transition of a pit into a crack, we assume that a pit immediately transforms into a crack if its depth exceeds some critical value $x_{tr.}$ As of now, the most widely accepted set of criteria for the transition of a pit into crack are the Kondo criteria.[120] According to these criteria, two conditions must be satisfied for crack nucleation to take place from a pit, namely,

$$K_I > K_{ISCC}(\text{forSCC}) \text{ or } \Delta K > \Delta K_{th}(\text{forCF}) \qquad [92]$$

and

$$V_{crack} > V_{pit} \qquad [93]$$

Here, K_I and K_{ISCC} are the stress intensity factor and critical stress intensity factor for propagation of a stress corrosion crack, respectively: ΔK is the stress intensity factor range, and ΔK_{th} is the threshold stress intensity factor range for fatigue crack propagation, respectively.

The first requirement defines the mechanical (fracture mechanics) condition that must be met for the prevailing stress and geometry, while the second simply says that the nucleating crack must be able to 'outrun' the pit.

It is important to note the following circumstance. At high values of corrosion potential, E_{corr}, the environmental component of crack propagation rate, V_{env}, could be much higher than the mechanical component, V_{mech} (see. **Figure 8**). Accordingly, the crack propagation rate, V_{crack} will practically coincide with its environmental part, V_{env}.

In accordance with the slip dissolution model, the crevice tip is partially blocked by the passive film. Accordingly, for a given set of tip conditions (metal potential, pH, etc.), V_{env}, must be smaller than that corresponding to the pit propagation rate, V_{pit}, with the tip surface being bare (see eqn [87]). This means that, under these conditions, the transition from a pit into a crack occurs if (1) the depth of the pit exceeds some critical length, x_{mech} (where $K_I \geq K_{ISCC}$ or $\Delta K_I \geq \Delta K_{I,th}$); and (2) the pit is passivated (when V_{pit} is very small). Thus, this latter criterion suggests that cracks will nucleate only from 'dead' pits. This example shows how important repassivation phenomena may be for predicting corrosion damage.

2.39.8 Statistical Properties of the Damage Function

There exists a close correspondence between damage function analysis (DFA), which has been described at some length above, and extreme value statistics (EVS). The latter technique has been used extensively to extrapolate damage (maximum pit or crack depth) from small samples in the laboratory to larger area samples in the field. Furthermore, DFA provides a means of calculating the central and scale parameters and their time-dependencies in EVS from first principles, and hence represents a unification of the two prediction philosophies.

From a statistical point of view, all distributed properties of the system are completely determined by the cumulative distribution function (CDF), $\Phi(x)$. By definition, $\Phi(x)$ is the probability that the depth of a randomly selected pit (crack) is $\leq x$. We postulate that the pit distribution on the metal surface is uniform. Accordingly, the total number of nucleated pits in the entire system is $SN(t)$, where S is the area of the system and, from the definition of the integral damage function, F, the number of pits that have the depth $\geq x$, is $S[N(t) - F(x,t)]$. Accordingly, from the definition of probability, we have

$$\Phi(x,t) = \frac{S[N(t) - F(x,t)]}{SN(t)} = 1 - \frac{F(x,t)}{N(t)} \qquad [94]$$

We see that the CDF for a given observation time, $\Phi(x,t)$, can be predicted if we know (can calculate) the integral damage function of the system (note that the number of nucleated, stable pits, $N(t)$ simply equals $F(0,t)$). This relationship can be regarded as being the bridge between the statistical and deterministic approaches for estimating the accumulation of localized corrosion damage on a surface.

As noted above, from the practical point of view, the most important value for characterizing corrosion damage is the failure probability, P_f, of the system. By definition, P_f, is the probability that at least one corrosion event in any form (pit, crevice, stress corrosion crack, or fatigue crack) reaches a depth, x, at a given observation time, t, where x, in this case, is the critical dimension. It can be shown[14] that

$$P_f(x,t) = 1 - \exp\{-SF(x,t)\} \qquad [95]$$

Equation [95] allows us to calculate the probability of failure if the integral damage function, F, is known. The latter function can be found as a solution of eqn [4]. However, here we will reject the equal velocities for all pits with a given depth assumption that was adopted previously (see eqn [9]). Thus, it is well known that the morphology of pits on any given surface can vary significantly, with some shapes favoring more rapid mass and charge transfer, and hence, a greater propagation rate.[121] In addition, some pits will initiate at metallurgical features that may favor more rapid propagation, for example, at MnS inclusion.[121] The distribution in pit propagation rate might also be explained, for example, by the spatial distribution in electrochemical activity of the anodic and cathodic sites on the corroding surface.[122]

Here, we will assume that the pits that propagate with initial rate V_0 are nucleated in accordance with the equation.

$$n(t) = \int_0^\infty \lambda(t, V_0)\mathrm{d}V_0 \qquad [96]$$

The function $\lambda(t,V_0)$ yields the number of pits (per square centimeters) that have initial propagation rates between V_0 and $V_0 + \mathrm{d}V_0$ and that nucleate in the period of time between t and $t + \mathrm{d}t$. Further propagation of the pits takes place in accordance with eqn [74]. It can be shown that, in this case, the expression for the flux of active pits can be expressed in general form as[123]

$$j_a(x,t) = \int_0^\infty \exp[-\gamma g(x)/V_0]\lambda(t - g(x)/V_0, V_0)\mathrm{d}V_0 \qquad [97]$$

where

$$g(x) = \int_0^x \frac{\mathrm{d}x'}{\varsigma(x')} \qquad [98]$$

(compare with eqn [9]). Accordingly, the expressions for the differential damage functions for active and passive pits and the integral damage function have the following forms[123] as solutions of eqn [4]

$$f_a = \int_0^\infty \frac{\exp[-\gamma g(x)/V_0]\lambda[t - g(x)/V_0, V_0]}{V_0\varsigma(x)}\mathrm{d}V_0 \qquad [99]$$

and

$$f_p = \int_0^\infty \frac{\gamma\exp[-\gamma g(x)/V_0)\Lambda[t - g(x)/V_0, V_0]}{V_0\varsigma(x)}\mathrm{d}V_0 \qquad [100]$$

Furthermore, for the case of pitting corrosion, the total integral damage function, F, is[123]

$$F = \int_0^{\infty} \exp[-\gamma g(x)/V_0)\Lambda[t - g(x)/V_0, V_0]\mathrm{d}V_0 \qquad [101]$$

where $\Lambda(t, V_0) = \int_0^t \lambda(t', V_0)\mathrm{d}t'$ yields the number of pits (per square centimeters) that have initial propagation rates between V_0 and $V_0 + \mathrm{d}V_0$ and nucleate in the period of time between 0 and t. F is the sum of the integral damage functions for active and passive pits for the given case.

Let us assume that the distribution in initial pit propagation rate does not depend on time, that is,

$$\lambda(t, V_0) = n(t)\psi(V_0) \qquad [102]$$

To move further, we must assume a distribution function, $\psi(V_0)$, for the pit growth rate, in order to account for those factors that result in a distribution in growth rate that is not captured by the (present) deterministic models. For our purposes, it is most convenient to approximate $\psi(V_0)$ by Laplace's distribution function

$$\psi(V_0) = \frac{\exp(-|V_0 - \bar{V}_0|/\beta)}{2\beta} \qquad [103]$$

where $\bar{V}_0$ is the mean initial pit propagation rate and $\sigma^2 = 2\beta^2$ is the dispersion.

If eqn [102] holds, we have, for the integral damage function and the cumulative damage function, the following expressions

$$F(x, t) = \int_0^{\infty} N[t - g(x)/V_0)\exp[-\gamma g(x)/V_0)\psi(V_0)\mathrm{d}V_0 \qquad [104]$$

and

$$\Phi(x, t) = 1 - \int_0^{\infty} N[t - g(x)/V_0)]\exp[-\gamma g(x)/V_0)\psi(V_0)\mathrm{d}V_0/N(t) \qquad [105]$$

respectively. In particular, for the case of instantaneous nucleation, we have

$$\Phi(x, t) = 1 - \int_{x/t}^{\infty} \exp[-\gamma g(x)/V_0]\psi(V_0)\mathrm{d}V_0 \qquad [106]$$

As has been shown experimentally, in many practical cases, the asymptotic behavior (for large values of x) of the CDF can be described by the exponential relationship[58,121]

$$\Phi(x, t) = 1 - \exp[-(x - u)/\alpha] \qquad [107]$$

where u is the central parameter (the most frequent value), and α is the scale parameter, which defines the width of the distribution. Accordingly, as follows from eqn [107], the extreme value distribution (EVD), $\psi(x,t)$ (the probability that the largest value of pit depth $\leq x$), is described by a double exponent (Gumbel Type I extreme value distribution) in the form[58]:

$$\Psi(x, t) = 1 - P_f = \exp[-\exp(-y)] \qquad [108]$$

where

$$y = (x - b)/\alpha \qquad [109]$$

and $b = u + \alpha \ln(SN)$.

Numerical calculation[14] shows that the dependencies of ln(1 − CDF) on pit depth, x, can be approximated by straight lines, at least for sufficiently large values of x, and therefore provide a theoretical basis for applying the Type I extreme value distributions to real corrosion systems. In some cases, parameters α and b can be expressed in analytical form. Thus, for the case of instantaneous nucleation and constant pit propagation rate, these parameters can be presented in the form[14]:

$$\alpha = \frac{\beta t}{1 + \gamma\beta t/\bar{V}_0} \quad \text{and} \quad b = \frac{[\bar{V}_0 + \beta Ln(0.5SN_0)]t}{1 + \gamma\beta t/\bar{V}_0} \qquad [110]$$

We see that, for small values of t, the parameters α and b can be described by $\alpha = \beta t$ and $b = [\bar{V}_0 + \beta Ln(0.5SN_0)]t$, that is, they are proportional to time. Such dependencies were actually observed, for example, for the case of the pitting corrosion of manganese steel in CO_2–acidified seawater.[122] On the other hand, for large observation times, the parameters α and b approach the following limits

$$\alpha = \bar{V}_0/\gamma \quad \text{and} \quad b = [\bar{V}_0 + \beta Ln(0.5SN_0)]t\bar{V}_0/\gamma \qquad [111]$$

that are independent of time. This fact can be regarded as being physically evident, because at sufficiently large times, all pits become passive (the damage function becomes 'frozen'), and further propagation of damage cannot occur, provided that $\gamma > 0$. This conclusion concerning the achievement of limiting pitting depth also follows from a formal statistical treatment of experimental data for underground carbon steel pipelines.[11] The 'freezing' of damage functions can also be seen in the propagation of corrosion damage on aluminum in tap water.[10]

Equation [110] with eqn [108] can be used for calculating the probability of the failure of a system. Unfortunately, very often it can be the case that not

all, or even none, of the kinetic parameters that determine the pit (crack) nucleation rate, pit propagation rate, and pit repassivation rate, are known. However, unknown parameters are determined by comparing the results of analytical or numerical calculations of the depths of the deepest pits (cracks) with the corresponding experimentally observed values for short-term experiments. After determining values for the unknown parameters, the depths of the deepest pit (crevice) for the total system (i.e., for the system with arbitrary surface area) for long-term exposure are predicted.

Usually therefore, in the case when none of the kinetic parameters are known, the following expression, based on the Gumbel type-I distribution is used for predicting perforation probability in the large systems with the total area, S, by using the results of measurements of the depth of the deepest pits on the series coupons with area, s[124]

$$P_f = 1 - \exp\{-\exp[-(d - u + \alpha\ln(S/s))/\alpha]\} \quad [112]$$

where d is the thickness of the wall.

For predicting probability of failure in long-term experiments by using the results of measurements in short-term experiments, it is necessary to assume some functional dependence of the location parameter, u, and scale parameters, α, on time. For the case when eqn [110] is valid, these parameters are described by hyperbolic relations of the form

$$u = \frac{a_1 t}{1 + a_2 t}, \quad \alpha = a_3 u \quad [113]$$

where, a_1, a_2, and a_3 are unknown parameters that must be fitted by using the short term experiments. The procedure of such fitting by using the maximum likelihood method is described in Laycock *et al.*[125] **Figure 11** illustrates the applicability of the hyperbolic dependencies for predicting the results of experiments measuring the depth of the deepest pits in the corrosion of aluminum Alloy 2S-O in Kingston tap water.[10] The mean depth of the deepest corrosion event, X_m, and standard deviation of this value, σ, can be calculated by using the following relations[125]:

$$X_m = h + \alpha[\mathrm{E} + \ln(S/s)] \quad \text{and} \quad \sigma = \pi\alpha/\sqrt{6} \quad [114]$$

We see that applicability of the hyperbolic functions can be considered as being very satisfactory. Note that only the data for $t = 1$ week and 1 month were used for calibration and hence, for prediction over a period of up to 1 year, providing accurate extrapolation over a range of 12 times that of the calibration range.

However, the typical time-relations that are usually used for estimating the location and scale parameters in extreme value distributions are of the power law type:

$$h = a_1 t^{a_2}, \quad \alpha = a_3 h \quad [115]$$

or logarithmic:

$$h = a_1 \log(t) + a_2, \quad \alpha = a_3 h \quad [116]$$

Figure 12 shows that application of the power law functions for fitting experimental data to the short

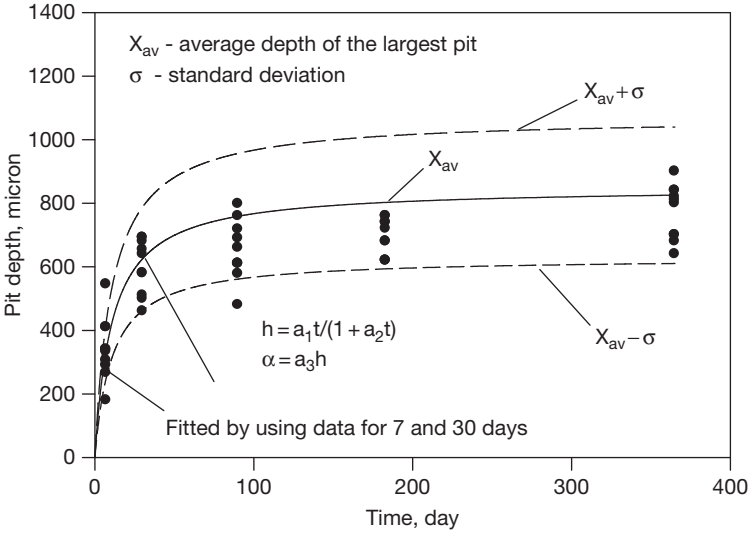


Figure 11 The mean depth of the deepest pit as a function of time. Fitting to experimental data (black circles) has been performed by using the hyperbolic dependencies [113].

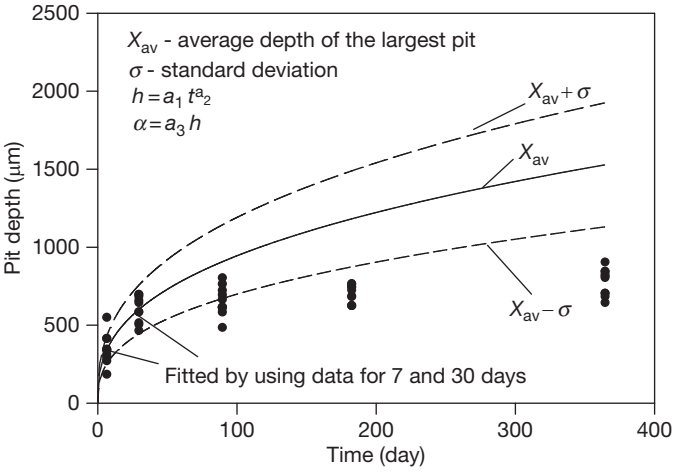


Figure 12 Fitting to experimental data (black circles) has been performed by using power law dependencies [114].

term experiments yields results that are not as good as application of hyperbolic law, at least for the considered experimental data. This happens, because the power law [115] does not take into account the repassivation of pits that plays a substantial role in determining the results of Aziz's experiments.

From the above, it is evident that the application of the logarithmic law yields much more acceptable results than does the application of the power law (**Figure 13**). However, at present, there is no physical model that could explain a logarithmic law of pit growth. In reality, if we assume that the growth law of individual pits can be described by a logarithmic function, it can be easily shown that the rate of the pit growth must decrease in accordance with the exponential law ($da/dt \sim \exp(-\zeta a)$). This dependence is substantially stronger that the maximum possible decrease of the rate of open pit growth under diffusion control when $da/dt \sim 1/a$. Accordingly, it might be suggested that the logarithmic dependence of pit growth is formally introduced for effectively describing pit propagation in the presence of repassivation phenomena.

A generalization of the hyperbolic function [113] in the form

$$h = \frac{a_1 t^{a_4}}{1 + a_2 t^{a_4}}, \quad \alpha = a_3 h \qquad [117]$$

appears to offer some advantages. The principal advantage of the general hyperbolic dependence [117] is that it coincides with the power law at short times and yields the accurate transition: $u, \alpha \rightarrow$ constant as $t \rightarrow \infty$ as it must, due to the repassivation. It is interesting to note that calculation

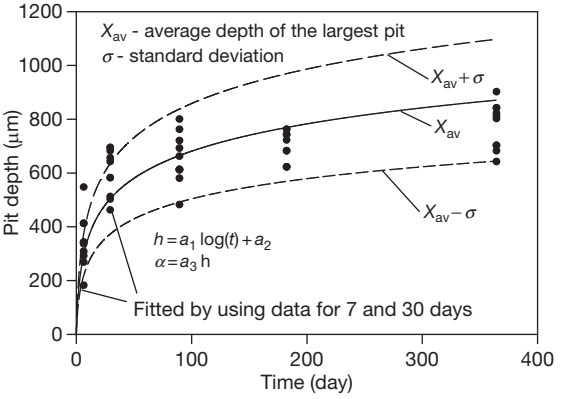


Figure 13 Fitting to experimental data (black circles) has been performed by using logarithmic dependencies [116].

shows that the parameter $a_4 = 1.0093 \approx 1$, that is, fitting by using the general hyperbolic dependencies yields, for the given case, practically the same results as the simple hyperbolic dependencies.

2.39.9 Monte Carlo Simulation

The existence of deterministic models to describe nucleation, propagation, and repassivation of localized corrosion processes allows us to replace the current DFA, which reduces to the direct solution of the balance eqns [2], with an alternative approach – Monte Carlo simulation of corrosion damage. However, it is worthwhile to develop this alternative

approach for describing corrosion damage for the following reasons. It is important to note that, frequently, under real conditions, only a few pits (or cracks), or sametimes only a single pit (or crack), may be alive (propagating) on the corroding metal surface (see the upper extremes in **Figure 1**). In this case, the differential equations for the damage function, which are equivalent to a balance equation for particles in discontinuous media, lose their strict physical meaning.

The main idea of the Monte Carlo method is to keep track of each stable pit (or crack) that nucleates, propagates and repassivates on the metal surface. A great advantage of this method lies in the fact that it allows us to take into account the interactions between particular individual pits (cracks) in an explicit manner. By doing so, it becomes possible to reduce the number of unknown parameters that describe the interaction between individual pits (cracks). Preliminary calculation shows that both approaches (DFA and Monte Carlo) yield the same results, if the same physical assumptions are made for describing the propagation of corrosion damage. At this point, it is important to note that, in the Monte Carlo method described here, the stages in the life of an individual pit or crack are described deterministically (i.e., in terms of models whose predictions are constrained by the natural laws), while the ensemble of pits is described statistically in terms of a Monte Carlo algorithm. Below, the current version of the deterministic Monte Carlo simulation (DMCS) algorithm that has been developed for predicting the accumulation of localized corrosion damage will be briefly described.

Let us denote by N_0 the total number (not per square centimeters) of stable pits that can nucleate on a metal surface having a total area of S. We number these pits by the index $k = 1, 2, \ldots, N_0$, and we will track each of these pits individually. Initially, we suggest that there are no stable pits on the metal surface. We proceed by calculating the probability that a pit will nucleate in stable form during the time step dt as

$$P_{\text{nucl}} = \frac{N(t + dt) - N(t)}{N_0 - N(t)} \qquad [118]$$

where, $N(t)$ is the number of stable pits that nucleate within the time interval between 0 and t. (Note that P_{nucl} is defined as a ratio of the number of pits that nucleate in the time interval between t and $t + dt$ to the number of available remaining sites at moment t).

For each pit, k, that has not yet been born, we generate a random number $0 \leq G_k \leq 1$. If $G_k \leq P_{\text{nucl}}$, a pit is proclaimed to have been born with a depth of penetration $x = 0$; otherwise, it is considered not to have been born and will be interrogated in the following step. In the case of instantaneous nucleation, all pits in the amount of N_0 are born during the first time step. In the more general 'progressive nucleation' case, the function $N(t)$ can be calculated by using, for example, the PDM, as noted above. The Cartesian coordinates of the centers of the nucleated pits, X_k and Y_k, are also established by using the random number generator, but this will be modified in the future to correspond to particular metallurgical, microstructural, and microchemical features on the surface (e.g., emergent precipitates or second phase particles) that may not be randomly distributed (e.g., $Cr_{23}C_7$ precipitates on emergent grain boundaries in stainless steels). This innovation will introduce, surface structural factors for the first time, in the deterministic prediction of localized corrosion damage. Practically, if the surface is rectangular with area $a \times b$, the Cartesian coordinates X_k and Y_k are declared to be $X_k = G_a$, and $Y_k = G_b$, where $0 \leq G \leq 1$ is a random number. If the surface is not rectangular, we insert the surface into a sufficiently large rectangle so that all points on the actual surface are included, and then repeat the described procedure. However, in this case, we declare X_k and Y_k as a center of a newly-born pit only if the point (X_k, Y_k) lies inside the actual surface. Otherwise, the procedure is repeated until success is achieved, that is, until (X_k, Y_k) lie inside the actual surface.

At each time step, the depth of the j-th stable, living pit, a_j, is calculated sequentially by using Faraday's law

$$a_j(t + dt) = a_j(t) + K_V i_j(t) dt \qquad [119]$$

We assume that i_j is described by Tafel's law for the active metal dissolution current density as a function of potential, that is,

$$i_j = i_0 \exp\left(-\frac{\alpha F \Delta\phi_j}{RT}\right) \text{ at } \Delta\phi_j < \Delta\phi_{\text{cr}} \text{ and}$$

$$i_j = 0 \text{ at } \Delta\phi_j \geq \Delta\phi_{\text{cr}} \qquad [120]$$

where $\Delta\varphi_j$ is the averaged potential drop on the active surface (relative to a point that is remote from the hemispherical pit of index j), α is the transfer coefficient of the metal dissolution reaction, i_0 is the corrosion current density at the corrosion potential, T is the Kelvin temperature, F is Faraday's constant, and R is

the gas constant. The physical meaning of the value $\Delta\varphi_{cr}$ is as follows. We can assume that the pit should continue to grow (be alive) if the metal potential at the bottom of the pit, E, is larger than the repassivation potential, E_{rp}, that is, at $E = E_{corr} - \Delta\varphi_j > E_{rp}$ or at $\Delta\varphi_j < E_{corr} - E_{rp}$. In other words $\Delta\varphi_{cr} = E_{corr} - E_{rp}$ is the difference between corrosion and repassivation potentials. Of course, it is assumed that $E_{corr} > V_{crit}$, which in turn is greater than E_{rp}, otherwise pitting corrosion cannot be initiated.

It can be shown[126] that the average potential drop on the active surface of a hemispherical pit (relative to a point far away from it), $\Delta\varphi_j$, can be estimated by the following relation

$$\Delta\phi_j = \frac{ba_j i_j}{\kappa} + \sum_{k\neq j} \frac{i_k a_k^2}{\kappa d_k} \qquad [121]$$

Here, the index k denotes the pits on the surface (the pit of interest is denoted by the index j), a_k is the radius of the hemispherical pit, i_k is the average current density inside the pit, κ is the conductivity, and $b \approx 2.1$ is a constant. Equation [121] can be obtained as an approximate solution of Laplace's equation for the electrical potential for the ensemble of hemispherical pits embedded in a plane under the condition that pits are far apart, that is, when the contribution of each pit can be considered independently. Noting that $\Delta\phi_j$ corresponds to the potential drop from some point within the pit and a remote point on the external surface, the first term in eqn [121] describes the contribution of the central pit to the total potential drop; this contribution can be found, for example, by using the method described in Pistorius and Burstein.[32] The second term describes the contribution of all other pits on the corroding surface.[50]

Accordingly, i_j is calculated by using eqn [121], where $\Delta\varphi_j$ is determined by a numerical solution of the equation

$$\Delta\phi_j = \frac{ba_j i_0}{\kappa}\exp\left(-\frac{\alpha F\Delta\phi_j}{RT}\right) + \sum_{k\neq j}\frac{i_0 a_k^2}{\kappa}\exp\left(-\frac{\alpha F\Delta\phi_k}{RT}\right) \qquad [122]$$

By solving this equation, we assume that all $\Delta\varphi_k$ in the second term on the right-hand side of eqn [122] are known and are equal to their previous (in the iteration sense) values.

After each time step, the repassivated (dead) pits are excluded from the population of actively growing pits. A pit is considered dead if $\Delta\varphi_j > \Delta\varphi_{cr}$. In addition, if two pits touch each other, that is, if

$$d_{i,j} \leq a_i + a_j \qquad [123]$$

where $d_{i,j}$ is the distance between the centers of the ith and jth pits, the smallest pit is declared repassivated (dead). A more correct assumption might be that, instead of two touching hemispherical pits, we have a single ellipsoidal pit with the depth of the deepest pit and a dimension obtained, for example, from the condition $\pi a^2 = \pi a_i^2 + \pi a_j^2$. Also, the entities that encroach on one another are the 'hemispheres of influence' (HOI), such that the pits begin to compete for the same resources (oxygen reduction). However, such an approach would substantially complicate the problem, but these factors will eventually be included in the analysis. To justify the adopted simplification, we note that the total area occupied by the active pits, S_a, must be much smaller than the area of the sample, S, in order to sustain differential aeration. Accordingly, overlap of the pits must be a relatively rare event. Moreover, our simplification can only decrease the value of $\Delta\varphi_j$ for any given pit j. Thus, if interactions between pits were substantial in the corrosion process, the conclusion would be all the more true than without the simplification.

In addition, we consider the probability that pits can repassivate accidentally (by chance). It is assumed that the probability of a pit repassivating during each time step is proportional to the magnitude of this step, that is, $P_\gamma = \gamma\, dt$, where γ is the delayed repassivation constant. Practically, for the jth living pit in each time step, we generate the random number $0 \leq G \leq 1$. If $G \leq P_\gamma$, the pit is declared to be dead (passivated) and is excluded from the further consideration; otherwise, the pit is considered to be alive on entering the next time step. We see that our Monte Carlo simulation method describes all three stages of the pit propagation damage – nucleation, propagation, and repassivation of stable pits.

Table 2 Parameter values for model calculations

i_0	1.5×10^{-2} A cm^{-2}
K_v	3.44×10^{-5} cm^3 C^{-1}
α	1
κ	$0.3 \times 10^{-3}\ \Omega^{-1}$ cm^{-1}
$\Delta\varphi_{cr}$	200 mV
T	25 °C
τ	3 days
a	3.65 days
b	1
γ	0.066 day^{-1}

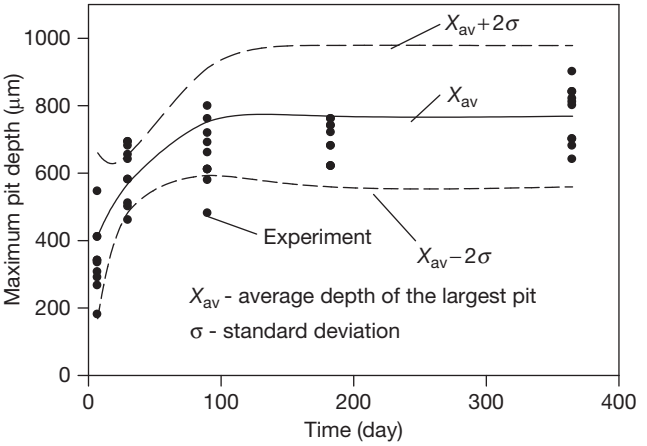


Figure 14 The mean depth of the deepest pit as a function of time. Experimental data are taken from Aziz[10] and are for aluminum alloy Alcan 2S-O in Kingston, Ontario tap water. Note that no calibration was performed on short term data, as was the case for **Figure 11**.

As an example of the application of this Monte Carlo simulation method within DFA, we consider again, the corrosion of aluminum alloy Alcan 2S_O in Kingston, Ontario tap water.[10] In this paper, for the purpose of numerically simulating the accumulation of pitting damage on the alloy surface, we will assume the values of various model parameters that are shown in **Table 2**. The parameters for the PDM (τ, a, and b) were chosen somewhat arbitrarily. The aim was only to ensure that, by the end of two weeks, practically all stable pits had nucleated, as indicated by the experimental data.[10] The value of K_V corresponds to that for aluminum and the value of κ corresponds to that typically assumed for tap water. The value for the delayed repassivation constant, γ, was chosen to ensure that, by the end of 2 months, the bulk of the pits represented by the bell-shaped curve have ceased to grow (see **Figure 1**). The value of α has been chosen as being typical for metal dissolution, and only the value of i_0 has been fitted to the experimentally measured values of maximum pit depth at $t = 1$ week after metal exposure.[10] Finally, the sample is considered to be rectangular with the dimensions 7×18 cm and $N_0 = 450$.[10]

Figure 14 shows a comparison of results of the Monte Carlo simulation with experimental data.[10] For a given observation time, t, we perform M simulations as described above and, accordingly, receive M values $x_1^*(t), x_2^*(t), \ldots, x_M^*(t)$ for the depth of the deepest pits. After that, we calculate the average (mean) value of the depth of the deepest pit as

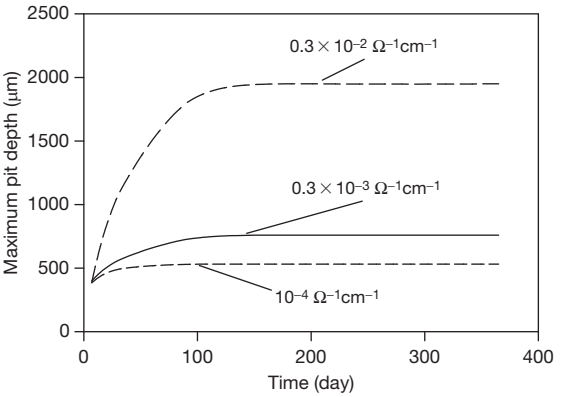


Figure 15 The mean depth of the deepest pit as a function of time for different conductivities of the electrolyte for aluminum alloy Alcan 2S_O in Kingston, Ontario tap water. Reproduced from Aziz, P. M. *Corrosion* **1956**, *12*, 35–46.

$$X_{av}(t) = \frac{1}{M}\sum_{m=1}^{M} x_m^*(t) \qquad [124]$$

and the standard deviation as

$$\sigma(t) = \sqrt{\frac{1}{M-1}\sum_{m=1}^{M}\left(X_{\mathrm{av}}(t) - x_m^*(t)\right)^2} \qquad [125]$$

For the sufficiently large value of M ($M \geq 40$, in our case), the results do not depend practically on the value of M. Because 95% of all experimental measurements usually lie in the interval ($X_{\mathrm{av}} - 2\sigma$, $X_{\mathrm{av}} +$

2σ), we posit that excellent agreement is observed between the experimental and simulation results.

It is important to emphasize that **Figures 11** and **14** were obtained by using absolutely different approaches. While **Figure 11** was obtained as a result of extrapolating data from short term experiments, without any assumption of kinetic parameters, **Figure 14** was obtained assuming that we have the complete set of required parameters without directly using data of short term experiments.

As noted previously, the main advantage of deterministic models over the statistics-based models lies in the fact that deterministic models allow us to make predictions for environmental conditions that lie beyond the available experimental (calibration) data, because the predictions are analytic and based upon a sound physical model and because the prediction is constrained by the natural laws to that which is physically real. As an example, **Figure 15** shows the influence of electrolyte conductivity on the mean depth of the deepest pits from multiple, identical sample areas on the surface as a function of time. It is seen that the mean of the deepest pits on multiple, but equal, areas on the alloy surface is a very sensitive function of the external environment conductivity; this probably reflects primarily an enhanced throwing power of the coupling current from the pit mouth to the external surface where it is consumed by oxygen reduction. Thus, increasing the conductivity of the external environment allows the coupling current to be thrown a greater distance from the mouth because of the lower IR potential drop in the solution, with the result that the pit can access a larger area for oxygen reduction. Because the pit growth rate is related to the magnitude of the coupling current through Faraday's law, the pit depth is also predicted to increase. It is also important to note that, in all cases, the mean of the deepest pits on the surface is predicted to become constant after ~100 days of exposure, simply reflecting that, at longer times, all the pits on the surface are 'dead' (repassivated). The critical question then is whether the deepest dead pits are sufficiently deep to nucleate cracks under the prevailing loading conditions.

The version of the Monte Carlo model that is described above deals only with the propagation of pitting corrosion damage. The example of Monte Carlo simulation that considers the transition of a pit into a crack can be found, for example, in Turnbull *et al.*[118] This model is based on mechanistic equations for pit and crack propagation rates (see eqns [70] and [90]), with statistically distributed input parameters, and no models were assumed for the pit or crack growth processes. These unknown parameters are fitted to the available experimental data, and accordingly, it will be impossible to use this model for predicting corrosion damage under different environmental conditions.

It is important to note that, in the models that consider deterministically the propagation of ensemble of cracks, it would be necessary, in the general case, to take into account the possible mechanical interactions between multiple cracks.[127]

2.39.10 Examples of Deterministic Prediction of Corrosion Damage in Complicated Industrial Systems

In this section, we present two examples of the deterministic prediction of corrosion damage in complex industrial systems. The two systems chosen are both from the electric power industry, where corrosion has proven to be an important factor in determining the availability of the power generating facilities. Indeed, it has long been recognized that the management of

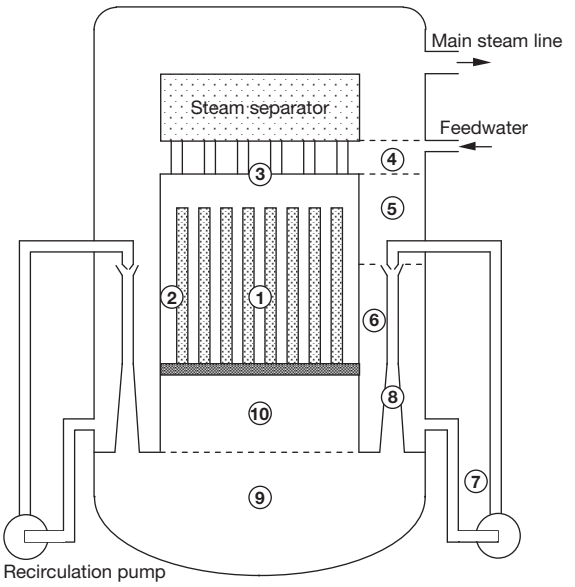


Figure 16 Schematic of the primary coolant circuit of a BWR having external coolant pumps. The regions are identified as (1) core channels; (2) core bypass; (3) upper plenum; (4) mixing plenum; (5) upper downcomer; (6) lower downcomer; (7) recirculation system; (8) jet pump; (9) bottom of lower plenum; (10) top of lower plenum. Reproduced from Macdonald, D. D.; Balachov, I.; Engelhardt, G. R. *Power Plant Chem.* **1999**, *1*, 9–16.

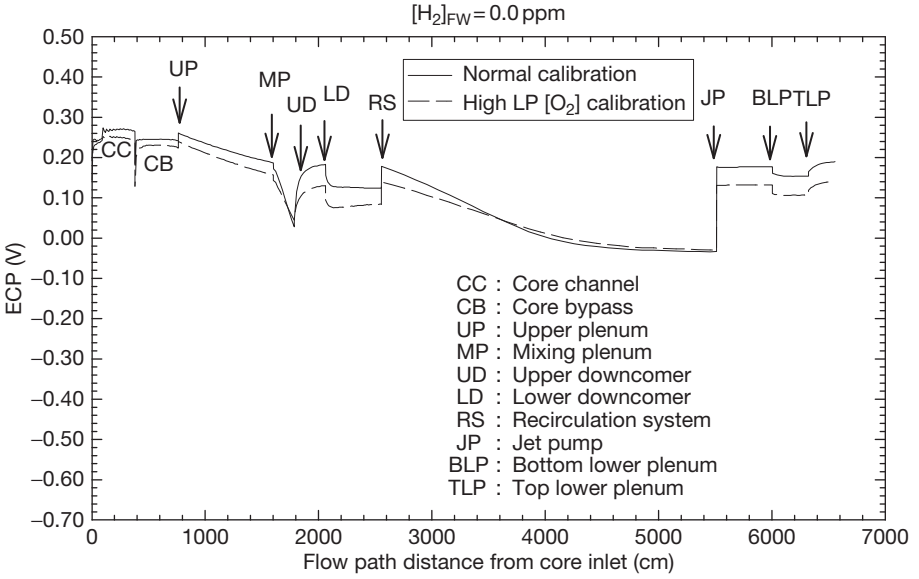


Figure 17 Calculated ECP (versus SHE) for the primary coolant circuit in the Leibstadt BWR operating under 'normal water chemistry' conditions (no hydrogen added to the feedwater). Reproduced from Macdonald, D. D.; Urquidi-Macdonald, M. *Corrosion* **1992**, *48*, 354–366.

the accumulation of corrosion damage, such that maintenance can be performed during scheduled outages, is an effective strategy for enhancing availability and for life extension. The primary purpose, here, however, is to convince the reader that the deterministic prediction of damage is practical and that such models may form the basis for the effective management of damage.

2.39.10.1 Cracking in Nuclear Reactors

The first example of the application of deterministic modeling of damage localized corrosion that we have chosen to describe in our work on modeling the accumulation of damage due to intergranular stress corrosion cracking (IGSCC) in sensitized Type 304 SS in the primary (liquid water) coolant circuit of a BWR.[23] This case is important because it illustrates the integration of the damage over the CEP taken by the reactor during normal operation. The reader will recall that the CEP is the path taken by the system in terms of those independent variables that have a significant impact on the rate of accumulation of damage as the system changes from the present state to the future state.

A schematic of the primary coolant circuit of a BWR with external coolant pumps is shown in Figure 16, and the various components and regions in the flow circuit are identified in the caption.[128]

Plotted in Figure 17 is the ECP calculated using the MPM (see Section 2.39.4) by first calibrating the radiolysis model (not the MPM) on Dresden II (a completely different reactor in Illinois) field data for the oxygen content of the coolant during operation and on another reactor (Duane–Arnold); these two reactors define the extremes of the US fleet of BWRs in their response to hydrogen added to the reactor feedwater.[20,31,99] These calculations predict that the ECP varies by about 250 mV around the entire coolant circuit when no hydrogen is added to the feedwater.

On the other hand, if significant hydrogen is added to the feedwater (e.g., 0.5 ppm), the corrosion potential is predicted to vary by about 0.9 V around the circuit. The Nuclear Regulatory Commission (NRC) has proposed a critical potential for intergranular stress corrosion cracking (IGSCC) in sensitized Type 304 SS of $-0.23\ V_{SHE}$; if the corrosion potential lies below this value, no cracking will occur but for more positive values, cracks will nucleate and grow, and the utility cannot take credit for a 'no cracking environment.' In the case of 'normal water chemistry (NWC)' (Figure 17, zero feedwater hydrogen), all components in the primary (liquid water) coolant circuit are predicted to be susceptible to IGSCC. Under 'hydrogen water chemistry (HWC)' (Figure 18, $[H_2]_{FW} = 0.50$ ppm), significant regions of the coolant circuit are protected by the added hydrogen, including much of the recirculation piping system

and the Lower Downcomer, while other regions are not. This example demonstrates how modern modeling techniques (i.e., reaction kinetic analysis, RKA) can be used to specify the conditions that must be achieved in a complex chemical system in order to protect components against corrosion.

As indicated above, the CEP describes the path taken by the system in terms of the independent variables that have a significant impact on the process, resulting in corrosion damage as the system changes from the present state to the future state. Once this path is defined, the corrosion rate is integrated along the path to yield the integral damage. This concept is illustrated below with reference to intergranular stress corrosion cracking in the core shroud of a boiling water nuclear reactor. (The shroud is a stainless steel drum that surrounds the core containing the fuel rods. The shroud contains a circumferential weld (the 'H-3 weld') at the top, and the steel adjacent to the weld, in the so-called heat-affected zone, HAZ, is sensitized and hence is susceptible to IGSCC. The phenomenon of IGSCC has been responsible for billions of dollars of losses in the nuclear power industry during a period extending over 30 years.) For IGSCC in sensitized Type 304 SS, the important independent variables in determining the crack growth rate are stress, degree of sensitization (DOS), crack length, temperature, conductivity of the water, oxygen, hydrogen, and hydrogen peroxide concentrations (all produced by the radiolysis of water or added, as in the case of H_2, for a reactor operating on hydrogen water chemistry), and water flow velocity. In turn, the concentrations of the redox species (H_2, O_2, and H_2O_2) determine the electrochemical corrosion potential (ECP), upon which the crack growth rate is exponentially dependent. Accordingly, the independent variables defining the CEP for the crack in the shroud ID surface within the HAZ of the weld are temperature, degree of sensitization (DOS) of the steel, crack length, tensile stress (residual and operational), water conductivity, ECP, and flow velocity. Thus, the CEP is defined in terms of the variations in each of these quantities as the reactor changes during the ten years of operation.

Figure 19 shows the operating history of the reactor in terms of reactor power (which defines the temperature and the bulk coolant flow rate) and hydrogen added to the feedwater (H_2 has a large impact on the ECP – see **Figures 17** and **18**). This generic reactor was operating on HWC, with hydrogen being added only when the reactor is operating, not during shutdown, and the damage, in terms of the crack length, was calculated by integrating the crack growth rate over the CEP for a period of 10 years.

The predicted damage is summarized in **Figure 20**, in which are plotted crack length (in centimeters) versus time curves for three operating protocols:

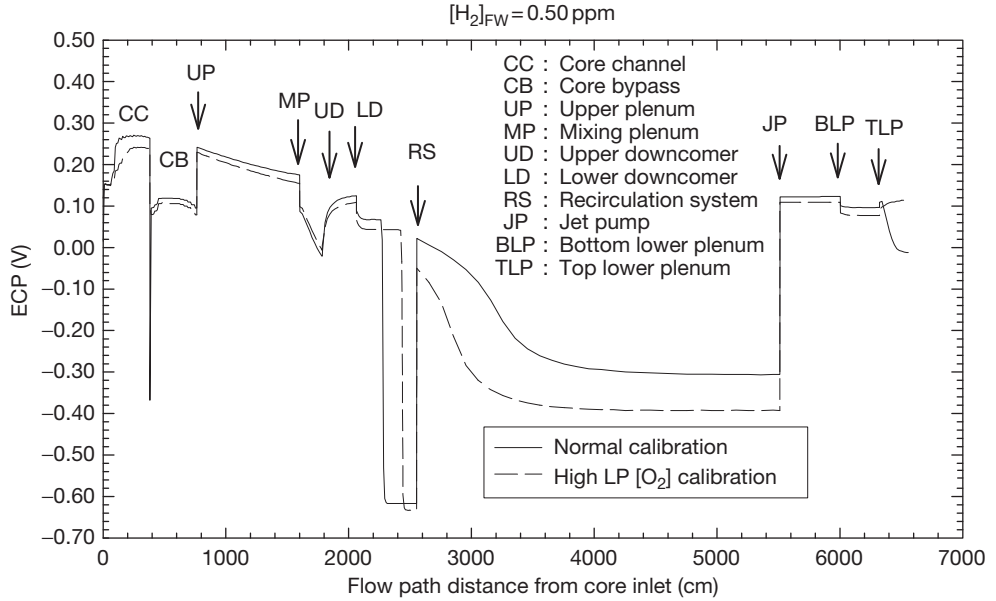


Figure 18 Calculated ECP (vs. SHE) for the primary coolant circuit in the Leibstadt BWR operating under 'hydrogen water chemistry' conditions (0.50 ppm hydrogen added to the feedwater). Reproduced from Macdonald, D. D.; Balachov, I.; Engelhardt, G. R. *Power Plant Chem.* **1999**, *1*, 9–16.

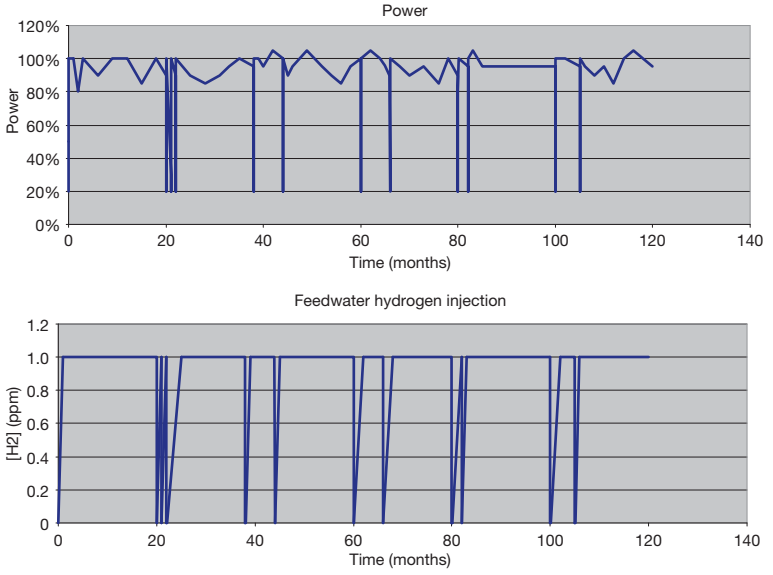


Figure 19 Corrosion evolutionary path (CEP) defined in terms of the prime variables of reactor power and hydrogen added to the reactor feedwater. Reproduced from Macdonald, D. D.; Balachov, I.; Engelhardt, G. R. *Power Plant Chem.* **1999**, *1*, 9–16.

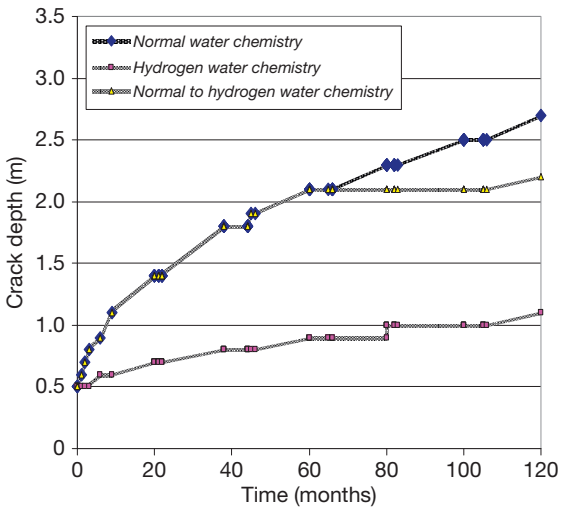


Figure 20 Integrated damage functions (crack length) versus reactor operating time for three operating scenarios: (1) normal water chemistry, (2) hydrogen water chemistry (1 ppm H_2 in the reactor feedwater), (3) NWC for 5 years followed by HWC for the remaining 5 years. Reproduced from Macdonald, D. D.; Balachov, I.; Engelhardt, G. R. *Power Plant Chem.* **1999**, *1*, 9–16.

(1) normal water chemistry (NWC), where no hydrogen is added to the reactor coolant (diamonds), (2) hydrogen water chemistry (1 ppm H_2 in the reactor feedwater over the entire 10-year operating period, squares), and (3) NWC for 5 years followed by HWC for the remaining 5 years (triangles). The various 'jigs and jogs' in the data are not artifacts caused by imprecise calculation, but represent changes in crack growth rate with the changing conditions along the CEP, particularly when the reactor is shut down and started up. During entering and exiting a shutdown period, some of which occur for refueling of the reactor, the hydrogen concentration, coolant conductivity, temperature, coolant flow velocity, and stress (although the stress on the crack was assumed to be primarily residual and hence, is assumed to be constant in the simulations described above) change markedly, but the impact on the overall damage is only small, because the shutdown and start up periods are short. It is seen from **Figure 20** that the crack growth rate (gradient of the curve) decreases with time. This is the consequence of coupling between the internal and external environments of the crack, as postulated in the coupled environment fracture model (CEFM) (see Macdonald *et al.*[23] and citations therein and the corresponding section of this article). Thus, as the crack grows in length under constant potential conditions, a larger IR potential drop occurs down the crack, resulting in a lower potential drop being available on the external surface for the reduction of oxygen or the evolution of hydrogen, thereby resulting in a reduction of the coupling current and hence, a lower crack growth rate. (Note that the crack cannot grow faster than the coupling current can be consumed on the external surface by oxygen reduction and/or hydrogen evolution, and it has been established experimentally

that the crack growth rate is linearly related to the coupling current.)

Under NWC, the crack is predicted to grow by about 2.2 cm over the 10 years of operation (note that we assume an initial crack length of 0.5 cm). If HWC (where 1 ppm of H_2 is added to the reactor coolant water only during operation, see **Figure 19**), the crack is estimated to grow by only 0.6 cm over the 10-year operating period, a substantial reduction in the damage. If, on the other hand, the reactor operated for the first 5 years on NWC (no added hydrogen) and then switched to HWC, the damage (increase in crack length) is predicted to be 1.7 cm. The progression of damage is clearly governed by the 'law of decreasing returns,' in that the damage avoided in the last 5 years by implementing HWC (0.5 cm) is substantially less than that incurred under NWC in the first 5 years (1.6 cm). This is entirely due to the dependence of the crack growth rate on crack length, a dependence that was never previously recognized in crack growth-rate studies. Clearly, predictions of this type are of considerable value, because they allow the benefits to be defined in a cost/benefit analysis. The cost of installing HWC in a BWR is significant because of the need to store considerable amounts of hydrogen on site and because of the need to shield personnel against the production of radioactive $^{16}NH_3$ that forms by neutron bombardment of oxygen in water ($^{1}n_0 + {}^{16}O_8 \rightarrow {}^{16}N_7 + {}^{1}p_1$) under the reducing conditions that exist in the coolant circuit under HWC operating conditions. Note that $^{16}N_7$ is a strong γ-photon emitter.

At this point, it is worthwhile to enquire how accurate the prediction might be for a system that is as complex as an operating BWR. Data published by Tang *et al.*[129] (see also Macdonald *et al.*[23]) for a crack adjacent to the H-3 weld in a reactor operating in Taiwan affords an opportunity to assess the accuracy, as depicted in **Figure 21**. The depth of the crack was assessed by inspection at times of 10 and 20 months after Outage 11 and the ALERT code was fit to the datum for the shorter time, essentially by adjusting the time of nucleation of the crack, such that the calculated and measured crack depth coincided. ALERT was then used to calculate the crack depth at the longer, 20-month time after Outage 11, and the comparison with the measured value is displayed in the figure. Excellent agreement is obtained, with the small residual difference being attributed to uncertainty in the operating history of the reactor.

This example has been presented because it illustrates the application of damage function analysis (DFA) to the deterministic prediction of damage in a complex industrial system. It also illustrates the role of, and the need for, careful characterization,of the CEP. It is the opinion of the authors that a water-cooled nuclear reactor is comparable in complexity to, say, an oil production system, and that meaningful damage prediction calculations are possible on the basis of DFA (or any other deterministic protocol), provided that the CEP is carefully defined.

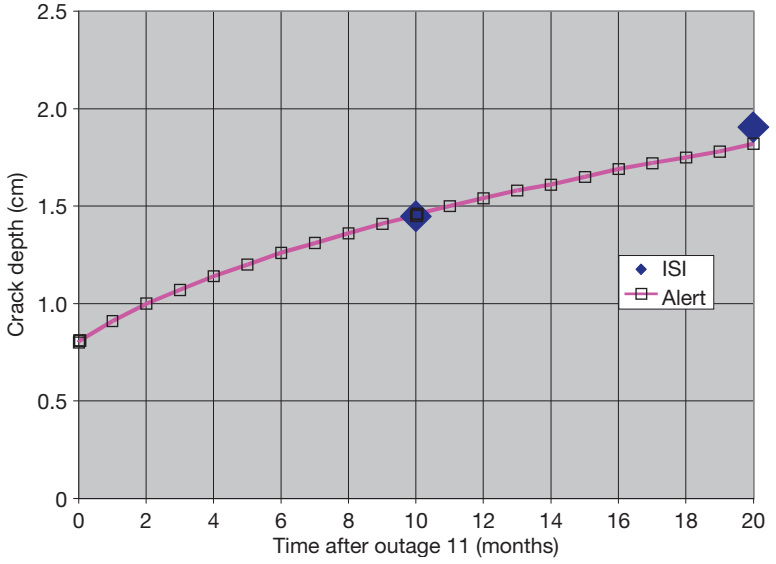


Figure 21 Depth of a crack at the H3 weld in the shroud as a function of time after Outage 11. (♦) Values calculated by ALERT.[19,23] Experimental data reported by Tang *et al.*[129]

Of course, the reward for being able to make such calculations is that it then becomes possible to answer the all-important 'what if' questions, such as: 'What damage will I incur if I operate in this manner?' or 'If I take this corrective measure, can I operate with reasonable certainty to the next outage?' Answers to questions of this type can have significant economic implications.

2.39.10.2 Low Pressure Steam Turbines

Another complex industrial system that has been modeled deterministically with regard to the accumulation of localized corrosion damage is low pressure steam turbines (LPSTs). Failure of LPSTs generally occurs via the cracking of the rotor, disk, or blades, from cracks that nucleate at pits.

Below we will show some results of calculating the probability of failure, P_f, of LPSTs by using the deterministic Monte Carlo simulation method that has been described above.[130] By definition, P_f is the probability that at least one corrosion event in any form (pit, stress corrosion crack, or fatigue crack) reaches a depth, a_{cr}, at a given observation time, t, where a_{cr}, in this case, is the critical dimension (e.g., the stress corrosion crack length at which unstable, rapid mechanical fracture and sudden failure occurs). Using the Monte Carlo simulation method, we can calculate P_f in a straightforward way. Let us assume that we make M calculations of the depth of the deepest corrosion event, x^*, and, in M_f cases ($M_f \leq M$), $x^* > a_{cr}$ for a given observation time, t. In this case, $P_f(a_{cr}, t)$ is defined as

$$P_f = \frac{M_f}{M} \qquad [126]$$

for sufficiently large M.

In particular, it was found that the failure probability is a very sensitive function of the conditions that exist in a low pressure steam turbine (LPST) during shutdown, including the oxygen concentration, chloride concentration, and the fraction of the time spent under shutdown versus operation, if the shutdown environment is not deaerated, and the steel

Table 3 Assumed operational cycle parameters for the development of corrosion fatigue in low pressure steam turbine blades

Shutdown	*Operation cycle*
$t = 100$ h	$t = 500$ h
$\sigma = 0$	$\sigma = \sigma_m + 0.5\ \Delta\sigma\sin(2\pi ft)$
	$\sigma_m = 84$ ksi, $\Delta\sigma = 4$ ksi, $f = 60$ Hz
$T = 25\,^{\circ}$C	$T = 95\,^{\circ}$C
$[O_2] = 8$ ppm	$[O_2] < 1$ ppb
$[Cl^-] = 3500$ ppm	$[Cl^-] < 100$ ppm
pH = 6	pH = 6

Table 4 Assumed operational cycle parameters for the development of corrosion fatigue in low-pressure steam turbine discs

Shutdown	*Operation cycle*
$t = 100$ h	$t = 500$ h
$\sigma = 0$	$\sigma = 95$ ksi
$T = 25\,^{\circ}$C	$T = 160\,^{\circ}$C
$[O_2] = 8$ ppm	$[O_2] < 1$ ppb
$[Cl^-] = 3500$ ppm	$[Cl^-] < 100$ ppm
pH = 6	pH = 6

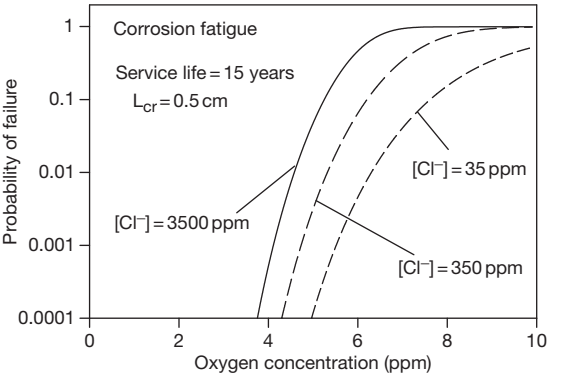


Figure 22 Calculated failure probability for LP steam turbine blades as a function of oxygen concentration for different chloride concentrations in the electrolyte film during the shutdown period. Other parameters are the same as in Table 3.

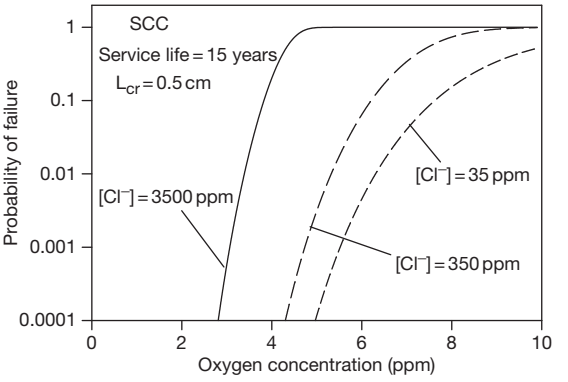


Figure 23 Calculated failure probability for LP steam turbine discs as a function of oxygen concentration for different chloride concentrations in the electrolyte film during the shutdown period. Other parameters are the same as in Table 4.

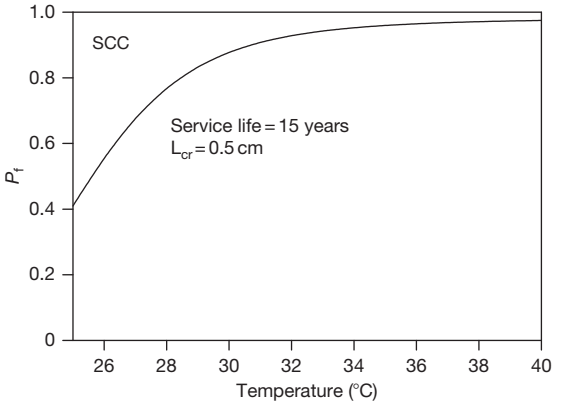


Figure 24 Probability of failure of a turbine disk as a function of temperature during the shutdown period. The oxygen concentration during shutdown period is 4.5 ppm. Other parameters are the same as in **Table 4.**

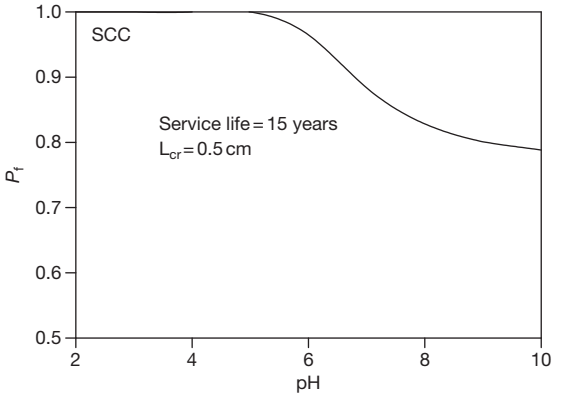


Figure 25 Probability of failure of a turbine disk as a function of pH during the shutdown period. The oxygen concentration during shutdown period is 5.2 ppm. Other parameters are the same as in **Table 4**.

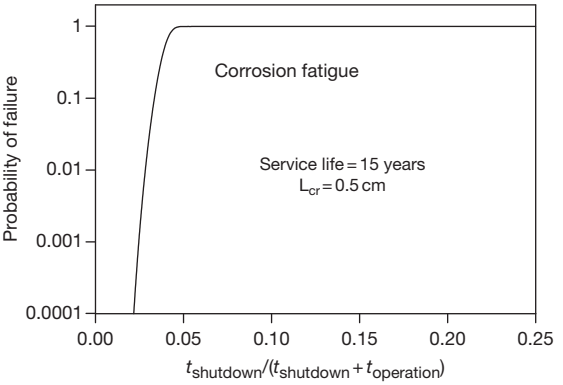


Figure 26 Calculated failure probability for LP steam turbine blades versus the fraction of the time spent in shutdown.

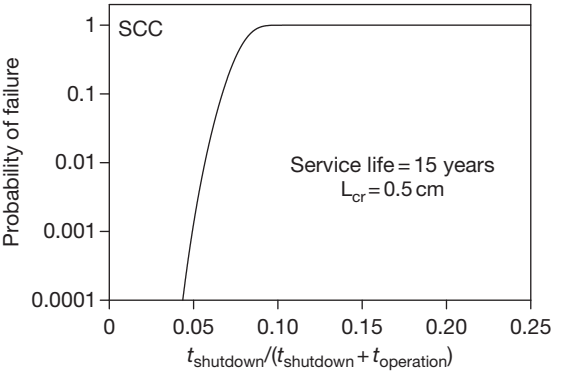


Figure 27 Calculated failure probability for LP steam turbine discs versus the fraction of the time spent in shutdown.

surfaces are not washed free of chloride ion. It is important to note that the calculations were made by using experimentally determined electrochemical parameters for Type 403 SS in the case of determining failure in blades (where CF takes place) and A470/471 steels in the case of determining failure in discs and rotors (where SCC takes place).[131]

For illustrative purposes, we present below only some examples of the influence of environmental parameters on the probability of failure of blades and discs in low pressure steam turbines. The assumed operating cycle parameters for the case of stress corrosion cracking can be found in **Tables 3** and **4**. This information defines the CEP.

The failure probabilities plotted in **Figures 22–27** as a function of $[O_2]$, $[Cl^-]$, temperature, pH, and fraction of the time spent in shutdown under aerated conditions all during the shutdown period, were calculated for 219 shutdown–operation cycles over 15 years, assuming instantaneous nucleation (all pits nucleating during the first shut down) and that all cycles are identical. This particular CEP was chosen for modeling convenience; more complicated CEPs, in which the cycle time and other parameters are varied over realistic ranges, are easily enacted. The critical crack length was assumed to be 0.5 cm, and crack growth was determined to occur primarily by mechanical fatigue during operation.

The calculated failure probabilities plotted in **Figures 22** and **23** are clearly sensitive functions of both the oxygen concentration and the chloride concentration in the condensate on the blade or disk surface left over from operation. Increasing the oxygen concentration under open shutdown (i.e., opening the turbine to the atmosphere) displaces the corrosion potential in the positive direction and hence

increases the pit nucleation and propagation rates, thereby resulting in enhanced pitting damage. Upon subsequent startup, the pits act as stress-raisers for the nucleation of cracks (SCC and corrosion fatigue) and hence lead to an increase in the failure probability. The important lesson to be learned from this plot, and from many simulations of this type, is that deaeration of the turbine upon shutdown should be an effective method of protecting low pressure steam turbines against blade and disc failure after subsequent start-up.

Figures 22 and **23** also show that the failure probability, P_f, decreases sharply to a very low value (essentially zero) as $[Cl^-]$ is lowered below 35 ppm. This provides a second strategy for reducing, or even eliminating the failure of discs and blades, reducing the chloride concentration of Cl^- on the surfaces immediately upon shutdown. This would be most easily instituted by washing the blade and disc surfaces with pure (chloride-free) water immediately upon shutdown of the turbine. In fact, an even more effective strategy would be to combine turbine surface washing with nitrogen blanketing. Given that turbine failure costs anywhere between \$10 million and \$100 million per event, depending upon the nature of the failure, an incentive certainly exists for the utilities to explore the strategies identified above.

The temperature of shutdown is also predicted, in certain situations, to have a visible effect on the failure probability (see **Figure 24**), primarily through the pit growth rate. While this is a substantial reduction, the investment in the necessary facilities to cool a turbine after shutdown is possibly too great to be cost effective, but the desired effect might be achieved through conventional air conditioning.

The remaining environmental variable that is amenable to change is the pH. This variable may be changed by doping the boiler water of the steam with pH buffers, including ammonia, various amines, and systems such as boric acid–hydroxide systems. However, our predictions are that pH has relatively little impact on the failure probability (see **Figure 25**). Thus, the gain that might be achieved in turbine service life possibly would not warrant instituting a major chemistry change upon the basis of enhancing turbine blade and disc life alone.

Figures 26 and **27** clearly show that the probability of failure increases rapidly with the time spent under aerated conditions during shutdown. That is why the reduction of this time could be an effective method of protecting low pressure steam turbines against blade and disc failure.

It must be emphasized that prediction of the probability of failure in LPTS can also be done on the basis of damage function analysis. Examples of the corresponding calculations can be found in Engelhardt *et al.*[132] and Engelhardt and Macdonald.[131]

2.39.11 Conclusions

The purpose of this chapter is to introduce the reader to deterministic methods that have been developed to predict the accumulation of corrosion damage in complex industrial systems. That such predictions can be made is now well beyond question, and it is expected that determinism will rapidly develop as the philosophy of choice in making damage predictions. Because practicing scientists and engineers have seldom taken a course on the philosophies of prediction, few have been exposed to the modeling choices that are available, ranging from empiricism, to 'mechanistic models,' to models that are characterized by varying degrees of determinism. This latter caveat recognizes that, in practice, 'determinism' is an ideal concept that is seldom, if ever, achieved in modeling in pure form. From a practical viewpoint, however, determinism offers a methodology that means minimal calibration, because the dependent/independent variable relationships are captured within the constitutive equations of the model, rather than by calibration, as is the case for an empirical model. As noted above, failure in industrial systems is, by and large, a very rare event, so that almost no chance exists for establishing reliable dependent–independent variable relationships by direct calibration.

Another great advantage of deterministic models is that they draw upon the wealth of scientific experience by constraining the solutions of the constitutive equations to that which is 'physically real' via the natural laws. This feature cannot be overemphasized, as it greatly reduces the need for calibration, provides a condition that must be satisfied and does not arise from the model itself (the model is described by the constitutive equations). If the constraints are absent, or are not imposed, the model may still be 'mechanistic,' but it clearly lacks the predictive power offered by determinism.

In making predictions with any model, be it empirical or deterministic, it is essential to define the CEP, which is the path taken by the system in terms of those independent variables that have a significant impact on the damage accumulation rate as the system changes from the present state

to the future state. Because, the future cannot generally be predicted with certainty, the CEP is most often synthesized to present 'what if' scenarios of operation. These scenarios coupled with economic analyses, may be used to identify the most cost-effective modes of operation and hence, are a valuable financial planning tool. Similarly, the CEP may be designed to ascertain the probability that failure will not have occurred by the time of the next scheduled inspection, which, given the deleterious cost of unscheduled outages, again becomes a valuable financial planning tool. Finally, it is evident that the tools discussed in this chapter are capable of providing valuable input to the design of complex industrial systems, particularly in the light of the fact that corrosion issues are seldom, if ever, given more than cursory examination.

One of the significant advantages of integrating the damage along the CEP to yield the damage function (DF) is that the process may be applied to historical data from which the CEP may be defined with considerable accuracy. In this mode, the predicted damage function becomes a means of accessing the veracity of the model and the damage prediction process. In fact, this type of calculation should always be performed, if for no other reason than that it affords a means of customizing the algorithm to the system of interest. Customization may require the adjustment of values for poorly known parameters in the model or may even require modification of the code itself. In any event, the code becomes the 'alter ego' of the system and provides a means of accurately predicting the accumulation of damage under conditions where the CEP can be defined with historically accuracy.

References

1. Macdonald, D. D.; Urquidi-Macdonald, M. *Corrosion* **1992**, *48*, 354–366.
2. Kowaka, M.; Tsuge, H.; Akashi, M.; Masamura, K.; Ishimoto, H. *Introduction to Life Prediction of Industrial Plant Materials: Application of the Extreme Value Statistical Method for Corrosion Analysis*; Allerton Press, 1994.
3. Harlow, D. G.; Wei, R. P. *AIAA J.* **1994**, *32*, 2073–2079.
4. Harlow, D. G.; Wei, R. P. *Eng. Fract. Mech.* **1998**, *59*, 305–325.
5. Harlow, D. G.; Wei, R. P. *Fatigue Fract. Eng. Mater. Struct.* **2001**, *22*, 427–436.
6. Harlow, D. G.; Wei, R. P. *Fatigue Fract. Eng. Mater. Struct.* **2001**, *24*, 523–535.
7. Harlow, D. G.; Wei, R. P. *Model. Simul. Mater. Sci. Eng.* **2005**, *13*, R33–R51.
8. Macdonald, D. D.; Liu, C. In *Proceedings of the 12th International Corrosion Congress*; NACE International, Houston, TX, 1993; Vol. 5A, pp 3446–3459.
9. Liu, C.; Macdonald, D. D. *J. Press. Vessel Technol.* **1997**, *119*, 393–400.
10. Aziz, P. M. *Corrosion* **1956**, *12*, 35–46.
11. Freiman, L. I. In *Progress in Science and Technology. Corrosion and Corrosion Protection* VINITI, 1985; Vol. 11, pp 3–71 [in Russian].
12. Macdonald, D. D. *Proceedings of Corrosion and Prevention-2000*; Australasian Corrosion Association: Auckland, NZ, Nov 2000; pp 19–22; Paper **42**.
13. Engelhardt, G. R.; Macdonald, D. D. *Corrosion* **1998**, *54*, 469–479.
14. Engelhardt, G. R.; Macdonald, D. D. *Corros. Sci.* **2004**, *46*, 2755–2780.
15. McCartney, L. N. *Int. J. Fract.* **1979**, *15*, 477–487.
16. Turnbull, A.; McCartney, L. N.; Zhou, S. *Corros. Sci.* **2006**, *48*, 2084–2105.
17. Polyanin, A. D.; Zaitsev, V. F.; Moussiaux, A. *Handbook of First Order Partial Differential Equations*; Taylor & Francis: London, 2002.
18. Wagner, C.; Traud, W. *Z. Elektrochem.* **1938**, *44*, 391–402.
19. Macdonald, D. D.; Urquidi-Macdonald, M. In *Encyclopedia of Electrochemistry*; Bard, A. J., Stratmann, M., Eds.; Wiley-VCH: Weinheim, 2006; *Electrochemical Engineering*, Vol. 5, Macdonald, D. D., Schmuki, P., Eds., pp 665–720.
20. Macdonald, D. D. *Corrosion* **1992**, *48*, 194–205.
21. Macdonald, D. D. *Pure Appl. Chem.* **1999**, *71*, 951–978.
22. Selman, J. R.; Tobias, C. W. *Advan. Chem. Eng.* **1978**, *10*, 211.
23. Macdonald, D. D.; Balachov, I.; Engelhardt, G. R. *Power Plant Chem.* **1999**, *1*, 9–16.
24. Anderko, A.; Young, R. D. *Corrosion* **2000**, *56*, 543–555.
25. Anderko, A.; McKenzie, P.; Young, R. D. *Corrosion* **2001**, *57*, 202–213.
26. Sridhar, N.; Brossia, C. S.; Dunn, D. S.; Anderko, A. *Corrosion* **2004**, *60*, 915–936.
27. Nordsveen, M.; Nešić, S.; Nyborg, R.; Stangeland, A. *Corrosion* **2003**, *59*, 443–456.
28. Nešić, S.; Cai, J.; Lee, K.-L. *Corrosion/2005*; NACE International: Houston, TX, 2005; Paper no. 05556.
29. Davydov, A. D.; Shaldaev, V. S.; Engelhardt, G. R. *Russian J. Electrochem.* **2006**, *42*, 121–128.
30. Davydov, A. D. *Russian J. Electrochem.* **2008**, *44*(7), 835–839.
31. Stewart, J.; Williams, D. E. In *Advances in Localized Corrosion*; Isaacs, H. S., *et al.* Eds.; NACE International: Houston, TX, 1990; pp 131–136.
32. Pistorius, P. S.; Burstein, G. T. *Philos. Trans. R. Soc. Lond. A* **1992**, *341*, 531–559.
33. Williams, D. E.; Stewart, J.; Balkwill, P. H. *Corros. Sci.* **1994**, *36*, 1213–1235.
34. Williams, D. E.; Westcott, C.; Fleischmann, M. *J. Electrochem. Soc.* **1985,** *132,* 1796, 1804.
35. Lin, L. F.; Chao, C. Y.; Macdonald, D. D. *J. Electrochem. Soc.* **1981**, *128*, 1194.
36. Macdonald, D. D.; Urquidi-Macdonald, M. *Electrochim. Acta* **1986**, *31*, 1079.
37. Macdonald, D. D.; Urquidi-Macdonald, M. *J. Electrochem. Soc.* **1987**, *134*, 41.
38. Macdonald, D. D.; Urquidi-Macdonald, M. *J. Electrochem. Soc.* **1989**, *136*, 961.
39. Macdonald, D. D.; Urquidi-Macdonald, M. *J. Electrochem. Soc.* **1992**, *139*, 3434–3449.

40. Uhlig, H. In *Passivity of Metals*; Frankenthal, R. P., Kruger, J., Eds.; The Electrochemical Society: Princeton, NJ, 1978; p 1.
41. Engell, H. *Electrochim. Acta* **1977**, *22*, 987.
42. Macdonald, D. D. *J. Electrochem. Soc.* **2006**, *153*, B213–B224.
43. Kong, D-S.; Chen, S-N.; Wang, C.; Yang, W. *Corros. Sci.* **2003**, *45*, 747.
44. Tsuchiya, H.; Fujimoto, O.; Chinara, O.; Shibata, T. *Electrochim. Acta* **2002**, *47*, 4357.
45. Kaesche, H. *Z. Phys. Chem. (Neue Folge)* **1962**, *34*, 87.
46. Böhni, H.; Uhlig, H. H. *J. Electrochem. Soc.* **1969**, *116*, 906–910.
47. Shibata, T. *Trans ISIJ* **1983**, *23*, 785.
48. Fratesi, R. *Corrosion* **1985**, *41*, 114.
49. Reuter, M.; Heusler, K. E. *Electrochim. Acta* **1990**, *35*, 1809–1814.
50. Carslaw, H. S.; Jaeger, J. C. *Conduction of Heat in Solids*, 2nd ed.; Oxford Press: London, 1959.
51. McGeough, J. A. *Principles of Elecrochemical Machining*; Chapman and Hall: London, 1974.
52. Newman, J.; Thomas-Alyea, K. E. *Electrochemical Systems*, 3rd ed.; Prentice Hall: Englewood Cliffs, NJ, 2004.
53. Chang, H.-Y.; Park, Y.-S.; Hwang, W.-S. *J. Mater. Process. Technol.* **2000**, *103*, 206–217.
54. Bird, B. R.; Stewart, W. E.; Lightfoot, E. N. *Transport Phenomena*, Wiley: New York, 2002.
55. Davis, J. T. *Turbulence Phenomena*; Academic Press: London, 1972.
56. Popov, Yu. A.; Alekseev, Yu. V.; Kolotyrkin, Ya. M. *Sov. Electrochem.* **1979**, *14*, 1260.
57. Walton, J. C.; Grangolino, G.; Kalandros, S. K. *Corros. Sci.* **1996**, *38*, 1–18.
58. Turnbull, A. *Br. Corros. J.* **1993**, *28*, 297–308.
59. Lasaga, A. C.; Kirkpatrick, R. J. *Kinetics of Geochemical Processes*; Mineralogical Society of America: Washington, 1981.
60. Walton, J. C. *Corros. Sci.* **1990**, *30*, 915–928.
61. Sharland, S. M. *Corros. Sci.* **1992**, *33*, 183.
62. Engelhardt, G. R.; Strehblow, H.-H. *Corros. Sci.* **1994**, *34*, 1171–1725.
63. Laycock, N. J.; White, S. P. *J. Electrochem. Soc.* **2001**, *148*, B264–B275.
64. Yao, L. S.; Prusa, J. *Adv. Heat Transfer* **1989**, *19*, 1.
65. Engelhardt, G. R.; Davydov, A. D. *Russian J. Electrochem.* **1994**, *30*, 865.
66. Engelhardt, G. R.; Strehblow, H.-H. *J. Electroanal. Chem.* **1994,** *365,* 7.
67. Engelhardt, G. R.; Strehblow, H.-H. *J. Electroanal. Chem.* **1995,** *394,* 7.
68. Pillay, B.; Newman, J. *J. Electrochem. Soc.* **1993**, *140*, 414.
69. Verbrugge, M. V.; Bakker, D. R.; Newman, J. *J. Electrochem. Soc.* **1993**, *140*, 2530.
70. COMSOL software, www.comsol.com.
71. OLI Systems software, www.olisystems.com.
72. Papavinosan, S.; Revie, W.; Friesen, W. *Corrosion Reviews* **2006**, *24*, 173–230.
73. Isaacs, H. S. *J. Electrochem. Soc.* **1973**, *120*, 1456.
74. Doig, P.; Flewitt, E. J. *Metall. Trans. A* **1981**, *12*, 927.
75. Ateya, D. G.; Pickering, H. W. *J. Appl. Electrochem.* **1981**, *11*, 453.
76. Gravano, S.; Galvele, J. R. *Corros. Sci.* **1984**, *24*, 517.
77. Sharland, S. M.; Tasker, P. W. *Corros. Sci.* **1988**, *28*, 603.
78. Sharland, S. M.; Jackson, C. P.; Diver, A. J. *Corros. Sci.* **1989**, *29*, 1149.
79. Sharland, S. M. *Corros. Sci.* **1992**, *33*, 183.
80. Alkire, R.; Sittary, D. *J. Elecrochem. Soc.* **1979**, *126*, 15.
81. Galvele, J. R. *J. Electrochem. Soc.* **1976**, *123*, 464.
82. Galvele, J. R. *Corros. Sci.* **1981**, *21*, 551.
83. Newman, J.; Hanson, D. B.; Vetter, K. *Electrochim. Acta* **1977**, *22*, 829.
84. Harb, J. N.; Alkire, R. C. *J. Electrochem. Soc.* **1991**, *138*, 2594.
85. Evans, U. R. *J. Inst. Met.* **1923**, *30*, 239.
86. Manahan, M. P.; Macdonald, D. D.; Peterson, A. J. *Corros. Sci.* **1995**, *37*, 189.
87. Zhou, X.; Balachov, I.; Macdonald, D. D. *Corros. Sci.* **1998**, *40*, 1349.
88. Yeh, T-K.; Tsai, C-H.; Chang, C-Y. In Proceedings of the 11th International Symposium on Environmental Degradation of Materials in Nuclear Power Systems – Water Reactors : ANS, WA, USA, 10–14, August 2003; pp 500–511.
89. Macdonald, D. D.; Urquidi-Macdonald, M. *Corros. Sci.* **1991**, *32*, 51.
90. Macdonald, D. D.; Lu, P.-C.; Urquidi-Macdonald, M.; Yeh, T.-K. *Corrosion* **1996**, *52*, 768.
91. Engelhardt, G. R.; Urquidi-Macdonald, M.; Macdonald, D. D. *Corros. Sci.* **1997**, *39*, 419–441.
92. Engelhardt, G. R.; Macdonald, D. D.; Urquidi-Macdonald, M. *Corros. Sci.* **1999**, *41*, 2267–2302.
93. Vankeerberghen, M.; Macdonald, D. D. *Corros. Sci.* **2002**, *44*, 1425.
94. Engelhardt, G. R.; Macdonald, D. D. In *Corrosion/ 2000,* Orlando, FL, March 2000, Paper no. 00227, pp 1–12.
95. Engelhardt, G. R.; Macdonald, D. D.; Millet, P. *Corros. Sci.* **1999**, *41*, 2165–2190.
96. Turnbull, A. *Corros. Sci.* **1999**, *41*, 2267.
97. Gavrilov, S.; Vankeerberghen, M.; Deconinck, J.; Vereecken, J. In 4th Workshop LWR Coolant Water: Radiolysis and Electrochemistry Avignon: 2002, 26 April France.
98. Ford, F. P. In *Corrosion sous contrainte-phénoménologie et mécanismes*; Desjardins, D., Oltra, R., Eds.; Bombannes, 1990.
99. Ford, F. P.; Taylor, D. F.; Andresen, P. L.; Ballinger, R. G. Corrosion-assisted cracking of stainless and low-alloy steels in LWR environments, EPRI Final Report NP-5064M, Project 2006-6; Palo Alo: Electric Power Research Institute, CA, 1987.
100. Shoji, T. Quantitative prediction of environmentally assisted cracking based on crack tip strain rate *American Society of Mechanical Engineers, Pressure Vessels and Piping Division (Publication) PVP* **1985**, *Vol. 99*, 127–142.
101. Shoji, T.; Moriya, S.; Arai, H.; Higashi, M. In *Fracture Mechanics Applications;* American Society of Mechanical Engineers, Pressure Vessels and Piping Division (Publication) PVP, 1994; Vol. *287*, pp 107–113.
102. Vankeerberghen, M.; Gavrilov, S. In EUROCORR 2001, Riva del Garda: Italy, 30 September–4 October 2001.
103. Peng, Q. J.; Kwon, J.; Shoji, T. *J. Nucl. Mater.* **2004**, *324*, 52.
104. Macdonald, D. D.; Lvov, S. N.; Kelkar, G.; Magalhaes, J. F.; Kwon, H.-K.; Wuensche, A.; Biswas, R.; Ahmad, Z.; Engelhardt, G. R.; Urquidi-Macdonald, M. The development of deterministic methods for predicting corrosion damage in water-cooled nuclear reactors, Final Report EP93–33 Empire State Electric Energy Research Corporation (ESEERCO): New York, NY, November 1996.

105. Wuensche, A.; Macdonald, D. D. In CORROSION 2001, Houston, TX, 2001, Paper no. 01236.
106. Andresen, P. L. *Corrosion* **1993**, *49*, 714.
107. Wilkinson, D. S.; Vitek, V. *Acta Metall.* **1982**, *30*, 1723.
108. Macdonald, D. D. In Proceedings of the Chimie 2002 (French Nuclear Society) Avignon: France, April 2002.
109. Szklarska-Smialowska, Z. *Pitting Corrosion of Metals*; National Assosiation of Corrosion Engineers: Houston, TX, 1986.
110. Engell, H. J.; Stolica, N. D. *Z. Phys. Chem., NF* **1959**, *20*, 113.
111. Macdonald, D. D.; Lui, C.; Urquidi-Macdonald, M.; Stickford, G. H.; Hindin, B.; Agrawal, A. K.; Krist, K. *Corrosion* **1994**, *50*, 761.
112. Engelhardt, G. R.; Macdonald, D. D. *Corros. Sci.* **2004**, *46*, 1159–1187.
113. Wei, R. P. In *Fracture Mechanics: Microstructure and Micromechanics*; Nair, S. V., *et al.* Ed.; ASM International: Metals Park: OH, 1989.
114. Turnbull, A. *Corros. Sci.* **1993**, *34*, 921–960.
115. Unger, D. J. *Analytical Fracture Mechanics*; Academic Press: San Diego, 1995.
116. Peng, O. J.; Kvon, J.; Shoji, T. *J. Nucl. Mater.* **2004**, *324*, 52–61.
117. Turnbull, A.; Zhou, S. *Corros. Sci.* **2004**, *46*, 1239–1264.
118. Turnbull, A.; McCartney, L. N.; Zhou, S. *Corros. Sci.* **2006**, *48*, 2084–2105.
119. Anderko, A.; Sridhar, N.; Dunn, D. S. *Corros. Sci.* **2004**, *46*, 1583–1612.
120. Kondo, Y. *Corrosion* **1989**, *45*, 7–11.
121. Marsh, G. P.; Bland, I. D.; Tailor, K. J. *Br. Corros. J.* **1988**, *23*, 157.
122. Strutt, J. E.; Nicholls, J. R.; Barbier, B. *Corros. Sci.* **1985**, *25*, 305.
123. Engelhardt, G. R.; Macdonald, D. D.; Zhang, Y.; Dooley, B. *PowerPlant Chem.* **2004**, *6*, 647.
124. Shibata, T.; Akoshi, M.; Ikematsi, K.; Nakajima, H.; Tsuge, H. *Corros Eng.* **1988**, *37*, 699–705.
125. Laycock, P. J.; Cottis, R. A.; Scarf, P. A. *J. Electrochem. Soc.* **1990**, *137*, 64–69.
126. Engelhardt, G. R.; Macdonald, D. D. In CORROSION/2008, New Orleans, April 2008; Paper no. 08270.
127. Kamaya, M.; Totsuka, N. *Corros. Sci.* **2002**, *44*, 2333.
128. Yeh, T. K.; Macdonald, D. D.; Motta, A. T. *Nucl. Sci. Eng.* **1995**, *121*, 468–482.
129. Tang, J. R.; Kao, L.; Shiau, D.-Y.; Chao, L.-Y.; Yao, C.-C.; Chiang, S. C. *Nucl. Technol.* **1998**, *121*, 324–336.
130. Engelhardt, G. R.; Macdonald, D.; Dooley, B. *PowerPlant Chem.* **2007**, *9*, 454–462.
131. Engelhardt, G. R.; Macdonald, D. D. Development of code to predict stress corrosion cracking and corrosion fatigue of low pressure turbine components, EPRI Report 1004190, 2004.
132. Engelhardt, G. R.; Macdonald, D. D.; Zhang, Y.; Dooley, B. *PowerPlant Chem.* **2004**, *6*, 647–662.