



# **Guidelines for use of statistics for analysis of sample inspection of corrosion**

Prepared by **TWI Limited**  
for the Health and Safety Executive 2002

## **RESEARCH REPORT 016**



# **Guidelines for use of statistics for analysis of sample inspection of corrosion**

**TWI Limited**  
Granta Park  
Great Abington  
Cambridge  
CB1 6AL  
United Kingdom

This document is intended to advise plant engineers and inspection personnel on methods for analysing and extrapolating inspections for large plant items including vessels, pipework and pipelines, taking into account the statistical nature of corrosion. The document is intended to introduce the methods of statistical analysis of corrosion inspection data. Before the methodology is included in standards, practical experience of industrial applications is needed to identify the most relevant distributions and statistical techniques.

This report and the work it describes were funded by the Health and Safety Executive (HSE). Its contents, including any opinions and/or conclusions expressed, are those of the author alone and do not necessarily reflect HSE policy.

© Crown copyright 2002

Applications for reproduction should be made in writing to:  
Copyright Unit, Her Majesty's Stationery Office,  
St Clements House, 2-16 Colegate, Norwich NR3 1BQ

First published 2002

ISBN 0 7176 2554 0

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording or otherwise) without the prior written permission of the copyright owner.

## SUMMARY

Leakages of hydrocarbons provide both a serious risk of fire and explosion, and a loss of plant availability. Corrosion has been shown to cause in the region of 15% of the leakage occurrences. Inspection is carried out, particularly for internal corrosion by means of non-destructive test methods which give values of the pipe or vessel wall thickness. Typically these methods only sample the overall area of a plant. There are risks associated with this. The sampling may lead to inaccurate estimates of corrosion rate, especially where pitting has occurred, or may not address the fact that a sample needs to be extrapolated over the whole area of plant in order to give a realistic estimate of the minimum wall thickness in that area.

Statistical methods to improve the estimation of corrosion rate or to estimate the minimum thickness over a larger area have been suggested for over 50 years, and have been applied in a few isolated cases. These methods, when combined with reliability methods, offer a potential for obtaining better information from inspections by further analysis of the data collected and can produce predictions of future probability of leakage. However widespread application is not common, largely because the use of statistics requires specialist knowledge, and no reference standards exist. These guidelines are intended to provide an introduction to the techniques and capabilities of the statistical methods with view to their wider application in industry.



## TABLE OF CONTENTS

1	SCOPE OF DOCUMENT.....	1
2	ACKNOWLEDGEMENT .....	1
3	BACKGROUND.....	2
4	INTRODUCTION – STATISTICS AND CORROSION.....	3
4.1	FREQUENCY DISTRIBUTIONS .....	3
4.2	DETERMINATION OF DISTRIBUTION (UNKNOWN CORROSION).....	11
5	THE USE OF INSPECTION DATA .....	14
5.1	GENERAL COMMENTS ON DATA COLLECTION.....	14
5.2	DETERMINATION OF CORROSION RATE - CURRENT PRACTICE.....	14
5.3	SUGGESTED ANALYSIS METHODS – NORMAL DISTRIBUTION .....	16
5.4	EXTREME VALUE FITTED DATA (TYPE 1 DISTRIBUTION).....	18
5.5	DETERMINATION OF CORROSION RATE .....	20
5.6	NUMBER OF SAMPLES NEEDED (EACH CASE) .....	21
5.7	CHOICE OF LOCATION.....	21
6	CASE STUDY: FAILED OIL PIPE .....	23
7	CONCLUSIONS.....	25
8	RECOMMENDATIONS .....	26
9	REFERENCES .....	27
10	APPENDIX A CALCULATION OF CORRELATION DISTANCE.....	28
10.1	APPLICATION .....	28
10.2	HANNING WINDOW.....	28
11	APPENDIX B CALCULATION OF INHERENT AND SAMPLING ERRORS.....	29
11.1	APPLICATION .....	29



# 1 SCOPE OF DOCUMENT

This document is intended to advise plant engineers and inspection personnel on methods for analysing and extrapolating inspections for large plant items including vessels, pipework and pipelines, taking into account the statistical nature of corrosion. The document is intended to introduce the methods of statistical analysis of corrosion inspection data. Before the methodology is included in standards, practical experience of industrial applications is needed to identify the most relevant distributions and statistical techniques.

The analyses contained refer to areas where corrosion conditions are known to be alike. In general, to achieve these conditions the following must be similar:

- Materials
- Corrosion product/chemistry
- Temperature
- Flow rate
- Presence of inhibitor
- Fluid composition
- Presence of contaminants

It should be noted that small changes in these parameters can cause wide changes in corrosion rate. Where conditions do change in the area to be inspected (e.g. welds), this can be handled by collecting data from these specific locations and treating them separately for analysis.

100% inspection of a large area of plant is not a practical proposition in most cases; the majority of applications use sample inspection. The majority also use ultrasonic thickness measurement for data collection, and this method is assumed for the most of this document. It is recognised that new methods for measurement of corrosion are becoming available, and a short review of the possible applications of these is included.

The document also does not address:

1. the choice of locations for inspection made on a risk-based or experience-based methodology, targeted at 'hot spots', or
2. analysis of corrosion localised to specific locations in a plant

It should be noted that the use of extreme values is particularly suitable for the evaluation of the potential for leakage, rather than bursting. For further guidance on the latter refer to the RACH (Reliability Assessment for Containers of Hazardous materials) project report (1).

If the data does not fit the distributions described here then analysis by more advanced techniques may be necessary (for example Type 3 Extreme Value Distributions). Additional issues to be considered include the correlation between adjacent data points (Appendix A) and the estimation of inherent and sampling errors (Appendix B).

TWI has drawn up these guidelines, at HSE's request, specifically in order to improve the sampling ultrasonic inspection of pipe, by making data collection, analysis and extrapolation available to non-specialists. However it should be pointed out that the guidance given here is introductory and more application experience is needed before standards can be developed.

# 2 ACKNOWLEDGEMENT

The Health and Safety Executive, together with TWI, acknowledge the input from Mitsui Babcock Technology Centre in preparation of this document.



### 3 BACKGROUND

Corrosion has been shown to be the cause of 15-20% of leakages of hydrocarbons from offshore plant (2). Leakages can lead to more disastrous consequences if subsequent ignition was to result in fire or explosion. In order to reduce the number of leakages from this source, the defects that lead to failure need to be detected and mitigating action taken before failure occurs.

Corrosion is in essence a statistical effect (3) governed by a number of variables. For example, microscopic variations in a surface tend to cause different forms of corrosion and also variations in the corrosion rate over either a wide or small area (pitting). In these areas the simple assumption that corrosion rate is uniform across an area is unlikely to be accurate, and sample thickness measurements are unlikely to be representative of the whole component.

Studies and applications of the statistical nature of corrosion, and its relationship to inspection, have been carried out since the 1950's, but have never been commonly applied in routine inspections. No standards exist for the analysis of inspection data for corrosion.

Initial work using extreme values was carried out by Gumbel (4). He used the theory to estimate the condition of pipelines with external corrosion.

Hawn (5) also used the extreme value method for external pits on pipelines, and extrapolated to 5820 times the inspected area to estimate the probability that a maximum pit size would be no greater than a certain value. For example, there was 97% probability that the pit size would be no greater than 139 mils (3.5mm).

Manley (6) described the use of an ultrasonic thickness gauge, and a computer, to log and record data from an erosion/corrosion survey. Essentially the method used was linear interpolation of sample readings.

Joshi et al (7) use the extreme value analysis method to extrapolate from small inspection patches in an above ground storage tank to the whole tank. They noted that the method particularly applied to pitting corrosion. Sparago (8) shows how the underlying thickness distribution can be used to estimate the probability of a wall thickness being below a certain level from ultrasonic thickness gauge data.

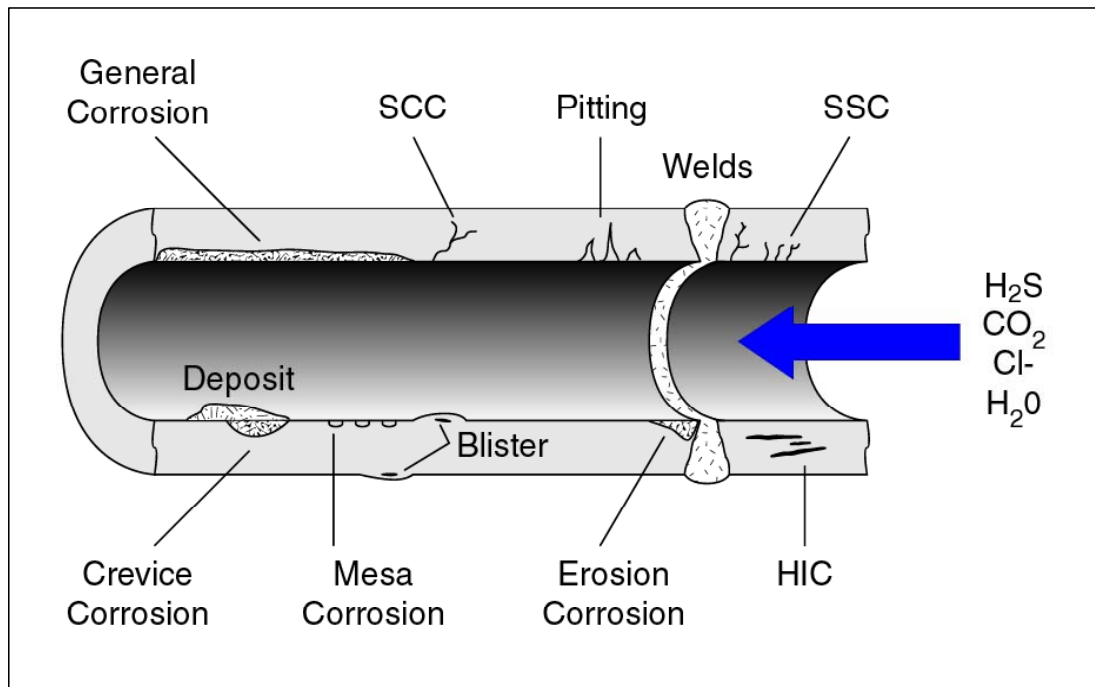
Kowaka (9) gives a useful overall text to the statistical method of analysing corrosion data. The use of these methods for analysing corrosion data has been referred to consistently in Japan since the 1980's, however little of the work refers to the strategies of data collection by NDT methods.

More recently Mitsui Babcock (10) conducted a group sponsored project which included some analysis of corrosion data by the extreme value method.

Thus it can be seen that although the statistical methods have been developed, the application of these methods is not carried out routinely apart from possibly in Japan. The papers above, although describing the methods used, do not generally validate the results obtained by comparing a sample with the whole population.

## 4 INTRODUCTION – STATISTICS AND CORROSION

Corrosion can take many forms (Figure 1), and the statistics of each will be different. The statistics arise from the measurements taken of the wall thickness of a component or the pit depth (when a surface is accessible). The morphology of the corrosion (the shape of the surface) will affect these measurements and form them into distributions of data. To understand the process of converting the series of data points into the statistical distribution, a useful step is to construct a frequency distribution (see Section 4.1) However the frequency distribution needs to be converted to a mathematical form for further analysis.

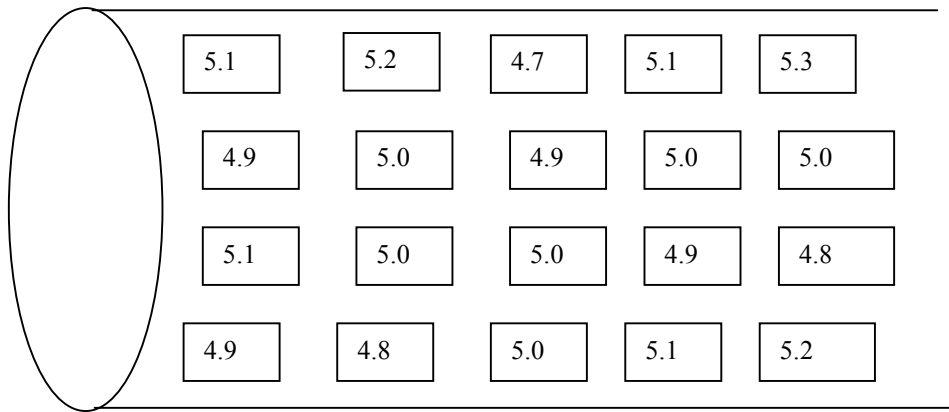


**Figure 1**  
**Different Forms of Corrosion**

### 4.1 FREQUENCY DISTRIBUTIONS

The inspection data can be shown in a graphical form, which is useful to show the overall variations. One example of this is a frequency distribution. To construct a frequency distribution, the data is ordered in size, and grouped into size ranges of equal intervals (usually for UT thickness data size groups of 0.1 or 0.2 mm are reasonable). The number of data points in each size group is plotted against the size group.

For example the following thickness data may have been taken of a pressure vessel or pipe:



**Figure 2**  
**Example Readings from Pressure Vessel/Pipe**

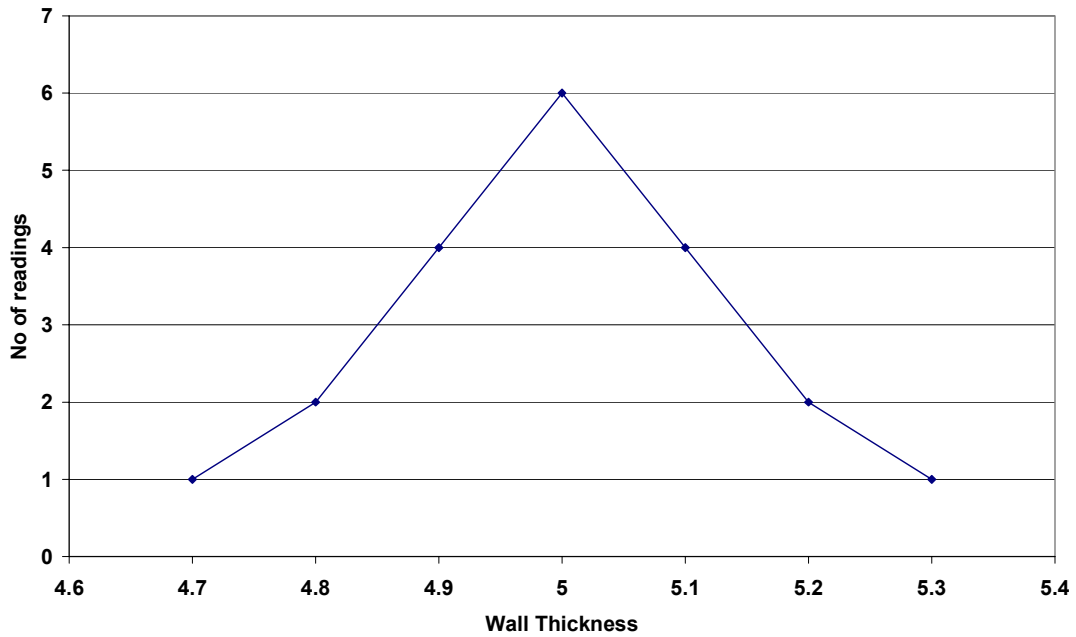
When ordered the data becomes:

4.7, 4.8, 4.8, 4.9, 4.9, 4.9, 4.9, 5, 5, 5, 5, 5, 5, 5.1, 5.1, 5.1, 5.1, 5.2, 5.2, 5.3

The values are grouped in intervals as tabulated below:

Wall Thickness	Frequency
4.7	1
4.8	2
4.9	4
5	6
5.1	4
5.2	2
5.3	1

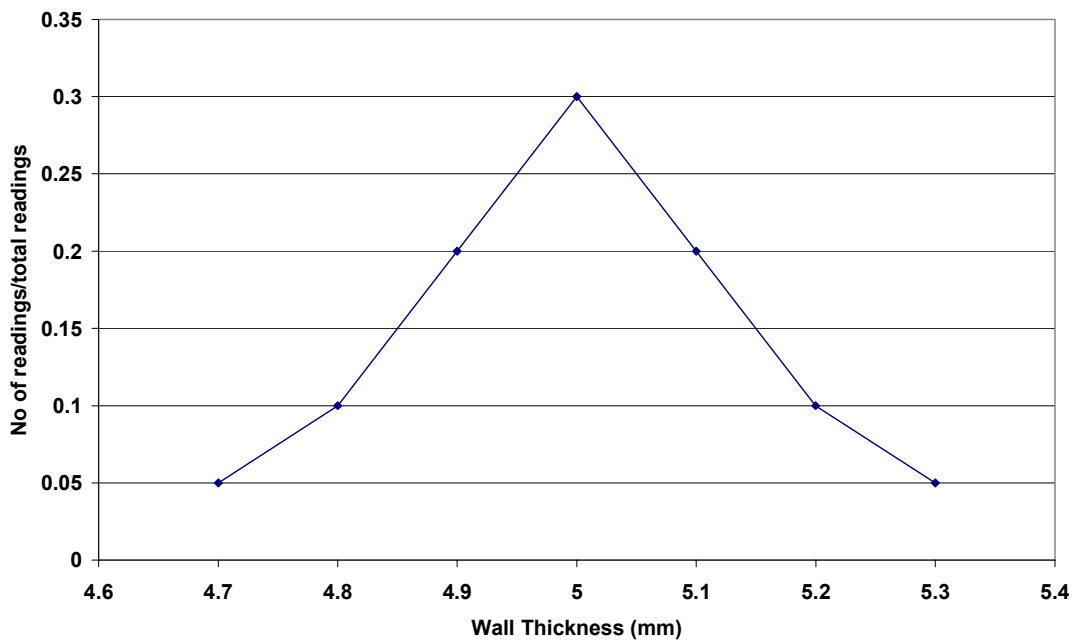
The frequency distribution is a plot of the data in this table as Figure 3.



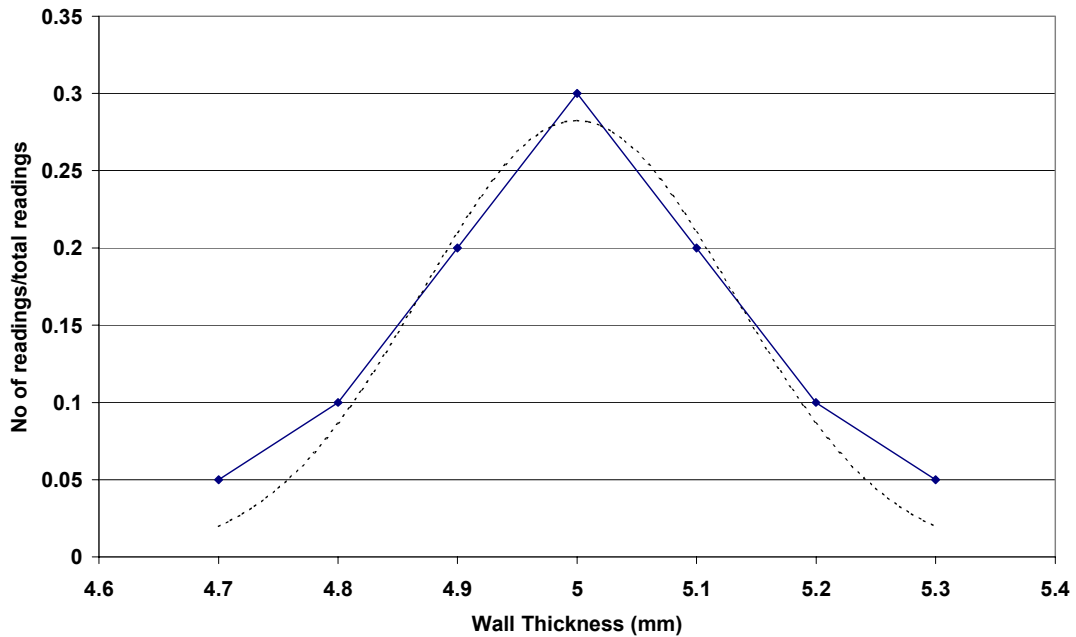
**Figure 3**  
**Frequency Distribution**

This frequency distribution appears to be symmetrical and the central tendency is 5.0.

In order to make use of this pattern in the data it is necessary to convert it into a mathematical form. The mathematical form must describe the data in Figure 3 or must be a close fit. Usually one step in this process is to normalise the data by dividing the frequency by the total number of readings (Figure 4). This is called a normalised frequency distribution (frequency being the number of readings). Basically the process of fitting the data is one of trial and error, although tools are available to assist in this process as described below.



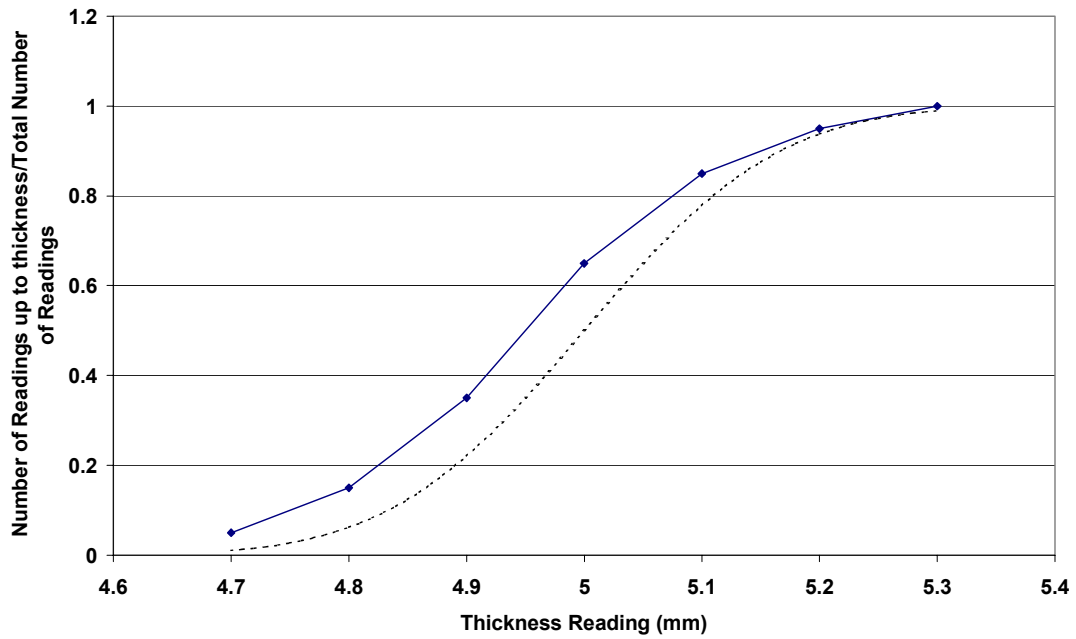
**Figure 4**  
**Normalised Frequency Distribution**



**Figure 5**  
**Fitting a Curve (in this case a normal distribution)**

There are a number of standard forms of frequency distribution. These have mathematical equations to describe them, but for the purposes of this document only the general shapes and the graphical display will be considered. A frequency distribution also has associated with it a cumulative distribution, which is the sum of the number of readings up to a certain point and is equivalent to the area under the frequency distribution curve.

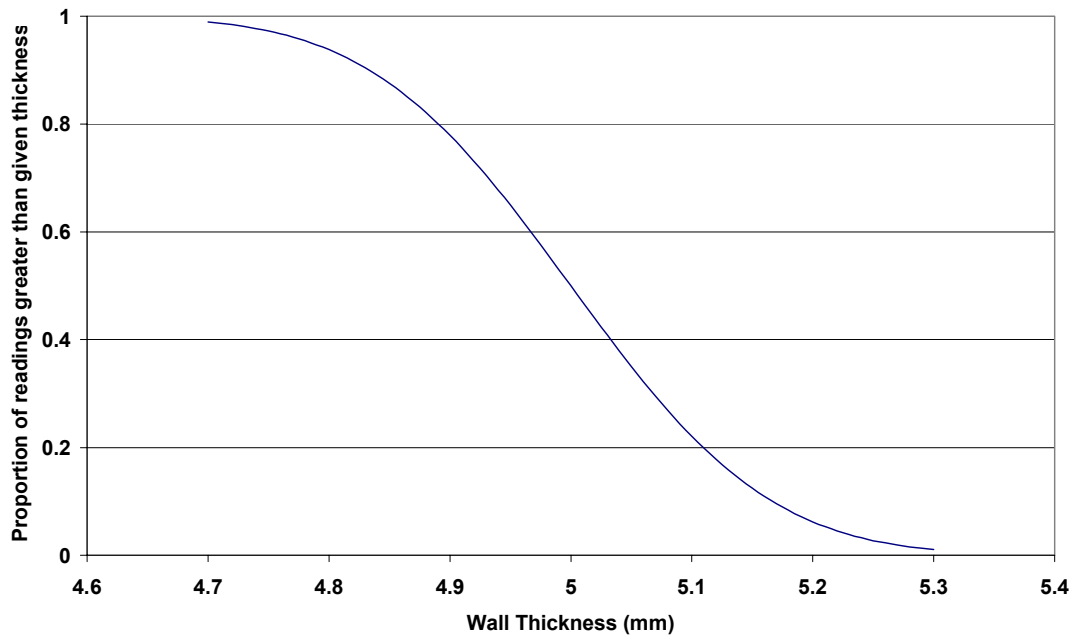
Figure 5 shows a normal distribution with a mean of 5.0 and standard deviation of 0.13 adjusted to fit the same scale as the previous frequency distribution. If we believe that this fit is good enough it is possible to make calculations on the data, for example to estimate the proportion of the vessel with a wall thickness less than 4.7mm. This is calculated from the area of the frequency distribution below 4.7mm. This is more easily calculated from the value of the cumulative distribution at this point. The cumulative distribution is plotted by successively adding in the number of points/total number. It always reaches a maximum of 1, for the largest measurement in the data set.



**Figure 6**  
**Cumulative Distribution of data and normal distribution curve**

The proportion of the wall thickness less than 4.7mm (for example) can be read directly from this graph.

A function derived from the cumulative distribution that will be used in later analyses is called the survivor function. This is simply the cumulative distribution subtracted from 1. Figure 7 shows the survivor function for the above normal distribution.

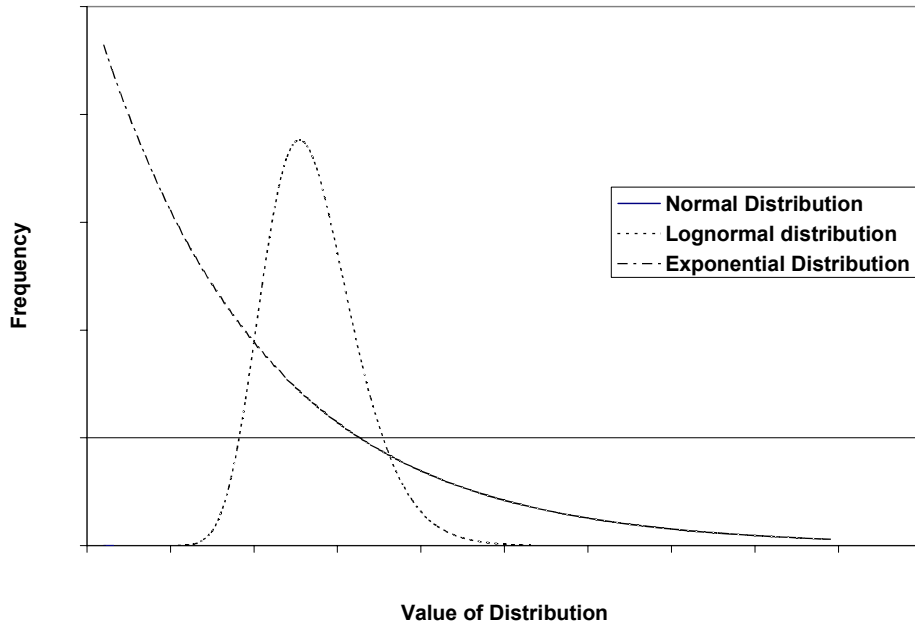


**Figure 7**  
**Survivor Function for Normal Distribution**

The choice of distribution for the calculations is crucial. Each distribution has a different shape and the wrong choice can lead to erroneous conclusions. For example two shapes which may represent certain corrosion types are the lognormal and exponential distributions. The basic shapes of the frequency distributions for these are given in Figure 8. It is also important to note that the accuracy of the parameters of the distribution, and the confidence limits eventually produced, depend on the number of data points used.

When an entire population of a defect is available, for example given a detailed (100%) surface scan covering an entire vessel, then statistical analyses of the data generally produce a fundamental or underlying distribution pattern. In this case all data values are fully defined. Typical fundamental distribution patterns include those shown in Figure 8, but could also include Poisson or Gaussian distributions.

In practice complete data population sets rarely occur. For example 100% scanning of a vessel would be considered impractical. Accordingly statistical treatments and procedures have arisen over the years to allow predictions to be made on the basis of limited sample information, where the sample is used to infer the greater population behaviour. N.B. A statistical sample which contains less than an entire population of data may follow the distribution of its parent population and this may be a fundamental distribution.



**Figure 8**  
**Different Distribution Shapes**

One special example that is important in the context of inspection is the family of extreme value distributions. Instead of being constructed from all the readings taken, they are constructed from the extreme values of groups of the data. Usually for inspection the data will be grouped in areas. So for example if the results from Figure 2 are in circumferential bands, the lower extreme values will be 4.9, 4.8, 4.7, 4.9, 4.8 and these will form a distribution of their own called an extreme value distribution.

#### **4.1.1 Extreme Value Distribution**

The methodology that underpins the extreme value statistical analysis of measurements of NDT inspection is defined fully in Kowaka (9). For the purposes of explanation the following text refers to damage due to corrosion. However the technique has applicability to the analysis of pitting defects resulting from some other factors.

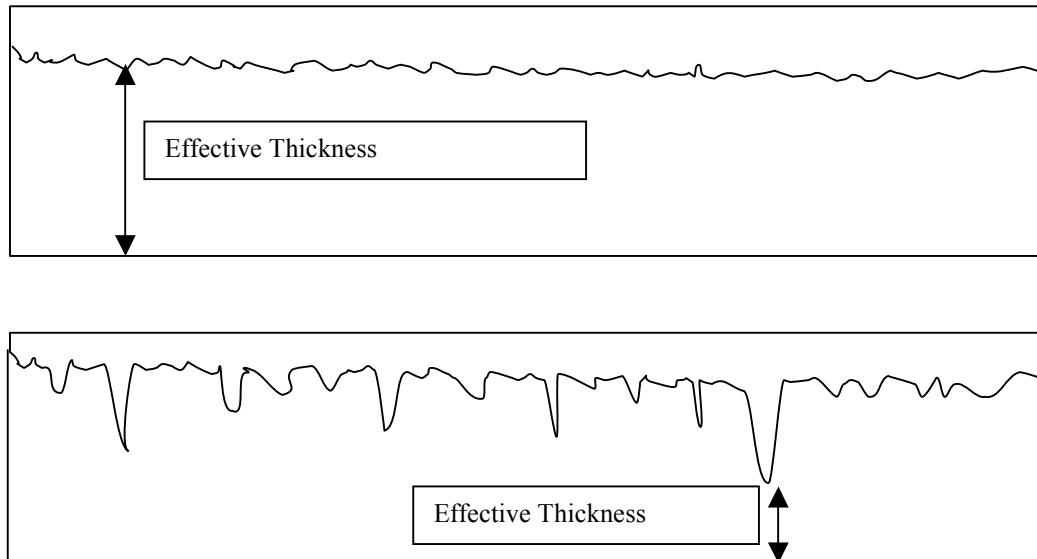
Corrosion damage may be classified according to their morphologies to be uniform (general) corrosion or non-uniform (pitting) corrosion. When a material deteriorates with uniform corrosion its life may be defined as the point at which the average thickness reaches a minimum allowable threshold (see Figure 9).

The upper example in Figure 9 shows an example of uniform corrosion. Due to the uniformity of the defect, fundamental statistical distributions can be used to predict the average wall thickness loss.

The lower surface in Figure 9 shows an example of non-uniform corrosion displaying more localised defect penetrations. In this case considerations of the average pit depth are inappropriate since loss of containment will result as soon as one extreme defect perforates the material. Fundamental statistical distributions are not suitable for analysis of such cases, and extreme value calculations are required in order to predict a maximum expected pit depth from what will generally be sample information. The need for the use of the extreme value distribution will be evident when the NDT data is analysed.



Extreme value data sampling differs from fundamental data sampling in that the former only considers a set of extreme values extracted from a larger sample (as described above for each circular ring above). Statistically the effect of this filtering (i.e. using only part of the distribution) allows the tail of the resulting distribution to more accurately model the potential defect extremes which may exist in the material. In practice to allow statistical integrity each extreme value must be collected from a subset of a larger sample which in itself (i.e. the subset) contains sufficient data to infer a parent fundamental distribution population.



**Figure 9**  
**Top: uniform corrosion, Bottom: non-uniform corrosion**

Kowaka (9) has attempted to fit several types of corrosion data to different distributions, some examples of which are in Table 1. However in most inspection situations there is little prior knowledge of the distribution type, and therefore a full fitting operation is needed prior to any assumptions about the distribution.

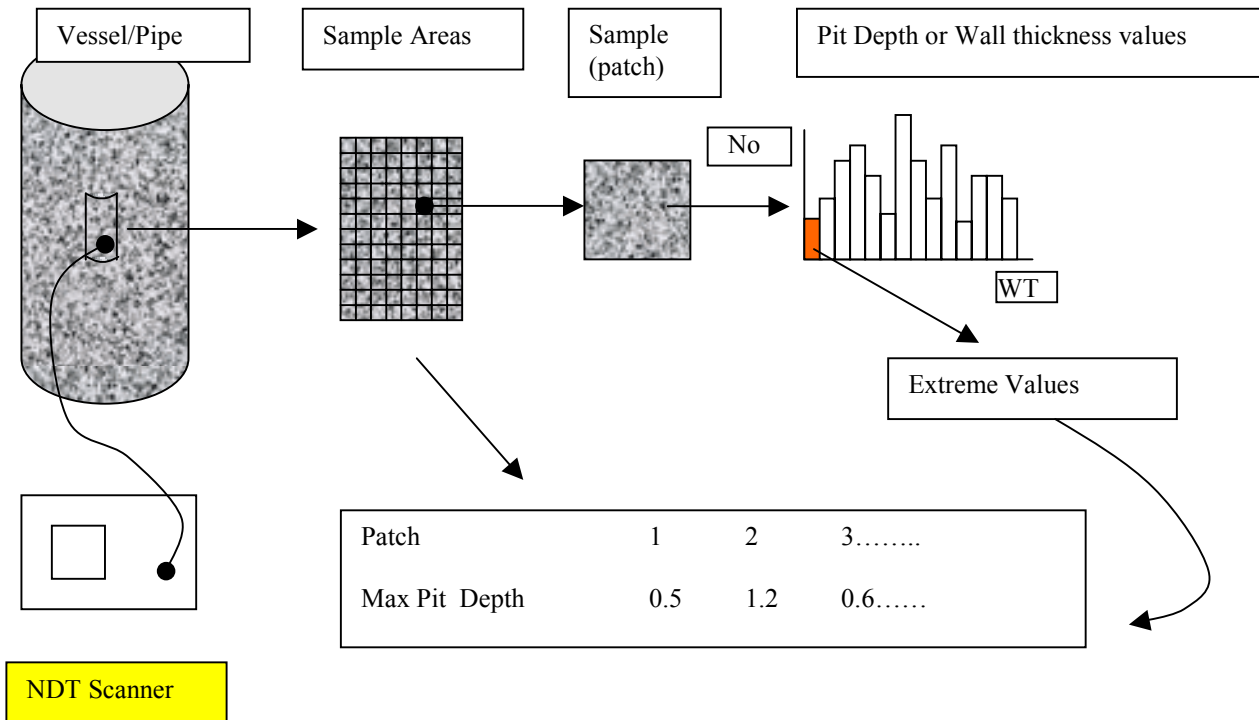
**Table 1**  
**Types of Distribution for Corrosion (from Kowaka)**

Corrosion Type	Distribution
Pit Depth for Carbon Steel fresh water supply pipe	Normal
Maximum Pit Depth for Petroleum Tank Bottom Plate	Extreme Value (Type 1)
Maximum Corrosion Depth beneath corrosion product layer for carbon steel tubing for mineral dressing slurry	Extreme Value (Type 1)

## 4.2 DETERMINATION OF DISTRIBUTION (UNKNOWN CORROSION)

### 4.2.1 Collecting and Sampling the Data

The manner in which data is collected and later analysed is represented in Figure 10. The vessel/pipe is first divided into sample areas for inspection (often determined by the inspection device scanner). An individual sample area (known sometimes as a patch) contains a number of individual values of thickness or pit depth taken by the NDT device.



**Figure 10**  
Data Collection from a vessel/pipe to produce extreme values

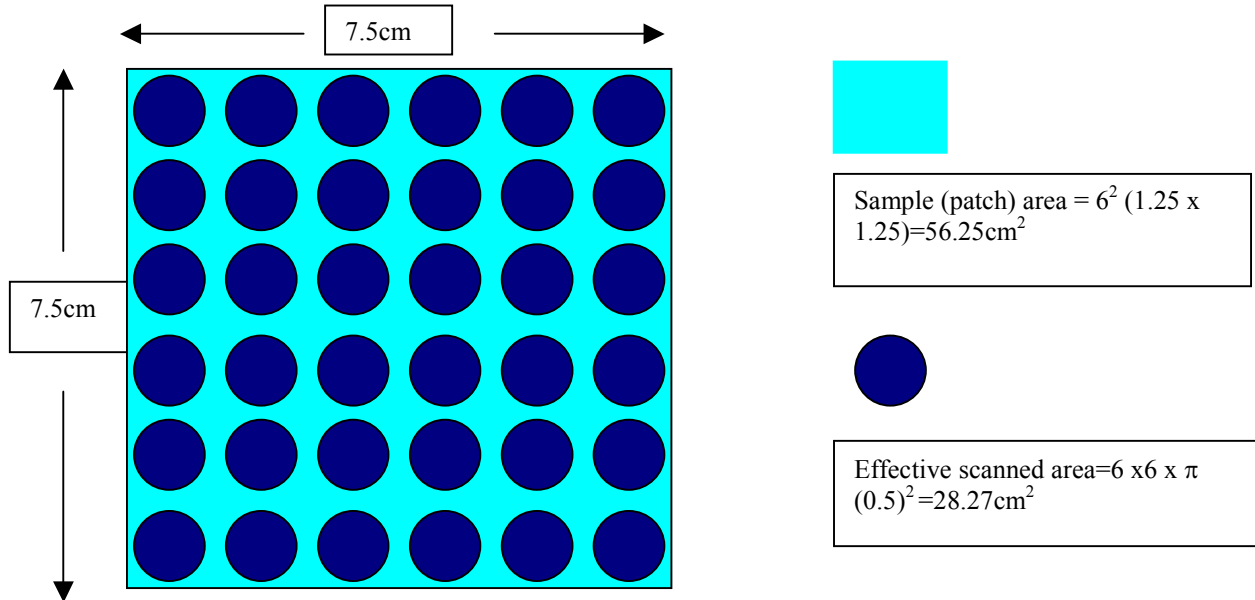
### 4.2.2 Identifying homogeneity within the vessel/pipework

As an initial step, the homogenous areas (i.e. those with similar corrosion properties) must be identified. Statistical treatment, such as extrapolation, can only be applied with integrity to samples which are considered to be representative of the parent populations (i.e. a larger area). Accordingly this may result in several analyses being conducted for the vessel as a whole.

#### Scanning Method

Data collection is typically (though not always) performed in a uniform manner, typically as equally spaced measurements on a grid as depicted in the schematic below. It is important to draw the distinction between the sample area represented by a patch of data and the effective area scanned by the NDT probe. The sample area can be defined as the surface area of that portion of the vessel under examination. The effective scanned area is the sum of all the areas actually covered by the probe within the boundary of the sample area. If calculations such as extrapolation are carried out, then the effective area should be used.

Clearly there is a direct relationship between the effective scanned area, the scan mechanism and the resulting number of data points. By way of illustrating the relationship, Figure 11 shows a sample area of approximately 56 cm<sup>2</sup> in which 36 spot data points have been collected using a probe of 1cm diameter. The resulting effective area scanned is about 28cm<sup>2</sup> representing about 50% of the patch area



**Figure 11**  
**Effective scanning Area**

#### 4.2.3 Extracting Extreme Values

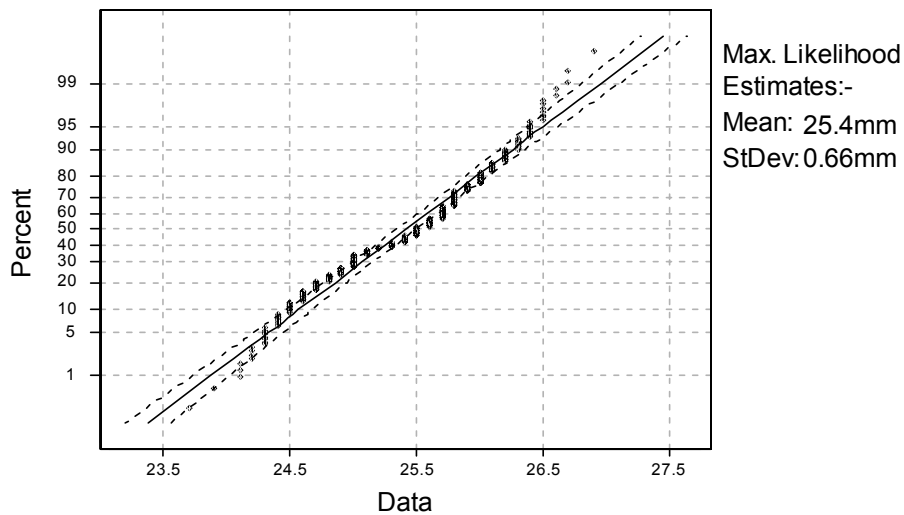
Based on the NDT Equipment, its resolution (i.e. the effective area actually scanned in one measurement recorded by the probe) and ideally, the confidence level required by the analyst, a set of sample areas is defined and scanned. Often, in practice, this may be all of the accessible areas, in which case the confidence level is set by the number of data points which can be taken. The measurements are then recorded and analysed. An example of extreme value selection of the data (which may be undertaken if prior knowledge of the corrosion indicates pitting) is shown in Figure 10 selected from each patch. When attempting to use extreme value analysis, it is imperative that each patch of data complies with the following criteria:

- Each individual patch contains sufficient measurements within itself to reflect a representative sample from the fundamental distribution. This is necessary to support the selection of an extreme value from that patch. Literature indicates that at least 5 readings are needed.
- the total number of extreme values that result across all patches (i.e. making up the extreme value data as a whole) are sufficient to enable a doubly exponential (or other extreme value) distribution to be adequately defined.

To fit distributions by matching the inspection data plotted as frequency or cumulative distributions to the mathematical plots by trial and error (as described in 4.1) could be extremely time consuming, and the task can be made much easier by the use of probability

plots. These are specially designed graphs which have an ordinate scale from 0 to 1, but the intervals on the scale depend on the distribution function. If the data matches a distribution then it will appear as a straight line. To use this method, the first step is to order the data as above (Section 4.1). The next step is to insert this data directly on to the probability plot. Many statistical packages have probability plots for a range of distributions included as standard.

When the data is plotted, a best fit regression line of the data is also plotted. If the two correspond, then the data matches the distribution described by the probability plot. If it doesn't fit then another distribution must be tried. The parameters of the distribution (i.e. those numbers which represent the spread and location, or central tendency) are determined from the slope of the line and its intercept with the horizontal axis.



**Figure 12**  
**Probability Plot showing fit to Normal Distribution**

An immediate application of the plot is to estimate the proportion of the plant below a certain thickness. In the probability plot this is simply read off the horizontal axis (for Figure 12, 1% of vessel is below about 24mm thick)

## **5 THE USE OF INSPECTION DATA**

### **5.1 GENERAL COMMENTS ON DATA COLLECTION**

#### **5.1.1 Manual Ultrasonic Thickness Gauge/A-Scan**

Where the data collected by ultrasonic thickness gauge or A-Scan equipment is sampled from pre-determined locations on a component, the data may be analysed for an underlying distribution. One exception to this is when the data is in a very close grid, in which case it is advisable to carry out a spatial analysis similar to that for scanned UT systems (see below). Alternatively, where the ultrasonic procedure is that of scanning individual portions of a vessel and reporting only the minimum thickness seen in each portion, then extreme value data is being collected and the extreme value approach can be applied directly.

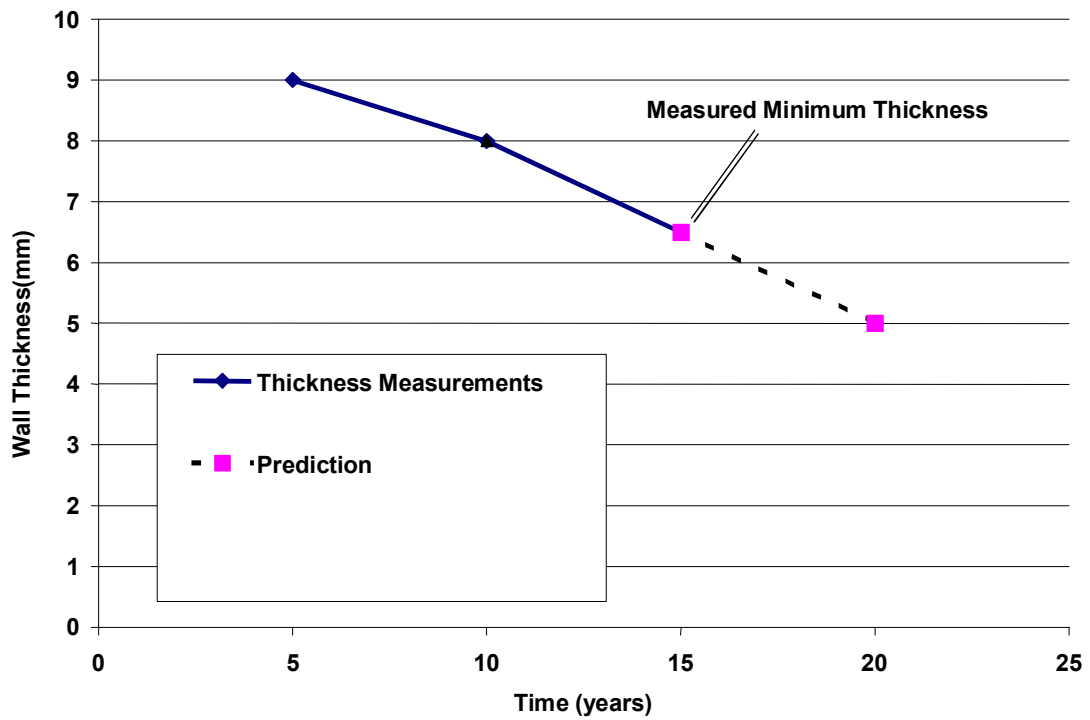
#### **5.1.2 Automatic Scanned Systems**

Scanning UT systems collect a large amount of data in a grid pattern. This data is taken offline of the NDT system for analysis. The raw data is suitable for underlying distribution analysis, subject to the individual readings being independent of each other. This can be tested by applying a spatial autocorrelation function to the data. This process is described in Appendix A.

If the autocorrelation distance is greater than the distance between readings then the readings used must be separated by at least the former distance. This involves grouping the data. An extreme value analysis can be implemented by taking the minimum value of each group.

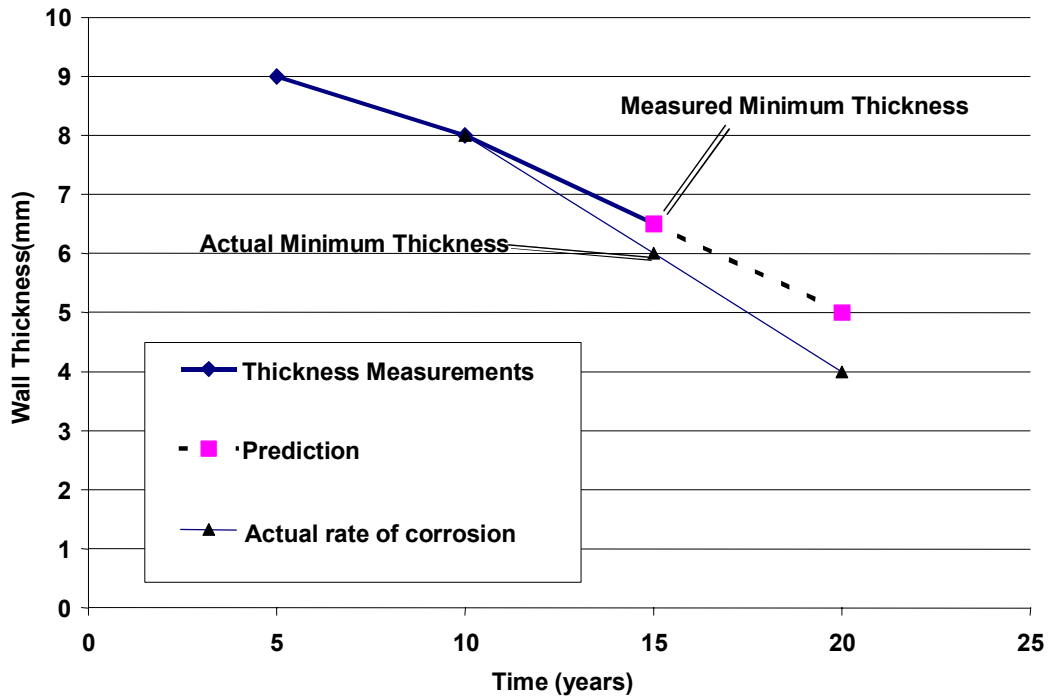
### **5.2 DETERMINATION OF CORROSION RATE - CURRENT PRACTICE**

Current methods of corrosion rate determination and life prediction tend to rely on a simplistic treatment of thickness gauge readings. Usually the minimum thicknesses or the thickness at particular locations at the last two inspections are taken and the prediction is a linear extrapolation (as in Figure 13).



**Figure 13**  
Simple corrosion rate method

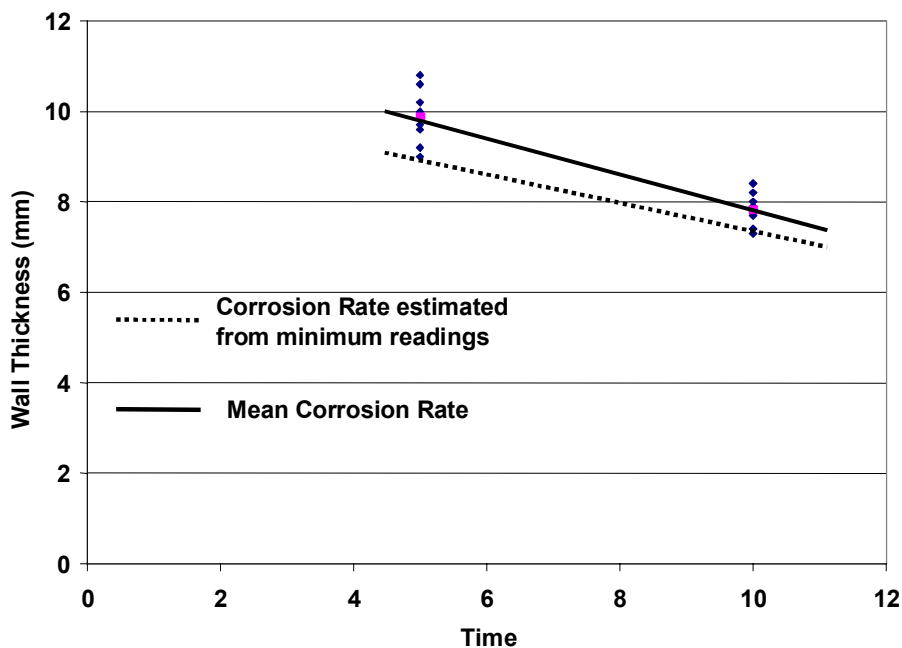
If the appropriate measurement is not made in each case, perhaps because the actual minimum has not been found in a scan, then the estimation of corrosion rate may be inaccurate. Figure 14 shows what can happen in this case.



**Figure 14**  
Illustration of Errors in Corrosion Rate Estimation caused by Inspection Errors

Where many readings are taken in an inspection survey, the use of minimum thicknesses or even individual thickness readings at a particular location can result in similar errors in the estimation of corrosion rate.

Figure 15 shows how this error can occur.



**Figure 15**  
**Methods for estimating corrosion rate**

These problems are exacerbated if there is pitting corrosion, since individual readings are unlikely to be representative of the minimum thickness, and since errors are much more likely in the inspection method (11,12). A source of error in the corrosion rate might also occur if any random scatter in the inspection method is reduced in successive inspections (for example by using improved techniques). In Figure 15 a lower corrosion rate is estimated by using the minimum thickness measurements than by using the mean value.

Statistical methods can help to solve these problems; sometimes even the simplest method will improve the likely accuracy of corrosion rate estimation.

### 5.3 SUGGESTED ANALYSIS METHODS – NORMAL DISTRIBUTION

#### 5.3.1 Determination of Corrosion Rate

It is currently common practice to estimate the corrosion rate from the minimum wall thickness measurement from each inspection, rather than from the mean. When the corrosion is uniform this procedure is especially prone to NDT measurement errors, because the minimum wall thickness simply reflects the worst undersizing of the wall thickness rather than any genuine variation due to the corrosion itself. For uniform corrosion, regression analysis is preferable. It is therefore recommended that multiple readings be taken for each

case, the average calculated, and this value used to estimate the corrosion rate. This will then yield not only a best estimate of the trend in the data, but can also be used to:

- check that the magnitude of any random measurement errors about this trend is as expected
- if required, estimate confidence limits on both the present and future corrosion.

Simple regression analysis is now available, not only within statistical software, but also within most spreadsheet packages. Failing this the rate can be estimated from the difference between the maximum loss of wall in the most recent inspection and the minimum loss of wall in the previous inspection(s). This will, in most cases, lead to a conservative estimate of corrosion rate.

### **5.3.2 Proportion below thickness**

The proportion below a certain thickness can be estimated from the probability graph. The thickness required is obtained from the horizontal axis, and the percentage of the area below this thickness is given by the probability reading corresponding to this thickness and the best fit line.

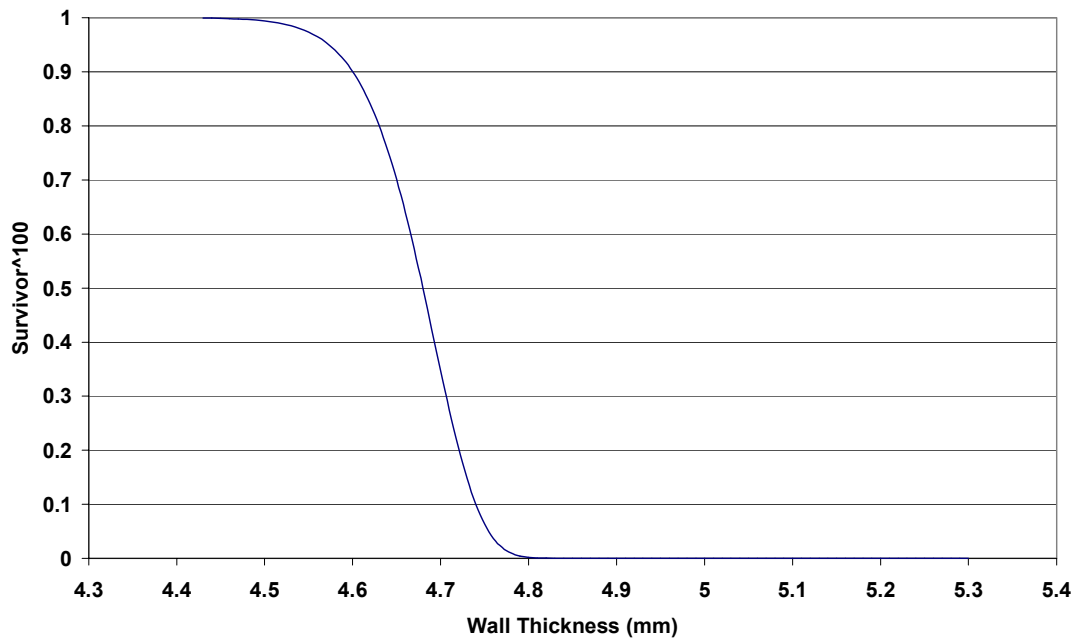
In Figure 12 for example the solid line illustrates a fitted distribution, while the dashed lines represent the 95% confidence limits. In this case, it can be seen that very little corrosion has occurred. The plot indicates, for instance, that ~1% of the thicknesses fall below 24mm. In general, the normal distribution gives a reasonably good fit to the data, but there appears to be a trend for the data to fall slightly below the fitted line for thicknesses less than about 24mm. This suggests that any predictions based on the fitted line within this thickness range will tend to be slightly conservative.

### **5.3.3 Area Extrapolation for Minimum Thickness Estimation (underlying distributions)**

This is achieved by the use of the survivor function (Figure 7). It can be established that the survivor function raised to the power of the ratio of total extrapolated area to inspection area gives the survivor function of the minimum thickness over the extrapolated area. Suppose that the survivor function in Figure 7 represents data from, for example, 1/100 of the total plant area, then in order to estimate the minimum value over the whole area, the probability values in Figure 7 are taken to the power of 100. This produces the distribution of minimum thickness in Figure 16, which shows that the probability that the minimum wall thickness over the whole area being less than 4.6mm is around 10%.

Where this ratio gets very large, the distribution of minimum thickness approximates to an extreme value distribution (see, for example, Gumbel (4))





**Figure 16**  
**Survivor Function of Minimum Thickness over 100 times Area as Figure 7**

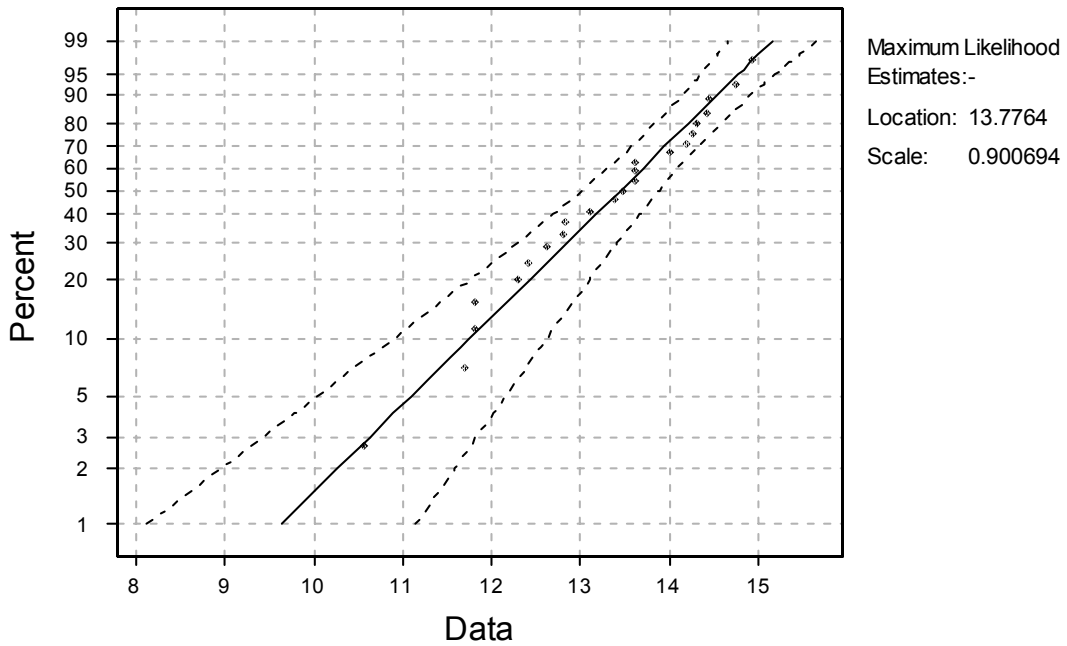
## 5.4 EXTREME VALUE FITTED DATA (TYPE 1 DISTRIBUTION)

### 5.4.1 Probability Plot

If the underlying distribution is an extreme value one, then it is possible that the corrosion is pitting. In this case the probability graph can be used directly to obtain an estimation of minimum thickness or maximum pit depth. The extreme values can be obtained by taking data in patches (either by scanning an area and getting the minimum wall thicknesses in each area, or by collecting all the data from an area, dividing it into patches and using the minima from these patches, or by selecting the maximum pit sizes from a given area).

Figure 17 shows an example. The inspection area or patch was 0.03m<sup>2</sup> and the minimum thicknesses from a number of them are shown as plotted points. The fitted line indicates, for instance, that if a 0.03 m<sup>2</sup> patch is randomly selected from the whole area then there is a 5% probability of the minimum thickness in this patch being less than 11mm. The plotted 95% confidence limits (shown dashed in Figure 17) can also be used as an aid to judgement. In this case, the data show a good fit to an extreme value distribution.

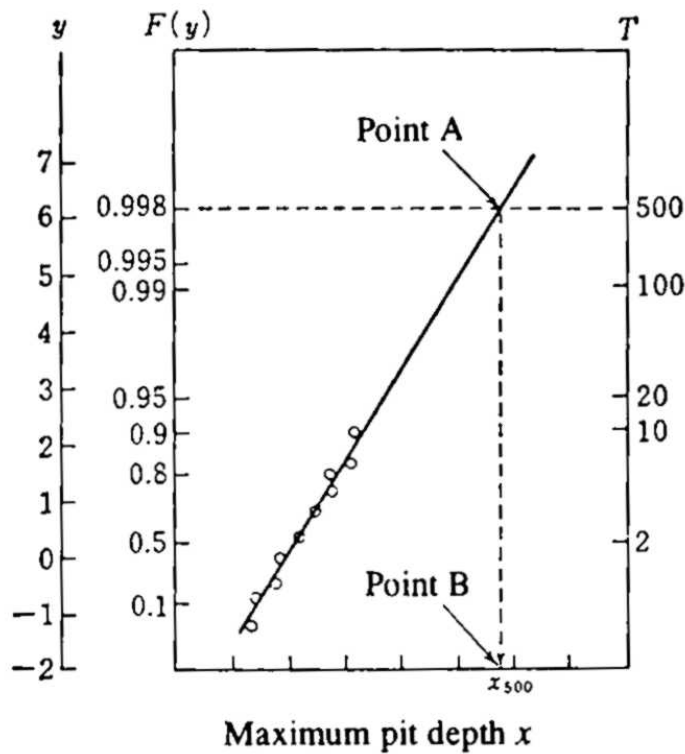
Statistical software can then be used to estimate the location and scale parameters ( $\mu$  and  $\sigma$ ) of the distribution, as shown in the top right of Figure 17.



**Figure 17**  
**Example of Extreme Value Distribution**

#### 5.4.2 Area Extrapolation (Extreme Value Type 1)

The method described by Gumbel gives an estimation of maximum pit size. The maximum pit sizes from a given area are ordered, then plotted on the extreme value probability paper. An example is given in Figure 18. The right hand axis shows the number of data points (i.e. the number of maximum pit depths each of which corresponds to a given area). In this case 10 points have been measured. If we wished to estimate the maximum pit depth for 500 measurements we refer to the 500 on the right hand axis and the line at Point A. Point B corresponds to the minimum thickness over the extrapolated area.

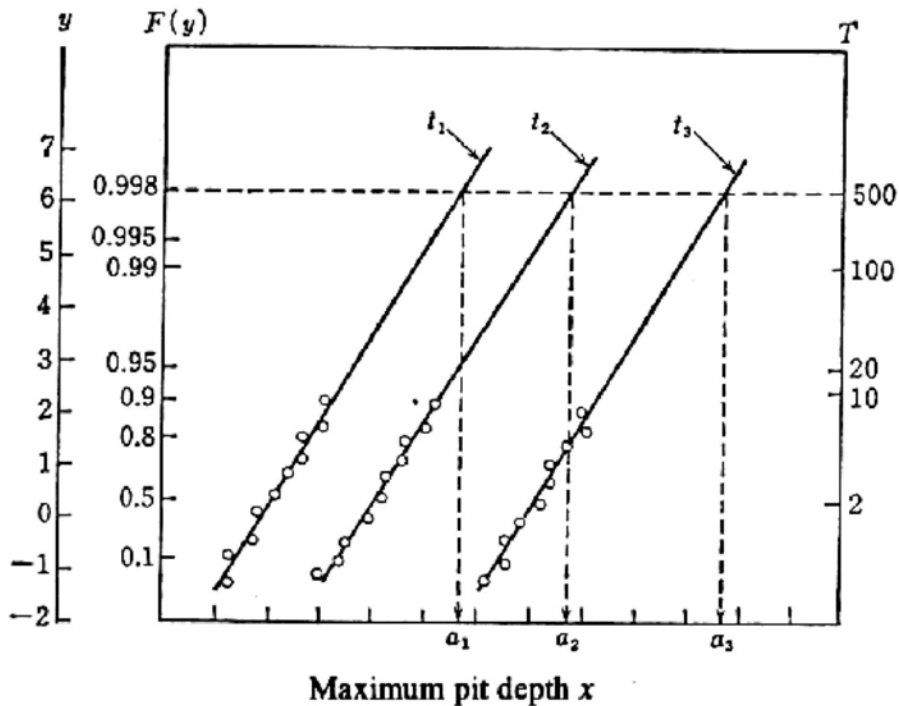


**Figure 18**  
**Return period method of extrapolating over area**  
**(from Kowaka (9))**

Note that this method does not include confidence limits on the estimation, nor does it produce the distribution of minimum thicknesses that is needed for reliability analysis. To achieve these the method described by Schneider et al (13) should be used, and the work of Reiss and Thomas (14) and Laycock (15) from which this is derived also studied.

### 5.5 DETERMINATION OF CORROSION RATE

When a Type 1 Extreme Value distribution has been identified, estimation of corrosion rate can be carried out using the probability plot, as described by Kowaka (9) (Figure 19). In this case it can be seen that the graph shifts across the thickness axis in time. The rate of corrosion can be estimated from the shift. The method works provided that the slope of the graph does not change between inspections. Should that occur, different parts of the component are corroding at different rates. Modelling of the distribution of the corrosion rate in this situation is beyond the scope of this document.



**Figure 19**  
**Corrosion Rate from Extreme Value Probability Plot**  
**(from Kowaka (9))**

## 5.6 NUMBER OF SAMPLES NEEDED (EACH CASE)

### 5.6.1 UT Methods

UT methods fall into three categories for corrosion inspection as mentioned in 5.1 above.

1. Sample readings with thickness gauges
2. Sample area scans with A-Scan instruments for minima within an area
3. Automated scans with recorded readings in a grid.

The choice of UT method depends on cost and the quality and quantity of data required. The cost will increase in particular with the deployment of automated systems, although all methods require the removal of coatings.

It should be pointed out that standard UT measurements are the only ones which will give a direct thickness reading of a small area, and are therefore the only ones suitable for measurement of pitting (although there are some restrictions) (11,12). They are also usually significantly more sensitive to variations in wall thickness (e.g. ~0.1mm) than the alternative methods.

## 5.7 CHOICE OF LOCATION

When using ultrasonic inspection in an unknown situation it is generally preferable to determine at some stage whether pitting corrosion is occurring. This can usually be identified

by an experienced operator with manual equipment, or more positively by a scanning system. The inspection locations should ensure that possible different cases of corrosion are inspected and checked to see if different statistics apply. Each Case corresponds to a potentially different corrosion environment, and may result from differing:

- Materials
- Corrosion product/chemistry
- Temperature
- Flow rate
- Presence of inhibitor
- Component geometry, e.g. orientation, level or clock position (for horizontal pipe)
- Fluid Composition
- Presence of Contaminants

Each Case must be treated separately to determine the sample numbers needed. In principle within those areas the data collection can be at any randomly selected points or areas.

The choice of locations for inspection made on a risk-based or experience-based methodology, targeted at 'hot spots', is not within the scope of this document.

### 5.7.1 Other methods

While this document is primarily associated with the use of standard ultrasonic methods, it is recognised that other methods do exist for corrosion inspection and it may be necessary to consider these as options for planning an inspection. A table of the alternative methods is given below together with their main characteristics and difficulties. For statistical analysis the other methods may sometimes produce thickness measurements which appear similar to UT. It should be noted that techniques which have a large footprint are not capable of providing the detailed information necessary for sampling. For generic statistical reasons, the data is more likely to fit a normal distribution for general corrosion whatever the surface morphology.

Method	Main application area	Possible restrictions
Creeping Head Wave	Under clamp inspection	Coatings / qualitative measurement / lack of internal-external discrimination
Long Range Low Frequency Ultrasonics	Screening / road crossings / large areas / difficult access	Some coatings / sensitive to loss of overall cross section / lack of internal-external discrimination
Pulsed Eddy Current	Under insulation	Large footprint / proximity of other objects / lack of internal-external discrimination Not welds
Magnetic Flux Leakage	Fast screening / some special methods for small diameter pipe	Thickness limitation / internal-external discrimination difficult Not welds
SLOFEC	Fast measurement	Scaling output to thickness Thickness variation signals from other causes
X Ray	Under insulation	Radiation hazard

## 6 CASE STUDY: FAILED OIL PIPE

A 100mm dia. pipe carrying crude oil had been supplied to the project by an operator. This had been removed from service due to a leakage failure. The total length of the pipe in service was 25metres. A sample of the pipe (Figure 19a) was investigated. This gives a total of 16 inspected areas, which is extrapolated to 1250 areas.

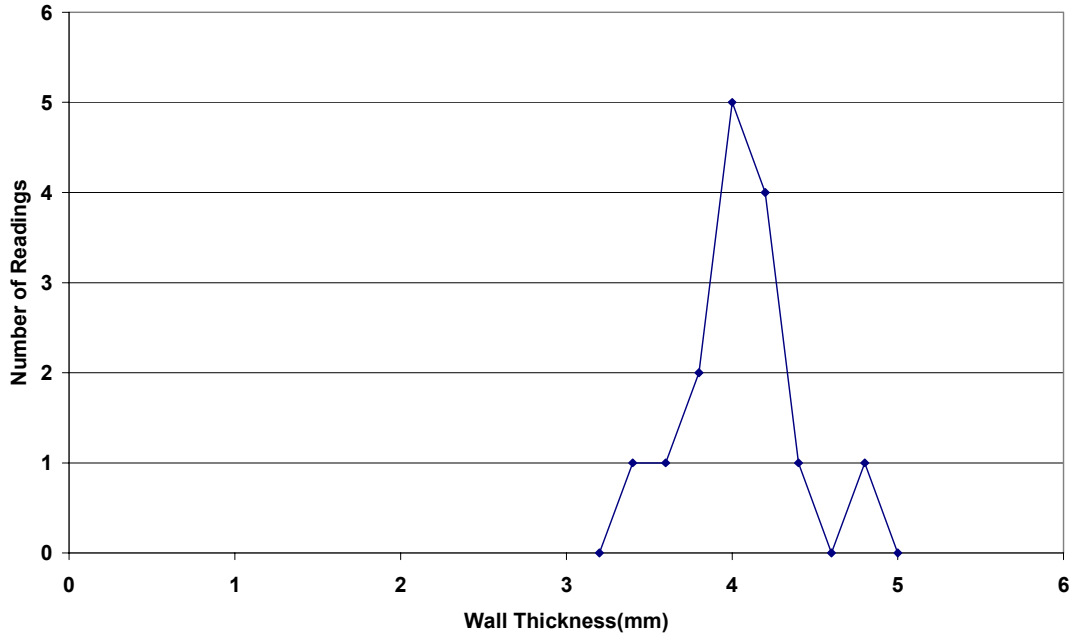


**Figure 19a**  
**Corroded Pipe Sample**

The pipe sample shown was divided into 16 areas and the minimum ultrasonic thickness (measured from the outside) was ascertained in each area by a manual search for the minimum thickness.

The ordered results were:

3.33	4.19
3.63	4.26
3.8	4.34
3.82	4.77
3.82	
3.92	
3.98	
3.98	
4.04	
4.06	
4.1	
4.13	

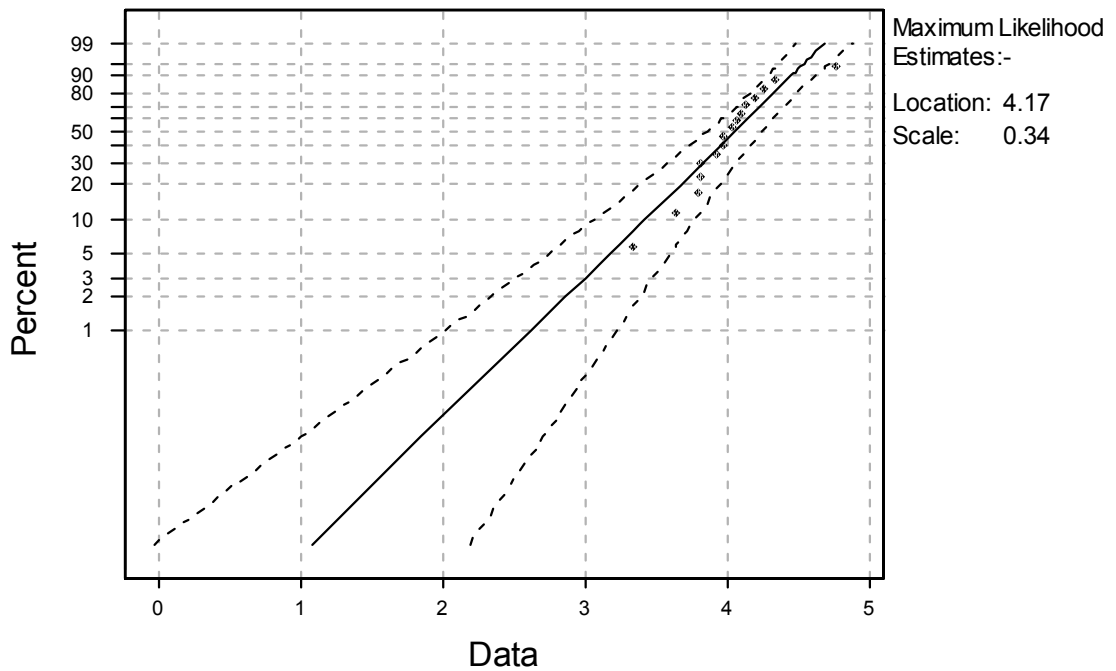


**Figure 20**  
**Frequency Distribution of Extreme Value Corrosion Data from Pipe**

The frequency distribution is shown in Figure 20. The data was plotted on extreme value probability paper as shown in Figure 21.

Note that in this example reproduction of the probability plot of the return period in Section 4.4.2 is not available (this is a function of the software used). The ordinates are plotted according to  $\{1-N/(N+1)\} \times 100\%$ . Therefore the minimum estimated wall thickness for 1250 areas will occur at an ordinate of 0.08% ( $=\{1-1250/1251\} \times 100\%$ ). The scale of this graph does not give this point directly, so it is necessary to refer internally to the software, which gives a central estimate of minimum wall thickness of 1.8mm and a one-sided 95% lower confidence limit of 1mm. Let us assume, for the moment, that leakage occurs once the minimum thickness reaches zero. Then the best estimate of the probability of leakage is 0.5%. However, the one-sided 95% upper confidence limit on this probability is about 9% (partly because of the relatively small sample size).

These values suggest that, once sampling errors are taken into account, and the possibility of inspection errors is also added, there could have been an appreciable probability that leakage would have occurred somewhere within the whole pipe area, which was in fact the case. Certainly it is unlikely, given this result, that an operator would choose to continue using the pipe.



**Figure 21**  
**Extreme Value Probability Plot for Pipe in Figure 19a.**

## 7 CONCLUSIONS

- A literature search has revealed that statistical methods are available for analysis of corrosion data, however these are not widely used, because the methods are not readily understood by engineers, nor are they conveniently available in simplified form.
- Although a standard exists for statistical analysis of laboratory corrosion test data, no such standard exists for the analysis of inspection data relating to corrosion measurement.
- Current data analysis methods for inspection for corrosion exhibit some weaknesses in their application. Statistical analysis can address some of these weaknesses.
- Statistical methods can also be used to extrapolate from the sample inspection data the proportion of a plant component that is below a specified wall thickness, and can be used to estimate the overall probability of leakage in the future.
- For pitting corrosion extreme value analysis is considered most appropriate and extrapolations from a small inspected area to a large area of plant can be made with this method.
- A case study has shown that inspection data from part of a pipe section, which had come from a failed pipe, indicated that there may have been an unacceptably high probability of leakage from somewhere within the pipe.



## **8 RECOMMENDATIONS**

- The application of statistical methods to the analysis of inspection data for corrosion should be encouraged.
- Software tools should be developed to enable easier application of the statistical methods.
- Further validation of the statistical methods should be attempted.

## 9 REFERENCES

- 1 Reliability Assessment for Containers of Hazardous Materials (RACH) Final Report, September 1999. EC Thermie Contract No OG/112/95-FR/UK. Available from TSC Inspection Systems, 6 Mill Square, Featherstone Road, Wolverton Mill, Milton Keynes MK12 5RB.
- 2 Patel R and Rudlin J: 'Analysis of erosion/corrosion incidents in offshore process plant and the implications for NDT'. *Insight* 2000 **42** (1) 17-21.
- 3 ASTM G16 'Standard guide for applying statistics to analysis of corrosion data'.
- 4 Gumbel E J: 'Statistical theory of extreme values and practical applications'. US Dept of Commerce Applied Mathematics series 33 (1954).
- 5 Hawn D E: 'Extreme value prediction of maximum pits on pipelines'. *Materials Performance* 1977 (March) 29-32.
- 6 Manley R B: 'An evaluation of the erosion-corrosion survey program in the Vermilion Block 14 field'. Offshore Technology Conference, Houston, 1972, Offshore Technology Conference, 6200 North Central Expressway, Dallas Texas 75206, paper OTC 1576.
- 7 Joshi N R: 'Statistical analysis of UT corrosion data from floor plates of a crude oil aboveground storage tank'. *Materials Evaluation* 1994 (July) 846-849.
- 8 Sparago M: 'Statistical tools for ultrasonic thickness data analysis, Part 2 – remaining life estimates'. *Inspectioning Journal* 1999 (March/April) 1-4.
- 9 Kowaka M: 'Introduction to life prediction of industrial plant materials – application of extreme value statistical method for corrosion analysis'. Allerton Press Inc, New York, 1994, ISBN 0 89864 073 3.
- 10 Mitsui Babcock: Group Sponsored Project 6748 Non Invasive Inspection Within An Asset Risk Management Strategy, Phase 1 Report. Mitsui Babcock plc, Renfrew, Scotland
- 11 Drury J C: 'Corrosion monitoring and thickness measurement – what are we doing wrong?'. *Insight* 1997 **39** (1) 17-20.
- 12 Porter R: 'Ultrasonic thickness surveys on tankers and bulk carriers'. *Insight* 1997 **39** (1) 21-23.
- 13 Schneider C.R.A, Muhammed A and Sanderson R.M. 'Predicting the Remaining Lifetime of In-Service Pipelines based on Sample Inspection Data'. *Insight* 2001 **43** No 2.
- 14 Reiss R-D and Thomas M: 'Statistical analysis of extreme values'. Birkhäuser Verlag, Basel, 1997.
- 15 Laycock P J, Cottis R A and Scarf P A: 'Extrapolation of extreme pit depth in space and time'. *J. Electrochem. Soc.* 1990 **137** (1) 64-69.

## **10 APPENDIX A CALCULATION OF CORRELATION DISTANCE**

### **10.1 APPLICATION**

The correlation distance needs to be taken into consideration if there is any risk that adjacent data points are related to each other.

The correlation distance is estimated by investigating the dependence between the data points at different distances from each other. This can be done by computing the two-dimensional auto-correlation function (2D ACF), for each area inspected. It is convenient to compute the 2D ACF as the two-dimensional inverse Fourier transform of the two-dimensional power spectrum (invoking the auto-correlation theorem for two-dimensional transforms). The power spectrum is, in turn, computed as the squared magnitude of the Fourier transform of the raw data. Software packages will generally compute these transforms using Fast Fourier transform (FFT) algorithms.

### **10.2 HANNING WINDOW**

Although a raw data set covers a finite area, a minor technical difficulty arises on taking its Fourier transform, because the data set is treated as though it was infinite and periodic, with periods equal to the dimensions of the sample area. This introduces apparent discontinuities at the edges of the sample area, which manifest themselves as spurious high frequency components in the Fourier domain. This is a well-recognised feature of finite Fourier transforms. It is overcome by pointwise pre-multiplication of the data by a so-called 'window function' that gradually drops to zero as the edges of the sample area are approached. A typical choice is to use the Hanning window, which is the product of two simple sinusoidal functions (having periods equal to the dimensions of the sample area).

## 11 APPENDIX B CALCULATION OF INHERENT AND SAMPLING ERRORS

### 11.1 APPLICATION

Consideration of the above errors is necessary to investigate the confidence limits of the data analysis.

Uncertainties in estimating the minimum thickness fall naturally into two categories. Firstly, for (unmeasured) thicknesses drawn from a population of a given 'underlying' distribution, there is a statistical uncertainty in the minimum thickness over that area of uninspected pipe. This uncertainty depends only on the extent of the uninspected region and the form of the underlying distribution. It cannot be reduced by increasing our knowledge of the underlying distribution, by (for instance) inspection of other pipework or by improved understanding of the corrosion process. We shall therefore call this the 'inherent' uncertainty in estimating the minimum wall thickness. Having taken account of this inherent uncertainty, the lower limit on the minimum wall thickness is called a 'confidence limit'.

The second source of uncertainty in predicting the minimum thickness is due to the so-called 'sampling errors' in estimating the parameters of the underlying distribution, based on a sample inspection. If the sample inspection comprises  $n$  statistically independent measurements, then, in most cases, the sampling errors are approximately proportional to  $1/\sqrt{n}$ . Once this extra sampling error is accounted for, the lower limit on the lower confidence limit on the minimum wall thickness is called a 'tolerance limit'.





**MAIL ORDER**

HSE priced and free  
publications are  
available from:

HSE Books  
PO Box 1999  
Sudbury  
Suffolk CO10 2WA  
Tel: 01787 881165  
Fax: 01787 313995  
Website: [www.hsebooks.co.uk](http://www.hsebooks.co.uk)

**RETAIL**

HSE priced publications  
are available from booksellers

**HEALTH AND SAFETY INFORMATION**

HSE InfoLine  
Tel: 08701 545500  
Fax: 02920 859260  
e-mail: [hseinformationservices@natbrit.com](mailto:hseinformationservices@natbrit.com)  
or write to:  
HSE Information Services  
Caerphilly Business Park  
Caerphilly CF83 3GG

HSE website: [www.hse.gov.uk](http://www.hse.gov.uk)

**RR 016**

**£10.00**

ISBN 0-7176-2554-0



9 780717 625543