4th Construction Specialty Conference
4e Conférence spécialisée sur la construction

Montréal, Québec
May 29 to June 1, 2013 / *29 mai au 1 juin 2013*

Predictive risk- based model for oil and gas pipelines

L. Parvizsedghy[1], T. Zayed[1]
[1] Dept. of Building, Civil, and Environmental Engineering, Concordia University,

Abstract: Among different means of oil and gas transportation, pipelines are considered to be the safest and reasonably efficient. However, reviewing the incidents recorded on oil and gas pipelines in the United States of America (1992 – 2012) proved that the consequences of failures in these pipelines have been considerable: over 5.5 billion dollars of asset damage, 380 fatalities and 1,500 injuries, in addition to more than 5.5 million barrels of product loss. The domain certainly requires attention. In addition, there is a lack of research available in this crucial field. Therefore, the objective of this research is to develop a risk model for oil and gas pipelines. Data from USA incidents recorded on oil and gas pipelines from 1992 to 2011 are utilized to build the intended model and identify risk factors leading to oil and gas pipeline failure while considering their consequences of failure. An artificial neural network (ANN) with two hidden layers through applying back propagation approach is trained to develop the model which will help decision makers to estimate how significant would be a failure on a pipeline for the identified risk factors.

1 Introduction

According to the United States' Department of Transportation (DOT1, 2012) more than 63 percent of US energy is provided through oil and gas products; and the only practical way of transporting them is through pipelines. They carry raw materials from wellheads to the processing facilities and then transport the final product to the customers. The operation of more than 2.5 million miles of oil and gas pipelines in the USA by around 3,000 companies has not been without hazards despite the fact that they are considered the most trustworthy way of transporting petroleum products. The statistics prove the necessity of regular inspections and repair or replacement of pipelines. On the other hand, many inspection tools have been developed to survey different types of failure sources as well as several types of repairing pipelines for various sources of failure. Repair manuals and guidelines recommend which technique to apply at specific situations. The examples of the inspection techniques include different types of inline inspection tools (intelligent pigs) to inspect internal and external corrosions and find out the criticality of each defect. Moreover, different types of sleeves as well as various clamps have existed to maintain those pipelines in critical situations. However, these are expensive to run regularly or at specific times and require careful selection of required inspection and repair technique. Accordingly, failure modelling is required to estimate how critical the situation of the pipelines is based on their specifications. Risk assessment is a tool that can help recognise possible sources of failure within a pipeline and measure its consequence of failure. Fortunately, data on failures of pipelines which have been recorded since 1970 is of help in guesstimating risk of pipeline failures, though they are not very complete. In this paper, the authors develop a model to estimate the consequences of failure of pipelines. The model is developed by applying artificial neural network (ANN) to forecast direct monetary consequences of each failure type of oil and gas pipelines. This learning machine is a type of pattern recognition method which is applied to identify a non-linear relationship between inputs and outputs of historical data. In this research, ANN is applied for the estimation of the probable consequences of pipelines' failure. The model is developed to predict monetary consequence of each failure type for pipelines considering their special attributes. The model would be useful for oil and gas operators to forecast the failure consequences of their pipelines and prioritise the pipelines of a network to do maintenance actions in the case of budget deficit.

2 Background

There are few examples of research on oil and gas pipelines that have tried to develop failure and risk models of these types of infrastructure. Dey et al. (2004) have done research in this field aiming to develop a risk-based maintenance model for offshore pipelines. After introducing likelihood and consequence loops of risks, experts' opinions were applied to calculate the relative weights of each factor of loops through analytical hierarchy process (AHP). Also, range of scores from 1 to 10 presents the effect values of each factor. The model calculated the risk score of each pipeline by summing up the score of effect value multiplied by related weights of each factor. Finally, one of the results was prioritising assets of a network or pipelines under control of one operating company. Most of the factors such as corrosion were scored in a subjective manner; though, the research tried to minimise the subjectivity of the decision making process in this problem. This research did not recognise the severity of different risks of failures: as a result, inspection tools were proposed through an experience-based process. Similar research by Al-Khalil et al. (2005) ranked a group of cross country pipelines with the benefit of AHP. It classified risks of failure in seven groups: corrosion, mid wall defect, external interference, structural defects, operation problems and loss of ground support. Then, experts scored probability and cost of failure for each pipeline against identified risk factors to calculate the overall expected cost of failure for each pipeline. These scores were later used to prioritise pipelines against the budget. This research tried to offer a "systematic risk- based approach" to prioritise a group of pipelines, yet it lacks a way to objectively prioritize the pipeline segments for repair. Zeng and Ma (2009) developed a risk model for underground pipelines which applies two sets of variables: general and inspection. The variables were correlated to five major types of failures: shape, seam, structural failures, pipe alignment, and blockage. Then, the model considered consequences of failure, cost, performance, interruption, and safety. And finally offers a maximum average method to maximise effect of severe consequences in risk score of pipelines. As the author has described, this model lacks any rating index to calculate probability of failure and only proposes an ordinal table of scales for different consequences; the absence of objectivity is apparent in this model just as in the previous ones. An artificial neural network(ANN) is used in Abdrabou's(2012) model to predict the most probable source of failure for pipelines as a binary value. There are three types of failures in this model: corrosion, mechanical or third party. The accuracy of the model is acceptable; although, except for age the other factors of the model including type of product, location, land use and diameter remain constant over the life of a pipeline. Consequently the model does not represent the changes that may happen in the environment and pipe itself. In addition, as the author has mentioned, it uses a limited number of factors that can be developed to forecast the failure rate of other types of failures. These limitations, as identified by the author are mostly due to the model's reliance upon the Concawe database (Davis et al. 2011) which has recorded data on a limited number of factors.

There are several predictive learning approaches, which can be employed to recognise a pattern among input variables and output(s). Considering properties of data in this case and the existing uncertainty, the artificial neural network is identified by the authors as an efficient solution of the aforementioned problem. Neural networks can employ a considerable range of learning models; here we apply back-propagation approach, which is very useful in construction management research. Christodoulou (2004) applied neural networks for research on optimum construction cost markup calculation, Hegazy (1993) applied neural network for bid preparation, Siqueira (1999) for cost estimating, Attalla and Hegazy (2003) for "Predicting Cost Deviation in reconstruction Projects", Al-Barqawi and Zayed (2006) in condition rating of water mains, and Achim et al. (2007) predicted remaining life of water pipes, applying neural networks. Based on all of this research, it can be seen that the approach has a wide application in construction problems, such as cost estimating to predict performance of engineering activities (Maged et al. 2004), and develop a model to estimate the productivity of pile construction (Zayed & Halpin, 2005). However, ANN is still new in predicting risk severity in pipelines and it will assist pipeline operators predict the consequences of identified risks to their assets.

2.1 Proposed Methodology

Historical data are the inputs of neural networks and a network is developed for each risk factor. The neural network trains itself through data entries and finds the relationship between inputs to forecast the output which is the monetary consequence of pipeline failures. ANN imitates the function of a human brain and is very “fault tolerant” and is able to generalise, hence these properties make it suitable for construction management issues. This technique provides a better platform for risk management research, since construction problems carry a lot of uncertainty. Although the concept is very easy to understand, the method has a complicated mathematical approach. On the other hand, MATLAB has developed a toolbox to apply neural network and we have used this software to develop our model.

2.2 Overview of ANN

Two main functions exist in ANNs. First, a function to find the relationship between variables is applied in the learning phase which is controlled based on the error of produced network. The second function is called recalling network which inserts the inputs to the trained network and creates predictive responses. Moreover, if the entry data includes output in the training phase it is called supervised, otherwise it is called unsupervised. (Zayed & Halpin, 2005) Artificial neural networks have different layers and there are some processing elements (PE) in each layer, which mimic the act of neurons, thus it is called neural network. It is very important to design the architecture of the network. The simplest network would have one input, one middle and one output layer as presented in figure 1. Middle layer is usually called hidden layer and its number may increase based on the complexity of the problem.

Neurons of each layer of ANNs are connected to the neurons of the next layer through connection lines. Each connection has a weight which is multiplied by the inputs transferred from the previous layer. In the end, they are summed up with a constant value called bias. (Moselhi et al. 1991) A transfer or activation function is used to create non-linear relationships between inputs and outputs. Sigmoid (logistic), hyperbolic tangent (tanh), the sine or cosine and linear function are the most frequently used transfer functions. Among them, the sigmoid function is the most commonly applied in construction problems. (Zhang et al.1998) Equations number one to five present these functions respectively. X represents values of the nodes. Outputs of these functions are transferred to the next layer. Performance of a network would be enhanced if the learning process is stopped sooner. Therefore, the network checks the pattern at the stopping points, called epochs, in order to stop training at the point that the learning rate starts to increase.

[1] Sigmoid function: $\mathrm{f(x)} = (1 + \exp(-\mathrm{x}))^{-1}$ (Zhang et al. 1998)
[2] Hyperbolic tangent: $\mathrm{f(x)} = (\exp(\mathrm{x}) - \exp(-\mathrm{x}))/(\exp(\mathrm{x}) + \exp(-\mathrm{x}))$ (Zhang et al. 1998)
[3] $\mathrm{f(x)} = \sin(\mathrm{x})$
[4] $\mathrm{f(x)} = \cos(\mathrm{x})$
[5] $\mathrm{f(x)} = \mathrm{x}$

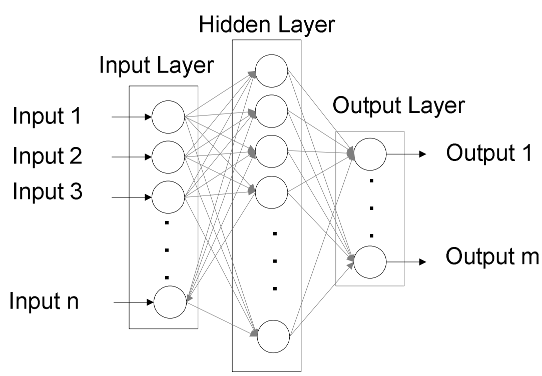


Figure 1: Typical Artificial Neural Network (ANN)

3 Risk Factors

Different classifications of risk factors have been considered in the literature. Various databases also have classified failures form different points of view including time dependency, sources of failures and frequency of happening. In this research sources of failure have been considered as the basis of grouping risks in the identification stage of risk assessment. Integrating the most frequent risks in different research and databases and regrouping them, we classified risks in three major groups: physical, external and operational. Figure 2 depicts the sub-categories of the risk factors identified in this research. Physical risks include four types of risks:

1. External corrosion: Interaction of pipeline's external surface with environment, changes iron to iron oxide and leads to external corrosion and finally structural disintegration. Cathodic protection and pipeline coating decrease this type of risk. (DOT2, 2012)
2. Internal corrosion: Internal corrosion happens when corrosive products, whether water or other chemicals, transfers through the pipeline, and leads to internal loss of pipe material. There are a number of mitigating actions to prevent pipelines from this risk such as injection of inhibitors and internal coatings. (DOT2, 2012)
3. Material and weld defect: Although production of steel has developed over time, some impurities remain in pipes and may lead to some defects in pipes resulting in failure. Generally the later the pipeline has been constructed the more reliable it is; nevertheless, inconsistencies persist to remain in materials and welds applied to join pipes.
4. Non-welded joints: Beside the defects of material and welding, some failures may happen over non-welded joints of pipes. These joints include flanges, fittings, etc.

The second group of risks takes account of external risks which may happen as a result of an external party or natural forces and consists of the following risks:

1. Earth movement: Movement of earth not due to heavy rains or floods may cause failure of pipeline.
2. Natural hazards: Heavy rains/floods, lighting, temperature and high winds that are grouped as natural hazards may be a source of failure of pipelines.
3. Sabotage: Intentional damages comprising vandalism, terrorism, theft of transported commodity and theft of equipment are classified under this group of risks.
4. Third party activities: Failures caused by third party activities either through excavation vehicles or automobile crashes into a pipeline are grouped under the category of third party activities.

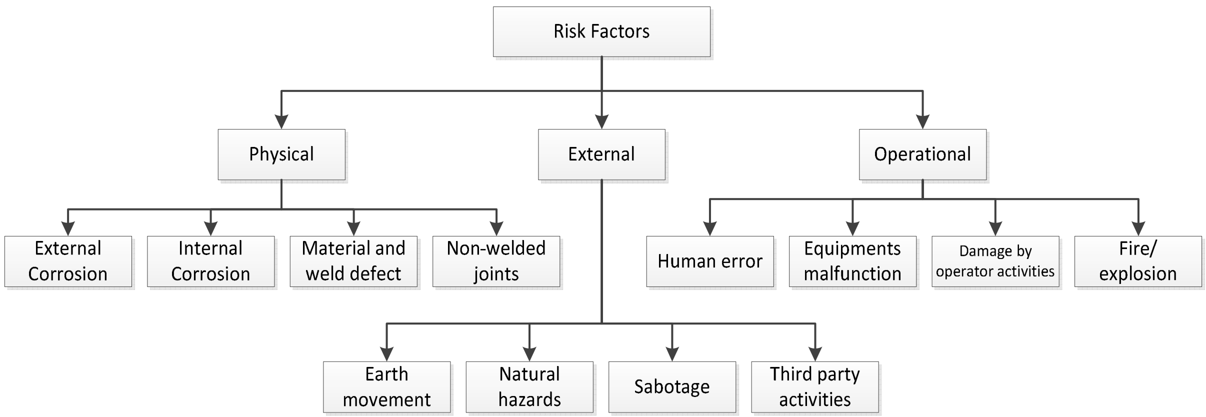
Finally, the last group is called operational risks. These are caused by human errors or activities of operating company of the pipeline and are comprised of four types:

1. Human error: Improper operations and activities by operator or contractor's personnel operating to failure.
2. Equipment malfunction: The breakdown or malfunction of equipment containing pumps and compressors, metering equipment, block, control, or relief valves, and tanks is considered as an equipment malfunction.
3. Damage by operator activities: Activities of operator around the pipelines to excavate or employ motorised vehicle for a special purpose may bring a failure.
4. Fire/ Explosion: Any fire or explosion as a primary cause of failure is taken into this type of failure.

4 Research Methodology and Model Development

Failures of oil and gas pipelines have monetary consequences as well as health, safety and environmental effects. The Department of Transportation (DOT1, 2012) has recorded five types of cost consequences for each incident, which includes "cost of public and non-operator private property damage", "cost of commodity released", "cost of operator's property damage & repairs", "cost of operator's emergency response", and the other costs. Among them, "cost of operator property damage" is directly related to the severity of failure, which we call "monetary consequences" in this research. If the monetary consequence is predicted, then estimating the severity of a failure in case of its occurrence would not be much harder to calculate. Thus, we strive to find a pattern between the severity of each risk factor, which may happen on a special pipeline, and identified variables. Moreover, we would like to optimise the variables contributed to each failure in order to have more efficient results. For the other types of consequences, it is very hard to find a similar pattern as it needs more data. For instance, to estimate the "cost of commodity released" we should forecast the spillage duration and product's flow rate. Spillage duration is dependent on lots of factors that make it more complicated and is hardly sensitive to the situation. In this research, direct monetary consequence has been the target of research and the methodology to discover pre-mentioned pattern will be clarified.

Figure 2: Risk factors classification



In this research we have developed a code in MATLAB 2010 software to compare variations of the architecture of the network. The code causes the software to be run repetitively to cover a range of neuron numbers and variables in order to optimise the number of neurons and variables. Khaw et al. (1995) propose to have (2n+1) neurons in the first hidden layer and (2n+1)/3 in the second one. Based on this recommendation the range of neuron numbers is defined in the code to be run for several times to cover this range. Moreover, the training process is repeated for the combinations of the variables to optimise their number in order to exclude non efficient ones and keep the variables which are more efficient. Results of the trained networks including their performance are saved after running. The best network is selected based on the performance of the network which was defined through minimum square error (MSE). Also, the network is adjusted by changing learning rate, the activation function, and the number of epochs, in order to obtain the least error on the generated pattern. Zhang et al. (1998) propose to standardise data of each set before training, since non-linear transfer functions restrict data to a limited range. Consequently, we have normalised the data to alter all to the range of zero to one.

After reviewing the literature in this field and identifying risks and consequences, we have developed a method to assess pre-defined risks based on their severity or consequences. Effective variables on risk factors of oil and gas pipelines are identified in the relevant research literature. However, to obtain a more effective model, logical combinations of variables are tested based on the error of network from each set of variables. As can be seen the logic in figure 3, for all of the possible combinations, the network has been tested to select the best combination and remove ineffective ones. Results of the developed model are presented in this paper to describe how it will be applied on each risk factor. Here, we have implemented the proposed model on external corrosion failures in hazardous liquid pipeline systems including crude oil products. The variables that have been identified through literature review and expert opinions are Coating type, Cathodic protection efficiency and existence (CP), Supervisory Control and Data Acquisition System (SCADA), Computational Pipeline Monitoring (CPM), area around the pipeline as well as age, diameter, pipe wall thickness, Operating Pressure (OP), and Specified Minimum Yield Strength (SMYS). The first five variables are non-continuous variables and the others are continuous. Primary identified factors are optimised based on minimising the mean square error (MSE) of trained networks. Based on a suggestion of Muhlbauer (2004), the values of diameter of the pipe and its wall thickness values have been combined. The variable “D/Th” is created through this combination and it has been considered a variable that can be related to the potential of the failure of pipelines.

For external corrosion risk factor, the model is run for 256 sets of variables (combinations of five out of nine pre-defined variables). Also, the networks with one and two hidden layers are tested with several numbers of neurons. However, two hidden layer networks result in higher performances based on the comparison of MSE amounts. For the sets of five variables there are five input neurons, therefore proposed number of neurons are eleven neurons in the first and four neurons in the second hidden layer. These are computed based on literature recommendation. (Khaw et al. 1995) Similarly, networks with nineteen and six neurons are suggested for the nine variables sets, for the first and second hidden layer respectively. For this purpose training phase of ANN is run through the code in the range of ten to nineteen neurons in the first hidden layer and four to ten neurons in the second hidden layer to cover the proposed numbers of neurons.

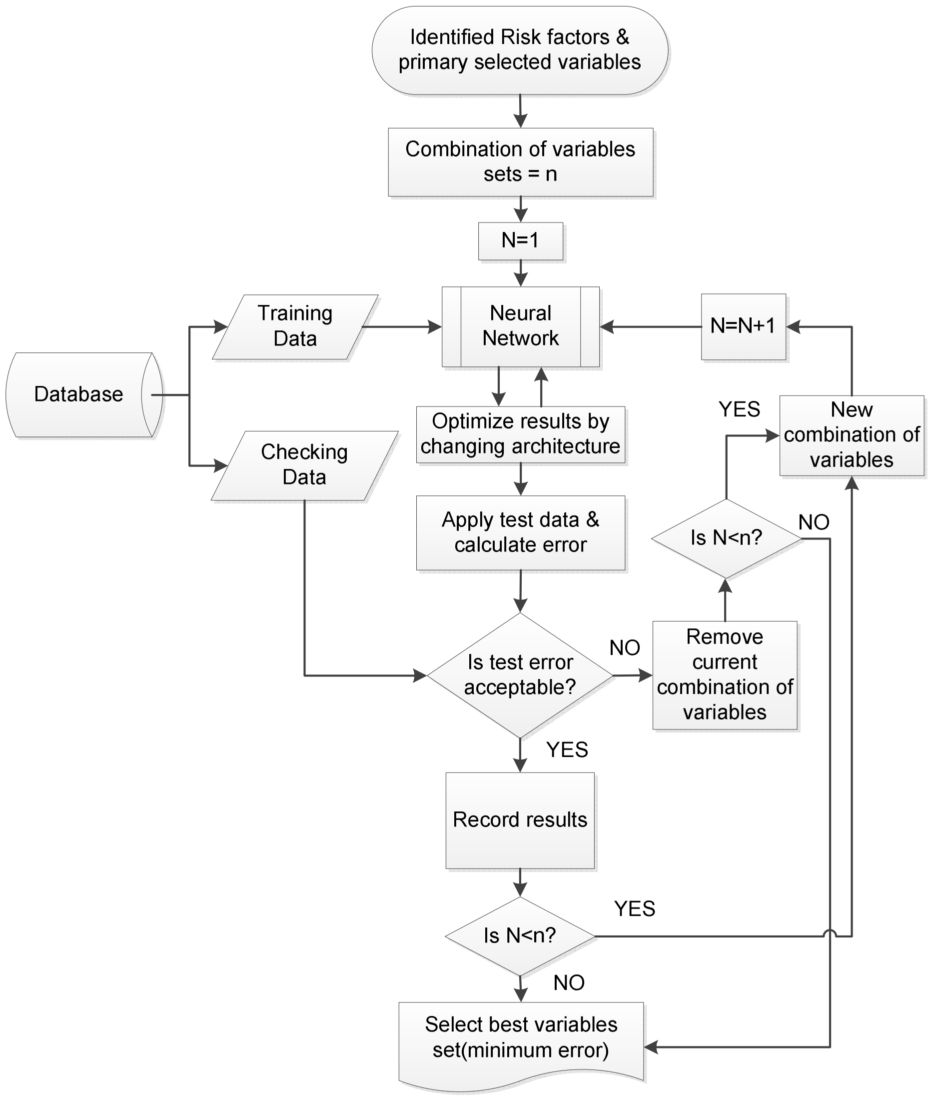


Figure 3: Overall Model Development Flowchart

5 Data Collection

Different database have recorded data of the failures of pipelines and a few have published data. Some have reported processed data on the causes and consequences of failures periodically and a few have published raw data on each failure. Among published databases, the authors found the database of the US Department of Transportation (DOT) to be the most complete one. DOT's record of failures of pipelines has been the main database used in our research to find a pattern between related variables and their consequences. This database includes data on more than 13,000 failures since 1986. Data on oil, distribution gas and transmission gas pipelines must be considered and analyzed separately, as they have different specifications and may behave distinctly in the case of a failure. In this paper, we have used the dataset from 2010-2012, which provides the most significant amount of data. In this manner, we could test more variables to develop an enhanced model. We have removed the data points that do not have values related to the variables considered in the model. Data is divided randomly into training and checking datasets with the ratio of nine to one. The dataset containing ten percent of data has been put aside to be used in the model validation. Then, training data is divided again in neural network, into training, testing and validation sets and the training process continues with the training dataset until the error in the testing and validation data is acceptable.

DOT has recorded general data of the pipelines, exact points of incidents, and information related to the operators. Furthermore, it provides detailed data about each failure: the cause of the failure, the cost and the environmental consequences of the incident, and the overall inspections that have been done during the pipeline's operation. Also, it includes data on the variables which have been selected to test the model on identified risk factors. Installation year of pipelines, date of failure, maximum allowable operating pressure, SMYS have been recorded exactly as a numerical value and there are some linguistic or binary values for some inspectional variables. For the variables related to inspection processes several items are integrated to produce one value representing the efficiency of related inspectional process. As a result, binary values are translated to numerical ones in several variables. Also, the same process is recurred for coating type with the purpose of translating different types of coatings to numbers. These variables include the inputs to the training phase of ANN. Output for each data point contains the monetary consequence of each failure type. For each failure type a network is developed in which the output for the training data contains the actual cost in the current dollars. For data points with a different type of failure, the consequence is changed to zero to represent ineffective failure type in that data point for the specific failure type. Table 1 summarizes ranges of values and categories attributed to each variable. "Con" is used to represent the continuous category of variables.

Table 1: Ranges of values and categories of defined variables

Variable	Age	Area	D/Th.	Coating type	SCADA	CPM	CP	OP	SMYS
Category	Con.	Label	Con.	Label	Label	Label	Label	Con.	Con.
Range of Value	0-112	0-3	4-122	1-6	0-3	0-4	0-4	10-2,220	52-80,000

6 Results of Model Implementation to Case Study

Table 2 summarizes a sample of actual values of the data embedded to the neural network to be trained. External normalization method is used in this paper which alters data to the range of zero to one. Equation 6 is used for normalisation of data, V represents values of the variable that is intended to be normalised and V_{min} and V_{max} represent minimum and maximum values of the same variable respectively.

[6] $\text{Normalised value} = \frac{\text{V}-\text{Vmin}}{\text{Vmax}-\text{Vmin}}$

Table 2: Sample of data entry

Age	Area	D/Th.	Coating type	SCADA	CPM	CP	OP	SMYS	PROP-DAMAGE Ex. Corrosion
33	2	72	6	3	0	1	285	24,000	150,000
40	1	64	2	3	0	0	275	24,000	0
56	1	31	2	3	0	1	188	24,000	35,000
56	3	64	1	3	0	0	275	24,000	0

The best performance is created by a network with two hidden layers. Table 3 presents the best MSE amounts for each group of subsets. Two subsets from the group of six variable networks and one from the group of five variable networks are selected for validation. Then the authors validate the model with the checking dataset comprised of 20 points of failures in the selected networks. Table 4 summarises selected subsets with containing variables as well as the amounts of MSE, root mean square error (RMSE) and correlation coefficient r for each one.

Table 3: Minimum MSE for each group of subsets

No of Variables	Min. MSE
5	5.62E-06
6	4.77E-06
7	1.25E-05
8	1.23E-05
9	0.000634

Among the chosen subsets, the authors selected second subset including six variables, which results in a lower RMSE and considerably higher r value. RMSE is a measure of error of predictive models that sums up squares of errors of forecasts through checking dataset and is calculated from equation 3. The closer the amount of RMSE is to zero, the more accurate is the model and here the authors have found a lower value of RMSE for the second subset. Correlation coefficient r indicates the relationship between predicted and actual amounts, and closer values to one present more fitted models. Architecture of trained network for the selected subset is presented in figure 4. This network includes six neurons in the input layer containing age, area type, D/Th, coating type, CPM efficiency and SMYS. Two layers are included in the hidden layer, the first one contains seventeen and the second one includes six neurons. The output layer is consisted of one output layer that is direct monetary consequence of pipelines' failures of external corrosion. For validation purposes the selected trained network is recalled with a code defined in MATLAB, then the checking dataset excluding values on selected variables is embedded into the trained network. The network predicts a value as the monetary consequence of each data point representing an actual pipeline from the database. Comparing the estimated consequences through the model with actual values from historical data resulted in RMSE equal to 0.001 and correlation coefficient r equal to 0.58 which is acceptable for this type of model.

[3] $$\mathrm{RMSE} = \frac{\sqrt[2]{\sum_{i=1}^{n}(\mathrm{Estimated\ i} - \mathrm{Actual\ i})^2}}{n}$$

Table 4: Variables, MSE, RMSE and R values of best subsets

Subset	Variables	MSE	RMSE	r
1	Area, D/Th., Coating type, CPM & SMYS	5.62E-06	0.003	0.42
2	Age, Area, D/Th., Coating type, CPM, SMYS	6.30E-06	0.001	0.58
3	Age, Area, Coating Type, SCADA, CPM, CP	4.77E-06	0.013	0.21

The checking dataset is comprised of 20 data points with two failures of external corrosion. Figure 5 compares the actual outputs from the checking dataset with the predicted values for the same data points after recalling the trained network. The checking dataset is embedded into the trained network and the predicted results are calculated. As it is evident from the graph, the model is able to forecast the pipelines with external corrosion failure; nevertheless, with a lower amount. Similarly, it has also recognised the points with lower expectations of this failure type with a little difference. The only problem is at points sixteen and seventeen on which the model has forecasted a negative consequence which is not reasonable. In conclusion, the objective of the model to predict the points with higher effects of monetary consequences from this failure type has been met. However, in the future the model should be enhanced with more data to obtain more reliable results.

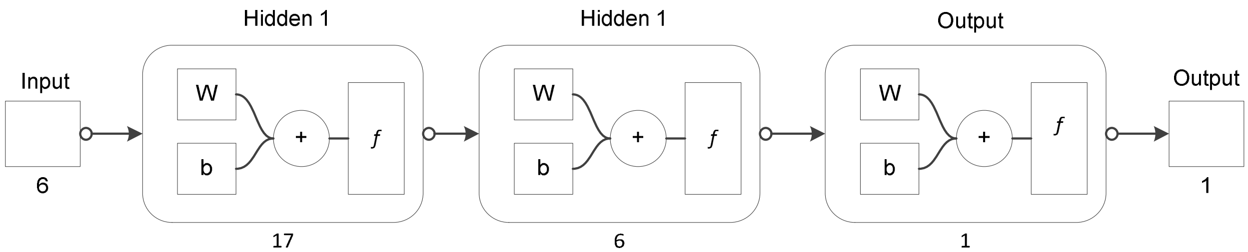


Figure 4: Architecture of selected subset's network

Conclusion

This research proposes an objective failure risk prediction model for oil and gas pipelines based on historical data on incidents of oil and gas pipelines in the USA. First, it classifies the risks in three major groups of physical, external, and operational. Then, it develops a methodology for a predictive model of risk based on a supervised artificial neural network learning machine. Finally, it presents results of the model for one of the risk factors which is external corrosion. The model predicts monetary consequences of the failure of pipelines. It applies a set of nine variables through neural network training and tests the network results with possible combinations of variables and changing architecture of the network and records values of errors for each network. Finally, the selection of best subset of variables is done based on a few factors: MSE, RMSE, and r values. It proves that a subset of age, area and pipeline diameter divided by wall thickness as well as coating type, efficiency of CPM monitoring system, and SMYS produces a more efficient model. Results of validation prove efficiency of the model and its accuracy. Moreover, the model is able to forecast the pipelines with the potential of external corrosion failure and recognises pipelines with an acceptable reliability which do not have the potential for this type of failure. Models for the other types of risk factors will be developed in the future to make a comprehensive risk assessment model to predict the most effective failure type on each pipeline. Also, the authors will try to enhance the model's performance to predict more reliable results.

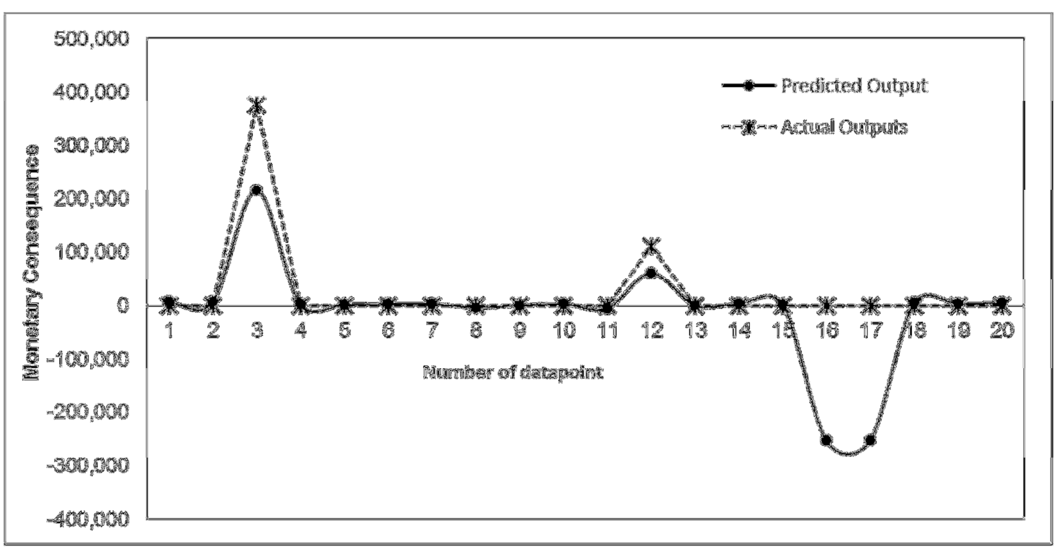


Figure 5: Actual and predicted outputs of checking dataset

References

Abdrabou, B. 2012. Failure Prediction Model for Oil Pipeline. *Thesis*. Montreal, Quebec, Canada: Concordia University.

Achim, D., Ghotb, F., & McManus, M. J. 2007. Prediction of Water Pipe Asset Life Using Neural Networks. *Journal of Infrastructure Systems*, 26-30.

Al- Barqawi, H., & Zayed, T. 2006. Condition rating Model for Underground Infrastructure Sustainable Water Mains. *Journal of Performance fo Constructed Facilities*, 126-135.

Al-Khalil, M., Assaf, S., & Al-Anazi, F. 2005. Risk- Based Maintenance Planning of Cross- Country Pipelines. *Journal of Performance of Constructed Facilities, Vol. 19*, 124-131.

Attalla, M., & Hegazy, T. 2003. Predicting Cost Deviation in Reconstruction Projects: Artificial Neural Networks versus Regression. *Journal of Construction Engineering and Management*, 405-411.

Christodoulou, S. 2004. Optimum Bid Markup Calculation Using Neurofuzzy Systems and Multidimensional Risk Analysis Algorithm. *Journal of Computing in Civil Engineering*, 322-330.

Davis, p., Dubois, J., Gambardella, F., Sanchez-Garcia, E., & Uhlig, F. 2011. *Performance of European cross country oil pipelines.* Brussels: CONCAWE.

Dey, P., Ogunlana, S. O., & Naksuksakul, S. 2004. Risk Based Maintenance Model for Offshore Oil and Gas Pipelines: A case Study. *Journal of Quality in Maintenance Engineering*, 169-183.

DOT1. 2012. *Pipeline & Hazardous Materials Safety Administration*. Retrieved January 10, 2013, from US. Department of Transportation: http://primis.phmsa.dot.gov/comm/reports/safety/sida.html?nocache=4372

DOT2. 2012. *Pipeline & Hazardous Materials Safety Administration*. Retrieved January 10, 2013, from US. Department of Transportation: http://primis.phmsa.dot.gov/comm/PipelineLibrary.htm?nocache=2941

DOT3. 2012. *Pipeline & Hazardous Materials Safety Administration*. Retrieved January 10, 2013, from US. Department of Transportation: http://primis.phmsa.dot.gov/comm/glossary/index.htm?nocache=2760#SCADA

DOT4. 2012. *Pipeline & Hazardous Materials Safety Administration*. Retrieved January 10, 2013, from US. Department of Transportation: http://primis.phmsa.dot.gov/comm/glossary/index.htm?nocache=2760#ComputationalPipelineMonitoring

Hegazy, T. M. 1993. Integrated Bid Preparation with Emphases on Risk Assessment Using Neural Networks. *Thesis*. Montreal, Quebec, Canada: Concordia University, Department of Building, Civil, and Environmental Engineering.

Hopkins, H., P., & Andrew. 2006. *Updated pipeline repair manual.* Arlington: Pipeline research council international.

Khaw, J. F., Lim, B., & Lim, L. E. 1995. Optimal design of neural networks using the Taguchi method. *Neurocomputing*, 225-245.

Maged, E. G., Chang, L.-M., & Zhang, L. 2004. Prediction of Engineering Performance: A Neurofuzzy Approach. *Journal of Construction Engineering and Management*, 548-557.

Moselhi, O., Hegazy, T., & Fazio, P. 1991. Neural Networks as Tools in Construction. *Journal of Construction Engineering and Management*, 606-625.

Muhlbauer, K. W. 2004. *Pipeline risk management manual.* Burlington: Elsevier.

Siqueira, I. 1999. Neural Network- Based Cost Estimation. *Thesis*. Montreal, Quebec, Canada: Concordia University.

Zayed, T. M., & Halpin, D. W. 2005. Pile Construction Productivity Assessment. *Journal of Construction Engineering and Management*, 705-714.

Zeng, C., & Ma, B. 2009. Analytic Mode for Predicting the Failure Risk of Pipeline System. *ICPTT* (pp. 17-27). Shanghai: American Society of Civil Engineers.

Zhang, G., Patuwo, B. E., & Hu, M. Y. 1998. Forecasting with artificial neural networks: The state of the art. *International journal of forecasting*, 35-62.