Malaysian Journal of Civil Engineering 20(1): 47 - 57 (2008)

THE EFFECT OF EXTREME CORROSION DEFECT ON PIPELINE REMAINING LIFE-TIME

Norhazilan Md Noor, Nordin Yahaya, Siti Rabe'ah Othman

Reliability Engineering & Safety Assessment Research Group (RESA), Department of Structures and Materials, Faculty of Civil Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor

**Corresponding author*: Norhazilan@utm.my

Abstract: Inspection data obtained from in-line inspection can be used to assess present integrity as well as to predict future integrity of corroding pipeline by using a combination of probabilistic analyses and simulation process. However, numerical errors arise when all the inspection data including non-extreme and extreme measured defects were used. In this study, a combination of an extreme value distribution and peaks-over-threshold method was adopted to analyse the effect of extreme data upon pipeline reliability. The paper focuses on the elimination of the "low-risk" data in the analysis in order for extreme values to be significantly emphasized. The study found that the high threshold value would lead to the high failure probability. The optimum threshold value is constrained by the number of remaining data. The selection of threshold and extreme distribution value must be cautiously carried out to minimise the possibilities of over predict the pipeline's time to failure.

Keywords: *Peak-Over-Threshold, extreme data, corrosion, pipelines, probabilistic*

1.0 Introduction

This paper deals with the effect of extreme data upon the current pipeline reliability, i.e. the estimated remaining lifetime subject to internal corrosion. A prediction of pipeline's time to failure using Monte Carlo simulation method data has been carried based on different threshold (Peak-Over-Threshold Method) selection and extreme distribution. The parent distribution also commonly known as the actual or initial distribution covers the whole matched data whereas the extreme distribution produced from parent distribution only covers data under tail region. Based on the 'weakest link' principle, extreme corrosion data can provides more precise information required for prediction of structural integrity. The extreme corrosion pit is more likely to give greater likelihood of perforation through pipelines wall thickness in the future. Thiruvengadam [1972] explained that the corrosion resistance of a particular element, denoted as pipelines in this research, distinctly determined by the largest pits and the largest concentrations of chemical agents. Moreover, small and intermediate pits and concentrations do not affect the corroded strength of the element (pipelines). Thus, a consideration of extreme data is necessary in order to evaluate the structural integrity more accurately.

In general, the modeling of extreme data from its parent distribution is always mislead by the deficiency of low rank data (low-risk data) under tail region on the left side of the distribution (Figure 1). Applying to the case in the present study, these numbers of low rank metal loss data is hardly measurable by inspection tools owing to limitation of tool resolution or the pits covered by wax, hence not visible during inspection. The uncertainties that govern the quality of low rank data may jeopardize the reliability of the chosen distribution for metal loss data especially on the right tail region in which extreme data are located. The research was carried out to determine the effect of threshold selection upon pipeline failure probability. The threshold or cut-off point was proposed to eliminate the low rank data in order to centralize the focus towards extreme data. By eliminating the low rank data which are theoretically governed by uncertainties, the modeling of extreme data can be properly done.

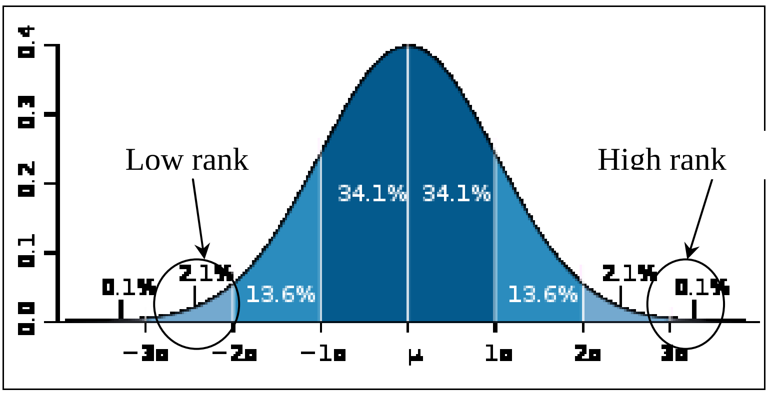

Figure 1: The location of low rank and high rank data under probability distribution line

2.0 Techniques to Eliminate Low Rank Data

The research applies two techniques namely Peaks-Over-Threshold (POT) and Extreme value theory to improve the modeling of extreme data by eliminating the low rank data from its parent distribution. The purpose of this elimination is to focus more on the high rank data from the centre of the distribution to the right tail and to increase the probability of the extreme data being randomly selected by computer program during simulation process without bias.

2.1 Peaks-Over-Threshold Method

Principally, POT is used to extract extreme values from data by taking specific exceedence over higher thresholds. The peaks-over-threshold (POT) method is also known as the Partial Duration Series method, and is generally used in reliability analysis to estimate out-of-sample quantiles (observed data) or extreme value beyond the largest available observation [Onoz and Bayazit, 2001 ; Falka and Reissb, 2001]. Many practical studies involve the estimation of values beyond the largest or smallest observation (the so-called extreme values) from collected data in order to measure the probability of rare events. Several studies have used POT to estimate extreme quantiles. For example Smith [1991] and Moon *et al*. [1993] for applications in Hydrology, Coles and Tawn [1996] for applications in Meteorology, Emmer *et al*. [1998], Longin and Solnik [2001] and Embrechts *et al*. [1997] for applications in Finance and Yahaya [1999] particularly for pipeline assessment using pigging data.

2.2 Extreme Value Theory

Extreme value theory was exercised to limit the selection of corrosion depth from its parent distribution (Figure 2). This is to ensure the extreme depth data are properly selected during simulation. The extreme value distribution of corrosion depth is produced from its parent distribution (initial or actual distribution) depending on the total numbers of the observed variable, *n*. Based on the general domain equation of Cumulative distribution function (CDF) of actual Weibull distribution for instance, the CDF of extreme Weibull distribution can be derived from Equation 1 as follows:

$$F_{Yn}(x) = \left[1 - \exp\left[-\left(\frac{x-\delta}{\theta}\right)^{\beta}\right]\right]^{n} \quad (1)$$

By differentiating the CDF of the Weibull extreme distribution, the PDF of this extreme distribution can be calculated as:

$$f_{Yn}(x) = n\left[F_{Yn}(x)\right]^{n-1}\left[\frac{\beta(x-\delta)^{\beta-1}}{\theta}\exp\left[-\left(\frac{x-\delta}{\theta}\right)^{\beta}\right]\right] \quad (2)$$

or:

$$f_{Yn}(x) = n\left[F_{Yn}(x)\right]^{n-1} f_{x}(x) \quad (3)$$

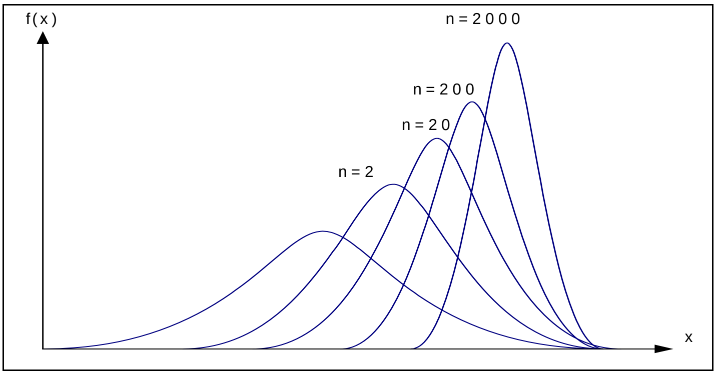

Figure 2: Parent and Extreme Normal distribution

3.0 Reliability Assessment Methodology – The Monte Carlo Simulation

The determination of time to failure in this study is based on the exceedence of a chosen annual target probability. A simulation program based on the Monte Carlo simulation procedure has been written using language software MATLAB. The selection of the Monte Carlo simulation for the assessment procedure is based on its capability in estimating failure probability from various types of distributions. An evaluation of pipeline reliability subjected to extensive internal corrosion is carried out for current and future conditions. The future growth of corrosion defects is based on linear model for simplicity sake due to lack of information pertaining to environmental and material properties which might govern the growth rate of the defect.

A POT was applied on corrosion depth data so that only a proportion of samples at the end tail above a certain threshold level were considered during simulation analysis. The inspection data is obtained from pigging inspection on hydrocarbon pipelines located in North Sea area. The fitting in the tail region is important as the extreme value distribution depends only on the tail region [Wolfram and Yahaya, 1999; Yahaya and Wolfram, 1999]. The threshold levels range from a depth to wall thickness ratio (*%wt*) of 5*%wt* (original threshold) to 35%*wt* with the number of extreme data remaining after cut-off has becoming smaller as the threshold level is increased. The most suitable threshold level depends on the amount of data left after the cut-off has been applied. Yahaya and Wolfram [1999] suggested that the selection of threshold value depends on the severity of the data and the appropriate number of samples. A threshold over 30%*wt* was proposed with a minimum number of samples of 200 data based on the consistency of pipeline failure probability [1999].

3.1 Statistical Parameters

The distribution parameters required in the simulation process have been verified using a probability plot and Chi square goodness-of-fit test. The material properties were provided by the manufacturers and pipeline operators, while the defect parameters were defined from the statistical and probabilistic analysis on real inspection data.

3.2 Material Properties

Table 1 shows the complete statistical parameters of the pipeline material properties. The pipelines are made from hydrocarbon steel. All of the material properties are assumed to follow a Normal distribution.

Table 1: Statistical value of material properties

Variables	***Distribution***	***Statistical Values***	
		Mean	**%COV**
Pipe Diameter, (*mm*)	Normal	914.2	5
Pipe Thickness (*mm*)	Normal	22.2	5
Specified minimum yield stress, SMYS (*MPa*)	Normal	459	5.6
Specified Ultimate tensile stress, SUTS (*MPa*)	Normal	573	5

3.3 Defect Properties

Defect properties represents corrosion depth, corrosion length and corrosion growth rate. In current practices, corrosion depth of less than 10% of wall thickness (*%wt*) is technically assumed not to have a significant effect on the pipeline integrity [Batte *et al*., 1997]. The smallest reported corrosion depth the pipelines was 5*%wt* based on the applied threshold by the inspection company. There is no threshold level applied to limit the selection of extreme data except the 5*%wt* threshold set by inspection vendor. Corrosion length is assumed independent of corrosion depth, hence a separate calculation of corrosion growth rate has been carried out. Table 2 lists the value of defects properties.

Table 2: Corrosion defect properties

Variable	**Unit**	**Distribution**	**Mean**	**COV%**
Depth, d_{B95}	mm	Extreme Weibull	6.950	28.67%
Length. L_{B95}	mm	Exponential	14.493	100%
Corrosion rate (length)	mm/year	Normal	0.656	33%
Corrosion rate (depth)	mm/year	Normal	0.174	33%

3.4 Failure Models

To estimate the probability of failure based on the residual pressure of corroded pipelines, the recommended practice of DNV RP-F01 model has been applied. However, the partial safety factors adopted in these equations were not used in the present work. The partial safety factor in these equations is generally used to replace the uncertainties with one single value. Instead, the existing uncertainties associated with each estimated variable are taken into account by using the simulation technique. The variation of each parameter from its corresponding distribution clearly covered the uncertainties.

DNV RP-F101-based failure model [DNV, 1999]

$$P_p = \frac{2tSMTS\left(1-\frac{d}{t}\right)}{(D-t)\left(1-\frac{d}{t}\right)Q^{-1}} \tag{4}$$

where:

D	=	outer diameter
d	=	depth of corrosion defect
L	=	measured length of corrosion defect
P_p	=	maximum allowable operating pressure
Q	=	length correction factor
$SMTS$	=	specified minimum tensile strength
t	=	nominal pipe wall thickness

3.5 Limit State Function

To describe the failure condition of the pipeline system, the limit state expression is based on the allowable and applied operating pressure as expressed by the equation 5. The primary failure mode is bursting under internal pressure. Bursting can be defined as the point at which uncontrolled tearing of the pipe wall occurs [Hellevik and Langen, 2000*a* and 2000*b*]. This is likely to occur at a section in the line weakened by corrosion.

The failure is preceded by local bulging accounted for in the models by the Folias bulging factor that involves terms in both the defect depth and length.

$$G(x) = P_p - P_a \tag{5}$$

where:
P_a = maximum applied fluid pressure
P_p = calculated allowable pressure using a suitable failure model equation.

Pipelines are considered to fail if the estimated maximum pressure capacity, P_p is less than the applied pressure, P_a. If this happens, the limit state failure, *G(x)* will be either zero or negative. If, however, the pipeline is operated below the estimated maximum pressure, a positive value of *G(x)* will define a safe condition.

3.6 Limit State Failure

The target reliability, otherwise known as limit state failure, is the maximum acceptable failure probability level for a particular limit state [Melchers, 1987]. Thus, the probability of exceedence for each relevant limit state can be estimated from existing databases of pipeline failures. A suitable target reliability level for this case study was taken in range of 1×10^{-4} to 1×10^{-3} pertaining to the Normal safety class and for the Ultimate limit state. The reliability target of 1×10^{-3} has been chosen to determine the failure probability. This annual target probability was included in the Submarine Pipeline Reliability Based Design Guideline known as The SUPERB Project [Jiou *et al.*, 1997; Sotberg *et al.*, 1997]. It was also adopted in the DNV RP-F101 corroded pipeline assessment procedure.

3.7 Effects of Threshold Selection on Pipeline Reliability Assessment

Five different thresholds have been chosen to threshold the data including the original threshold applied by the inspection operator. The selected thresholds are 5%*wt*, 20%*wt*, 25%*wt*, 30%*wt* and 35%*wt* which are similar to the selection of thresholds for pigging data done by Yahaya [1999]. From the POT analysis, it is found that the data distribution approaches the Exponential distribution (when the shape parameter, β approaches one) as the threshold value increases (see Table 3). The Weibull probability plot indicates straight fitting for all thresholds even though the amount of remaining data reduces as the threshold value is increased. Five simulations to predict the future remaining life-time of corroding pipeline have been carried out based on the different threshold selections.

Table 3: Threshold level of corrosion depth

Numbers of data, *n*	**Threshold (*%wt*)**	**Weibull Parameter**	
		β	θ
627	All data (>5%)	1.931	4.440
268	>20%	1.268	3.022
195	>25%	1.108	2.290
47	>30%	1.087	1.859
90	>35%	1.001	1.413

4.0 Simulation Results

Based on the result presented in Figure 3, it is apparent that different thresholds will give a variation of failure probability of pipelines upon corrosion attack. The most conservative results are the prediction based on a threshold of 20*%wt*, in which the pipelines exceeds the limit state function in year 1997. Thresholds over 25*%wt* and 35*%wt* give almost similar result with time to failure predicted in the middle of year 1998. For thresholds over 30*%wt* and 35*%wt*, the remaining data are quite small and possibly not reliable for prediction purposes due to the fact that a small sample will increases the level of uncertainties in fitting the data into the probability distribution.

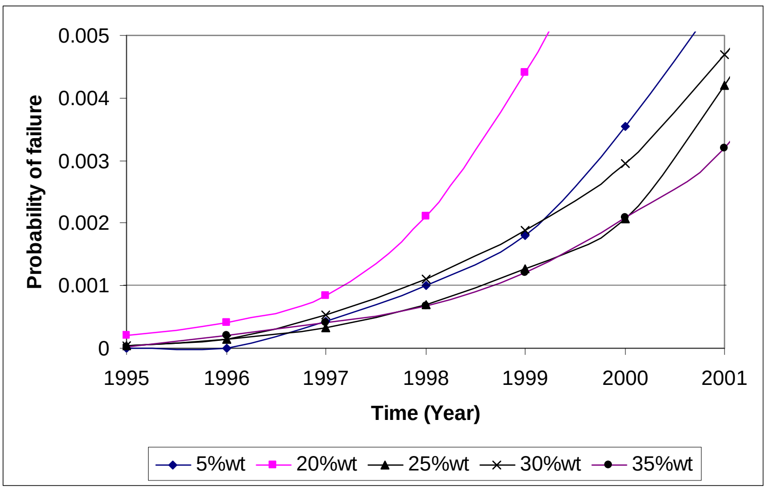

Figure 3: The effects of different threshold selections upon probability of failure using Monte Carlo simulation

Simulations for all threshold values give a prediction of failure probability between years 1997 and 1999 which indicates that the Peak-over-threshold method does predict a different time to failure for different thresholds but that the variation is relatively small. There is no definite way to prove which threshold selection is the best and give the most accurate prediction at this stage because of the limited availability of repeated inspection data Nevertheless, with only a 2 year variation between predicted times to failure, the effect of selecting data over a certain threshold at this point is not significant.

5.0 Discussion and Conclusions

Theoretically, the higher the threshold, the better the representation of the extreme data since most of the low rank data has been eliminated using POT method. However, the elimination of low rank data can as well reduce the overall numbers of remaining data, *n*. If *n* is too small, this may affect the selection of extreme data. Based on the extreme value theory, high value of *n* is required so the likelihood of selecting randomly extreme data from its parent distribution is greater. The elimination of low rank data based on high cut-off value (threshold) may not benefit the extremity of analysis results since the severity of corrosion data is reduced by the low numbers of remaining data, *n*. Moreover, the fitting of data to the hypothesized distribution is becomes more intricate when the remaining data, *n* are in small quantity. The so called 'cross-over' effect is clearly depicted in Figure 3. The failure probability of pipelines for threshold values of 25%*wt* and above are supposed to be lower or at least remain the same than the case study with threshold of 5%*wt* and 20%*wt* due to the fact that the higher the threshold, the higher the extremity of the data. Instead, the failure probability becomes less conservative owing to the effect of low *n* value.

There is no definite way of fitting a probability distribution to long-term data for prediction purposes. However, the Peak-Over-Threshold (POT) method can be considered as an alternative [Ferreira and Guedes Soares, 1998]. Even if the POT method is used in practice to extrapolate outside the observation range, no theoretical result can guarantee the quality of the estimation. One way of avoiding the problem of fitting the whole distribution is to concentrate on the tails. It is clear from the simulation results that different threshold values of corrosion depth distribution produced a variability of time to failure. However, the selection of an appropriate threshold level is difficult to determine. As a guide, the selection of threshold level depends on the numbers of the remaining data. Yahaya and Wolfram [1999] suggested that as a general guidance, the selection of threshold value adopted in POT analysis should be at an intermediate range, where it is supposed to be neither too low to include the immaterial data nor too high to exclude the important data and consider the small samples in the analysis. Based on the reliability assessment of corroding pipelines, this research suggested that the original sample be used for assessment purposes, since the gap between times to failure of corroding pipelines based on a different selection of thresholds was relatively small.

Acknowledgement

The authors would like to express deep gratitude to MOSTI, MOHE and UTM for providing the research group with fund and facilities to facilitate the research activities.

References

Batte A.D., Fu B., Kirkwood M.G. and Vu D. (1997), *New Methods for Determining the Remaining Strength of Corroded Pipelines*, Pipeline Technology, V, pp. 221-228.

Coles S. and Tawn, J. (1996), *A Bayesian Analysis of Extreme Data Rainfalls*, Applied. Statistic, 45(4), pp. 463–478.

DNV (Det Norske Veritas) (1999), *Recommended Practice RP-F101 for Corroded Pipelines* 1999, Norway: Det Norske Veritas.

Embrechts P., Kluelberger C. and Mikosch T. (1997), *Modelling Extreme Events*. Berlin: Springer. 10-36.

Emmer S., Kluelberger C. and Trustedt M. (1998), *VAR-A Measure of the Extreme Risk*, Solutions, 2, pp. 53–63.

Falka M. and Reissb R.D (2001), *Estimation of Canical Dependence Parameters in a Class of Bivariate Peaks-Over-Threshold Models*, Statistics & Probability Letters, 52, pp. 233– 242.

Ferreira J.A. and Guedes Soares C. (1998), *An Application of the Peaks Over Threshold Method to Predict Extremes of Significant Wave Height*, Journal of Offshore Mechanics and Arctic Engineering, 120(3), pp.165–176.

Hellevik S.G. and Langen I. (2000a), *In-Service Inspection Planning of Flow lines Subjected to CO_2 Corrosion*, *Seattle*, USA: ISOPE 2000, Proceedings of the 10th International Offshore and Polar Engineering Conference, May 28-June 2. Vol IV, pp. 372-379.

Hellevik S.G. and Langen I. (2000b), *Optimal Design-Phase Inspection and Replacement Planning of Pipelines Subjected to CO_2 Corrosion*, International Journal of Offshore and Polar Engineering, 10(2), pp. 123-130.

Jiou G., Sotberg T., Bruschi R. and Igland R.T. (1997), *The SUPERB Project-Line pipe Statistical Properties and Implications in Design of Offshore Pipelines*, ASME Pipeline Technology, V, pp.45-51.

Longin F. and Solnik B. (2001), *Extreme Correlation Structure of International Equity Markets*, Journal of Finance, December Issues, pp.20-23.

Melchers R.E (1987), *Structural Reliability Analysis and Prediction*. 2nd Edition, Chichester, UK: Ellis Herwood.

Moon Y., Lall U. and Bosworth K.(1993), *Comparison of Tail Probability Estimators for Flood Frequency Analysis*, Journal of Hydrology, 151(2-4), pp.343 –363.

Onoz B. and Bayazit M. (2001), *Effect of the Occurrence Process of the Peaks Over Threshold on the Flood Estimates*, Journal of Hydrology, 244. pp.86-96.

Smith, J. (1991), *Estimating the Upper Tail of Good Frequency Distributions*, Water Resources, 23(8), pp.1657–1666.

Sotberg T., Moan T., Bruschi R., Guoyang J. and Mork K.J. (1997), *The SUPERB Project , Recommended Target Safety Levels for Limit State Based Design of Offshore Pipelines*, ASME Pipeline Technology, V, pp.71-77.

Thiruvengadam, A. (1972), *Corrosion Fatigue at High Frequencies and High Hydrostatic Pressures*. Stress Corrosion Cracking of Metals-A State of the Art, Philadelphia: ASTM Special Technical Publication.

Wolfram J. and Yahaya N. (1999), *On the Effect of Some Uncertainties on the Structural Integrity Assessment of Corroding Pipelines*, St Johns, Newfoundland, Canada: Proceedings of 17th International Conference Offshore Mechanics and Arctic Engineering. Paper OMAE-99-6032, 11-16th July, pp.1-10.

Yahaya N. (1999), *The Use of Inspection Data in the Structural Assessment of Corroding Pipelines*, Heriot-Watt University, Edinburgh: PhD Thesis.

Yahaya N. and Wolfram, J. (1999), *The Application of Peaks-Over-Threshold Approach on the Structural Reliability of Corroding Pipeline*, Jurnal Teknologi, 30, pp.51-68.