Extrapolation of Extreme Pit Depths in Space and Time

P. J. Laycock, R. A. Cottis, and P. A. Scarf

Department of Mathematics, University of Manchester Institute of Science and Technology,
Manchester, England M60 1QD

ABSTRACT

A four parameter model is proposed for data collected on maximum pit depths enabling simultaneous extrapolation into the future and over large areas of exposed metal. This model is based on the generalized extreme value distribution whose use in this context is here justified mainly on statistical, rather than metallurgical, reasoning. Those aspects of the model which allow for extrapolation in time rely on reported power law dependencies for mean pit depths. Use of the model for predicting means, standard deviations, percentiles, bounds, order statistics, hole counts, and size of perforated areas, is demonstrated, both in general and for a particular data set. Comparisons are made with other reported techniques.

This paper is concerned with modeling the statistical distribution of maximum pit depths, and relies on results from experimental work and field observations directed toward such measurements, rather than work directed toward the underlying metallurgical theory of individual pit initiation and propagation.

The measured variate z is here defined to be the depth of the deepest pit in an exposed area of metal a after an exposure time t, and we are interested in extrapolating such values out to a larger area A and for a longer exposure time T. The number of potential sites for pitting is typically very large. Theoretical considerations suggest a value of at least 100 cm² for mild steel, while experimental data presented by Bhakta and Solomon (1) show pit densities for low carbon steel in the range 100-350 cm². Hence, the measured pits can often be regarded as the largest of a very large number of pits, even if this fact is not noted, nor further measurements made. This background leads naturally to the statistical theory of extreme value probability distributions, as first presented in an engineering context by Gumbel (2,3). For a typical example on the use of these techniques for prediction of maximum pits, see Hawn (4).

Extreme Value Distributions

Suppose a random variable Z has distribution function \( H(z) \), so that \( P[Z \leq z] = H(z) \), and suppose X is the maximum of a random sample of n observations on Z. Then \( P[X \leq x] = H(x)^n \), which will typically have a complex form taken from a different family of distributions than H. However, if H is one of the three classic extreme value distributions, conventionally known as Types I, II, and III, then we can (uniquely) write \( H(x) = H(a_n + b_n x) \), for some known constants \( a_n \) and \( b_n \). Apart from the simplicity of this particular relationship when taking successive extremes, a practical justification for using extreme value distributions is that this linear form can be shown to hold asymptotically (i.e., for large n) for essentially all the commonly used probability distributions. In particular, extremes from the upper tails of normal, exponential, and Weibull distributions are all asymptotically Type I, which has the (standardized) distribution function

\[
H(y) = e^{-e^{-y}} \quad -\infty < y < \infty
\]

and is often called the extreme value distribution. This distribution has been widely used in corrosion engineering. The Type II, or three-parameter Cauchy distribution, has a lower bound for maxima, while the Type III, or three-parameter reversed Weibull distribution, has an upper bound for maxima. Both the Type II and Type III are easily mapped into the Type I distribution via simple transformations. But these transformation functions depend crucially on knowledge of the appropriate bound, which bound is effectively the third, extra, parameter for these distributions. Hence, the existence of these transformations is of little practical importance in any situation where this bound is unknown and must be estimated from data. However, the generalized extreme value (GEV) distribution as introduced by Jenkinson (6) subsumes all three types into one formula, with the sign of a shape parameter, \( k \), indexing the types (zero for Type I, negative for Type II, and positive for Type III). In the absence of any a priori reason for selecting a particular sign for \( k \) it would seem sensible to work, at least initially, with the GEV distribution when using extreme value statistics on a given body of measured maxima (or equivalently, minima). In its most convenient form for this paper, its cumulative distribution function can be expressed as

\[
F(x) = \exp\left[-\frac{1 - k(x - u)/\alpha}{\alpha^k}\right] \quad kx = u + \alpha k
\]

where \( k \) is a shape parameter, \( u \) is a location parameter, and \( \alpha \) is the scale parameter. We write GEV(\( u, \alpha, k \)) for the distribution [2]. A representative selection of plots for this distribution can be found in Fig. 1, 2, and 3. An important observation from Fig. 1 is that for small positive values of \( k \), the finite upper bound to the distribution may well not be
visually evident in the data, although its presence may be crucial for the purposes of extrapolation. It is also evident from these plots that the GEV distribution offers a unimodal but otherwise very flexible family of distributions, via its scale location and shape parameters, for general bodies of data, not necessarily arising directly from a process of maximization. Hence, in the absence of external reasons for selecting an alternative distribution, it may be difficult in practice to refute the assumption of a GEV distribution for pitting data at any level of the maximization process.

**Time Dependence**

The dependence of pit depth on time is usually reported as a simple growth in mean pit depth, typically as a power of $t$

$$\bar{x} = at^b$$  \[3\]

Values of $b$ ranging from 0.33 through 0.5 (common) or 0.6 have been reported. The classic data reference here is Ramonoff (6). The value $b = 1$ is sometimes used by default, in particular when reference is made to an (implicitly constant) "pitting rate" of, say, "1 mm per year." Other values of $b$ give a nonconstant monotone decreasing pitting rate. There is a mechanistic justification for the value $b = 0.5$. Finley (7) suggests a logarithmic dependence on time, justified by reference to the aluminum pitting data in Aziz (8). Note that both these papers use Type I extreme value distributions for which $k = 0$. The parameterization

$$\text{GEV}(u_0, \alpha_k, k)$$  \[4\]

with $u_t = u_0 + \alpha_k t$ will lead to [2] having a population mean $\mu_t$ of the requisite form implied by [3], namely

$$\mu_t = \left[ \xi - \frac{\alpha}{k} \Gamma(1+k) \right] t^b, \quad k > -1$$  \[5\]

where $\xi = \xi_t = u_0 + \alpha_k k$ is the moving (upper or lower, depending on the sign of $k$) bound on pit depths implied by [2], and $\Gamma()$ is the (standard) gamma function. Cottis et al. (9) have found experimental evidence for $k > 0$ with pitting in austenitic stainless steel. Also, Eldredge [Ref. (10), p. 75] observes that a "Type I distribution is not exactly followed" for oil welb tubing pitting corrosion, while the deviations he remarks on in his figures "10" and "11" are precisely those expected for data following a Type III ($k > 0$) distribution.

Our assumption of a constant value for the "shape" parameter $k$ is firstly driven by statistical necessity, since this is an inherent property of the GEV distribution as applied to maxima. However, when we have fitted the three-parameter distribution [2] at separate time-points we have found that while $k$ is not always constant for small values of $t$ (i.e., time measured in hours or days rather than weeks or years), it soon rises to a constant value. And since this paper is directed toward large-scale extrapolation we have felt the assumption of constant $k$ to be a reasonable compromise.

The parameterization [4] implies a standard deviation for extreme pit depths given by

$$\sigma_t = \left( \frac{\alpha_k}{k} \left[ \Gamma(1 + 2k) - \Gamma(1 + k)^2 \right] \right)^{1/2}$$  \[6\]

Note that this implies a common growth rate parameter, $b$, for the mean and the standard deviation. This is consistent with the "good correlation" between means and standard deviations found by Masamura and Matsushima (11) for pit depths in carbon steel pipes from a variety of industrial environments and for time periods ranging from 3 to 20 years. Their figure "7" which underlies the above conclusion (quoted from the English abstract of their paper) is a plot of mean pit depths against standard deviation. If mean and standard deviation follow a common power law within any one environment, then their ratio will be a constant independent of time, and if this ratio is independent of the environment, then a plot of mean against standard deviation for a variety of times and environments would produce a straight line, as in their figure "7." Conversely, plotting means or standard deviations separately against time for a variety of environments could produce an indecisive result if the (power law) rate parameter depends on the environment. This could explain the appearance of their figures "5" and "6," which support their conclusion, for their data, that "the law of pit growth is unclear."

**Area Dependence**

To extrapolate the statistical behavior of the leading (or extreme) pits, over time, we need some assessment of $N_0$, the variation of pit count with time. In the Stahl and Miller (12) report, a Weibull distribution with uniformly decreasing pit generation rate is used, implying

$$N_t(A) = \lambda(A)(1 - e^{-\lambda A})$$  \[7\]
in the mean, where \( N(A) \) is the number of pits in an area \( A \) after time \( t \), and \( \lambda \) is the final pit density per unit area. Hence \( N(A) = A \lambda \).

Assuming each pit counted by \( t \) has distribution function \( F(x) \), for its depth \( x \), it follows that the leading pit over the extrapolation area \( A \), \( A_{\text{max}} \) say, has cumulative distribution function \( F(x)^A \), which if \( F \) is given by [2], produces another GEV distribution, namely

\[
\text{GEV}(ut^b + \frac{\alpha t^b}{k} (1 - N(A)^{-b})^{-1}, \alpha t^b N(A)^{-b}, k) = \text{GEV}(uA^b, \alpha A^b, k) \quad [8]
\]

The corresponding implied distribution for the maxima over experimental coupons with areas \( A \) can then be written as \( \text{GEV}(uA^b, \alpha A^b, k) \), using the same notation. We therefore have

\[
a_k = \alpha N(A)^{-b} = \alpha M^k
\]

where \( M = A/a \), from [7], and hence \( \alpha_k = \alpha M^{-b} \). Also observe that \( \xi = \xi_k = u_k + \alpha_k/k = \xi_k = u_k + \alpha/k \), and hence \( u_k = \xi - \alpha M^{-b}k/k \). So that knowledge of the GEV parameters for experimental coupons of size \( A \) leads directly to the corresponding parameters for extrapolation out to any multiple \( M \) of \( A \). Note that

\[
a_k = \epsilon(a\alpha(1 - e^{-\alpha})^{-1} - \alpha(1 - \alpha)^{-1}) \quad \text{as } t \to \infty
\]

Insofar as we are here relying on the asymptotic convergence of an underlying arbitrary distribution for pit depth maxima to the GEV form, it will only be important for this to be a good approximation when the maxima are being extracted from coupons of size \( A \) or more. And, as implied above, for a typical mild steel coupon of 100 cm² in a pitting environment, there may be a nominal count of 10⁴, or more, pits from which to extract this maximum.

There is experimental evidence that once pitting has commenced \( N \), it will reach its limiting value \( N \) fairly rapidly. See for example Bhakta and Solomons (1) or Azzan (8) where \( \text{~20-30 days} \) appears to be the relevant time constant. The mean of this extrapolation distribution can then be written unequivocally as

\[
\mu_{\text{max}} = \frac{\xi^b - \frac{\alpha M^{-b}k}{k} \Gamma(1 + k)}{k} \quad [9]
\]

For fixed \( t \) and \( k \neq 0 \), this implies a power law relationship on area, unbounded when \( k < 0 \) and bounded above when \( k > 0 \); while for \( k = 0 \), the limit of [9] is

\[
\mu_{\text{max}} = \frac{\alpha t^b}{\gamma + \ln M} = 0 \quad [10]
\]

where \( \gamma = 0.5772157... \) is Euler’s constant, showing an unbounded logarithmic dependency on area. The standard deviation is

\[
\sigma_{\text{max}} = \frac{\alpha M^{-b}k}{k} \left( \Gamma(1 + 2k) - \Gamma(1 + k)^2 \right)^{1/2} \quad [11]
\]

With limit at \( k = 0 \) of

\[
\sigma_{\text{max}} = \frac{\alpha m}{\sqrt{2}} = 0 \quad [12]
\]

Note that for \( k > 0 \), these formulas imply that this particular standard deviation decreases as the extrapolation area increases. This is a consequence of the upper bound to pit depths when \( k > 0 \) so that successive realized values for maxima are “squeezed-up” to this bound. However, the individual pit depths, measured over the whole area without censoring by any maximization process, will exhibit a sample standard deviation increasing with area sampled, since the (mean) range of randomly sampled depths will increase, by virtue of relationship [9] for the maxima. These area relationships as modeled here are principally a statistical phenomenon driven by the random characteristics of pitting processes. Conversely, the time dependence of the bound \( \xi \) and associated parameters \( u_k \) and \( \alpha_k \) is determined by the electrochemical characteristics of pit propagation, even though we have justified the particular form chosen, principally by reference to historical data on pit depths.

The \( \text{th} \) percentile, \( x_p \), defined by \( Pr(x \leq x_p) = F(x_p) = p \), and found by inverting [2], is

\[
x_p = \xi^b - \frac{\alpha p}{k} \left[ \frac{\ln p}{-\xi} \right]^{1/b} \quad [13]
\]

with limit at \( k = 0 \) of

\[
x_p = ut^b - \alpha t^b \ln(-\ln p) \quad [14]
\]

**Time to First Perforation**

If the metal wall thickness is \( d \), we can now deduce that \( t_d \) the predicted mean time to first penetration is

\[
t_d = \left[ d \left( \frac{\xi - \alpha M^{-b}}{k\Gamma(1 + k)} \right) \right]^{1/b} \quad [15]
\]

Observe that when \( k > 0 \) we must have \( t_d > (d/\xi)^{1/b} \).

**Perforation Count**

To assess the number of holes due to pitting, or the “penetration count,” \( i \), at time \( t \), we have utilized a result due to Nagaraja (13) which relates extreme values to record values and a particular random walk. Let \( x_{\text{max}} = x_1 \geq x_2 \geq ... \geq x_n \) be the largest pit depths over the extrapolation area \( A \). Now define a rescaled version of these depths by setting

\[
y_j = \begin{cases} 
\xi - \sigma \sqrt{\ln j} & j = 1 \ldots i \\
\xi - \sigma \sqrt{\ln n} & j = n
\end{cases}
\]

Then for large \( N \) \((\text{~large } A \text{ or } M, \text{ here})\), the joint distribution of \( \{y_j\} \) is asymptotically equal to the joint distribution of \( \{-\text{sgn}(kS_j)^b\} \), or \( \{-\text{log } S_j\} \) when \( k = 0 \); where \( S_j = \sum_{k=1}^{j} E_k \), and the \( E_k \) are independent exponential variates with rate parameter 1. Each state in the random walk \( S_j \), therefore, has a gamma probability distribution for which there are standard results on fractional moments (k, here) [see Johnson and Kotz (14)]. Hence, we can deduce that the mean and variance of \( x_i \) are, respectively

\[
\mu_i = \xi^b - \frac{\alpha M^{-b}k}{k} \left( \Gamma(1 + k) \right) \quad [16]
\]

and

\[
\sigma_i^2 = \frac{\left( \frac{\alpha M^{-b}k^2}{k^2} \right)^2}{k^2} \quad [17]
\]

Note that [16] should be used with care, since it will take negative values for large enough \( \text{~} k/M \), which is physically meaningless and also invalidates the limiting process by which [16] was derived. By solving [16] for \( t \) when \( \mu_i = d \), we can extend [15] to give an estimate for the predicted mean time to the \( i \)th penetration

\[
t_i = \left[ d \left( \frac{\xi - \alpha M^{-b}}{k\Gamma(1 + k)} \right) \right]^{1/b} \quad [18]
\]

Alternatively, by using the approximation \( \left( \Gamma(1 + k)\Gamma(t) \right)^b = t^b \), we can solve [16] for \( i \) when \( \mu_i = d \) to give an estimate for the penetration count at time \( t \)

\[
i = M \left[ \frac{(d - \xi)\alpha M^{-b}k^2}{k^2} \right]^{1/b} \quad [19]
\]

This particular result can alternately be derived by regarding the metal base as an absorbing barrier, appropriately scaled, in the above random walk \( S_j \), and then applying standard random walk approximations for the step count to absorption.

Equations [16]-[19] are nominally only valid for \( k \neq 0 \), but this value can be catered for by considering the limit as \( k \to 0 \) in each case.
Area of Perforations

To predict the total area of holes due to pitting at time $t$, $H_i$, say, we need to make some assumptions about the shape of a pit and about pit growth after penetration. It might be supposed that pits stop growing after penetration due to the loss of the occluded cell, but this would imply holes of zero area for hemispherical pits. Presumably the rate will slow down at first, rather than stop. An upper bound to hole area can be obtained by assuming unchecked pit growth.

If a hemispherical pit has depth $d$ in a wall of thickness $d$, then the hole area will be $\pi(d^2 - d^2)$. Summing this formula over the predicted number, $i$, of pits which have penetrated at time $t$ and applying standard results on the statistical moments of the above random walk, we find

$$H_i = \pi t \left( \frac{\alpha M^{\gamma - \frac{1}{2}}}{k f(j)} \right) \cdot \left[ \Gamma(j) \Gamma(j + 2k) - \Gamma(j + k) \right]^2$$

$$+ \left[ \xi - \frac{\alpha M^{\gamma - \frac{1}{2}}}{k f(j)} \Gamma(j + k) \right]^2 - \pi t \left( \xi^2 d^2 - d^2 \right)$$

as $A \to \infty$ (20)

The nearest integer to $i$, as given by [19], would have to be used for $i$, in the ‘finite A’ form of the above expression.

An alternative rectangular model for pit geometry is suggested by Sharland and Tasker (15). If such a pit has base area $A$, then

$$H_i = A \cdot i$$

which is stochastically equivalent to the ‘finite A’ form of (20). The stochastic properties of $H_i$ then come from $i$, whose stochastic behavior, in turn, can be deduced from standard results in random walk theory.

Model Fitting and Data Collection

The simplest data set to which these techniques could be applied would consist of paired values $(x_i, t_i), i = 1, \ldots, n$, where $x_i$ is the depth of the largest pit over a standard area $a$ of metal exposed to a pitting environment for time $t_i$. The separate areas a could be either distinct coupons from a designed experiment, or else a random sample at various times from regions of metal thought to be representative of the whole region over which extrapolations are to be made. The likelihood function for the above model, given this data, is then

$$\prod_{i=1}^n f(x_i | a, k, t_i)$$

where

$$f(x_i | a, k, t_i) = \frac{\alpha t_i^2}{k f(j)} \cdot \exp \left[1 - \frac{a M^{\gamma}}{k f(j)} \Gamma(j + k) \right]$$

is the GEV density function, derived by differentiating (2) with respect to $x$.

The method of maximum likelihood is known to be asymptotically efficient, subject to regularity conditions which are known to hold for the three-parameter GEV provided $k < 1/2$, and this will usually be the case in practice as can be seen from Fig. 1, 2, and 3. This is also pointed out by Hosking et al. (16) who suggest an alternative to maximum likelihood, based on probability-weighted moments, which they show to be simpler than maximum likelihood and more efficient for small samples. Unfortunately, these properties do not extend to the above four-parameter model, and for this reason we recommend and have utilized maximum likelihood when fitting this model to data. The FORTRAN program we have written to implement this technique for the above four-parameter GEV is a development of the algorithm described by Prescott and Walden (17) for the basic three-parameter model.

### Table I.

<table>
<thead>
<tr>
<th>Time in solution (hours)</th>
<th>Number of coupons</th>
<th>Maximum pit depth (mils)</th>
<th>Fitted mean</th>
<th>Fitted upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>40.50</td>
<td>1</td>
<td>30.5</td>
<td>28</td>
<td>38</td>
</tr>
<tr>
<td>144.17</td>
<td>2</td>
<td>52.2</td>
<td>43</td>
<td>57</td>
</tr>
<tr>
<td>215.33</td>
<td>2</td>
<td>40.8</td>
<td>50</td>
<td>67</td>
</tr>
<tr>
<td>292.50</td>
<td>2</td>
<td>35.9</td>
<td>56</td>
<td>57</td>
</tr>
<tr>
<td>331.00</td>
<td>2</td>
<td>46.2</td>
<td>59</td>
<td>79</td>
</tr>
<tr>
<td>378.50</td>
<td>2</td>
<td>63.5</td>
<td>62</td>
<td>82</td>
</tr>
<tr>
<td>453.25</td>
<td>2</td>
<td>82.7</td>
<td>66</td>
<td>88</td>
</tr>
<tr>
<td>477.00</td>
<td>3</td>
<td>79.9</td>
<td>70</td>
<td>90</td>
</tr>
<tr>
<td>528.00</td>
<td>3</td>
<td>67.5</td>
<td>70</td>
<td>93</td>
</tr>
</tbody>
</table>

*Partly extracted from Pierpoline et al. Ref. (18).

b 1 mil = 0.001 in.

To use the program, initial estimates are required for the four parameters $\alpha$, $u$, $k$, and $b$. These estimates have to be feasible, in the sense that

$$k \alpha x_i < b \leq \alpha + u \alpha k_0 = k_0^\alpha$$

for all $j = 1, \ldots, n_i; i = 1, \ldots, r$

where $\alpha_0$, $u$, $k_0$, and $b_0$ are the initial estimates and \{t_i = 1, \ldots, r\} are the exposure times at which the extremes \{x_i = 1, \ldots, n_i; i = 1, \ldots, r\} were measured. They can be found by utilizing the following three-step procedure:

1. Fit the basic three-parameter GEV distribution to \{x_i = 1, \ldots, n_i\} for each i where feasible (i.e., $n_i > 2$), obtaining estimates ($\alpha_i$, $u_i$, $k_i$).
2. Regress log $u_i$ and log $a_i$ on log $t_i$ to find $u_0$, $a_0$, and $b_0$.
3. Put $k_0$ equal to the average of the $k_i$.

This procedure can be implemented provided $n_i > 2$ for at least two distinct exposure times $t_i$. If this is not the case, the following two-step procedure can be tried:

1'. Regress log $x_i$ on $t_i$, giving an initial estimate for $b$, since $E(x_i) = a \exp t_i$.
2'. Fit the three-parameter GEV to transformed data $y_i = \ln(x_i) - \ln b - j - 1, \ldots, n_i; i = 1, \ldots, r$. Note that the log likelihood for these transformed data differs from that for the original data only by a factor $\beta \log t_i$, implying that the results from step 2' will be at the true optimum whenever $b$ from step 1' is optimal.

These initial estimates can then be used by the four-parameter program to produce final estimates, along with

![Fig. 4. Extreme pit depths for stainless steel; area extrapolation for three exposure times.](image)
their estimated standard errors derived from the information matrix in the usual way.

**Stainless Steel Example**

Pierpoline et al. (18) exposed a series of 316L coupons (2 x 2 x 1/2 in. thick) in a 10% ferric chloride solution at 50°C. The four largest pits were measured on removed coupons, and the extracted single maxima from these coupons, along with the associated exposure times, are given in Table I.

The transformation method described above gave starting values

\[ a_0 = 0.875 \quad u_0 = 5.502 \quad k_0 = 0.408 \quad b_0 = 0.460 \]

On running the four-parameter program with these as initial estimates, convergence was achieved after six iterations to give final estimates

**Standard error**

\[ \tilde{a} = 1.004 \quad 0.317 \]

\[ \tilde{u} = 6.322 \quad 1.721 \]

\[ \tilde{k} = 0.401 \quad 0.162 \]

\[ \tilde{b} = 0.378 \quad 0.047 \]

Rescaling to a time base measured in days rather than hours, and area measured in square inches rather than multiples of the coupon size 4 in.², implies

\[ \tilde{a} \rightarrow \tilde{a} = 24^{0.378} \cdot 4^{0.401} = 5.791 \]

Substitution of these values in [9], [11], and [13] gives extrapolated mean for the largest pit depth

\[ p_{\text{max}} = [29.16 - 14.404/A^{0.401}]^{0.376} \]  

plus standard deviation

\[ \sigma_{\text{max}} = [1.664/A^{0.401}]^{0.376} \]  

pth percentile

\[ x_p = [29.16 - 2.50(-\ln(p)/A^{0.401})^{0.376} \]  

and moving upper bound (independent of area)

\[ \xi = 29.16^{0.376} \]  

The mean depth of the 2nd largest pit is

\[ \mu_2 = [29.16 - 20.18/A^{0.401}]^{0.376} \]

All for t measured in days, area A measured in square inches, and depth x measured in mils. A plot of [23] as a function of area for selected values of t can be found in Fig. 4. Note that the asymptotic value of [23] for large A is given by [26], and that at point the associated standard deviation, given by [24], is zero. A Gumbel Type I fit to this data (with suitable modification for the time dependence) would have had no such upper bound. It would also have a constant (over area) standard deviation. A plot of [23] as a function of time for selected plate sizes can be found in Fig. 5.

along with the concurrent upper bound given by [26]. While a plot of the limiting form of [20] expressed as the proportion of surface penetrated (by using the hole count per unit area, \( i_t / A \), from [9]) against time in days can be found in Fig. 6.

The time to first penetration through an exposed area measuring \( A \) in.² of half-inch thick metal is

\[ t_d = (500[29.16 - 12.812/A^{0.401}]^{0.376} \rightarrow 5.26 \text{ years as } A \rightarrow \infty \]

while the predicted number of holes for \( t > t_d \) is

\[ i_t = A \cdot (2.02 - 34.63 A^{0.376})^{0.376} \]

This result can be interpreted as indicating a fixed limiting density (\( = i_t / A \)) of 5.77 "large" pits per sq. in. However, the mathematical limit already used to derive the formula for \( i_t \) means that this inference should be treated with caution. This limit also depends crucially on the fact that we have modeled the time dependence of the location (u) parameter and the scale parameter (a) in the same way and with the same rate constant (i.e., ut² and at²).

**Discussion**

Standard techniques for extrapolating extreme pit depths over large areas rely heavily on the use of the Type I (Gumbel) extreme value distribution, with no built-in check on the possibility that one of the other types might be more appropriate. As this analysis shows, this can be of crucial importance, particularly when a Type III distribution is suggested, since this has a finite upper bound which cannot be exceeded no matter how large the extrapolation area, which property might be thought more in keeping with metallurgical intuition and overall theoretical constraints to pit growth. Extrapolation of pit growth measurements into the future is usually handled separately from area extrapolation and typically by fitting a power-law growth curve to plots of pit depths against time. The above analysis shows how this procedure can be merged into one technique which simultaneously tackles the problems of time and area extrapolation by using a family of probability distributions appropriate to measurements of maximum pit depths. The procedure also produces standard errors for all estimated parameters, thus providing some check on the validity and appropriateness of the model, and on the accuracy of forecasts.

Finally, we remark that although this paper is directed toward data measured on maximum pit depths, the technique is of wide applicability and in particular has been successfully applied by us to data collected on minimum wall thicknesses along a gas pipeline, which data also, and incidentally, provided a positive value for k, the shape parameter.
Acknowledgment

This work has been supported in part by a grant from the U.S. National Bureau of Standards.

Manuscript submitted Nov. 1, 1988; revised manuscript received June 6, 1989. This was Paper 240 presented at the Honolulu, HI, Meeting of the Society, Oct. 18-23, 1987.

REFERENCES

8. P. M. Aziz, ibid., 12, 495 (1956).