APPLICATION OF RECEIVER OPERATING CHARACTERISTIC CURVE FOR PIPELINE RELIABILITY ANALYSIS

Kong Fah Tee[1*], Lutfor Rahman Khan[1] and Tahani Coolen-Maturi[2]

[1]Department of Civil Engineering, University of Greenwich, UK

[2] Durham University Business School, Durham University, Durham, UK

ABSTRACT

Structural reliability analysis of buried pipeline systems is one of the fundamental issues for water and wastewater asset managers. Measuring the accuracy of a reliability analysis or a failure prediction technique is an effective approach to enhancing its applicability and provides guidance on selection of reliability or failure prediction methods. The determination of threshold value for a particular pipe failure criterion provides useful information on reliability analysis. However, this threshold value is not always known. In this paper, Receiver Operating Characteristic (*ROC*) curve has been applied where empirical and Nonparametric Predictive Inference (*NPI*) techniques are used to evaluate the accuracy of pipeline reliability analysis and to predict the failure threshold value. Multi-failure conditions, namely, corrosion induced deflection, buckling, wall thrust and bending stress have been assessed in this paper. It is hoped that choosing the optimal operating point on the *ROC* curve which involves both maintenance and financial issues, can be ideally implemented by combining the *ROC* analysis with a formal risk-cost management of underground pipelines.

Keywords: Receiver Operating Characteristic (*ROC*) curve; Nonparametric Predictive Inference (*NPI*); Reliability analysis; Underground pipelines; Corrosion.

*To whom correspondence should be addressed. Email: K.F.Tee@gre.ac.uk

Preprint submitted to Journal of Risk and Reliability

1. INTRODUCTION

Many challenges have been faced by water and wastewater industry during installation and maintenance of underground pipelines. The most common challenges are found as various pipe failure modes under loading, poor design detailing and installation practices, insufficient corrosion protection procedures, pipe material deterioration, scouring underneath the ground level, frost heave action and insufficient understanding of product limitations. In reality, a buried pipe's mechanical strength begins to decrease as soon as it is installed because of the environmental conditions surrounding the pipe [1-2]. For buried metal pipelines subject to both corrosion and external loading, a vital failure criterion is the loss of structural strength which is influenced by localised or overall reduction in pipe wall thickness. Due to their low visibility and lack of proper information regarding underground pipes condition, assessment and maintenance are frequently neglected until a disastrous failure occurs. The long-term planning for renewal of underground pipe distribution networks requires the ability to predict system reliability as well as assess the economic impact with good accuracy [3-6].

Structural reliability analysis of buried pipeline systems is one of the fundamental issues for water and wastewater asset managers. Methods of reliability analysis such as first order reliability method, second-order reliability method, point estimate method, Monte Carlo simulation, gamma process, probability density evolution method, subset simulation, dynamic reliability, etc. are available in literature [7-13]. The correlation coefficients between different failure modes show that all the failure modes are strongly correlated positively, i.e., where the failure modes might happen concurrently within a buried pipeline system [14]. The determination of threshold value for a particular pipe failure mode provides useful information on reliability analysis. However, this threshold value is not always known. When the actual value of pipe condition (such as deflection, buckling, bending, etc.) is greater than the threshold value or allowable limit, then this indicates a failure condition and if the actual value is smaller than the allowable limit, then it indicates a non-failure condition. However, in reality, pipelines may not follow the predicted pipe conditions and failure criteria which are estimated according to the proposed models. Gustafson and Clancy [15], Kettler and Goulter [16], Mailhot et al [17] showed that there were 10% to 20% discrepancies in the actual and the estimated pipe conditions measured by available models such as Cox's proportional hazards model, Weibull and exponential distributions, etc.

Classical reliability theory and methodologies rarely consider the actual state of a pipe system and therefore, these are not capable to reflect the dynamics of runtime systems and failure processes. Conventional methods are typically useful in design and prediction of long term pipe behaviour. However these are not good enough in pipe reliability evaluation with good accuracy. Measuring the accuracy of a pipe reliability analysis technique is an effective approach to enhancing its applicability and provides guidance on selection of reliability or failure prediction methods. One of the accuracy measurements for assessment methods is Receiver Operating Characteristic (*ROC*) curve which is a statistical approach with concepts like sensitivity and specificity to express the accuracy.

ROC curve has been commonly used for describing the performance of medical tests for parametric and non-parametric analysis. The *ROC* curve has also been used in many other areas, such as signal detection, radiology, machine learning, data mining and credit scoring [18-20]. In recent years, Nonparametric Predictive Inference (*NPI*) has been developed as an alternative and frequent statistical framework method based on few modelling assumptions and considers one or more future observations instead of a population [21]. It is a statistical method based on Hill's assumption [22], which gives direct probabilities for a future observable random quantity, given observed values of related random quantities [23]. *NPI* uses lower and upper probabilities for uncertainty quantification and has strong consistency properties within theory of interval probability [21]. From a statistical perspective, *NPI* is defined as a plot of results as true positive fraction (*TPF*) or sensitivity along y coordinate versus false positive fraction (*FPF*) or its 1-specificity along x coordinate. Normally, *ROC* curve is useful in evaluating the discriminatory ability of an analysis, finding optimal cut-off point and comparing efficacy of two or more assessment or tests results.

The authors Debon et al [24] and Arian et al [25] conclude by identifying a knowledge gap and research possibilities, mainly relating to data collection and how to best use the existing data for the development and calibration of predictive deterioration models, risk assessment methods, etc. In this study, a *ROC* curve has been applied in buried flexible metal pipeline network where classical (or empirical) and Nonparametric Predictive Inference (*NPI*) technique are used for assessing the accuracy of failure prediction and identifying failure-prone situations, i.e. the threshold value for different pipe failure modes. The multiple time-dependent failure modes for underground flexible metal pipelines, namely, corrosion induced

deflection, buckling, wall thrust and bending are considered. The loss of structural strength is due to corrosion through reduction of pipe wall thickness which then leads to pipe failure. Pipe wall thickness is considered as a key random variable and Monte Carlo simulation has been applied to generate the thickness data based on pipe material and soil parameters.

The contents of this paper are structured as follows. In Section 2, the formulations for pipe failure modes of corrosion induced deflection, buckling, wall thrust and bending are presented. The basic of *ROC* curve is studied in Section 3, where classical *ROC* and *NPI* for *ROC* curve are briefly discussed. In Section 4, a numerical example is considered for underground pipeline reliability prediction using *ROC* curve. The results and discussion are presented for different failure modes in Section 5. Finally, some conclusions are made on the basis of outcomes from this study in Section 6.

2. PIPE FAILURE MODES

For a buried pipe structure, the number of potential failure modes is very high. This is true in spite of the simplifications imposed by assumptions such as having a finite number of failure elements at given points of the structure and only considering the proportional loadings. It is, therefore, important to have a method by which the most critical failure modes can be identified. The critical failure modes are those contributing significantly to the reliability of the system. In this paper, the dominating failure criteria of flexible pipes are characterised by limit states as follows:

a) Excessive deflection;
b) Actual buckling pressure greater than the critical buckling pressure;
c) Actual wall thrust greater than critical wall thrust;
d) Actual bending stress greater than the allowable stress

The failure modes adopted here are due to loss of structural strength of pipelines and these failure criteria are influenced by corrosion through reduction of the pipe wall thickness over time.

2.1 *Corrosion of metal pipes*

Buried pipes are made of plastic, concrete or metal, e.g. steel, galvanized steel, ductile iron, cast iron or copper. Plastic pipes tend to be resistant to corrosion. Damage in concrete pipes can be attributed to biogenous sulphuric acid attack [26-27]. On the other hand, metal pipes are susceptible to corrosion. Metal pipe corrosion pit is a continuous and variable process. Under certain environmental conditions, metal pipes can become corroded based on the properties of the pipe, soil, liquid properties and stray electric currents. The corrosion pit depth can be modelled with respect to time [28-29] as shown in Eq. (1).

The corrosion pit depth, $D_T = kT^n$ (1)

where D_T is corrosion pit depth, T is exposure time and k and n are corrosion empirical constants which are determined from experiments and/or field data.

Due to reduction of wall thickness given by Eq. (1), the moment of inertia of pipe wall per unit length, I and the cross-sectional area of pipe wall per unit length, A_s can be defined as follows [30-31].

$$I = (t - D_T)^3/12 \text{ and } A_s = t - D_T \quad (2)$$

where t is wall thickness of pipe.

Deflection

The performance of a flexible pipe in its ability to support load is typically assessed by measuring the deflection from its initial shape. Deflection is quantified in terms of the ratio of horizontal (or vertical) increased diameter to the original pipe diameter. Normally, the allowable deflection for flexible pipe is 5% of its internal diameter [32]. The actual deflection for flexible pipes Δ_y can be calculated as follows [33].

$$\Delta_y = \frac{K_b(D_L W_c + P_s)D}{\left(\frac{8EI}{D^3} + 0.061E'\right)} \quad (3)$$

where K_b is deflection coefficient, D_L is deflection lag factor, D is mean diameter = $D_i + 2c$, D_i is inside diameter and c is distance from inside diameter to neutral axis, W_c is soil load, P_s is live load, E is modulus of elasticity of pipe material and E' is modulus of soil reaction.

Buckling Pressure

Buckling is a premature failure in which the structure becomes unstable at a stress level that is well below the yield strength of structural material [8]. The actual buckling pressure should be less than the critical buckling pressure for the safety of structure. The actual buckling pressure, p and the allowable buckling pressure, p_a can be calculated as follows [34].

$$p = R_w \gamma_s + \gamma_w H_w + P_s \tag{4}$$

$$p_a = \sqrt{\left(32 R_w B' E' \frac{EI}{D^3} \right)} \tag{5}$$

where R_w is water buoyancy factor = 1 – 0.33 (H_w/H), γ_s is unit weight of soil, γ_w is unit weight of water, H_w is height of groundwater above the pipe and B' is empirical coefficient of elastic support.

Wall Thrust

Wall thrust or wall stress on a pipe wall is determined by the total load acting on the pipe including soil arch load W_A, live load P_s and hydrostatic pressure P_w as shown in Eq. (6). Two wall thrust analyses are required: (a) accounts both the dead load and live load and employs the short term material properties throughout the procedure, (b) accounts only the dead load and employs the long-term material properties. Then, the most limiting value is used for wall thrust analysis [32, 35]. The actual wall thrust, T and the allowable wall thrust, T_a can be calculated as follows.

The actual wall thrust, $T = 1.3(W_A + P_S C_L + P_w)(D_0 / 2)$ (6)

where D_0 is outside diameter and C_L is live load distribution coefficient.

The allowable wall thrust, $T_a = F_y A_s \phi_p$ (7)

where F_y is the minimum tensile strength of pipe and ϕ_p is capacity modification factor for pipe.

Bending Stress

Under the effect of earth and surface loads, the buried pipe may bend through pipe wall. The allowable bending stress for flexible pipes is longitudinal tensile strength of pipe material. The bending stress is important to ensure that it is within material capability. Excessive bending will cause the pipe wall to collapse. The actual bending stress σ_b can be calculated as follows [32].

Bending stress, $\sigma_b = 2D_f E\Delta_y y_0 / D^2$ (8)

where D_f is shape factor, y_0 is distance from centroid of pipe wall to the furthest surface of the pipe and Δ_y is the pipe deflection which can be calculated from Eq. (3).

3. BASIC OF *ROC* CURVE

ROC curves are two-dimensional graphs that visually depict the performance and performance trade-off of a classification model [36]. *ROC* curves are originally designed as a tool to distinguish between the actual results and analytical results. Sensitivity and specificity, which are defined as the number of true positive decisions (the number of actually positive cases) and the number of true negative decisions (the number of actually negative cases), respectively, constitute the basic measures of performance of *ROC* curve. A *ROC* curve displays the full picture of trade-off between the true positive fraction (*TPF*) or sensitivity and false positive fraction (*FPF*) or 1 – specificity across a series of cut-off points. Area under the curve is considered as an effective measure of inherent validity of an analysis or experimental result. It is a very powerful tool to measure the accuracy of analysis results and commonly used in medical field but currently *ROC* curves are also using in other fields, such as engineering and agricultures.

A *ROC* curve is applicable only for continuous data or at least ordinal data. A classification model classifies each instance into one of two classes: a true and a false class. This gives rise to four possible classifications for each instance: (1) a true positive, (2) a true negative, (3) a

false positive, and (4) a false negative. The classifications that lie along the major axis x and axis y of the curve are the 100% correct classifications, that is, the true positives and the true negatives, respectively (Figure 1). For a perfect model, only the true positive and true negative fields are filled out, the other fields would be zero. A number of regions of interest can be identified in a *ROC* graph. The *ROC* curve illustrates the relationship between *TPF* and *FPF* at all possible cut-off levels. Therefore, it can be used to assess the performance of analysis results independently with respect to the decision threshold.

Area under *ROC* curve and the threshold value of reliability assessment can be predicted from *ROC* curve which are main concerns in this study. Let *D* be a variable describing the pipeline condition, where *D* = 1 for pipe failure condition and *D* = 0 for non-failure condition. Suppose that *Y* is a continuous random quantity which is related to the pipe condition (such as pipe wall thickness) and those large values of *Y* which are greater than threshold or allowable limit are failure states. Using a threshold, for example c, the result is called positive if $Y > c$, so it indicates a failure condition and if $Y < c$, i.e., negative, pipe condition is a non-fail condition, where $c \in (-\infty, \infty)$. Obviously, an accurate assessment will have both sensitivity and specificity close to 1. In ROC curve analysis, the aim is to find a cut-off point (threshold) of a classifier that minimizes the number of false positives and false negatives (maximizes the sensitivity and specificity). Based on the above conceptions, *FPF*, *TPF* and *ROC* curve can be estimated using Eqs. (9) – (11), respectively [23].

$$FPF = P(Y^0 > \mathrm{c} \mid D = 0) = S_0(c) \tag{9}$$

$$TPF = P(Y^1 > \mathrm{c} \mid D = 1) = S_1(c) \tag{10}$$

$$ROC = \{(FPF(c), TPF(c),\ c \in (-\infty, \infty)\} \tag{11}$$

Throughout this paper it is assumed that the two groups (failure and non-failure) are fully independent, i.e., no information about any aspect related to one group contains information about any aspect of the other group. If there are n_1 conditions data from a failure group and n_0 data from non-failure group, then these can be denoted by $\{y_i^1, i = 1,2,....,n_1\}$ and $\{y_j^0, j = 1,2,....,n_0\}$, respectively. For the classical (empirical) method, these observations per

group are assumed to be realisations of random quantities that are identically distributed as Y^1 and Y^0 with corresponding survival functions $S_1(y) = P[Y^1 > y]$ and $S_0(y) = P[Y^0 > y]$. According to Pepe [18], the empirical estimator of the *ROC* can be estimated as follows.

$$\hat{ROC} = \{(\hat{FPF}(c), \hat{TPF}(c)), c \in (-\infty, \infty)\} \tag{12}$$

$$\hat{TPF}(c) = \hat{S}_1(c) = \frac{1}{n_1}\sum_{i=1}^{n_1} 1\{y_i^1 > c\} \tag{13}$$

$$\hat{FPF}(c) = \hat{S}_0(c) = \frac{1}{n_0}\sum_{i=1}^{n_0} 1\{y_j^0 > c\} \tag{14}$$

where 1{A} is the indicator function which is equal to 1 if A is true or else. $\hat{S}_1$ and $\hat{S}_0$ are the empirical survival functions for Y^1 and Y^0, respectively. The empirical estimator of the *ROC* can also be written as shown in Eq. (18).

$$\hat{ROC}(c) = \hat{S}_1(\hat{S}_0^{-1}(c)) \tag{15}$$

3.1 Area under ROC curve

One of the important factors in *ROC* curve analysis is the area under the *ROC* curve, denoted as *AUC*. *AUC* has been used to predict the accuracy of failure prediction of pipeline in this paper. *AUC* can be estimated both parametrically and non-parametrically. The parametric estimation of *AUC* under the empirical *ROC* curve is the area under the curvature. On the other hand, the nonparametric estimation of the area under the empirical *ROC* curve is the summation of the areas of the trapezoids formed by connecting the points on the *ROC* curve. The nonparametric estimate of the area under the empirical *ROC* curve tends to underestimate *AUC* when discrete rating data are collected, whereas the parametric estimate of *AUC* has negligible bias except when extremely small case samples are employed. Therefore, for discrete rating data, the parametric method is preferred. For continuous or quasi-continuous data (e.g., a percent confidence scale from 0% to 100%), the parametric and nonparametric estimates of *AUC* will have very similar values and the bias is negligible [36].

A useful way to estimate the area under the *ROC* curve, *AUC*, can be expressed using Eq. (16) [37].

$$AUC = \int_0^1 ROC(t)dt \tag{16}$$

According to Zhou et al [36], the *AUC* is equal to the probability that the analytical results from a randomly selected pair of fail and non-fail group, as shown in Eq. (17)

$$AUC = P[Y^1 - Y^0] \tag{17}$$

The *AUC* measures the overall performance of the assessment. Higher *AUC* values indicate more accurate results, where *AUC* = 1 for perfect or ideal results and *AUC* = 0.5 for uniform results. So the *AUC* represents the ability to correctly classify a randomly selected individual as being from either the failure group or non-failure group. The empirical estimator of the *AUC* is the well-known Mann–Whitney U statistic which can be represented by Eq. (18) [23].

$$\hat{AUC} = \frac{1}{n_1 n_0} \sum_{j=1}^{n_0} \sum_{i=1}^{n_1} \psi(y_i^1 y_j^0) \tag{18}$$

where

$$\psi(y_i^1 y_j^0) = \begin{cases} 1 & if \quad y_i^1 > y_j^0 \\ 0.5 & if \quad y_i^1 = y_j^0 \\ 0 & if \quad y_i^1 < y_j^0 \end{cases}$$

The *AUC* value of 0.50 to 0.75 is fair, 0.75 to 0.92 is good, 0.92 to 0.97 is very good and 0.97 to 1.00 is considered as excellent result of an analysis [38].

3.2 Optimum threshold value in ROC curve

Another potential use of *ROC* curve is optimising the threshold value of an assessment. The optimum threshold values for pipe failure due to corrosion induced deflection, buckling, wall thrust and bending stress have been predicted in this study. The *ROC* curve comprises all possible combinations of sensitivity and specificity at all possible threshold values. This

offers the opportunity to assess the optimal threshold value to be used in critical decision practice.

In practice, choosing an optimal threshold value based on *ROC* analysis is practicable only for continuous data. For continuous data, all operating points on the curve correspond to realistic threshold values are considered. Different criteria are used to find optimal threshold point from *ROC* curve, such as points on curve closest to the (0, 1) and Youden index (*J*) etc, based on number of observed operating points (Figure 2). The Youden index (*J*) is the point on the *ROC* curve which is farthest from the line of equality [39].

Most of the operating points on the *ROC* curve consist of sensitivity and specificity combinations that do not correspond to realistic threshold values. Naturally, one would identify the threshold or optimal operating point as the point on the *ROC* curve that is closest to the ideal upper left-hand corner. The optimal range of the operating point will thus, shift towards the lower left hand corner of the *ROC* graph. Ideally, such decisions should be made by linking the constructed *ROC* curve in explicit decision analysis. If S_N and S_p denote sensitivity and specificity respectively, the distance between the point (0, 1) and any point on the *ROC* curve can be predicted by applying Eq. (19) as follows [39].

$$d = \sqrt{[(1-S_N)^2 + (1-S_p)^2]} \tag{19}$$

where *d* is the distance from top point (0, 1) to any point on curve. To obtain the optimal cut-off point, it is necessary to calculate this distance for each observed cut-off point and locate the point where distance is found minimum. The main aim of Youden index is to maximise the difference between *TPF* (S_N) and *FPF* $(1-S_p)$ and this yields $J = Max[S_N + S_p]$. The value of *J* can be located by doing a search of plausible value where sum of sensitivity and specificity is the maximum value [39].

3.3 NPI for ROC curve

In *NPI*, the uncertainty is quantified by lower and upper probabilities for events of interest. In effect, the optimal lower and upper bounds for the *ROC, AUC* can be derived. Suppose that $\{Y_i^1, i = 1,2,....,n_1, n_1+1\}$ are continuous and exchangeable random quantities from the failure

group and $\{Y_j^1, j = 1,2,....n_0, n_0 + 1\}$ are quantities from the non-failure group, where $Y_{n_1+1}^1$ and $Y_{n_0+1}^0$ are the next observations from the failure and non-failure groups following n_1 and n_0 observations, respectively. Let $y_1^1 < .... < y_{n_1}^1$ are the ordered observed values for the first n_1 pipes data from the failure group and $y_0^1 < .... < y_{n_0}^1$ for the first n_0 pipes data from the non-failure group. For ease of notation, let $y_0^1 = y_0^0 = -\propto$ and $y_{n_1+1}^1 = y_{n_0+1}^0 = +\propto$. Thus *NPI* can be used for reliability applications when the data represents failure and non-failure event which are non-negative. The *NPI* lower and upper survival functions for $Y_{n_1+1}^1$ and $Y_{n_0+1}^0$ can be determined as follows [20-21].

$$\underline{S}_1(c) = \underline{TPF}(c) = \underline{P}(Y_{n_1+1}^1 > c) = \frac{\sum_{i=1}^{n_1} 1\{y_i^1 > c\}}{n_1 + 1} \tag{20}$$

$$\overline{S}_1(c) = \overline{FPF}(c) = \overline{P}(Y_{n_1+1}^1 > c) = \frac{\sum_{i=1}^{n_1} 1\{y_i^1 > c\} + 1}{n_1 + 1} \tag{21}$$

$$\underline{S}_0(c) = \underline{FPF}(c) = \underline{P}(Y_{n_0+1}^0 > c) = \frac{\sum_{j=1}^{n_0} 1\{y_j^0 > c\}}{n_0 + 1} \tag{22}$$

$$\overline{S}_0(c) = \overline{FPF}(c) = \overline{P}(Y_{n_0+1}^0 > c) = \frac{\sum_{j=1}^{n_0} 1\{y_j^0 > c\} + 1}{n_0 + 1} \tag{23}$$

where $\underline{P}$ and $\overline{P}$ are *NPI* lower and upper probabilities. As the *ROC* curve clearly depends monotonously on the survival functions, therefore, it is easily seen that the optimal bounds, which is defined to be the *NPI* lower and upper *ROC* curves areas, are given as follows [37].

$$\underline{AUC} = \underline{P}(Y_{n_1+1}^1 > Y_{n_0+1}^0) = \frac{1}{(n_1 + 1)(n_0 + 1)} \sum_{j=1}^{n_0} \sum_{i=1}^{n_1} 1\{y_i^1 > y_j^0\} \tag{24}$$

$$\overline{AUC} = \overline{P}(Y_{n_1+1}^1 > Y_{n_0+1}^0) = \frac{1}{(n_1 + 1)(n_0 + 1)} [\sum_{j=1}^{n_0} \sum_{i=1}^{n_1} 1\{y_i^1 > y_j^0\} + n_1 + n_0 + 1] \tag{25}$$

Based on Eqs. (24) and (25), it is evident that the difference between upper and lower *AUC* can be expressed as follows.

$$\overline{AUC} - \underline{AUC} = \frac{n_1 + n_0 + 1}{(n_1 + 1)(n_0 + 1)} \tag{26}$$

Equation (26) indicates that it depends on the two sample sizes n_0 and n_1 only. Similarly for the partial area under *ROC* curve which can estimated using Eqs. (24) and (25) for any specific point of interest.

4. NUMERICAL APPLICATION

The proposed *ROC* approach has been applied to a steel buried pipe under a heavy roadway subject to external loading and corrosion. Four underground pipeline failure modes, namely corrosion induced deflection, buckling, wall thrust and bending stress have been used to illustrate the application of *ROC* curve in the accuracy of failure prediction and threshold value estimation. The loss of structural strength is due to corrosion through reduction of pipe wall thickness which then leads to pipe failure. In this study, pipe condition (i.e. pipe wall thickness) is considered as a classifier whereas the threshold is the cut-off point or limit to distinguish between the failure and non-failure conditions. The threshold obtained from the ROC curve is compared with the allowable limit from the limit state function. Due to lack of real data, 100 pipe wall thicknesses have been simulated at 100-year of service life using Monte Carlo method for each failure criterion based on soil and pipe material listed in Table 1 [29, 40, 41].

It is assumed that when actual pipe behaviour or pipe wall thickness exceeds the threshold value or allowable limit ($Y > c$), the result is positive (D = 1), i.e. failure condition; and when $Y < c$, the result is negative (D = 0), i.e. non-failure condition. However, there are 10% to 20% discrepancies in the actual and the estimated pipe conditions [17]. Therefore, it is assumed that, the predictions of pipe failure and non-failure conditions are not 100% accurate. The empirical and *NPI* lower and upper *ROC* curves have been applied for different failure modes with 10%, 20% and 30% noise which are introduced into the data to simulate the inaccurateness of failure predictions. Tables 2 to 5 show the pipe wall thickness with 10%

inaccurate prediction for the case of corrosion induced deflection, buckling, wall thrust and bending stress, respectively.

5. RESULTS AND DISCUSSION

The empirical *ROC* curves are applied for estimation of *AUC* and threshold value of pipe failure condition with 10%, 20% and 30% inaccurate failure prediction for different corrosion induced pipe failure modes. The performance of the *ROC* curve analysis is computed in terms of the true positive and false positive rates. This traces the curve from left to right (maximum ranking to minimum ranking) in the *ROC* graph. That means that the left part of the curve represents the behaviour of the model under high decision thresholds (conservative) and the right part of the curve represents the behaviour of the model under lower decision thresholds.

Empirical *AUC*, which is interpreted as the average value of sensitivity for all possible values of specificity, is a measure of the overall performance of the analysis for every failure case. The area under empirical *ROC* curve (*AUC*) is estimated using Eq. (18). *AUC* can take any value between 0 and 1, where a bigger value suggests the better overall performance of an analysis with 95% confidence level. Figures 3 to 6 show that *AUC* is higher for the case of 10% than that for 20% inaccurate prediction. Similarly, the case for 20% inaccurate prediction shows higher *AUC* than that for 30%. This indicates that the area under empirical *ROC* curve can be used to predict the reliability accuracy for different failure modes.

Table 6 indicates that different failure modes have different *AUC* for the same percentage of inaccurate prediction due to randomness of the data. The analysis shows that if simulated inaccurate prediction is 10%, the accuracy of the results is still fair enough for all the failure modes ($AUC > 0.75$). However if it is more than 10%, the accuracy of the results falls below the acceptable value ($AUC < 0.75$) which is implemented in practice as suggested by Huguet et al [38].

The allowable limit and the corresponding threshold pipe wall thickness for each corrosion induced failure modes, namely deflection, buckling, wall thrust and bending stress can be calculated using pipeline design formula as discussed in Section 2. For example, in the case of corrosion induced deflection, the allowable limit of deflection is estimated as 5% of initial

inside diameter of pipe. Then, the corresponding threshold pipe wall thickness is calculated using Eq. (3). Similarly, in the case of corrosion induced buckling, the allowable limit is estimated using Eq. (4) based on the assumption that the pipe fails when the actual buckling pressure is equal to the allowable buckling pressure and then the corresponding pipe wall thickness is calculated using Eq. (5). The same procedure is followed for other failure modes.

Besides that, the proposed approach has established a threshold at which a pipe can be considered in a high-risk condition. The threshold values of pipe wall thickness are predicted for the failure modes of deflection, buckling, wall thrust and bending stress. The optimum threshold value for each failure criteria predicted from the empirical *ROC* curve is obtained from Eq. (19) and the results are shown in Table 6 for comparison with the values obtained from pipeline design formulae. Both results are reasonably close in which the optimum threshold value of pipe wall thickness obtained from empirical *ROC* curve is more conservative. The results from Table 6 also show that the corrosion induced bending stress is the most dominating failure mode whereas buckling is the least susceptible failure mode.

Next, *NPI ROC* curves are applied to estimate the lower and upper bounds of *AUC* for all the failure modes and the results are shown in Figures 7 -10 and Table 7 with different percentages of inaccurate prediction. The *NPI* lower and upper areas under the *ROC* curves are calculated from Eqs. (24) and (25), respectively. As shown in Tables 6 and 7, the area under the upper bound of *NPI AUC* is always larger than empirical *AUC* for all the failure modes. It is clear that with increasing the percentage of inaccurate prediction, the areas under the upper and lower bounds of *NPI* are decreased. Therefore, the accurateness of the failure predictions is decreased as shown in Figures 7 to 10 and Table 7.

The performance of a prediction analysis should be judged in the context of the situation to which the data is applied. It can be seen that *AUC* for *NPI* is given in terms of upper and lower limits instead of a single curve. In this way it provides an interval of accuracy prediction which is more reasonable compared to classical *ROC*. Alternatively, the partial area estimation, where only a portion of the entire *ROC* curve needs to be considered, can also be used to predict the accuracy of an analysis when a particular *FPF* is useful indicator.

6. CONCLUSIONS

ROC curve has been applied in reliability analysis for underground pipelines due to corrosion induced deflection, buckling, wall thrust and bending stress. The *ROC* curve provides a performance assessment model for prediction of pipe failure state function. The analysis shows that *ROC* curve is a useful technique to predict the optimum threshold value and the accuracy of the results. The area under the curve provides an objective valuation for the accuracy of an analysis with combinations of sensitivity and specificity values. Thus two or more failure prediction methods can be compared using *ROC* curve. The results demonstrate that with increasing inaccurateness of failure prediction, the areas of the *ROC* curves (both classical and *NPI*) are decreased. Choosing the optimal operating point on the *ROC* curve which involves both maintenance and financial issues, can be ideally implemented in a formal risk-cost management process of buried pipeline network.

REFERENCES

1. Chughtai F and Zayed T. 2008. Infrastructure condition prediction models for sustainable sewer pipelines. *Journal of Performance of Constructed Facilities*, 22(5), 333-341.

2. Tee KF, Khan LR and Chen HP. 2013. Probabilistic failure analysis of underground flexible pipes, *Structural Engineering and Mechanics*, 47(2), 167-183.

3. McDonald S and Zhao J. 2001. Condition assessment and rehabilitation of large sewers. *Proc. of the International Conference on Underground Infrastructure Research*, Univ. of Waterloo, Waterloo, Canada, 361–369.

4. Tee KF and Li CQ. 2011. A numerical study of maintenance strategy for concrete structures in marine environment. *Proc. of the 11th International Conference on Applications of Statistics and Probability in Civil Engineering*, Zurich, Switzerland, August 1-4, 618-625.

5. Khan LR, Tee KF and Alani AM. 2013. Reliability-based management of underground pipeline network using genetic algorithm, *Proc. of the 11th International Probabilistic Workshop*, Brno, Czech Republic, November 6-8, 159-170.

6. Tee KF, Khan LR, Chen HP and Alani AM. 2014. Reliability Based Life Cycle Cost Optimization for Underground Pipeline Networks, *Tunnelling and Underground Space Technology*, 43, 32-40.

7. Baecher GB and Christian JT. 2003. Reliability and statistics in geotechnical engineering. *Wiley, New York*. USA.

8. Sivakumar Babu GL and Srivastava A. 2010. Reliability analysis of buried flexible pipe-soil systems. *Journal of Pipeline Systems Engineering and Practice,* ASCE, 1(1), 33-41.

9. Mahmoodian M, Alani AM, Tee KF. 2012. Stochastic failure analysis of the gusset plates in the Mississippi river bridge, *International Journal of Forensic Engineering*, 1(2): 153-166.

10. Fang Y, Chen J and Tee KF. 2013. Analysis of structural dynamic reliability based on the probability density evolution method, *Structural Engineering and Mechanics*, 45(2), 201-209.

11. Fang Y, Tao W and Tee KF. 2013. Repairable k-out-of n system work model analysis from time response, *Computers and Concrete*, 12(6), 775-783.

12. Tee KF, Khan LR and Li HS. 2014. Application of Subset Simulation in Reliability Estimation for Underground Pipelines, *Reliability Engineering and System Safety*, 130, 125-131.

13. Fang Y, Wen L and Tee KF. 2014. Reliability Analysis of Repairable k-out-n System from Time Response under Several Times Stochastic Shocks, *Smart Structures and Systems*, 14(4), 559-567.

14. Tee KF and Khan LR. 2014. Reliability analysis of underground pipelines with correlation between failure modes and random variables. *Journal of Risk and Reliability, Proc. of the Institution of Mechanical Engineers, Part O*, 228(4), 362-370.

15. Gustafson JM and Clancy DV. 1999. Modelling the occurrence of breaks in cast-iron water mains using methods of survival analysis. *Proc. AWWA Annual Conf.*, American Water Works Association, Denver.

16. Kettler AJ and Goulter IC. 1985. An analysis of pipe breakage in urban water distribution networks. *Canadian Journal of Civil Engineering*, 12(2), 286 – 293.

17. Mailhot A, Pelletier G, Noël J and Villeneuve J. 2000. Modelling the evolution of the structural state of water pipe networks with brief recorded pipe break histories: Methodology and application. *Water Resources Research*, 36(10), 3053 – 3062.

18. Pepe MS. 2003. The statistical evaluation of medical tests for classification and prediction. Oxford, Oxford University Press, United Kingdom.

19. Krzanowski WJ and Hand DJ. 2009. ROC curves for continuous data. Chapman & Hall/CRC.

20. Coolen-Maturi T, Coolen-Schrijner P and Coolen PA. 2012. Nonparametric Predictive Inference for Binary Diagnostic Tests. *Journal of Statistical Theory and Practice*, 6(4), 665-680.

21. Augustin T and Coolen FPA. 2004. Nonparametric predictive inference and interval probability. *Journal of Statistics Planning Inference*, Elsevier ltd, 124(2), 251-272.

22. Hill BM. 1968. Posterior distribution of percentiles: Bayes' theorem for sampling from a population. *Journal of the American Statistical Association*, 63(322), 677-691.

23. Coolen-Maturi T, Coolen-Schrijner P and Coolen FPA. 2011. Nonparametric predictive inference for diagnostic accuracy. *Journal of Statistical Planning and Inference,* 142(5), 1141-1150.

24. Debon A, Carrion A, Cabrera E and Solano H. 2010. Comparing risk of failure models in water supply networks using ROC curves. *Reliability Engineering and System Safety,* 95(1), 43 – 48.

25. Arian R, van Erkel, Peter M and Pattynama T. 1998. Receiver operating characteristic (ROC) analysis: Basic principles and applications in radiology, *European Journal of Radiology,* 27(2), 88-94.

26. Tee KF, Li CQ and Mahmoodian M. 2011. Prediction of time-variant probability of failure for concrete sewer pipes, *Proc. of the 12th International Conference on Durability of Building Materials and Components*, Porto, Portugal, April 12-15.

27. Alani AM, Faramarzi A, Mahmoodian M and Tee KF. 2014. Prediction of sulphide build-up in filled sewer pipes. *Environmental Technology,* 35(14), 1721-1728.

28. Ahammed M and Melchers RE. 1994. Reliability of pipelines subject to corrosion. *Journal of Transportation Engineering*, 120(6), 989-1002.

29. Sadiq R, Rajani B and Kleiner Y. 2004. Probabilistic risk analysis of corrosion associated failures in cast iron water mains. *Reliability Engineering and System Safety,* 86(1), 1-10.

30. Watkins RK and Anderson LR. 2000. Structural mechanics of buried pipes. *CRC Press, LLC*, Washington DC, USA.

31. Tee KF and Khan LR. 2012. Risk-cost optimization and reliability analysis of underground pipelines, *Proc. of the 6th International ASRANet Conference*, London, UK, July 2-4, Paper 49.

32. Gabriel LH. 2011. Corrugated polyethylene pipe design manual and installation guide, Plastic Pipe Institute, Irving, Texas, United States.

33. BS EN 1295:1-1997. 2010. Structural design of buried pipelines under various conditions of loading - General requirements. *British Standards Institution,* United Kingdom.

34. AWWA (American water works association). 1999. Buried pipe design, Fiberglass pipe design. *AWWA Manual M45, 35–53*, Denver, USA.

35. Hancor Inc. 2009. High density polyethylene (HDPE) pipe design. Drainage Handbook, Chapter 2, Findlay, OH, United States.

36. Zhou XH, Obuchowski NA and McClish DK. 2002. Statistical methods in diagnostic medicine. New York, Wiley, USA.

37. Coolen FPA. 1996. Comparing two populations based on low stochastic structure assumptions. *Statistics and Probability Letters*, 29 (4), 297–305.

38. Huguet EL, McMahon JA, McMahon AP, Bicknell R and Harris AL. 1994. Differential expression of human Wnt genes 2, 3, 4 and 7b in human breast cell lines and normal and disease states of human breast tissue. *Cancer Res,* 54(10), 2615–2621.

39. Indrayan A. 2012. Medical biostatistics. Third Edition, *Chapman & Hall/CRC Press*, 1008 pages.

40. Ahammed M and Melchers RE. 1997. Probabilistic analysis of pipelines subject to combined stresses and corrosion. *Journal of Engineering Structure*, 19(12), 988 – 994.

41. Sivakumar Babu GL, Srinivasa MBR and Rao RS. 2006. Reliability analysis of deflection of buried flexible pipes. *Journal of Transport Engineering*, 132(10), 829-836.

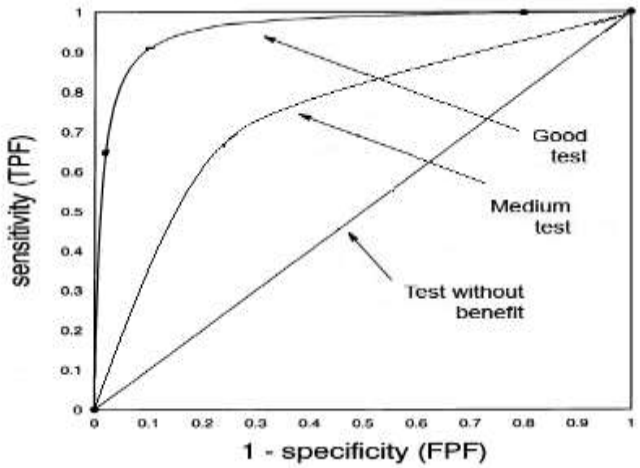

Figure 1: Basic of a typical *ROC* Curve

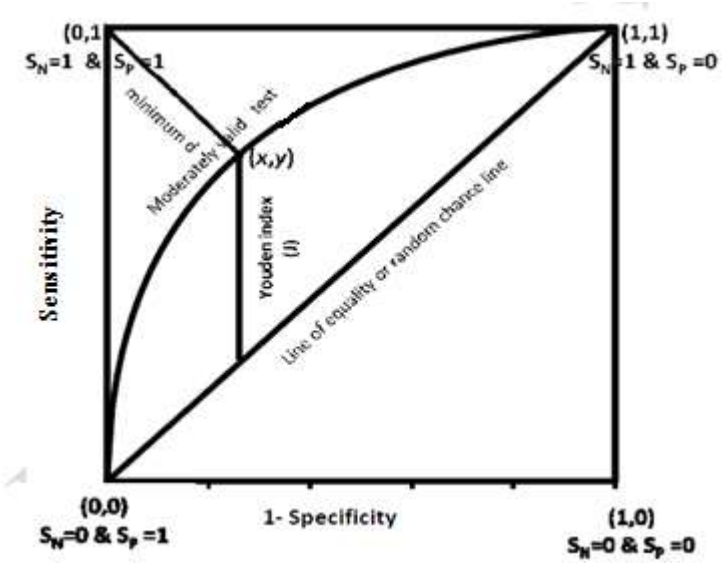

Figure 2: A typical best cut-off or threshold value in *ROC* curve

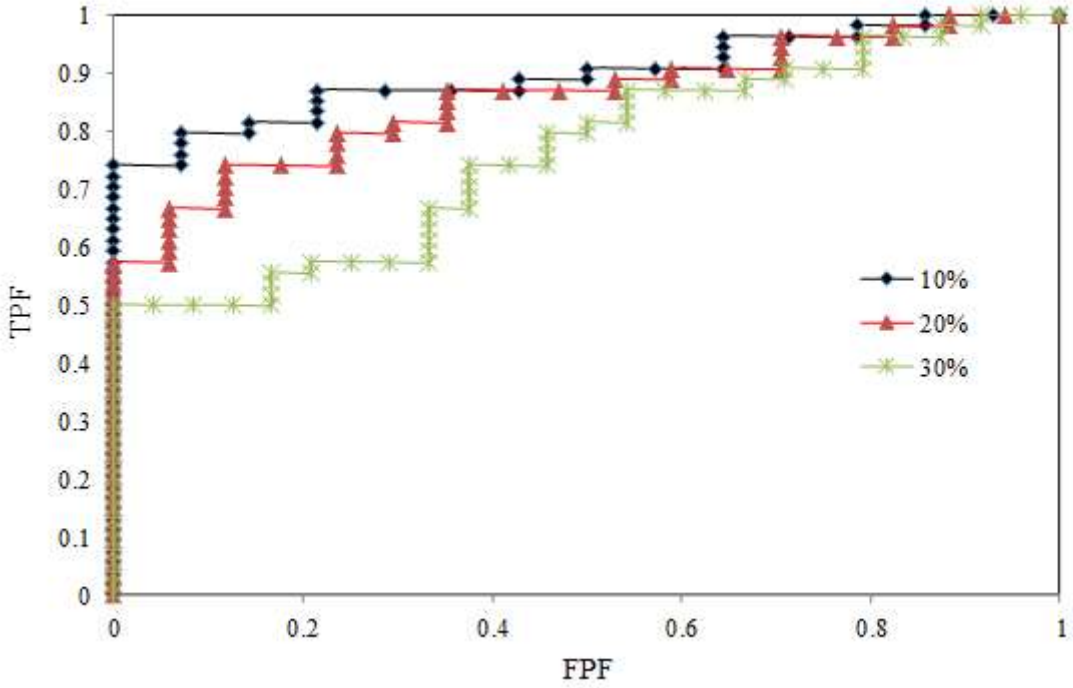

Figure 3: Empirical *ROC* curve for pipe failure due to corrosion induced deflection for different percentages of inaccurate prediction

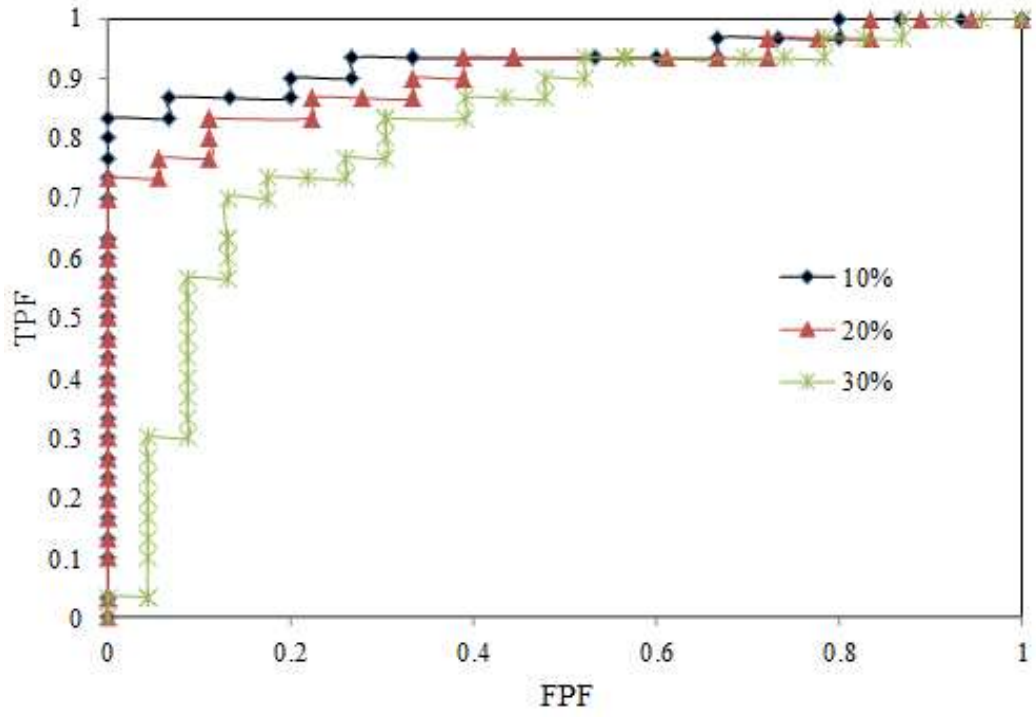

Figure 4: Empirical *ROC* curve for pipe failure due to corrosion induced buckling for different percentages of inaccurate prediction

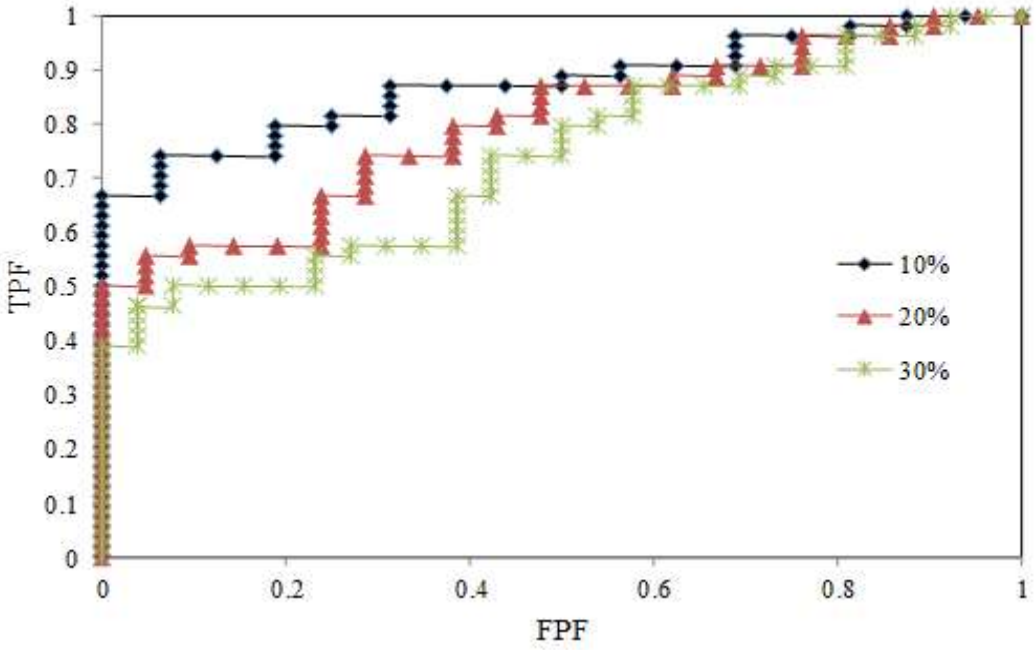

Figure 5: Empirical *ROC* curves for pipe failure due to corrosion induced wall thrust for different percentages of inaccurate prediction

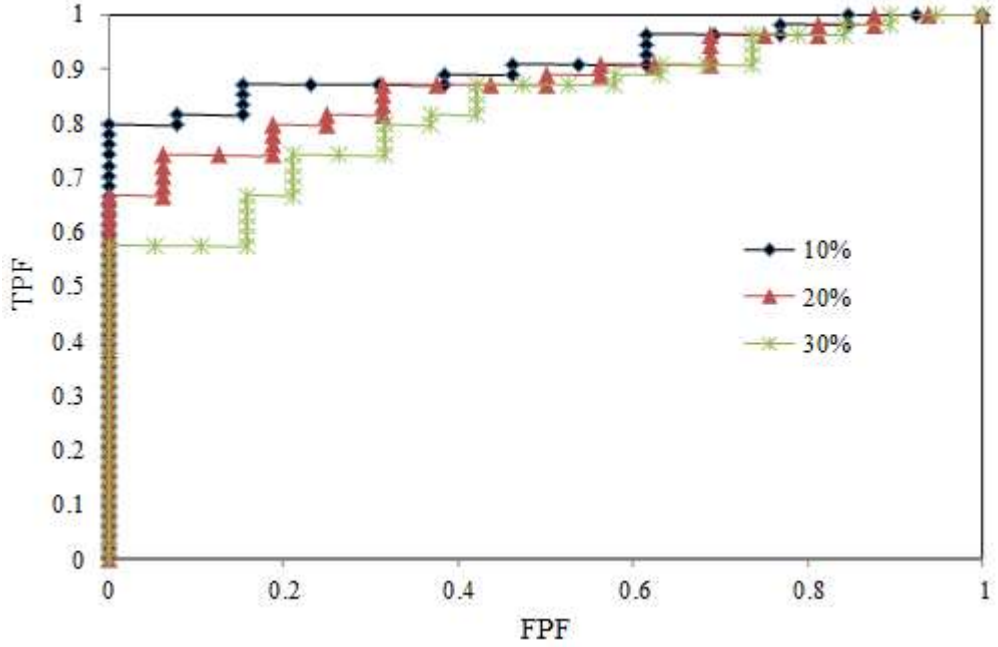

Figure 6: Empirical *ROC* curves for pipe failure due to corrosion induced bending stress for different percentages of inaccurate prediction

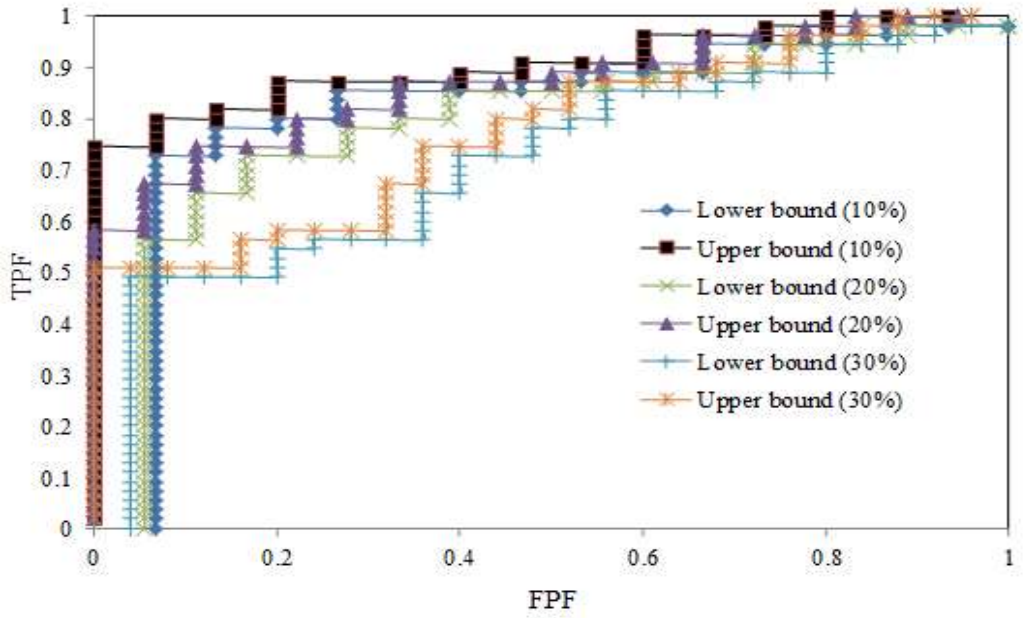

Figure 7: *NPI* lower and upper *ROC* curves for pipe failure due to corrosion induced deflection for different percentages of inaccurate prediction

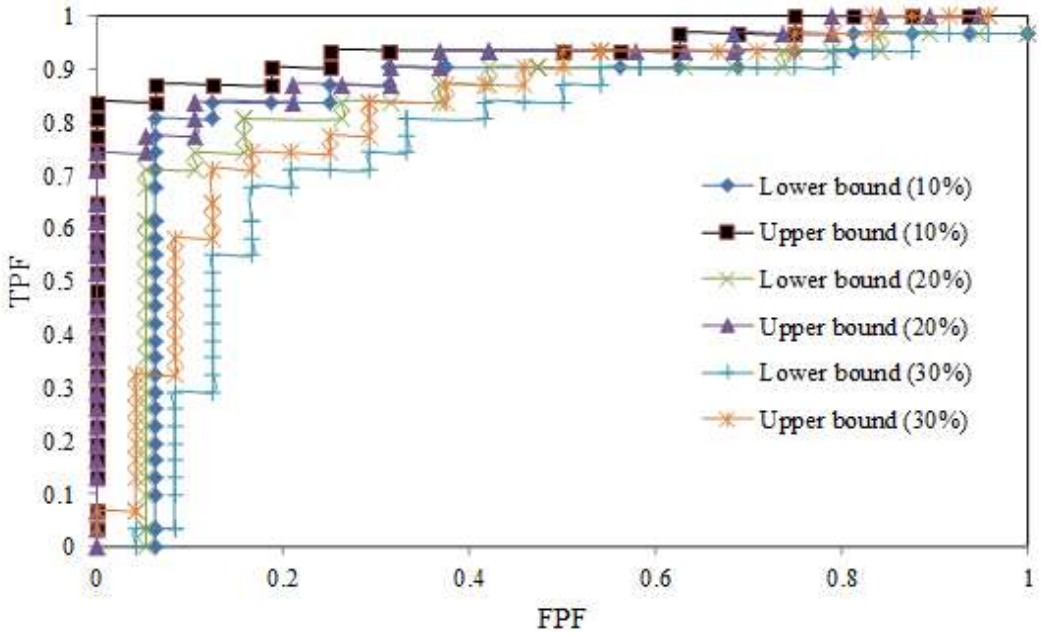

Figure 8: *NPI* lower and upper *ROC* curves for pipe failure due to corrosion induced buckling for different percentages of inaccurate prediction

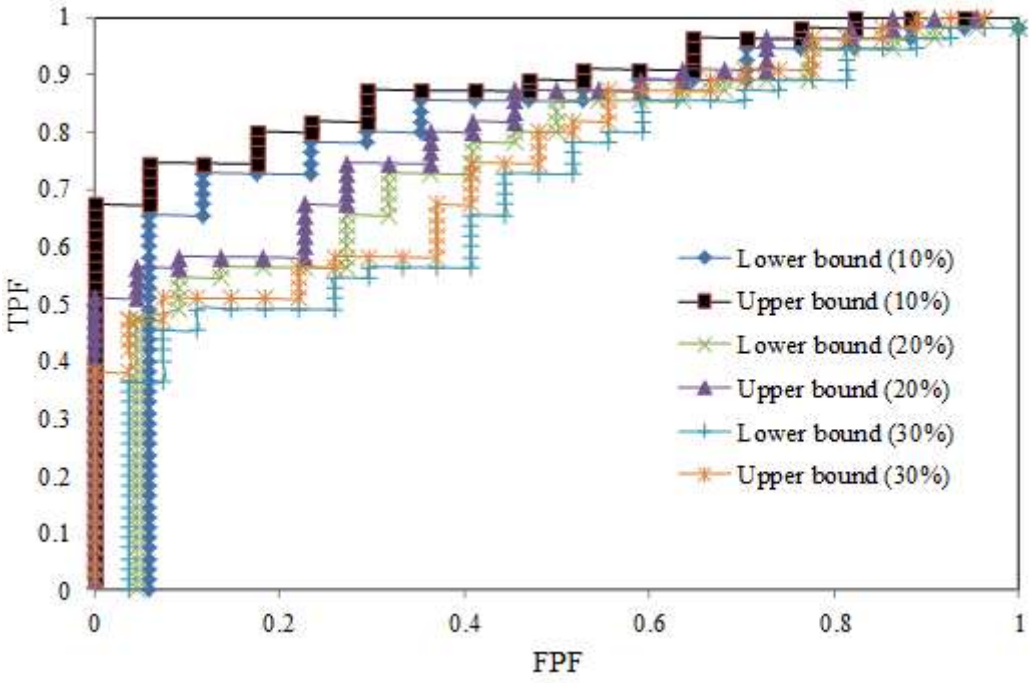

Figure 9: *NPI* lower and upper *ROC* curves for pipe failure due to corrosion induced wall thrust for different percentages of inaccurate prediction

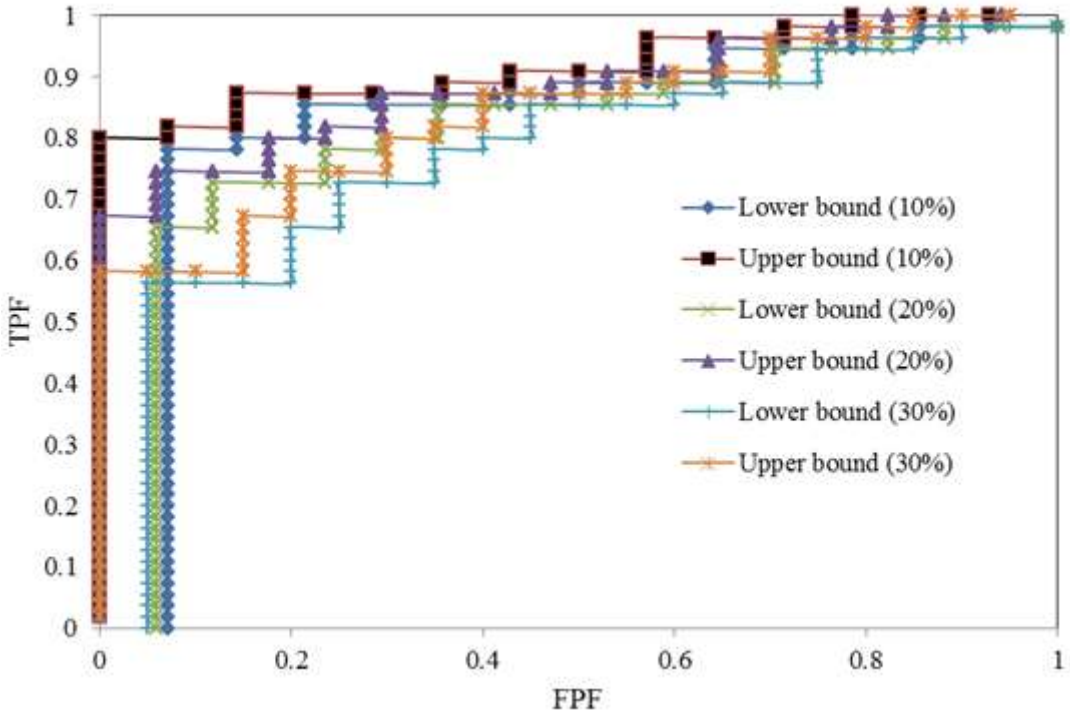

Figure 10: *NPI* lower and upper *ROC* curves for pipe failure due to corrosion induced bending stress for different percentages of inaccurate prediction

Table 1: Materials properties

Description	Symbol	Value	COV (%)	Distribution
Buoyancy factor	R_w	1.00	-	-
Trench width	B_d	2.00 m	-	-
Outside pipe diameter	D_o	1.231 m	-	-
Inside pipe diameter	D_i	1.189 m	-	-
Soil constrained modulus	M_s	2.02×10^3 kPa	-	-
Deflection Lag factor	D_L	1.0	-	-
Shape factor	D_f	4.0	-	-
Capacity modification factor for pipe	ϕ_p	1.0	-	-
Capacity modification factor for soil	ϕ_s	0.90	-	-
Allowable bending stress	σ_a	450 MPa	-	-
Poisson ratio	v	0.3	-	-
Elastic modulus of pipe	E	213.74×10^6 kPa	1.0	Normal
Backfill soil modulus	E_s	10^3 kPa	5.0	Normal
Unit of weight of soil	γ	18.0kN/m^3	2.5	Normal
Wheel load (Live load)	P_s	80.0 kPa	10.0	Normal
Deflection coefficient	K_b	0.11	1.0	Lognormal
Multiplying constant	k	2.0	10.0	Normal
Exponential constant	n	0.3	5.0	Normal
Initial thickness of pipe	t	0.021 m	1.0	Normal
Height of the backfill	H	3.75 m	1.0	Normal

COV = Coefficient of variation.

Table 2: Pipe wall thickness (m) with 10% inaccurate prediction for the case of deflection

Failure group									
0.013711	0.013717	0.013638	0.01367	0.012256	0.013659	0.013754	0.013056	0.014336	0.013639
0.013621	0.013749	0.013913	0.012942	0.013693	0.01367	0.01365	0.01395	0.013921	0.0138
0.013989	0.01699	0.013639	0.012865	0.0138	0.01376	0.0139	0.01361	0.013821	0.013755
0.013976	0.013431	0.014138	0.013709	0.013895	0.013147	0.013159	0.012774	0.012002	0.012245
0.013983	0.013866	0.013934	0.017792	0.01386	0.016665	0.012867	0.01744	0.013876	0.016101
Non-failure group									
0.011358	0.011579	0.0131	0.013198	0.013332	0.012482	0.012303	0.013431	0.012126	0.013077
0.012755	0.013135	0.012934	0.011323	0.012859	0.012523	0.01289	0.013035	0.013332	0.013018
0.012963	0.013181	0.013824	0.012724	0.012456	0.012408	0.012732	0.012675	0.014351	0.012753
0.014237	0.013091	0.012728	0.011857	0.013177	0.013711	0.013231	0.013534	0.012028	0.014094
0.013576	0.01348	0.013257	0.013538	0.014696	0.012475	0.013428	0.012847	0.012283	0.011654

Table 3: Pipe wall thickness (m) with 10% inaccurate prediction for the case of buckling

Failure group									
0.016711	0.016717	0.016638	0.016621	0.016749	0.016913	0.012942	0.016693	0.01667	0.01665
0.016989	0.01699	0.016639	0.012865	0.0168	0.01676	0.0169	0.01695	0.016921	0.0138
0.016976	0.016431	0.016738	0.016709	0.013895	0.016847	0.016859	0.01661	0.013821	0.016755
Non-failure group									
0.011358	0.011579	0.0131	0.013198	0.013332	0.012482	0.012303	0.013431	0.012126	0.013077
0.012755	0.013135	0.012934	0.011323	0.012859	0.012523	0.01289	0.013035	0.013332	0.013018
0.012963	0.013181	0.016824	0.012724	0.012456	0.012408	0.012732	0.012675	0.014351	0.012753
0.014237	0.013091	0.012728	0.011857	0.013177	0.016711	0.013231	0.013534	0.012028	0.017094
0.013576	0.01348	0.013257	0.013538	0.016696	0.012475	0.013428	0.012847	0.012283	0.011654
0.01367	0.012256	0.013659	0.013754	0.013056	0.014336	0.013639	0.013983	0.013866	0.013934
0.01744	0.013876	0.016101	0.013792	0.01386	0.013665	0.016867	0.012774	0.012002	0.012245

Table 4: Pipe wall thickness (m) with 10% inaccurate prediction for the case of wall thrust

Failure group									
0.013711	0.013717	0.013638	0.01367	0.012256	0.013659	0.013754	0.012056	0.014336	0.013639
0.013621	0.013749	0.013913	0.012942	0.013693	0.01367	0.01365	0.01395	0.013921	0.0138
0.013989	0.01699	0.013639	0.012865	0.0138	0.01376	0.0139	0.01361	0.013821	0.013755
0.013976	0.013431	0.014138	0.013709	0.013895	0.013147	0.013159	0.012774	0.012002	0.012245
0.013983	0.013866	0.013934	0.017792	0.01386	0.016665	0.012867	0.01744	0.013876	0.016101
0.014237	0.013091	0.012728	0.011857	0.013177	0.013711	0.013231	0.013534	0.012028	0.014094
0.013576	0.01348	0.013257	0.013538	0.014696	0.012475	0.013428	0.012847	0.012283	0.011654
0.014351	0.013824								
Non-failure group									
0.011358	0.011579	0.0129	0.012198	0.013332	0.012482	0.012303	0.013431	0.012126	0.013077
0.012755	0.013135	0.012934	0.011323	0.012859	0.012523	0.01289	0.012035	0.012332	0.013018
0.012963	0.013181	0.012724	0.012456	0.012408	0.012732	0.012675	0.012753		

Table 5: Pipe wall thickness (m) with 10% inaccurate prediction for the case of bending stress

Failure group									
0.013711	0.013717	0.013638	0.01367	0.011256	0.013659	0.013754	0.011056	0.014336	0.013639
0.013621	0.013749	0.013913	0.01142	0.013693	0.01367	0.01365	0.01395	0.013921	0.0138
0.013989	0.01699	0.013639	0.01165	0.0138	0.01376	0.0139	0.01361	0.013821	0.013755
0.013976	0.013431	0.014138	0.013709	0.01125	0.013147	0.013159	0.012774	0.012002	0.012245
0.013983	0.013866	0.013934	0.017792	0.01386	0.016665	0.012867	0.01744	0.013876	0.016101
0.014237	0.013091	0.012728	0.011857	0.013177	0.013711	0.013231	0.013534	0.012028	0.014094
0.013576	0.01348	0.013257	0.013538	0.014696	0.012475	0.013428	0.012847	0.012283	0.011654
0.012408	0.012732	0.012675	0.014351	0.012753	0.012963	0.013181	0.013035	0.013332	
Non-failure group									
0.011358	0.011579	0.0131	0.011198	0.013332	0.011482	0.011303	0.011431	0.011126	0.011077
0.011755	0.011135	0.011934	0.011323	0.010859	0.011523	0.01189	0.013824	0.012724	0.012456

Table 6: Threshold value and area under empirical *ROC* curve

		Failure modes			
		Deflection	Buckling	Wall thrust	Bending stress
Allowable limit using pipeline design formula		0.0605 m	1023.8 kPa	5867 kPa	450000 kPa
Threshold wall thickness using pipeline design formula		0.0137 m	0.0171 m	0.0136 m	0.0132 m
Optimum threshold wall thickness from empirical *ROC* curve		0.01357 m	0.0166 m	0.013 m	0.0128 m
Area under empirical *ROC* curve with inaccurate prediction	10%	0.89	0.78	0.80	0.76
	20%	0.68	0.67	0.68	0.70
	30%	0.55	0.63	0.58	0.56

Table 7: Area under *NPI ROC* curve

% of inaccurate prediction	*NPI* Area	Failure modes			
		Deflection	Buckling	Wall thrust	Bending stress
10%	$\overline{AUC}$	0.92	0.86	0.87	0.90
	$\underline{AUC}$	0.88	0.81	0.78	0.87
20%	$\overline{AUC}$	0.73	0.70	0.71	0.72
	$\underline{AUC}$	0.67	0.64	0.66	0.65
30%	$\overline{AUC}$	0.60	0.61	0.60	0.59
	$\underline{AUC}$	0.54	0.58	0.56	0.54