# AN ARTIFICIAL NEURAL NETWORK MODELING FOR PIPELINE CORROSION GROWTH PREDICTION

Mazura Mat Din[1], Norafida Ithnin[1] Azlan Md. Zain[1] Norhazilan Md Noor[2] Maheyzah Md Siraj[5] and Rosilawati Md. Rasol[2]
[1]Faculty of Computing, Universiti Teknologi Malaysia, Malaysia
[2]Faculty of Civil Engineering, Universiti Teknologi Malaysia, UTM Skudai, Johor, Malaysia
E-Mail: azura@utm.my

**ABSTRACT**

Corrosion defect assessment becoming a forte issue in pipeline reliability assessment to accurately predict the severity of its condition. Due to the uncertainties inherit from the pipeline inspection at present, statistical model use to model the corrosion growth apply a correctional methods to reduce the gap (means and variation) between predicted values and the actual data. This study aims to develop a time dependent corrosion growth model for oil and gas pipeline using Artificial Neural Network (ANN) as an alternative to the current method and to evaluate its applicability without enforcing data correctional methods. This model is formulated based on parameters of defect extracted from in-line inspection data (ILI) and quantified by statistical analysis. The develop model gives the prediction of the corrosion depth and length of the defect that can be used to calculate the corrosion rate or growth. The results and outcome of the present study can help pipeline operators to predict the reliability of the pipeline structure in terms of its probability of failure or lifetime estimation.

**Keywords:** artificial neural network, corrosion defect assessment, corrosion rate prediction, uncertainties modeling.

## INTRODUCTION

The corrosion process is time-variant and the amount of corrosion damage is normally defined by a corrosion rate with units of mm/year which representing the depth of corrosion increase per year (Paik and Thayambali, 2002). While the extent of corrosion presumably increases with time, it is not straightforward to predict the progress of corrosion. The only real alternative is then to pessimistically assume more corrosion extent than is likely (Paik and Thayambali, 2002). There are three models available to estimate the rate of corrosion growth: theoretical, empirical and artificial intelligence models (Tran, 2007). Generally, a theoretical model such as linear model is simple and practically available to predict the average growth rate based on metal loss evidence regardless the material and environment properties factors. In the other hand, empirical model is developed by defining the relationship between material and environment properties to predict the corrosion rate. Furthermore, the statistical use introduced another stage of processing namely correction methods. Thus the implementation of the artificial intelligence method was able to eliminate this extra stage of processing. This research is designed to establish a model for prediction of the corrosion rate of pipeline regardless of the environmental factors. The result of this study simplifies the works for future researches in assessing pipeline corrosion. With the ideal of network model in accessing the pipeline corrosion, future researchers able to put more effort on the analysis between environmental parameters and pipeline corrosion. Thus, this study improves the procedure for data analysis and the accuracy on predicting the corrosion growth in structural analysis regarding the condition of pipeline structure. Besides that, the proposed approach can also be used in propagation with the current approaches to provide a better vision of the corrosion assessment. This study was carried out as a sequel research by Din *et al*. (2009) and Noor (2010). The results describe provide a useful information for decision making on maintaining the structure.

## BACKGROUND

Researchers frequently use models in problem formation and solutions. The modeling is based on the physical, chemical or engineering science knowledge of the phenomenon (Tran, 2007). In such cases, the models can theoretical be tested with the experiments which are known as mechanistic models. However, when combine two or more factors that might not correlate or didn't give an effect to the response of interest coupled with complex probabilistic way which poorly understood will results an insignificant decision making. Based on these, it is necessary to build a model that correlate the contributing factors with the response based on real inspection data (ILI data). This corrosion defect assessment model which utilize a multiple inspection data are often used to predict the current and future condition of pipelines. An accurate prediction will leads to a better decision making on the maintenance of the structure.

Obviously, an appropriate modeling technique for corrosion defect assessment will increase the accuracy of the predictions. However, unavailability of data and the use of the deterministic methods hinder the progressive of a new modeling. The existing models can be categories into three models namely, deterministic models, statistical models, and artificial intelligence based models (Marcous et al. (2002). The deterministic model and statistical model were also known as model-driven, whereby the artificial intelligence model is a data-driven model (Dasu and Johnson, 2003). Looking into this classification and based

ARPN Journal of Engineering and Applied Sciences



on the current ILI data available, it is decided that an artificial intelligence will be the modeling method chosen for the assessment of the defect.

Artificial intelligence models were designed to mimic the operation of human brain of behavior which are learning and generalizing (Taylor, 1993). Many engineering models especially for corrosion defect assessment, use this learning and generalizing feature in their application where model outputs were classified from a set of input pattern by learning from the past data and generalizing the lessons to predict future targets (Sinha and Pandey, 2002, El-Abbasy *et al*., 2014a; Sumarni *et al*., 2012; Mabbutt *et al*., 2012). The artificial intelligence methods also concern only with the input and output data without specifying the underlying mechanism and been categorize as a 'black box' models.

Previous works has been carried out using artificial neural networks in, were mostly applied a multi-layered perceptron (MLP) network, primarily because it has been shown that the MLP can approximate any input-output mapping or function (Picton, 2000). The ANNs in corrosion defect assessment areas used various types of parameters in order to predict corrosion behaviour. For example, the uses of various environmental variables that cause non-linear behaviour were investigated by Kreilova and Haloma (2008). The results of loss data was compared from test at different geographical location. Other work based on atmospheric variables and using ANN to predict corrosion rates was carried out on metals used in high voltage power transmission in Brazil (Ramon et al., 2009). Results were comparable with those obtained from exposure specimens indicating that the ANN was a good indicator of corrosion rates in these circumstances. El Abbasy *et al*. (2014a) also use ANN for corrosion growth prediction but considering not only the metal loss data but also other physical parameters such as flow rate.

Corrosion growth modeling provide a proactive method of analyzing large quantities of ILI data, prioritizing pipeline repair programs, and optimizing re-inspection interval (Desjardins, 2002). Issues in corrosion growth calculation based on ILI data can be divided into three. Firstly, the analysis based on multiple match defect produce a negative corrosion rate. Theoretically, this condition is unrealistic due to the nature of corrosion that proliferates over time. Secondly, the usage of mathematical model i.e. linear model and correction method to simulate the corrosion growth which might cause an overly conservative results of the corrosion data (Noor, 2006). Finally, the limited amount of sample ILI data being used by the related works is arguable to produce a good prediction of the corrosion growth (Noor, 2006, Yahaya, 1999). To overcome the limitations discuss above, a computational method is introduced in this research to cater the non-linear growth of the corrosion and at the same time eliminate the extra task on correcting the measured data provided a number of ILI data available for the prediction.

Researchers like Liao et al. (2012) and Ren *et al*. (2012) applied ANN in corrosion growth study involving corrosion rate prediction of natural gas pipelines. Liao *et al*. (2012) also benchmark its result with other computational methods such as GA and PSO in this area but the finding shows that ANN outperformed both methods. Fuzzy-ANN model was developed by Peng et al. (2009) to predict the failure rate for oil and gas pipelines. On the other hand, Belsito (1998) utilize ANN for lead detention and sizing also in liquid pipeline. Problem with the previous mentioned model is unwrapped by Senouci (2014) which claims that it lacks of objectivity in predicting the different failure types of pipelines. As a result an ANN and regression were used to predict possible failure types besides corrosion for oil and gas pipelines i.e. natural hazard, third party and so on. Fuzzy logic is later tested on the same data as the previous research to measure the model validity (Senouci et al., 2014). Result shows that it can outperform the other method in term of model valisity. Another progressive research made by el-Abbasy et al. (2014a) in projecting the corrosion growth using Analytic Network Process and Monte Carlo simulation but fallback due to lack of collected data. To cater the problem El-Abbasy et al. (2014b) make another projection on this topic using regression analysis which give better validity percentage for up to 96%. Although the study conducted by these researchers provided sound results, the regression analysis as a computational model has still some limitations (El-Abbasy et al., 2014b). Several other computational model can be used to predict pipeline conditioned especially based on corrosion defect. The model includes ANN, Support Vector Machine (SVM), Naïve Bayes (NB), Decision Tree (DT), and K-Nearest Neighbor (K-NN) (El-Abbasy *et al*., 2014c).

Although ANN are considered among of the oldest methods, it is still found to be a competitive algorithm among the new ones especially in prediction modeling. The reliable use of ANN as the good predictors were proven via its accuracy measures in (Fathi and Aghakouchak, 2007; Ding *et al*., 2008; Caruana and Mizil, 2006; Mohana and Thangaraj, 2013; Prabhakar, 2013; Tabesh *et al*., 2009). Based in those research there is no specific computational model perform better or worse from the other depending on the nature as well as data format. Motivated by this observation, it was decide in this study to use ANN is explored in this study for predicting the growth rate of depth of corrosion defects on underground pipelines. El-Abbasy *et al*. (2014c) outline the advantages of ANN as follows:

a) Able to detect implicitly a nonlinear relationship between independent and dependent variables by adjusting its connection weights and undergo a nonlinear transformation at each hidden and output node.
b) The hidden layer can detect interaction or relationship between all the input variables, resulted a better prediction model.

## CASE STUDY: ILI DATA

In this study, databases of ILI data were gathered from three different pipelines, named Pipelines A, B and C. Pipelines A and B consist of three sets of data, recorded in Year 1990, 1992 and 1995 and for C in Year 1999 and 2001. All collected pigging data represent internal defects in the form of corrosion pits only. Thus, other types of corrosion defects were not considered in the data sampling procedure. Normally, pigging data provides valuable information on the corrosion defect geometry, such as depth and length, orientation, defect location and types of corrosion regions. The physical dimensions and other related information of these three pipelines are presented in Table-1 and Table-2.

**Table-1.** Number of recorded defects for each set.

| Set of data | Pipeline A | Pipeline B | Pipeline C |
|---|---|---|---|
| Number of data | 90 -1425<br>92 -2995<br>95 -3314 | 90 -1397<br>92 -1528<br>95 -4084 | 99 - 2581<br>01 - 4058 |

**Table-2.** A typical presentation of pigging data.

| **SL** (m) | **RD** (m) | **AD** (m) | **d% wt** | **l mm** | **w mm** | **O'clock** |
|---|---|---|---|---|---|---|
| 11.6 | 6.6 | 1016.5 | 18 | 32 | 42 | 6.00 |
| 11.5 | 11.5 | 1033.0 | 19 | 46 | 64 | 5.30 |
| 11.8 | 10.6 | 1043.6 | 12 | 18 | 55 | 5.30 |
| 11.7 | 1 | 1045.8 | 13 | 28 | 83 | 5.30 |

where
Absolute distance (AD) = Distance of corrosion from start of pipeline
d%wt = Maximum depth of corrosion in terms of percentage
l = Longitudinal extent of corrosion
Loc = Location of corrosion either internal or external.
O'Clock = Orientation of corrosion as a clock position of pipe wall thickness.
Relative distance (RD) = Relative distance of corrosion from upstream girth
Spool length (SL) = Length of pipe between weld (10m to 12m approximately)
w = Extent of corrosion around pipe circumference weld

## ANN MODELING

This section begins with the introduction of ILI data and its characteristics, discussed about the analysis of the results for the stages of the proposed Artificial Neural Network (ANN) model. The sequences of the ANN model started with the determination of input parameters, following by group of datasets, then the optimization of network parameters and ended with the results of testing datasets. First stage of ANN model was the determination of input parameters which decided the total number of input parameters was significant to the designed model. The second stage of ANN model was separating the collected samples into four distinct groups of samples according to year variable and split each samples group into training, validation and testing datasets by percentage. In optimization of network parameters stage, the weight vectors like number of hidden neuron, learning rate and momentum factor needed to be modified in order to ensure adequate generalization. The last stage of ANN model was the discussion of prediction results based on the testing datasets.

In the generalization performance evaluation phase, four indexes which adopted for comparison between ANN and SVM models were mean absolute error (MAE), relative absolute error (RAE), root mean square error (RMSE) and correlation coefficients ($R^2$). Their formulas were expressed by equation (1) till (4), respectively:

$$MAE = \frac{1}{m}\sum_{j=1}^{m}|\hat{y}_j - y_j| \quad (1)$$

$$RAE = \frac{|\hat{y}_j - y_j|}{|\bar{y} - y_j|} \quad (2)$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{j=1}^{m}(\hat{y}_j - y_j)^2} \quad (3)$$

$$R^2 = \frac{\sum_{j=1}^{m}(\hat{y}_j - \bar{y})^2}{\sum_{j=1}^{m}(y_j - \bar{y})^2} \quad (4)$$

where m denotes the number of test samples, $y_j$ represents the jth target value, $\hat{y}_j$ stands for the predicted value for the jth test sample and $\bar{y}$ is the mean target value for all test samples.

### The development of ANN model

The development of ANN model covered the whole procedures from the handling of the input parameters to the prediction results of corrosion rates based on testing datasets. There were four major procedures related to the development of artificial neural network model.

a) Determined the input factors as input parameters for the designed model.
b) Group of datasets and dividing samples into training, validation and testing data.

c) Optimization for the appropriate network parameters applied to the designed model.
d) Predicted the corrosion defect values using the testing datasets.

**Determination of input parameters**

A set of input parameters was identified based on the significant abilities towards the ANN model. Several test runs had been conducted using single set of samples to determine which input parameter as one of the critical factors in mapping to the output parameter. The results compared based on the value of model accuracy and the percentage errors. The highest value of model accuracy or the lowest value of differences errors indicated the ideal input parameter accepted as input for the ANN model.

In order to determine which input parameters are significant to the designed model. The test runs have been conducted on the 105 samples of dataset using the same network architecture and network parameters. The 105 samples been applied which collected from the year 1996 to year 2001. The 105 samples have been divided into three groups of dataset which are training, validation and testing datasets. In this study, the transfer function been applied was log-sigmoid and the training algorithm used was gradient descent. The training parameters for the network architecture were illustrated in Table-3.

**Table-3.** Training parameters and its values.

| Group No. | Year duration | Number of sample |
|---|---|---|
| 1 | 1990 – 1992 | 417 |
| 2 | 1992 – 1995 | 417 |
| 3 | 1990 – 1995 | 417 |
| 4 | 1996 - 2001 | 105 |

Five input parameters: odometer, orientation, depth, length and width of pipeline defects were selected as input instances in this training model and the results been compared based on the value of correlation coefficient ($R^2$), mean absolute error (MAE), root mean squared error (RMSE), relative absolute error (RAE) and root relative squared error (RRSE). Table-4 illustrated the results of comparison for determine the input parameters.

**Table-4.** Comparison of input parameters.

| Input Parameters | $R^2$ | MAE | RMSE | RAE (%) | RRSE (%) |
|---|---|---|---|---|---|
| d, l | 0.9989 | 0.0015 | 0.003 | 3.11 | 4.90 |
| d, l, w | 0.9986 | 0.0018 | 0.0033 | 3.68 | 5.34 |
| d, l, o | 0.9988 | 0.0016 | 0.0031 | 3.23 | 5.07 |
| d, l, w, o | 0.9991 | 0.0013 | 0.0027 | 2.62 | 4.45 |
| d, l, w, o, od | 0.9988 | 0.0016 | 0.003 | 3.28 | 4.95 |

*Depth (d), Length (l), Width (w), Orientation (o), Odometer (od)

The result showed that the ideal input parameters which were significant to the designed model were depth, length, width and orientation. This could be proved by the highest index of correlation coefficient ($R^2$): 0.9991 and the lowest indexes of mean absolute error (MAE), root mean squared error (RMSE), relative absolute error (RAE) and root relative squared error (RRSE) with its value 0.0013, 0.0027, 2.6227% and 4.4466%. Since the difference between the ideal input parameter and the first group of input parameter with depth and length was not so huge in term of correlation coefficient and root mean squared error, the depth and length were been identified as input parameters in this study. The depth and length input parameters were the most common and significant factors which contributed to the pipeline corrosion defects. As the reason of common factors, the depth and length parameters have been selected as input parameter throughout the whole process of this study.

**Group of datasets**

The data samples collected through in-line inspection (ILI) process on the same pipelines from year 1990 to year 2001. The samples gathered have been separated into four different groups of samples according to the year's factor. First, second and third groups of samples comprised of same instances of data which were 417 samples with the time duration of year 1990 to 1992 (first group), year 1992 to 1995 (second group) and year 1990 to 1995 (third group). Fourth group of samples contained of 105 samples which collected from year 1996 till year 2001. Table-5 showed the mapping of four different groups of samples.

**Table-5.** Training parameters and its values.

| Network parameters | Value |
|---|---|
| Hidden layers | 3 |
| Learning rate | 0.1 |
| Momentum factor | 0.1 |
| Number of epochs | 500 |

The data samples were split into training, validation and testing datasets which 70% of datasets were

VOL. 10, NO. 2, FEBRUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

presented to the network model for training purpose. The other 15% of datasets concurrent with the training set were used for cross validation which identify the stopping point for training process and the remaining 15% of datasets were used for testing the designed network model.

**Optimization of network parameters**

Multi-Layer Perceptron of network architecture was applied in this study for the prediction achievement of the ANN model. The weight vectors have been modified through continuous training adjustment in order to design the ideal model with the best accuracy prediction ability and to ensure adequate generalization. The weight vectors involved in this process were the number of hidden neurons, learning rate and momentum factor.

In this study, fourth group of sample with 105 samples of dataset which been divided into three groups of datasets: training, validation and testing datasets have been used on the training process. As consistency, the transfer function been applied was same is log-sigmoid and the training algorithm used was gradient descent. The number of epochs was set fixed at 500 based on the maximum number of gathered samples. The results have been compared based on correlation coefficient ($R^2$) and root mean squared error (RMSE) which been illustrated in Table-6. For simplicity results IN Table-6 shows a test no. 30 to 46 only which been observed give an optimal performance at run 43, but the discussion will involve a whole range of runs been tested. From test no. 1 to 4, the result showed that increase of learning rate did not amend the output of network architecture by observing the correlation coefficient become smaller. However, the increase of momentum the factor caused correlation coefficient increased as well based on result from the test no. 5 till 9 yet correlation coefficient stop increasing when momentum factor at 0.8. In test run with two of hidden neurons, the larger value of learning rate did showed quite critical amendment for the output with its peak stopped at 0.3 of learning rate based on test no. 10 to 13.

In test run with two of hidden neurons, the larger value of learning rate did showed quite critical amendment for the output with its peak stopped at 0.3 of learning rate based on test no. 10 to 13. There also produced some positive results that the momentum factor increased as well as correlation coefficient but halted to do so when the momentum factor at 0.5 which showed on test no. 14 till 21. From test no. 22 to 28, it proved that the increase of momentum factor caused the correlation coefficient become higher with three of hidden neurons. Like test before, it ceased to produce positive results when momentum factor at the range from 0.5 to 0.7. From the test run until now, it could be concluded that the ideal of learning rate will be between 0.1 till 0.2 and the momentum factor at range from 0.5 to 0.7.

**Table-6.** Comparison of network parameters.

| Test No. | HN | LR | MF | $R^2$ | RMSE |
|---|---|---|---|---|---|
| 30 | 4 | 0.1 | 0.5 | 0.9993 | 0.0025 |
| 31 | 4 | 0.2 | 0.5 | 0.9995 | 0.0023 |
| 32 | 4 | 0.3 | 0.5 | 0.9994 | 0.0025 |
| 33 | 4 | 0.1 | 0.7 | 0.9994 | 0.0025 |
| 34 | 4 | 0.2 | 0.7 | 0.9994 | 0.0025 |
| 35 | 4 | 0.1 | 0.8 | 0.9993 | 0.0027 |
| 36 | 5 | 0.1 | 0.7 | 0.9993 | 0.0029 |
| 37 | 5 | 0.2 | 0.7 | 0.9993 | 0.0026 |
| 38 | 5 | 0.3 | 0.7 | 0.9987 | 0.0034 |
| 39 | 6 | 0.1 | 0.7 | 0.9994 | 0.0026 |
| 40 | 7 | 0.1 | 0.7 | 0.9987 | 0.0032 |
| 41 | 8 | 0.1 | 0.7 | 0.9995 | 0.0023 |
| 42 | 9 | 0.1 | 0.7 | 0.9996 | 0.002 |
| 43 | 10 | 0.1 | 0.5 | 0.9997 | 0.0017 |
| 44 | 10 | 0.2 | 0.5 | 0.9996 | 0.0023 |
| 45 | 10 | 0.1 | 0.7 | 0.9997 | 0.0019 |
| 46 | 11 | 0.1 | 0.7 | 0.9997 | 0.0019 |

*HN-Hidden Neurons, LR-Learning Rate, MF-Momentum Factor

In order to prove that theory, additional test run with four of hidden neurons have been conducted which the results showed on the test no. 29 to 35. The value of correlation coefficient maintained consistent when the momentum factor value set at range 0.5 till 0.7 and the results became negative once the momentum factor set at

VOL. 10, NO. 2, FEBRUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences



out of the specific range. When the increase of learning rate till 0.2, the results still quite positive and became inverse with learning rate set at 0.3. By applying the proven theory, the only network parameter which had to be altered for the remaining test run was number of hidden neurons. From test no. 43 and so on, the correlation coefficient became consistent even increasing the number of hidden neurons. It showed that the results will be an ideal output with the settings of network parameters on test no. 43.

Based on the highest index of correlation coefficient (0.9997) and lowest index of root mean squared error (0.0017), the ideal of network parameters have been decided was network model with number of hidden layer of 10, 0.1 of learning rate and 0.5 of momentum factor. The network structure consisted of 3 input neurons in input layer, 10 of hidden neurons within single hidden layer and one neuron in output layer which been described as 3-10-1 structure.

**Results of testing datasets**

In the model training course for pipeline corrosion rate using ANN, defect depth and length in between two different years were employed as input variables while corrosion rate acted as output variable. The samples were collected from year 1990 till year 2001 and been divided into four distinct groups of training and test samples. First, second and third groups of samples comprised of same instances of data which were 417 samples with the time duration of year 1990 to 1992 (first group), year 1992 to 1995 (second group) and year 1990 to 1995 (third group). Fourth group of samples contained of 105 samples from year 1996 till year 2001. As the four sample groups, 70% of datasets were selected as the training samples accompanying with 15% of datasets used as validation samples and the other 15% of datasets acted as the test samples.

First, the ANN models were trained using the training samples, the network parameters were optimized through continuous training adjustment. Based on the result from Section 5.4, the optimal parameter subsets obtained by training via training samples were 10 numbers of hidden neurons, learning rate of 0.1 and momentum factor of 0.5. After that, the simulation was carried out by using test samples.

Comparison between experimental rates and predicted rates calculated by ANN for different training and test datasets were conducted and illustrated as small parts of the results as the data from Tables 7 and 8. The result comprised of the simulation using test samples only. The number of test samples for fourth group was 31 instances from total samples of 105 instances while the amount of test samples for other remaining groups was 125 instances from total samples of 417 datasets. But, for simplicity only one group of test sample is shown here.

As the output of the ANN models, the experimental rates acted as the target values for the designed models to achieve while the predicted rates was the outcome values that the designed models calculated out. The percentage error was the differences between the experimental rates and predicted rates in percentage. The smaller the percentage error was, the better the output produced from the designed models which indicated that the designed models able to act as an ideal models for prediction performance.

The prediction results for first samples group as shown in Table-8 with the accuracy of model was 0.9999 which near to perfect index, 1.0. In other terms of comparison, the designed model produced 0.0074 of mean absolute error and 0.0102 of root mean square error, accompanying with relative absolute error in percentage of 1.1899%.

**Evaluation of the ANN performance**

The graphs showed in Figure 1(a) - (d) presented the pattern of data intersection for the actual experiment values against the prediction values of designed models. Four graphs of data interpretation which each of them represented one of the samples group using the ANN model. The X axis represented as the actual experimental rates, while the Y axis represented as the predictive rates by the ANN models.

**Corrosion growth rate analysis**

The pattern of corrosion growth of a pipeline for each single defect could be calculated based on two or three sets of ILI data from the same pipeline segment. Normally, the corrosion growth rate was calculated using a simple linear equation (Noor, 2006). The linear equation is shown in Equation 5.

**Table-7.** Comparison between actual rates and estimated rates for first samples group (90-92).

| Instances | Experimental rate | Predicted rate | Percentage error |
|---|---|---|---|
| 1 | -0.77 | -0.767 | 0.003 |
| 2 | 0.77 | 0.785 | 0.015 |
| 3 | -0.11 | -0.106 | 0.004 |
| 4 | 2.42 | 2.390 | -0.030 |
| 5 | 0.22 | 0.230 | 0.010 |
| 6 | -0.44 | -0.436 | 0.004 |
| 7 | -0.33 | -0.327 | 0.003 |
| 8 | 0.77 | 0.787 | 0.017 |
| 9 | 0.33 | 0.341 | 0.011 |
| 10 | 1.98 | 1.983 | 0.003 |
| 11 | 0.00 | 0.007 | 0.007 |
| 12 | 0.00 | 0.006 | 0.006 |
| 13 | 0.33 | 0.341 | 0.011 |
| 14 | -0.33 | -0.329 | 0.001 |
| 15 | -0.11 | -0.108 | 0.002 |

VOL. 10, NO. 2, FEBRUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

**Table-8.** Prediction results for first samples group (90-92).

| | |
|---|---|
| Correlation coefficients $R^2$ | 0.9999 |
| Mean absolute error, MAE | 0.0074 |
| Root mean square error, RMSE | 0.0102 |
| Relative absolute error, RAE (%) | 1.1899 |

$$CR = \frac{d_{T2} - d_{T1}}{T_2 - T_1} \quad (5)$$

where

CR = corrosion growth rate
dT1 = corrosion depth in year T1
dT2 = corrosion depth in year T2
T1 = year of inspection T1
T2 = year of inspection T2

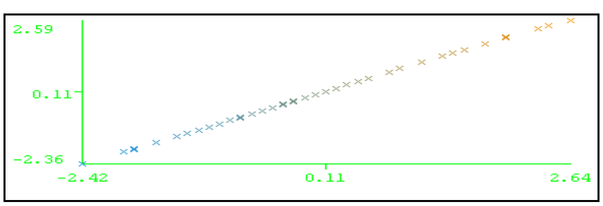
Moreover, the depth of located defect could also be calculated vice-verse using Equation 6.

dT2 = dT1 + CR x (T1 - T2) Equation 6

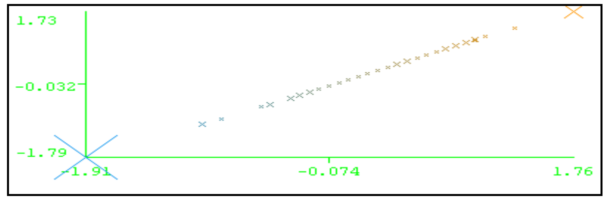
The analysis of corrosion growth rate can be presented in term of the average and standard deviation. These results were then use for reliability modeling of the pipeline, provided with other parameters such as the limit state functions or the safety factor and the failure pressure of that particular pipeline.

## CONCLUSIONS

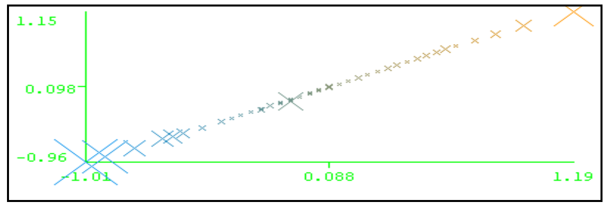
The study will contribute to a betterment of pipeline integrity assessment practiced among pipeline operators globally. It is an attempt to find an alternative solution towards predicting a corrosion growth rate based on available corrosion defect data (ILI) besides the deterministic and statistical methods that been widely employed. The application of ANN of this study is pioneering towards applying an artificial intelligence method in reliability based corrosion management systems to cater the uncertainties and vagueness possess by inspection methods and modeling. Moreover, the results can be used for fitness for service evaluation that contributes to future maintenance cost of the pipeline.



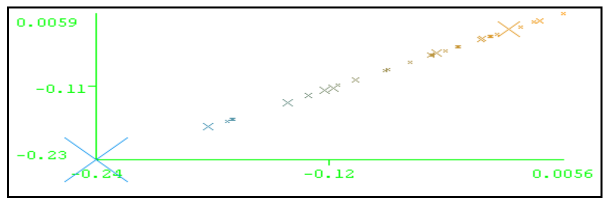

(a) Testing respond of first samples group




(b) Testing respond of second samples group




(c) Testing respond of third samples group




(d) Testing respond of fourth samples group

**Figure-1.** Comparison of the pattern of actual values against predicted values for ANN model.

## ACKNOWLEDGEMENTS

The authors are pleased to acknowledge Ministry of Science, Technology and Innovation, Malaysia (MOSTI), and the Ministry of Education (MOE) for the support of providing the research funds and scholarship.

## REFERENCES

Sumarni D., Supriyatman K. A., Sidarto R., Suratman R., Dasilfa. 2012. Artificial Neural Network for Corrosion Rate Prediction in Gas Pipelines. Proceedings Indonesian Petroleum Association.

Picton P. 2000. Neural Networks 2nd Edition. Palgrave, ISBN 0 333 80287 X.

Kreislova K., Haloma M. 2008. Atmospheric corrosion of structural metals - methods of prediction of corrosion attack. In Eurocorr 2008: The European Corrosion Congress, Edinburgh, 7th - 11th Sept. pp. 1-10.

Kenny D. E., Ramon S. C. 2009. Artificial neural network corrosion modelling for metals in an equatorial climate. Corrosion Science. 51: 2266-2278.

Cai J., Cottis R. A. and Lyon S. B. 1999. Phenomenological modelling of atmospheric corrosion using an artificial neural network. Corrosion Science. 41(10): 2001-2030.

Tran H. D., Ng A. W. M., Perera B. J. C. 2007. Neural Networks Deteriorations Models For Serviceability Condition of Buried Stormwater Pipes. Engineering Applications of Artificial Intelligence. 20: 1144-1155.

Sinha S. K., Pandey M. D. 2002. Probabilistic neural network for reliability assessment of oil and gas pipelines. Computer. Aided Civil Infrastructure. Eng. 17(5): 320-329.

El-Abbasy M. S., Senouci A., Zayed T., Mirahadi F., Parvizedghy L. 2014a. Artificial Neural Network Model for Predicting Condition of Offshore Oil and Gas Pipeline, Automation in Construction. 45: 50-65.

El-Abbasy M. S., Senouci A., Zayed T., Mosleh F. 2014b. A condition assessment model for oil and gas pipelines using integrated simulation and analytic network process. Journal of Structure and Infrastructure Engineering, Taylor and Francis Group, (in-press).

El-Abbasy M. S., Senouci A., Zayed T., Mirahadi F., Parvizedghy L. 2014c. Condition prediction models for oil and gas pipelines using regression analysis. ASCE J. Constr. Eng. Manag. 140(6).

Paik J. K. and Thayamballi A. K. 2002. Ultimate Strength of Ageing Ships. Journal Engineering for the Maritime Environment. 216: 57-77

Caruana R., Niculescu-Mizil A. 2006. An empirical comparison of supervised learning algorithms. Proceedings of the 23rd international conference on Machine Learning, ACM, Pittsburgh, PA, USA.

Prabhakar M. D. 2013. Prediction of software effort using artificial neural network and support vector machine. Int. J. Adv. Res. Comput. Sci. Softw. Eng. 3(3): 40–46.

Mohana R. S., Thangaraj P. 2013. Machine learning approaches in improving service level agreement-based admission control for a software-as-a-service provider in cloud. J. Comput. Sci. 9(10): 1283-1294.

Tabesh M., Soltani J., Farmani R., Savic D. 2009. Assessing pipe failure rate and mechanical reliability of water distribution networks using data-driven modeling. J. Hydroinf. 11(1): 1-17.

Marcous G., Rivard H., Hanna A. M. 2002. Case-Based Reasoning System for Modeling Bridge Deterioration. Journal of Computing in Civil Engineering. ASCE. 16(2): 104-114.

Mabbutt S., Picton P., Shaw P. and Black S. 2012. Review of Artificial Neural Networks (ANN) applied to corrosion monitoring. Journal of Physics: Conference Series. 364.

Dasu T., Johnson T. 2003. Exploratory Data Mining and Data Cleaning. John Wiley and Sons Inc. New York, USA.

Desjardins G. 2002. Improved Data Quality Opens Way for Predicting Corrosion Growth and Severity. Pipeline and Gas Journal, December. pp. 28-33.

Noor N. M. 2006. A Generic Assessment Approach to the Analysis of Corrosion Data and its Application to Reliability Assessment. Heriot-Watt University, Phd Thesis.