NOVA

University of Newcastle Research Online

nova.newcastle.edu.au

Melchers, Robert E. "Estimating uncertainty in maximum pit depth from limited observational data" Corrosion Engineering Science and Technology Vol. 45, Issue 3, p. 240-248 (2010)

Available from: <http://dx.doi.org/10.1179/147842209x12489567719581>

Accessed from: <http://hdl.handle.net/1959.13/933006>

Estimating uncertainty in maximum pit depth from limited observational data

Robert E Melchers*

Centre for Infrastructure Performance and Reliability, The University of Newcastle, Australia

* rob.melchers@newcastle.edu.au

Abstract

The expected maximum depth of corrosion pits in structural steel exposed to marine environments is important for estimating the reliability of engineered systems such as pipelines, tanks, ships and other containment structures. The conventional approach is to use observed data to determine the best fit extreme value (EV) distribution and to then estimate its parameters. In turn these may be used to estimate the maximum pit depth for other situations. However, in many cases there is insufficient data for this classical approach. Progress can be made by invoking a Bayesian statistics approach when a prior EV distribution can be invoked based on previous experience with similar data. The procedure is described herein for estimating the coefficients of variation for the maximum depth of pits on steel surfaces exposed to immersion, tidal, coastal atmospheric and marine inland atmospheric exposure conditions. For these situations previous experience for longer-term exposures has shown that Frechet is the most appropriate *prior* distribution.

Keywords: Pitting, extreme value analysis, data, uncertainty, Bayesian.

Introduction

The estimation of the probability of occurrence of a corrosion pit having a given, extreme, depth is of considerable engineering interest. Perforation of containment systems such as pipelines, tanks, ships and waste storage systems as a result of corrosion pitting can have significant environmental and economic costs. The corrosion processes involved in pitting are known not to be amenable to exact prediction and it has long been accepted that pitting can be modelled most appropriately as a stochastic process [1]. Since the parameters required for detailed corrosion pit development modelling are seldom

available, a somewhat simpler approach, based on using data, is to apply extreme value (EV) statistics for representing the uncertainty in pit depth. Indeed, the statistics of maximum pit depth is usually considered as one of the classical applications of extreme value theory [2, 3].

There are three classical theoretical extreme value probability distributions - Gumbel, Frechet and Weibull - although these can be transformed from one to the other by simple logarithmic transformations. The Gumbel has the longest history particularly for fitting to maximum depth pit data [4] and is still the most widely applied. The so-called Generalized Extreme Value distribution, also known as the von Mises-Jenkinson distribution, includes the classical distributions as special cases [5].

Application of extreme value theory requires a data set composed of a reasonable number of observations of the depths of extremal pits. Conventionally, that data set is then used to determine the extreme value probability distribution that best fits the data, the 'best fit' being judged by some criterion involving the fitted parameters for that distribution [2]. This is a purely empirical approach. It takes no account of the fact that each of the classical extreme value distributions has a different mathematical affinity (the so-called 'domain of attraction') with particular underlying distributions (that is, the probability distributions for all pits, not just the extremes). The domain of attraction suggests which extreme value distribution is likely to be the most appropriate for a given underlying distribution [2]. The concept of the domain of attraction has proved to be very useful in engineering applications [6] and also has been applied recently for modelling pitting statistics under long-term exposures [7]. There is obviously no unique domain of attraction for the Generalized Extreme Value distribution. It may be considered simply to be a distribution with greater flexibility for fitting to data.

The data required for extreme value analysis of pit depths ideally is obtained by measuring the depth of the deepest pit on a pitted surface, such as a corrosion coupon, and doing this for a sufficiently large number of surfaces of similar size. (The analysis can be extended to use, say, the 5 deepest pits rather than just the deepest – this refinement is not relevant in what follows – but see [5, 8].) Conventionally, the maximum depth of pit on each surface is assumed statistically independent of all other maximum depths, even when both the material and the exposure environment are similar. Obviously, a statistically adequate sample size requires a sufficient number of individual surfaces, each of sufficient size. Historically, only a few exposure test programs have been designed specifically with this in mind [9]. More commonly only two or three coupons (4 or 6 faces) are recovered at any one point in the

testing program. This provides 4 or 6 independent data sets, insufficient for conventional extreme value analysis. One classical approach to circumvent the problem is to combine data from different exposure periods [1, 4]. As will be described below, this may not be valid. The second approach is to divide a larger surface area into regular, smaller (sub-)areas. Each of these is then assumed to be an independent sample for pit depth. Evidently, this subdivision can be done only at the time the data is analysed.

An important and comprehensive collection of pit depth data is that for marine and freshwater corrosion in the Panama Canal Zone [10, 11, 12]. The data covers a period of 16 years and a wide variety of metals (and non-metals). For each exposure period there were just two coupons (four faces) recovered. For each face the deepest pit and the 5 deepest pits were observed. However, these data are not reported – instead only the maximum pit depth and the average depth of the 20 deepest pits are given for each recovery period. This does not provide sufficient data at each time point for a conventional extreme value analysis. The original observations appear no longer to be available.

Although the reported PCZ pit depth data is insufficient for a conventional extreme value analysis, it can be used if a Bayesian approach is used together with an appropriate *prior* extreme value distribution. This approach is described below. Although the analysis described herein is confined to carbon steel coupons exposed in the immersion, mid-tide, coastal and marine inland zones, it will be readily evident that the same approach can be used more widely.

Background

The data base for corrosion loss and pit depth under marine and fresh-water exposures in the Panama Canal Zone (PCZ) was the outcome of a very extensive and quite detailed experimental field study. It was designed to provide better data for corrosion under tropical conditions as a result of the very high rates of corrosion and deterioration experienced by military equipment in the Pacific region [13]. The PCZ data generally cover a period of 16 years and mostly are for duplicate coupons each 229 x 229 mm (9" x 9") in size and 6 mm (1/4") thick. The coupons were left undisturbed until recovered after 1, 2, 4, 8 and 16 years exposure. Both mass-loss and pit depth were measured, using two (duplicate) coupons recovered at each time point. The depths of the 5 deepest pits were recorded for each coupon face. The results were reported as (i) the deepest pit depth and (ii) the average of the 20 deepest pits.

Allowance was made for thickness loss due to general corrosion of the surfaces [10]. Evidently, this provides just two data points, insufficient for conventional extreme value analysis.

One conventional approach to the problem of insufficient data points (i.e. pit depths) is combine data from different exposure periods with the data from different periods of exposure brought back to one reference period at which the extreme value distribution parameters are estimated. The pit depth growth function usually employed has the form

$$d(t) = At^B \quad (1)$$

where t is time since first exposure, $d(t)$ is the pit depth at time t and A and B are constants estimated from the data. Eqn (1) is widely accepted as a reasonable approximation for the growth in pit depth, particularly over relatively short periods of exposure. However, for longer-term exposures the growth in maximum pit depth in marine environments has been shown to be rather more complex [14, 15]. Figure 1a-d give some examples of data reported for pit depths measured on coupons of carbon steel exposed to various environments in the Panama Canal Zone [10,11,12]. Figure 2 shows a schematic model of the behaviour observed in Figure 1 together with interpretations of the rate controlling processes as first given for immersion corrosion [16] but subsequently interpreted more generally [17].

Figure 1 Maximum and average of 20 deepest pit depths as a function of exposure period for coupons exposed to (a) marine immersion, (b) mean tide, (c) marine coastal and (d) inland marine environments in the Panama Canal Zone.

Figure 2 Idealized model for growth in maximum pit depth as a function of exposure period, with brief summaries of the rate controlling mechanism for each phase.

Figures 1a-d show immediately that the conventional pit depth growth model (1) is not applicable for describing longer-term pitting behaviour. It follows that it is inappropriate, in general, for direct use in merging data obtained for different exposure periods to provide a sufficient number of points for extreme value analysis. Note that this does not preclude the use of (1) as a local non-linear model, properly calibrated to local changes in pitting behaviour, for combining data obtained at neighbouring exposure times. Valid use of such a local model for data combination does, however, require detailed

knowledge of the characteristics of the model of Figure 2 to ensure that data blending occurs only within one or other of the phases shown in Figure 2.

The second approach for obtaining a sufficient number of samples for extreme value analysis is to divide the area of exposed surface into smaller areas that are then considered to be independent of each other. Conventionally a simple rectangular or square grid is used to divide the exposed area into smaller, equal, sub-areas. This is convenient since it saves having to account for differences in area relative to pit size. For each sub-area the maximum pit depth is measured and EV statistics applied to these sample pit depths [18]. It has been conventional (and perhaps just convenient) for sub-areas to be rectangular or square. However, in principle this need not be so. They may be of any shape - the only provisos being (i) that they can be considered to provide independent samples (see below) and (ii) that they are of similar surface area. The possibility of having arbitrarily shaped sample areas follows directly from the accepted theory that in a multiple random sampling scheme involving a number of random variables, the sampling for one of the variables can be taken as regular, as a sampling producing a regular, equally-spaced outcome is a special but plausible outcome of a random sampling [19]. For corrosion pit sampling this possibility was recognized implicitly earlier, although not specifically pointed out [20]. This means that if there are 'q' large pits, say, on a sample surface of area 'a', it is permissible to associate each pit with a sub-area 'a/q' and use these as 'q' samples for the EV analysis. (The actual sub-areas need not be explicitly identified – they can be notional).

As noted, an important proviso is that the individual surface (sample) areas must be such that the deepest pits can be considered independent events, or, more generally, that the sample surfaces may be considered, in the limit of large numbers, as independent samples. This is the so-called 'asymptotic' assumption [2]. It also requires that pit clusters be represented only by the deepest pit in the cluster (cf. clustering of data more generally in extreme value theory and in stochastic process theory [21]).

In previous analyses of pit depth for mild steel in seawater it was found that the Frechet distribution was consistent with the presumed underlying pit-generating processes for longer-term exposures, that is, in phases 3 and 4 of the model of Figure 2 [22]. This suggests that it is an appropriate *prior* distribution for modelling the statistics of the longer-term pitting for tidal, coastal and inland exposures. This distribution will be used in the following.

Application of Frechet EV Distribution

Frechet Distribution

The Frechet distribution for a random variable Y is given by

$$P(Y < y) = F_Y(y) = \exp[-(u/y)^k] \quad (2)$$

where u and k are parameters of the distribution, defined implicitly through the mean μ_Y and variance σ_Y^2 defined by

$$E(Y) = \mu_Y = u\Gamma(1 - 1/k) \quad k > 1 \quad (3)$$

$$Var(Y) = \sigma_Y^2 = u^2[\Gamma(1 - 2/k) - \Gamma^2(1 - 1/k)] \quad k > 2 \quad (4)$$

where $\Gamma(\cdot)$ is the Gamma function. In conventional applications a set of data can be fitted to the distribution in a number of ways [2]. One is to rank the data as an ‘ordered’ set, with each observation assigned an exceedence probability, defined in a number of possible ways but most commonly with the mean ranks approach [5] as $i/(n+1)$ where i is the number of the sample in the rank (with 1 being the shallowest pit and n the deepest pit in the rank). The assigned exceedence probability associated with each maximum pit depth can then be used to plot each data point on a Frechet plot. On such a plot, data that is truly Frechet distributed will plot as a straight line.

For the PCZ data this classical approach cannot be used directly, since any one exposure period provides just one extreme pit depth and just the mean value of the 20 deepest pits observed. (The individual values of the 20 deepest pits are not available in the reported data). This means there are only two points available for plotting on a Frechet plot. In a conventional extreme value statistics approach this would not be sufficient to determine whether the data was Frechet distributed. However, invoking a Bayesian approach permits the Frechet distribution to be the assumed *prior*, and then the two points are sufficient to define the distribution, i.e. to draw a straight line between the two points. The location and the slope of that line can then be used to estimate the distribution parameters. This approach will now be considered for the available PCZ data.

Use of the mean of 20 data points and the one extreme value

The PCZ data can be used to estimate the parameters of the Frechet distribution and the coefficient of variation by using the mean μ_Y of the $n = 20$ observations of maximum pit depth y as one point and the maximum extreme pit depth $y_{\max}$ as the other, together with the prior Frechet distribution assumption that allows these points to be connected with a straight line. However, the plotting process can be obviated by using the available data and the prior distribution to directly estimate the coefficient of variation $V_Y = \sigma_Y / \mu_Y$.

The exceedence probability associated with the one extreme value $y_{\max}$ can be estimated from order statistics. Using the mean ranks approach [5] it follows that it is $P(Y < y_{\max}) = 20/21$. Substituting into (2) produces

$$u = y_{\max} (0.04879)^\alpha \text{ with } \alpha = 1/k \quad (5)$$

Substituting (5) into (3) yields

$$\mu_Y = y_{\max} (0.04879)^\alpha \Gamma(1 - \alpha) \quad (6)$$

Figure 3 shows a plot of $\mu_Y / y_{\max}$ as a function of $\alpha = 1/k$ for 20 measurements. This allows α to be estimated as a function of $\mu_Y / y_{\max}$, i.e. from the known values of μ_Y and $y_{\max}$ at each exposure time t . Similarly, substituting (5) into (4) produces

$$\sigma_Y = y_{\max} (0.04879)^\alpha \left[\Gamma(1 - 2\alpha) - \Gamma^2(1 - \alpha) \right]^{1/2} \quad (7)$$

which is plotted in Figure 4 for 20 samples. This allows the ratio $\sigma_Y / y_{\max}$ to be estimated for the values of α at each time point. Further, dividing (7) by (6) yields the coefficient of variation for pit depth $V_Y = \sigma_Y / \mu_Y$.

Figure 3 Plot of equation (6) for ratio of mean pit depth to measured maximum pit depth ($\mu_Y / y_{\max}$) as a function of the Frechet parameter $\alpha = 1/k$ for 20 and for 60 pit depth measurements.

Figure 4 Plot of equation (7) for the ratio of pit depth standard deviation to maximum measured pit depth $\sigma_Y / y_{\max}$ as a function of the Frechet parameter $\alpha = 1/k$ for 20 and for 60 pit depth measurements.

Data for the mean and the maximum pit depths for immersion, mid-tide, coastal marine atmospheric and inland marine atmospheric exposures of carbon steel in the Panama Canal zone, as reported in the literature, are shown in Table 1. Also shown are the values estimated using the above approach for α and hence for V_Y . Figure 5 shows the calculated values of V_Y as a function of exposure period.

Evidently, the calculated values show very considerable variation. Only poor quality linear trends can be fitted. This is likely due to the relatively small sample size and to the reliance on just one data point to represent the upper tail of the extreme value probability distribution. Evidently, with a small sample size there will be a high degree of uncertainty attached to the probability to be associated with any one observed extremal depth pit.

Table 1 Pit depth data and parameters $\mu_Y / y_{\max}$, $\alpha = 1/k$, $\sigma_Y / y_{\max}$ and derived Coefficient of Variation.

Figure 5 Coefficient of Variation of the measured pit depths as a function of exposure period for each exposure condition, based on 20 samples.

Use of the mean of 60 data points and one extreme value

To throw some light on the effect of sample size, the analysis was repeated for a sample set consisting of 60 measurements, obtained by combining the reported mean and extreme pit depth data for carbon steel for three different surface treatments, each obtained from 20 pit depth measurements: (i) acid-cleaned coupons, as for Table 1 and Figure 5, (ii) coupons with mill-scale not removed prior to exposure, and (iii) coupons with machined surfaces. While surface finish is known to affect corrosion loss [23], and presumably therefore pit depth, it is known that this is most clearly seen only for relatively short-term exposures and to become insignificant for longer exposure periods. The extreme

pit depth was taken as the deepest of the three reported deepest pits, one for each surface finish. The exceedence probability associated with that maximum pit depth is now $P(Y < y_{\max}) = 60/61$. Equation (5) now becomes

$$u = y_{\max} (0.01653)^\alpha \quad (8)$$

and the subsequent expressions are modified accordingly. Curves relating $\mu_Y / y_{\max}$ and $\sigma_Y / y_{\max}$ as functions of $\alpha = 1/k$ for 60 measurements are shown in Figures 3 and 4. The resulting analysis for the coefficient of variation $V_Y = \sigma_Y / \mu_Y$ for pit depth yields the results shown in Figure 6.

Figure 6 Coefficient of Variation of the measured pit depths as a function of exposure period for each exposure condition, based on 60 samples.

It is evident that there is now much less variability in the data and therefore greater consistency in the trend lines between the different exposure environments.

Use of the mean of 60 data points and three extreme values

The above analysis can be refined a little further by assuming that the three maxima from each data set of 20 observations are in fact the three deepest pits in the set of 60 observations. This may not be correct if there are significant differences in the distribution of pit depths between the three data sets, but it may be a reasonable assumption if the data sets are comparable, as is the case here.

It is illustrative to consider the data as they would appear on a Frechet plot. This requires analysis for the exceedence probability to be attached to each mean value. Equivalently, use can be made of the well-known relationship between the Frechet distribution and the Gumbel distribution, analogous to that between the Normal and the Lognormal distributions. Thus if Y is Frechet distributed with parameters u and k , $\ln(Y)$ is Gumbel distributed with parameters $\ln(u)$ and a . This enables the Frechet distributed random variable to be plotted on a standard Gumbel plot, that is one that has the reduced variate w , defined as $W = (Y - u)/a$, with W having $u = 0$ and $a = 1$ provided the variate Y is plotted on a log-scale. On this plot, data following the Frechet distribution lies on a straight line. The slope of the line is a measure of the uncertainty (standard deviation) of the data.

Figure 7 shows the Frechet plots with the three deepest pit depths and the relevant mean rank probabilities ($p_i = i/(n + 1)$, $i = 58, 59, 60$, $n = 60$) as well as the mean value plotted at its estimated exceedence probability. In each case the left axis is the Gumbel parameter w . For each value of $\alpha = 1/k$ expression (8) can be used to estimate u for the relevant $y_{\max}$ at each exposure time t . Equation (2) may then be used to estimate $P(Y < \mu_y)$, the exceedence probability for the mean value at that time t . This has been done for each exposure case and for the time periods when phases 3 and 4 of the model (Figure 2) are likely to be applicable to the data. Thus, for immersion and tidal exposures the Frechet data and trend lines for years 2-16 are shown, while for the atmospheric conditions only years 4-16 are shown since in these two cases the Frechet model is unlikely to be applicable prior to year 4.

Figure 7 Frechet plots showing the three extremal pit depths and the mean pit depth based on 60 pit depth samples. The grey rounded rectangle highlights the region where the mean lies for each exposure period for coupons exposed to (a) marine immersion, (b) mean tide, (c) marine coastal and (d) inland marine environments in the Panama Canal Zone. *Note w on the left axis is the standard Gumbel/Frechet parameter corresponding to the probabilities shown at right.*

Evidently, the three data points, each representing an extremal pit depth, do not fall neatly on a single line drawn from the mean. This indicates immediately the sensitivity in uncertainty estimation resulting from the use of just one extreme value to estimate the Frechet line. Since the exceedence probability can be underestimated by sub-maximal pitting, it is appropriate to use that pit depth which produces the flattest slope of the Frechet line – i.e. the slope corresponding to the most extreme pit depth observed. This has been done in Figure 7. As noted, this slope is related directly to the uncertainty (standard deviation, variance or coefficient of variation) in the pit depth distribution [2]. The fact that the three points do not necessarily lie on the same Frechet line is one of the reasons for the variability seen in Figures 5 and 6 since for these only one points was used.

Using the median location

As noted above, the published data includes the sample mean μ_y (for 20 or 60 samples). It cannot be plotted immediately on a Frechet plot as the associated probability is not known *a priori* and must be estimated as described above from the data for the mean pit depth and the one reported value of the

maximum pit depth. This means that the estimated probability is highly dependent on the measured depth of the maximum depth pit and on the rank order probability assigned to it. An approximate alternative, not requiring any analysis, is to plot the mean μ_Y at the known probability of the median y_m (i.e. at $P(Y < \mu_Y) \approx P(Y < y_m) = 0.5$). In Figure 7 this means that the points within the shaded rounded rectangle are all moved downward to a probability of 0.5. Evidently, this involves small changes in slopes of the Frechet lines. Importantly it allows, without analysis, immediate extrapolation from the data base so as to permit a simple graphical estimate of the probability of occurrence of deeper pits.

Since the slope of the trend lines is a measure of the data uncertainty (such as measured by the coefficient of variation), it is of interest to ascertain the error involved in the above approximation. Let the approximate coefficient of variation based on the median be denoted $\text{COV}|_{y_m} = \sigma_Y / y_m$. From (2) it follows that for $P(Y < y_m) = 0.5$ the median is given by $y_m = u / (0.693)^\alpha$ and using (3) for the mean, the ratio between mean and median is given by

$$\frac{\mu_Y}{y_m} = (0.693)^\alpha \Gamma(1 - \alpha) \quad (9)$$

This is plotted in Figure 8 as a function of α . It follows immediately that although the slope of the Frechet lines, and hence the standard deviation, is reduced by the approximation, the $\text{COV}|_{y_m}$ estimated using y_m instead of the mean is overestimated and should be divided by (9) to obtain a better estimate for the true COV. As in most cases α varies between about 0.1 and 0.3 it is clear from Figure 8 that the error in COV is between 5 and 15%. In view of the variability seen in Figures 5 and 6 this may be considered an acceptable error.

Figure 8 Correction factor (ratio mean pit depth to maximum pit depth) $\frac{\mu_Y}{y_m} = (0.693)^\alpha \Gamma(1 - \alpha)$ as a function of Frechet parameter α .

Discussion

As noted in the Introduction, the procedure proposed here is based on the *a priori* assumption of a particular extreme value distribution being appropriate to represent the statistics of the pitting phenomenon and then using the available data to make estimates of the parameters for that distribution. This is the opposite of the conventional notion of finding a distribution that somehow ‘best fits the data’. Here the selection of the Frechet distribution was based on the earlier observation that it both fits longer-term pit depth data and that it is in the domain of attraction of the underlying Pareto distribution, that in turn itself can be derived from plausible assumptions about the underlying corrosion mechanisms for pitting of mild steels in seawater environments [7, 22]. This is a logical process.

Obviously, the Frechet model is not necessarily the most appropriate prior distribution in all circumstances for long-term pitting, particularly where significantly different corrosion mechanisms are known to be critical. As in any Bayesian statistical application, the choice of the prior distribution must be based on understanding of the underlying processes or on previous, preferable extensive, experience. Nevertheless, the same general approach and procedure as described above would be applicable if it were judged, for example, that the Gumbel extreme value distribution was the most appropriate prior distribution.

Figure 6 shows that the point estimates of the coefficients of variation have considerably less scatter than those in Figure 5. The linear trends drawn through the estimates also are more consistent. These trends indicate the variation of the coefficient of variation with increased exposure period and also show that the uncertainty is generally least for the inland marine exposure site and most for the immersion regime, although, given the scatter in the points, there actually may be little to separate them.

The scatter in the data and in the fitted linear trends is much more for the limited (20 sample) data set shown in Figure 5 than it is for the data using 60 samples. Following through the above theoretical analysis shows immediately that a weak link is the probability assigned by order statistics to the extremal pit depth in any one data set. For just 20 samples this implies a probability that a pit depth greater than that of the observed extremal pit is about 0.05. However, for any one sample extremal pit this may not be a realistic assessment of associated probability, which could be much less or even rather higher. Evidently, any such error in assigning a probability to the extremal pit using this procedure is carried right through the analysis. That this is the likely cause is indicated by increasing

the size of the sample, somewhat artificially, by combining data for carbon steel with three different surface finishes, providing the outcome shown in Figure 6 for 60 data points. In this case the extremal pit acquires an estimated probability of being exceeded in depth around 0.016.

The analysis using the three separate extremal points of each of the 20-point data sets as the three deepest pits for the combined 60-point data set shows clearly the arbitrariness of assigning an exceedence probability to the extreme observation. As seen in Figure 7, by comparing the pit depths with the trend line from the mean, the extreme pit depths are not always consistent with each other and may be less than expected. These problems are obviated to an increasing extent by having available a larger number of maximum pit depths obtained from independent observations and for independent pitting events. As noted earlier, this is likely to be available only from experimental schema specifically designed to estimate pit depth uncertainty [14]. Conversely, where this is not the case, the present analysis shows what is possible with more limited data. The cost, of course, is a greater level of uncertainty in the developed Frechet trend lines and hence greater uncertainty in the estimated uncertainty measures (standard deviation, variance or coefficient of variation) for pit depth.

From a practical point of view the high coefficients of variation attached to the extreme pit depths, for each of the exposure environments could be of some concern. Perhaps surprisingly, the results are generally consistent with estimates made previously for carbon steel under immersion exposure conditions [9]. Comparisons cannot be made at the present time for tidal and atmospheric exposure conditions. For these there appear to be no previous estimates of uncertainty in pit depth. More data will become available in due course as a result of a controlled exposure program currently being concluded at three locations on the Eastern Australian seaboard.

The plots of the mean of the 20 deepest pits measured on the coupons exposed in the Panama Canal Zone show, for each of the four environments considered (immersion, tidal, coastal and inland), a distinctly non-uniform behaviour (Figure 1). Further, the plots of the maximum measured pit depth show similar behaviour as a function of time. Taken together, this suggests that the corrosion model (Figure 2) previously proposed for pit depth is applicable also to data for the tidal and coastal atmospheric zones, as illustrated in Figure 1. The curves in Figure 1 indicate that there is a period of active pit growth immediately after exposure, followed by a gradual decline in growth rate and then a period during which the average pit depth increases only very slightly. This is followed, after about a

year or more, with a period of renewed pit depth growth. This behaviour is observed for the maximum pit depth but also for the average of 20 pits, indicating that if there are changes in pit development they are not such as to cause an increase in pit depth but rather in lateral growth of pits, consistent with the observations for pit development reported earlier [24]. These earlier observations showed that pits appear to coalesce through lateral growth, developing apparent plateaus before renewed new pitting appears. Previously the second development of pitting has been associated with bacterial activity [22]. It is of interest to note that this temporal development of pit depth, both extremal and average, is seen for each of the four exposure environments.

Conclusion

The analysis given herein allows the limited data for average and maximum pit depth, such as reported for the 16 year exposures of carbon steel in the tropical marine environments in the Panama Canal Zone (PCZ), to be employed to estimate coefficients of variation for extreme value distributions. The analysis is focussed on longer-term exposures and, based on previous observations, applies the Frechet extreme value distribution as appropriate for the statistics of extreme pit depth. However, the same procedure can be applied for other distributions. Immersion, tidal, coastal atmospheric and marine inland atmospheric exposure conditions were considered. Although the estimates of some coefficients of variation showed considerable variability, comparable estimates have not previously been available. The analysis can be applied also to provide extreme value probabilistic interpretations for other PCZ pit depth data.

Acknowledgements

The work reported herein is part of a project on the structural reliability of deteriorating structures. The support of the Australian Research Council is acknowledged.

References

- 1 Shibata T (2000) Corrosion probability and statistical evaluation of corrosion data, Chapter 22 (In) Uhlig's Corrosion Handbook (Ed) RW Revie, New York: John Wiley & Sons, 367-392.
- 2 Galambos J (1987) The asymptotic theory of extreme order statistics, Second Edition, Krieger, Malabar, FL.

- 3 Castillo E (1994) Extremes in engineering applications (in) Extreme value theory and applications (Ed) J Galambos, J Lechner, E Simiu and N Macri, Dordrecht: Kluwer Academic Publishers, 15-42.
- 4 Aziz PM (1956) Application of the statistical theory of extreme values to the analysis of maximum pit depth data for aluminum, Corrosion (NACE) 12 (10)495t-506t.
- 5 Kotz, A and Nadarajah, S (2000) Extreme Value Distributions - Theory and Applications, Imperial College Press, London, 59.
- 6 Castillo E, Galambos J and Sarabia JM (1989) The selection of the domain of attraction of an extreme value distribution from a set of data, Lecture Notes in Statistics, 51: 181-190.
- 7 Melchers RE (2008) Extreme value statistics and long-term marine pitting corrosion of steel, Probabilistic Engineering Mechanics 23: 482-488.
- 8 Scarfe PA, Cottis RA and Laycock PJ (1992) Extrapolation of extreme pit depths in space and time using the r deepest pit depths, J. Electrochem. Soc. 139(9) 2621-2627.
- 9 Melchers RE (2004) Pitting corrosion of mild steel in marine immersion environment - 2: variability of maximum pit depth, Corrosion (NACE) 60(10) 937-944.
- 10 Southwell CR, Forgeson BW and Alexander AL (1958) Corrosion of metals in tropical environments – Part 2 – Atmospheric corrosion of ten structural steels, Corrosion (NACE) 14(9) 53-59.
- 11 Forgeson BW, Southwell CR and Alexander AL (1960) Corrosion of metals in tropical environments, Part 3 – Underwater corrosion of ten structural steels, Corrosion (NACE) 16(3) 87-96.
- 12 Southwell CR, Bultman JD and Alexander AL (1976) Corrosion of metals in tropical environments – Final report of 16-year exposures, Materials Performance, 15(7) 9-26.
- 13 Southwell, CR, Bultman, JD and Hummer, CW (1979) Estimating service life of steel in seawater (in) Schumacher, M (Ed) Seawater Corrosion Handbook, Park Ridge, New Jersey; Noyes Data Corporation, 374-387.
- 14 Melchers RE (2004) Pitting corrosion of mild steel in marine immersion environment - 1: maximum pit depth, Corrosion (NACE) 60(9) 824-836.
- 15 Melchers RE (2009) Extreme value statistics for pitting corrosion of steel infrastructure in marine environments, ICOSSAR09
- 16 Melchers RE (2003) Modeling of marine immersion corrosion for mild and low alloy steels - Part 1: Phenomenological model, Corrosion (NACE) 59(4) 319-334.
- 17 Melchers RE (2007) The transition from marine immersion to coastal atmospheric corrosion for structural steels, Corrosion (NACE) 63(6) 500-514.
- 18 Provan JW and Rodriguez ES (1989) Development of a Markov description of pitting corrosion, Corrosion (NACE) 45 (3)178-192.
- 19 Kahn H (1956) Use of different Monte Carlo sampling techniques, (In) Proceedings Symposium on Monte Carlo Methods, (Ed) HA Mayer, John Wiley & Sons, New York, 149-190.
- 20 Scarf PA and Laycock PJ (1996) Estimation of extremes in corrosion engineering, Journal of Applied Statistics, 23(6) 621-643.
- 21 Leadbetter MR, Lindgren G and Rootzen H (1983) Extremes and Related Properties of Random Sequences and Processes, Springer, New York.
- 22 Melchers RE (2006) Pitting corrosion of mild steel under marine anaerobic conditions – Part 2: Statistical representation of maximum pit depth, Corrosion (NACE) 62(12) 1074-1081.
- 23 Melchers RE and Jeffrey R (2004) Surface roughness effect on marine immersion corrosion of mild steel, Corrosion (NACE) 60(7) 697-703.
- 24 Jeffrey R and Melchers RE (2007) The changing topography of corroding mild steel surfaces in seawater, Corrosion Science 49: 2270-2288.

Table 1

Exposure Period (years)	Pit depth (aver 20 deepest) μ_Y mm	Pit depth (max) $y_{\max}$ mm	$\mu_Y / y_{\max}$	α (from Fig 3)	$\sigma_Y / y_{\max}$ (from Fig 4)	Coefficient of Variation $V_Y = \sigma_Y / \mu_Y$
Immersion						
1	1.04	1.58	0.67	0.17	0.19	0.28
2	1.19	1.55	0.77	0.16	0.16	0.21
4	1.67	2.87	0.58	0.25	0.25	0.43
8	1.67	2.18	0.76	0.16	0.16	0.21
16	2.29	3.94	0.58	0.25	0.25	0.43
Mid-tide						
1	0.46	0.69	0.67	0.17	0.18	0.27
2	0.41	0.58	0.70	0.16	0.16	0.14
4	0.68	1.40	0.49	0.32	0.32	0.65
8	1.02	1.65	0.62	0.22	0.22	0.35
16	1.68	2.49	0.67	0.17	0.19	0.28
Coastal						
1	0.33	0.51	0.65	0.18	0.18	0.28
2	0.51	0.66	0.77	0.11	0.12	0.16
4	0.46	1.27	0.36	0.60		
8	0.86	1.70	0.51	0.31	0.31	0.61
16	1.29	3.12	0.42	0.46	0.74	1.7
Inland						
1						
2	0.23	0.48	0.47	0.37	0.40	0.84
4	0.25	0.38	0.67	0.17	0.19	0.28
8	0.43	0.69	0.63	0.20	0.20	0.32
16	0.56	0.84	0.67	0.17	0.19	0.28

PCZ figures

(a) plot10

(b) plot 11

(c) plot12cor

(d) plot13

Figure 1a-d

Figure 2 model

Figure 3 plot69

Figure 4 plot62

Figure 5 plot 50XX

Figure 6 Plot49XX

(a) FrechImmer60YYY

(b) FrechTidal60YYY

(c) FrechCoast60YYY

(d) FrechInland60YYY

Figure 7 a-d

Fig 8 Plot54CorrectFact