Journal of Engineering and Applied Sciences 13 (23): 9908-9913, 2018
ISSN: 1816-949X

Using Differential Evolution with Neural Networks Forecasting Model Creating for Pipeline Corrosion

Abdul Sttar Ismail Wdaa
Department of Mathematics, Faculty of Education for Pure Sciences, University of Anbar, Anbar, Iraq

Abstract: Pipeline corrosion is among the most critical and precarious causes of pipeline incidents which is observed year after year. As these pipeline incidents give rise devastating harms to people as well as to the economy and ecosystem of a country. Monitoring this component, pipeline operators have installed a more systematic and comprehensive program for pipeline inspection by different sensors for the attainment of data that may be helpful to gauge the existing pipelines state. However, in this corrosive process different factors are involved which cause erosion, therefore, current inspection methods are not sufficiently particular in the measuring process. Hence, a prediction model, capable to measure precise corrosion damage mechanisms is required to develop. The most apposite method to be adopted for such model is Artificial Neural Networks (ANN). Among the existing works on ANN, a critical research has proved the requirement to develop time effectiveness of the technique. A hybrid prediction model is developed in this research which can measure particular corrosive mechanisms. An elementary ANN Model is enhanced by incorporating the Differential Evolution (DE) algorithm in order to acquire an improved and ideal performance. The obtained hybrid model will be tested with industrial dataset of world to approve its time effectiveness as compared to the elementary ANN Model.

Key words: Artificial neural network, damage mechanism, corrosion, differential evolution, forecasting model, time

INTRODUCTION

The industry of oil and gas is among the biggest and the most concentrated industries in the world as indicated by the report of American Petroleum Institute. Presently, there are more than millions of kilometers of gas and oil pipelines being connected and are daily used in the world (Demma *et al*., 2004; Reber *et al*., 2002).

Mostly pipelines are prepared by steel as they transport the secure intends to deliver huge amounts of gas and oil products. These steel pipelines have a tendency to be deteriorated when presented to numerous corrosive mechanisms after some time (Reber *et al*., 2002; Lowe *et al*., 1998) in spite of the use of insulation. These corrosive mechanisms, i.e., sulfidation, cavitation and CO_2 corrosion ultimately bring about corrosion and the pipeline is prone to ruptures, breaks and leakages, leading to the financial loss to the operators and ultimately, cause extensive Health, Environmental and Safety (HSE) threats to the nearby ecosystem (Singh and Markeset, 2009; Hirao and Ogi, 1999). Observing this fact, the operators keep on inspecting pipelines for a long time to ensure their smooth operation and to minimize the risk of sulfidation, cavitation and CO_2 corrosion (Rose, 2004). Sensors are used for these observations that are added in definite parameters in pipelines and store them in a database. Prediction techniques are employed to foresee and observe the pipelines condition by using this corrosive data to regulate protective actions to be made in front of a potential occurrence.

Table 1: Existing prediction methods

Index	Years	Incidents
1	2000	390
2	2001	350
3	2002	650
4	2003	690
5	2004	690
6	2005	710
7	2006	650
8	2007	600
9	2008	650
10	2009	600
11	2010	600
12	2011	580
13	2012	605
14	2013	590
15	2014	700

Even though all-inclusive measures have been employed all through the years, pipelines are still undergoing corrosive process and incidents of pipeline are still happening in the whole world, producing fatal results. The annual incidents of gas and oil pipelines that happened from 2000-2014 are shown in Table 1 incidents

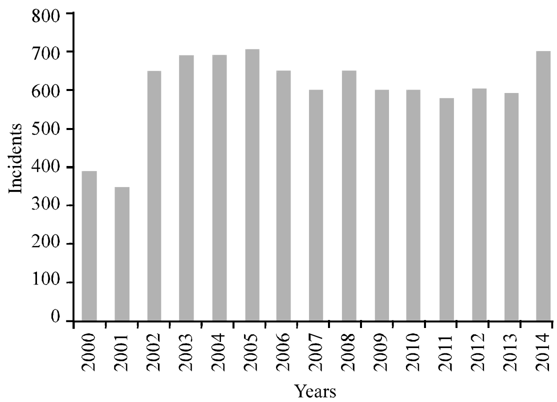

Fig. 1: Number of pipeline incidents from years 2000-2014 (Anonymous, 2016)

and Fig. 1, according to the report of the Pipeline and Hazardous Materials Safety Administration (PHMSA) (USDT., 2016). Even though, there is an irregular example of decay and increment but it can be observed that the pipelines episodes trends are generally increasing. It is imperative to monitor and foresee the condition of gas and oil pipelines to avoid the casualties of pipeline incidents.

The percentage of corrosive Pipeline incidents is around 25% of the total annual number of incidents in the world as reported by Transportation's Research and Special Programs Administration, Office of Pipeline Safety (RSPA/OPS) which is a Department of US (Anonymous, 2016). A Malaysian gas and oil company corrosion engineer agrees with the stats that the corrosion of pipeline is more serious and records for more than 35% of pipeline incidents and failures in Malaysia. The stats show that the existing corrosion prediction techniques employed in the gas and oil space have not yet possessed the capacity to adapt the corrosive problems. Therefore, an enhanced prediction corrosion model is required that fulfills the gaps in the present strategies.

Literature review

Corrosion data: The industry of oil and gas contrary to most industries has dealt with large quantities of data for a long time to make technical assessments, like the monitoring of pipeline fitness, hence, the ultrasonic waves are used for this purpose. Ultrasonic sensors are inserted at particular sectors of pipelines that determine the current corrosion rate and measurements on the pipeline wall thickness (Anonymous, 2016; ILLC., 2016). A part from this, several other parameters of environment in which the pipeline is presented to are additionally gathered (Veiga *et al*., 2005). Conjointly, they can be alluded to as corrosive data and can be encouraged into algorithms which foresee the corrosion rate (Krautkramer and Krautkramer, 2013).

Corrosion forecasting methods: Models are still incapable to judge particular damage mechanisms that cause pipeline corrosion, despite of possessing extensive data. Resultantly, recognition of damage mechanisms and current analysis still depend entirely on the experience and understanding of human (Singh and Markeset, 2009; Veiga *et al*., 2005). The recent models possess limited accuracy, since, the related parameters that lead to particular damage mechanisms are not observed specially. A part from this certain organizations only have slight.

Information related to the actual properties of damage mechanisms and are recently making suppositions about its nature (Supriyatman *et al*., 2012). High priority" or "Risk" and even a slight error in the expected result can lead to most important results (Black and Baldwin, 2012).

ANN is selected to be the central model to be applied for this study after studying and comparing the current prediction methods in Table 2. The intricate nature of corrosive process creates complications in the modeling of damage mechanisms, as the study shows, hence, the capacity of ANN Model to complex model relations is very helpful to the study.

The dataset will be acquired from a Malaysian oil and gas company and University Technology Petronas (UTP) centre of corrosion for this strategy, consequently, the dependency of this plan on dependable input data should not be of supreme value. So, ANN, performing the best with a dependable dataset is appropriate to be employed. Nonetheless, the ANN Model is required to be backed up employing numerous other algorithms or optimization techniques to overpower its weaknesses in terms of poor time efficiency and long training time.

Current research on ANN Model in the domain of gas and oil: The correlation among three current models on the ANN are shown by Table 2, like the prediction in oil and gas domains. Supriyatman, Sumarni, Sidarto and Suratman's recent research in 2012 (Supriyatman *et al*., 2012) has verified the suitability of ANN to be applied in the gas and oil space as it exhibits a high precision in predicting complex associations. The finding of this study is likewise helped by Ren *et al*. (2012) and Sinha and Pandey (2002) whose results have verified that nonlinearity in the variables employed in their experimentations are precisely represented by ANN.

Table 2: Existing prediction methods comparison in the gas and oil domain

Forecasting model	Fault tree analysis	Mechanistic Model	Artificial neural network
Training time	Low	Very high	High
Reliability on data	Low	High	High
Accuracy	High	Very high	Very high
Advantages	Exploits finding of skilled human operators	It is the outcome of detailed and all-inclusive experimentation	Capable of signifying complex forms of relationships Good to signify non-linear relationships
Disadvantages	Limitations of the human knowledge	Needs perception of the fundamental chemical, electrochemical and transport processes	Dependent on trustworthy input data
Disadvantages	Much time is required to scheme the fault tree	Much time is required to produce a consistent model for a single DM	Long training time

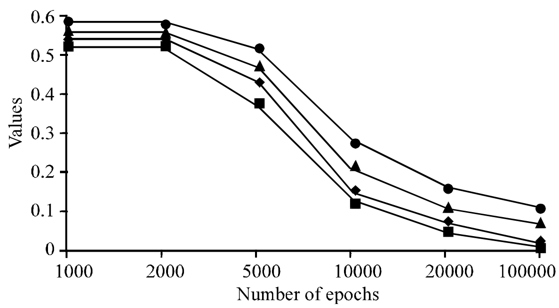

Fig. 2: Comparison of average network error with different network topologies (Tong, 2015)

These studies (Ren *et al*., 2012; Sinha and Pandey, 2002) clarify that the choice of neural network topology is completed physically. The ANN Model's short time efficiency or prolong training time is another problem (Supriyatman *et al*., 2012; Sinha and Pandey, 2012). ANNs show prolongs training time because of its learning nature.

From past dataset. The training epoch's number is indirectly proportional to the network error, the network error decreases by increasing the number of training epochs and the prediction correctness increases for an extended training time.

Both complications demonstrate that ANN model area can be upgraded. The research of Krautkramer and Krautkramer (2013) has proved that diverse network topologies influence on the efficiency that how quick the network learns. Different error rates of various topologies after a specific epochs number are displayed in Fig. 2.

Every one of the four lines denotes a distinctive topology. It can be realized from Fig. 2 that the 5-node entirely linked topology attains less error rates at the identical number of epochs while relating to the other three topologies. As various topologies attain a smallest error rate at dissimilar epochs number, it shows the required time to prepare the model's difference. Hence, it is conceivable either to apply an optimization algorithm or form of selection that is capable to perform choice of an optimum network topology employing ANN as the fitness function in place of manual choice (Supriyatman *et al*., 2012).

The investigators have suggested numerous recommendations for upcoming research, for example, an improved investigation that classifies the concerned parameters to be sustained into the neural network (Supriyatman *et al*., 2012) and to Decrease the required training time (Sinha and Pandey, 2002). The classification of the associated input.

Parameters in first proposal differentiates the motivation for this study which is to see particular corrosive mechanism by concentrating on the precise parameters that identify with them. The decrease in the training time is the second proposal, it can be explained by applying the algorithm of optimization to choose an optimum ANN topology.

Algorithms of optimization: Genetic Algorithm (GA) and Differential Evolution (DE), the two most generally used optimization algorithms have been observed. The comparative results of DE and GA are shown in the following Table 3.

On the basis of Table 4 by Ren *et al*. (2012), it is observed that DE overtakes GA in terms of the following four parameters: capacity to approach good solution outside local search, population under the effect of optimal solution, stability of search space and solution time under the influence of population size which are related to this study. First of all, DE shows a developed capability to approach a fine solution without doing local search, contrasting to GA, therefore, pulling down the time desired for the search. It is owing to the lesser continuity of search area in GA which might reduce its incompetency of creating all optimum results in the search area.

Besides DE, the optimum result has a greater effect on the populace when comparing to GA, assisting to protect search time as the consequent iterations will turn around the noted optimum result. The superiority of computational efficiency of DE over the GA having 99% confidence level is proved in the research (Ren *et al*.,

Table 3: Current works on ANN Model with comparisons

Researchers	Forecasting output type	Gap of study	Future recommendations
Supriyatman *et al*. (2012)	Numerical	Topology selection is manually completed Long training time/low time efficiency	To have an optimal research that identifies the related parameters
Ren *et al*. (2012)	Numerical	Topology selection is manually completed	-
Sinha and Pandey (2012)	Probability	Topology selection is manually completed	Decrease training time
Long training time/low time efficiency			

Table 4: DE and GA algorithms comparison

Parameters	DE	GA
Capacity to approach good solution outside local search	Higher	Lower
Population under the influence of optimal solution	Higher	Lower
Stability of search space	Higher	Lower
Solution time under the influence of population size	Linear	Exponential

2012). This has presented to be in accordance with the results acquired by Supriyatman *et al*. (2012) as the effect of population estimate on the solution time is exponential in GA as related to DE. Thus, the GA time consumed will dependably be longer and may not be appropriate for this development wheretime proficiency is an issue that requires to be upgraded.

Research novelty: From the literary study of current methods and exploration, it is proved that the novelty for this study is dual. Primarily, there are no current prediction models that are capable to gauge specific damage mechanisms that influence the pipeline erosion and further current techniques make their prediction without concerning to these damage mechanisms. Secondly, I will try to create a novel algorithm that optimizes both ANN topology selection along with ANN training to develop time efficacy of the whole model.

MATERIALS AND METHODS

The methodology for this project is showed by Fig. 3. Primarily, the attained erosion dataset will be prepared to fit the ANN Model. The corrosive data is first regularized and afterward separated into validation data and training data under data-preprocessing. The ANN Model is trained with the help of training data. The ANN Model performance is certified by using the validation data. Then, the data will be added into the DE, employing training time of ANN as the fitness function to choose an optimal topology for ANN. An optimal topology for ANN denotes a topology that shows a greater time efficacy than a topology that is choosing physically. The ANN will be processed employing the optimal topology together with more parameters of ANN Model. The optimal ANN model will be accomplished using the similar training data accomplished using the similar training data till a specific condition is seen. Initially stopping a technique to save the network from over fitting will be employed by preventing the training once the error of network noted is greater than its former reading (Table 5). The ANN Model will be validated in contrast to the validation data to limit its time efficacy if effectively trained. The final result will be in the form of the hybrid prediction model capable to note specific damage mechanisms for pipeline corrosion.

Table 5: Synthetic dataset in terms of some sample rows (Tong, 2015)

D1	D2	D3	States
0.4027	3.1099	69.1401	Corrosion
0.1497	3.2726	51.3000	Normal
0.4919	1.4049	53.5009	Acceptable
0.3493	2.4791	71.3176	Corrosion
0.0805	3.4467	50.2971	Normal
0.5633	1.2386	52.7457	Acceptable
0.4073	2.0708	72.8467	Corrosion
0.1297	3.0922	56.1964	Normal
0.5491	1.2265	58.6375	Acceptable
0.3759	2.2113	68.1399	Corrosion
0.1276	3.4726	52.6193	Normal
0.5074	1.3124	57.5068	Acceptable
0.2875	1.9848	70.7309	Corrosion
0.0982	3.6992	52.0072	Normal
0.4544	1.6525	53.9283	Acceptable
0.6421	2.3433	71.9723	Corrosion

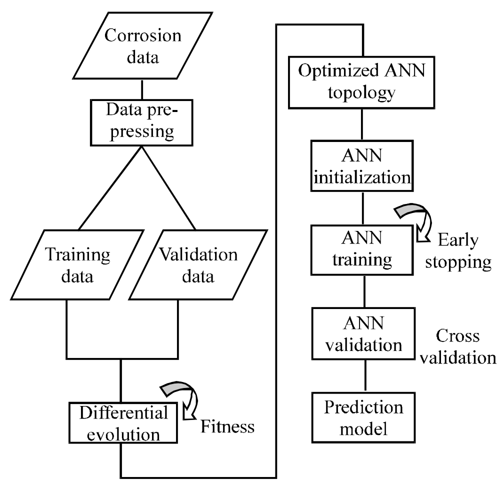

Fig. 3: The flow of research methodology

Initial proof of concept

Assessment Model: The synthetic dataset has been acquired by a literary survey (Black and Baldwin, 2012; Ren *et al*., 2012; Sinha and Pandey, 2012) and has been certified by a staff and senior corrosion expert to achieve

Table 6: Values of initial parameters with justification

Parameter	Values	Justification
Nodes for output	3	The output has 3 possible results
Nodes for input	3	The data comprises 3 predictors
Hidden nodes	3	Based on literary study where the number of hidden nodes is formatted between the number of 0utput and input nodes
Initial weights	Between 0-1	Based on literary study
Hidden layers	1	Based on literary study

Table 7: Results obtained for the proof of concept

No. of epochs	Total number of validation data	Total number of training data	Number of wrong prediction	Number of correct prediction	Average training time (sec)	Accuracy of prediction(%)
50	270	31.50	81	189	01.98	70.00
500	270	315.00	7	263	21.33	97.41
1.50	270	945,000	4	266	63.27	98.52

real principles industry and validated for use in this verification of concept. The dataset comprises 300 cross-sectional rows of data.

Synthetic dataset in terms of some sample rows displays in Table 5-7. There are three input parameters in every row of the data set: flow velocity (m/sec), CO_2 partial pressure (MPa), temperature (°C) and a classifier output that categorizes the strictness of erosion, it is possible that it is in severe "Corrosion" or "Acceptable" range or "Normal" range. The connection between the outputs and inputs denote the real nature of corrosion in subsea pipelines (Tong, 2015).

A basic ANN Model was established for the evidence of conception. It is a feed-forward Multi Layered Perceptron (MLP) and is trained using Back-Propagation (BP). The model consists of a hidden layer an output layer and an input layer, all have three nodes. Every node in each layer is related to every node in the subsequent layer, making a topology with 18 connections or weights. The ANN Model parameters are shown in Table 4 and validates the reasoning behind the formatted values.

RESULTS AND DISCUSSION

It is observed from Fig. 4 that the ANN is effective by creating high accuracy predictions on nonlinear erosion data. The model prediction accuracy has reached 98.52% after 1,500 epochs of training and it is essential due to the implication of ANN Model on real industrial data. It also shows nonlinearity and prediction of a high accuracy that verifies that ANN Model is a reasonable model for this study. It is observed that an increment in the number of epochs results in the linear rise of training time. Therefore, the training time shows time efficacy in this study, it may be observed that an increment in the number of epochs reduced the time efficacy linearly. Moreover, it is seen that the model prediction correctness rises logarithmically by increasing the number of epochs. A logarithmical development is showed by a development that begins quickly, succeeded by slow development that goes on increasing at a low rate. Henceforth, it can be decided from Fig. 4 that the time efficacy declines as the prediction

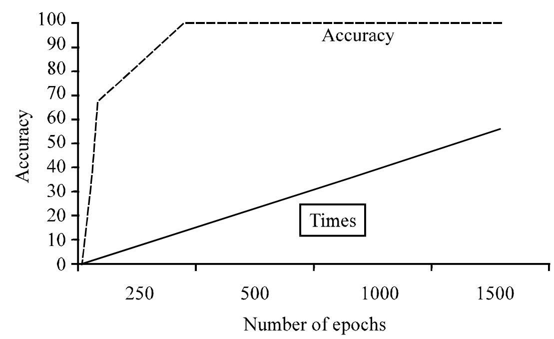

Fig. 4: Graph shows an increase in the number of epochs gives increase in prediction accuracy and training time

accuracy of ANN rises. This shows that ANN Model is capable to better its prediction correctness by exchanging off its time efficacy. Though, it is seen that the ANN training time is not important from the outcomes of the proof, it is notable that the synthetic data size is much lesser than the real industrial dataset. Each epoch takes more data in an actual industrial dataset which is translated to an enlarged training time (ILLC., 2016).

Hence, the suggested technique is optimum for the neural network topology employing DE to develop its time efficacy is recommended strongly. The logarithmic development revealed by the rise in prediction correctness approves the motivation for network optimization due to the slow rise in prediction accuracy, the next phases of training will assume more training time when endeavoring to train the ANN Model to proposed accuracy level.

So, the proposition of this study is that DE can be employed to discover an optimum topology for network that learns quicker by utilizing capitalize on the initial fast development of logarithmic curve. Thus, only a slight time will be used on the DE whereas preserving additional time on the real training of ANN Model (Tong, 2015).

CONCLUSION

This review study has explored complications concerning to the current prediction models and strictness of pipeline corrosion. The advantages and disadvantages of some well-known models utilized in the supposition of corrosion in the gas and oil field are presented. ANN Model is adopted and is selected as the model to be upgraded in this research.

A primary proof of idea is showed to prove the probability of ANN to be employed in this study. The collective outcomes from the proof and literary study have exposed that an optimal topology would permit the model to be trained quicker, simultaneously rendering a correct prediction associated to the pipelines state. It is significant to target the correct parameters that relate to particular damage mechanisms. The ANN will have a capacity to make predictions of a high accuracy and confidence level with input and the correct parameters. The next stage of study will comprise of ANN Model optimization by applying DE and detecting an appropriate method to use DE to choose a topology for network which increases ANN time effectiveness. Finally, the developed hybrid model can be set to training with the industrial data. It can also be used to make a prediction model that could be very useful for the gas and oil industry.

REFERENCES

Anonymous, 2016. Pipeline failure causes. U.S. Department of Transportation's Research and Special Programs Administration, Washington, USA. http://www.corrosion-doctors.org/Pipeline/Pipeline-failures.htm

Black, J. and R. Baldwin, 2012. When risk-based regulation aims low: A strategic framework. Regul. Governance, 6: 131-148.

Demma, A., P. Cawley, M. Lowe, A.G. Roosenbrand and B. Pavlakovic, 2004. The reflection of guided waves from notches in pipes: A guide for interpreting corrosion measurements. NDT. E. Intl., 37: 167-180.

Hirao, M. and H. Ogi, 1999. An SH-wave EMAT technique for gas pipeline inspection. NDT. E. Intl., 32: 127-132.

ILLC., 2016. Damage mechanisms. Inspectioneering, LLC., Ross Township, Pennsylvania. https://inspectioneering.com/topics.

Krautkramer, J. and H. Krautkramer, 2013. Ultrasonic Testing of Materials. Springer, Berlin, Germany,.

Lowe, M.J.S., D.N. Alleyne and P. Cawley, 1998. Defect detection in pipes using guided waves. Ultrasonics, 36: 147-154.

Reber, K., M. Beller, H. Willems and O.A. Barbian, 2002. A new generation of ultrasonic in-line inspection tools for detecting, sizing and locating metal loss and cracks in transmission pipelines. Proceedings of the IEEE Symposium on Ultrasonics Vol. 1, October 8-11, 2002, IEEE, Munich, Germany, ISBN:0-7803-7582-3, pp: 665-671.

Ren, C., W. Qiao and X. Tian, 2012. Natural Gas Pipeline Corrosion Rate Prediction Model Based on BP Neural Network. In: Fuzzy Engineering and Operations Research, Cao, B.Y. and X.J. Xie (Eds.). Springer, Berlin, Germany, ISBN:978-3-642-28591-2, pp: 449-455.

Rose, J.L., 2004. Ultrasonic guided waves in structural health monitoring. Key Eng. Mater., 273: 14-21.

Singh, M. and T. Markeset, 2009. A methodology for risk-based inspection planning of oil and gas pipes based on fuzzy logic framework. Eng. Fail. Anal., 16: 2098-2113.

Sinha, S.K. and M.D. Pandey, 2002. Probabilistic neural network for reliability assessment of oil and gas pipelines. Comput. Aided Civ. Infrastruct. Eng., 17: 320-329.

Supriyatman, D., K.A. Sidarto and R. Suratman, 2012. Artificial neural networks for corrosion rate prediction in gas pipelines. Proceedings of the International SPE Asia Conference on Pacific Oil and Gas and Exhibition, October 22-24, 2012, SPE Publisher, Perth, Australia, pp: 1-9.

Tong, H.C., 2015. A corrosion prediction model for subsea oil pipeline using enhanced associative classification. Proceedings of the International Conference on Computer and Information Sciences, Jun 28-July 01, 2015, IEEE, Las Vegas, Nevada, pp: 175-181.

USDT., 2016. Pipeline and hazardous materials safety administration. U.S. Department of Transportation, Washington, USA. https://www. phmsa. dot. gov/pipeline/library/data-stats/pipelineincidenttrends

Veiga, J., A. Carvalho, I.D. Silva and J. Rebello, 2005. The use of artificial neural network in the classification of pulse echo and TOFD ultrasonic signals. J. Braz. Soc. Mech. Sci. Eng., 27: 394-398.