ADVANCES IN MATERIALS SCIENCE, Vol. 24, No. 4 (82), December 2024

DOI: 10.2478/adms-2024-0020

Ashref Ahmaid[1], Fuad Khoshnaw[2*]

[1] *Sirte Oil company, Brega, Libya*
[2] *School of Engineering and Sustainable Development, De Montfort University, UK*

** fuad.khoshnaw@dmu.ac.uk*

CORROSION RATE PREDICTION FOR UNDERGROUND GAS PIPELINES USING A LEVENBERG-MARQUARDT ARTIFICIAL NEURAL NETWORK (ANN)

ABSTRACT

This study addresses the challenge of accurately predicting corrosion rates and estimating the remaining life of underground gas pipelines, which is complicated by the complex interaction of physical factors and environmental conditions. Traditional models are inadequate in capturing these variables, leading to less reliable predictions, which this study aims to address by developing a more accurate and optimized artificial neural network (ANN) model. This study focuses on predicting corrosion rates and estimating the remaining life of underground gas pipelines using ANNs implemented in MATLAB. It incorporates both physical factors, such as maximum corrosion depth and pipe thickness, and environmental variables such as moisture, soil resistivity, and chloride concentration. The analysis identified corrosion depth and wall thickness as significant contributors, influencing material integrity by 20% and 16%, respectively. The optimal ANN model, with a Levenberg-Marquardt structure and one hidden layer of 10 neurons, achieved superior accuracy, with an MSE of 0.038 and R^2 of 0.9998. The study addresses the challenge of accurately predicting corrosion rates and remaining life in underground gas pipelines by developing an optimised ANN model. Its contribution lies in creating a highly accurate prediction tool that outperforms traditional models and enables more informed decisions for pipeline maintenance and safety.

Keywords*: pipeline integrity; corrosion, ANN; remaining life estimation*

INTRODUCTION

Pipelines are essential for safe, swift petrochemical transport, supporting economic and environmental needs by connecting extraction sites to industrial hubs and fostering growth [1]. Their underground deployment negates high-emission above-ground transport methods, lessening the ecological impact and advancing environmental conservation efforts [2]. Underground gas pipelines prioritize safety, shielded from weather and human interference. Regular monitoring detects anomalies early, enabling swift intervention. This proactive approach enhances safety standards, fostering community confidence in pipeline integrity and environmental protection measures [3]. External corrosion, particularly affecting steel underground pipelines, arises from

Received: 2024.09.02; Received in revised form: 2024.10.16; Accepted: 2024.10.16

chemical interactions with soil, posing risks to safety and integrity. Environmental damage, reputational risks, and inspection complexities further compound these issues, necessitating costly restoration efforts and risking public trust in operational integrity [4].

Corrosion rate, often measured in millimetres or inches per year, indicates how quickly metal degrades due to chemical reactions with its surroundings. Factors such as material properties, chemical compositions and environmental elements such as temperature and presence of corrosive substances, influence this rate significantly. Remaining life estimates the safe operational duration of metallic structures, considering factors like corrosion rate, material properties, and acceptable metal loss, ensuring continued safe operation before repair or replacement is needed [5, 6].

To ensure the integrity of pipelines and minimize the risk of failure, it is essential to understand the factors contributing to corrosion and implement effective strategies for its prevention and control. The primary objective was to maximise the predictive power of Artificial Neural Networks (ANN) for corrosion rate forecasting. By developing a dependable and accurate ANN-based method, the objective is to improve the management and maintenance of these vital assets, thereby reducing the risk of catastrophic failures.

Abubakirov et al [7] indicated that soil moisture content is affected by factors such as rainfall, soil type, temperature, humidity, soluble salts, aeration, and drainage characteristics. They categorized moisture content into four levels: None (0%), Low (0-1%), Medium (1-5%), High (5-20%), and Extremely High (above 20%). While Kim et al [8] demonstrated that soil resistivity - which is influenced by salinity, water content, and texture, plays a crucial role in assessing soil corrosivity - categorized soil corrosivity into several ratings based on resistivity: essentially non-corrosive (>20,000 ohm.cm), mildly corrosive (10,000 to 20,000 ohm.cm), moderately corrosive (5,000 to 10,000 ohm.cm), corrosive (3,000 to 5,000 ohm.cm), highly corrosive (1,000 to 3,000 ohm.cm), and extremely corrosive (<1,000 ohm.cm). Moreover, they found that chloride concentration in soil, influenced by factors such as geographical location, brackish groundwater, offshore proximity, mine drilling, and weather conditions, helps categorize soil corrosivity. They outlined corrosivity categories based on chloride concentrations: severe (greater than 5000 ppm), considerable (1500 – 5000 ppm), moderate (500 – 1500 ppm), and low (less than 500 ppm).

Pipeline operators combat external corrosion using coatings, cathodic protection, and inhibitors. Coatings serve as barriers, while cathodic protection and inhibitors provide additional protection. Regular inspections are crucial for identifying weak spots, employing advanced techniques such as Magnetic Flux Leakage (MFL) and ultrasound. However, these methods are both costly and time-consuming. Current failure prediction models often depend on subjective expert opinions or are too narrowly focused. There is an urgent need for a detailed, data-driven model based on past incidents to guide preventative measures and avert disasters [9].

Utilizing ANN models for predicting corrosion rates in subterranean gas pipelines offers significant technological and economic benefits to oil firms. These models surpass traditional methods in predictive accuracy, enhancing pipeline integrity evaluations and safety protocols. Their adaptability to new data allows them to excel in dynamic environments, and seamless integration with other inspection instruments provides a comprehensive approach to corrosion detection. Economically, accurate predictions reduce costs by minimizing manual inspections and enabling proactive maintenance, thereby avoiding costly repairs or replacements. Targeted maintenance based on corrosion rate predictions prolongs pipeline lifespans, ensuring consistent production and supply while optimizing return on investment [10].

Biezma et al. [11] introduced a fuzzy logic method for forecasting pipeline external corrosion rate, vital for oil and gas pipeline safety. Their approach involves field data acquisition, integrating it with a Fuzzy Logic Expert System (FLES) to map corrosion rates along the pipe, enhancing inspections and maintenance. Compared to previous methods, their tool proves more effective, offering cost-effective corrosion prediction for pipeline safety. Shaik et al. [12] analyzed an oil pipeline dataset from Sudan to predict metal loss defects using Support Vector Machines (SVM). The quadratic SVM model exhibited high predictive accuracy (93.0%) evaluated via a confusion matrix. With a fault detection rate of 90.4% and a misclassification rate of 9.6%, SVM proves effective for proactive pipeline maintenance, enhancing lifespan and reducing degradation rates. Shaik et al. [13] introduced an ANN-based method to forecast pipeline lifespan and classify metal loss issues in Sudan's oil and gas industry. The model achieves high accuracy, aiding efficient maintenance scheduling and preventing degradation and financial losses. It provides insights into the remaining useful life and type of metal loss, facilitating cost-effective planning. Further case studies could enhance and refine this method for broader industry applications.

ANNs replicate the brain's structure, comprising input, hidden, and output layers of interconnected artificial neurons. Data is received by the input layer, processed through hidden layers, and outputted. Neurons adjust interconnections' weights during learning to optimize performance, enabling ANNs to excel in pattern identification for applications like image recognition and forecasting. Figure 1a illustrates the structure of an artificial neuron, which computes inputs (x1, x2,..., xm) multiplied by corresponding weights (w1, w2,..., wm), adding a bias term (b). The combined value passes through an activation function (f) to produce the output (y), introducing nonlinearity into the model [14].

Equation 1 used for calculating the output of an artificial neuron is:

$$y = f(w1 * x1 + w2 * x2 + \cdots wm * xm + b) \quad (1)$$

where: y is the output of the artificial neuron
f is the activation function
w1, w2, ..., wn are the weights corresponding to the inputs x1, x2, ..., xm
x1, x2, ..., xm are the input values
b is the bias value

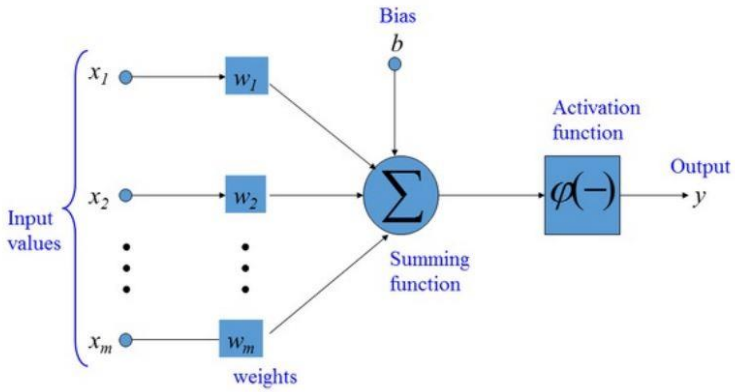
Existing literature on pipeline corrosion prediction lacks comprehensive investigation into methodology, scope, and outputs. Studies often focus on specific corrosion severities, neglecting interactions and influential operational factors, which reduces model accuracy. Outputs typically include only corrosion rate or remaining life, overlooking indicators like failure probability or damage extent.

Despite these advancements, there remains a need for more accurate and efficient models tailored specifically to underground gas pipelines. Existing models often lack the precision required for real-world applications, and many fail to integrate the full range of physical and environmental variables influencing corrosion. This study aims to address this gap by developing an optimized ANN model, using a Levenberg-Marquardt algorithm, for accurately predicting corrosion rates and remaining life in underground gas pipelines. The novelty of the research lies in the model's ability to achieve superior prediction accuracy, with an MSE of 0.038 and an R^2 of 0.9998, outperforming traditional approaches and offering a more reliable tool for pipeline management.

DATA AND METHODOLOGY

This study focuses on a quantitative analysis of the 34" underground gas pipeline managed by Sirte Oil Company in Libya, which supplies gas to vital establishments for over 25 years. Employing an intelligent pig device in 2009, data on corrosion locations and dimensions was gathered, augmented by soil analyses including resistivity, moisture, and chloride concentration. Correlating these with inspection data provides a comprehensive pipeline status understanding, enabling precise maintenance strategies and proactive corrosion point identification. Historical data usage is justified for cost-effectiveness, availability, reliability, and incorporating soil analyses, enhancing analysis comprehensiveness and maintenance strategy precision. Increased data samples aid ANN training by improving generalization, accuracy, robustness, complexity handling, and bias reduction, contingent upon data quality and representativeness assurance.

(a)



(b)

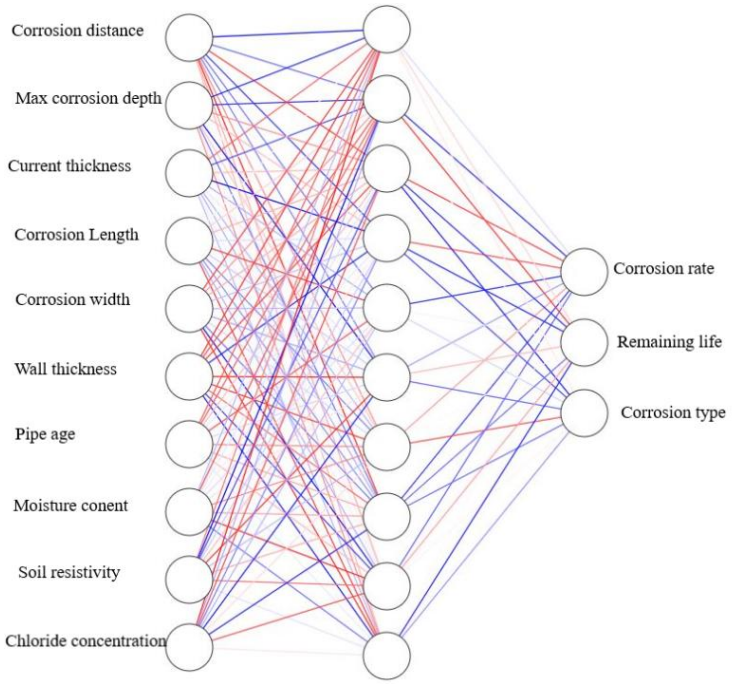


Fig. 1. The ANN structure (a) and architecture (b)

The methodological process for developing and testing a model started with algorithm selection and consideration of system factors. Data preparation, model training, and validation follow, with iterative adjustments until meeting validation criteria, ensuring a tailored and accurate model for the piping system [15, 16].

Input and output parameters

The ANN model incorporates inputs such as corrosion distance, depth, current, age, moisture, resistivity, and chloride, predicting corrosion rate, remaining life, and corrosion type. This aids proactive maintenance, optimizes repair schedules, and prevents failures. Figure 1b illustrates the input and output parameters in the ANN.

Corrosion distance measures the location from a reference point, while maximum corrosion depth indicates severity. The current thickness compares to the original wall thickness. Corrosion length and width represent dimensions, and pipe age affects susceptibility. Moisture content in soil, resistivity, and chloride concentration influence corrosion. Corrosion rate quantifies material loss over time while remaining life estimates operational lifespan considering current corrosion. In intelligent pig surveys, these parameters assist in identifying various types of corrosion, such as pinhole, pitting, and general corrosion, as well as in detecting circumferential and axial grooving. These manifest as small holes, pits, uniform surface erosion, or grooves running along or encircling the pipe's length, respectively [17].

Data Calculations

Two essential parameters derived from the inputs of the ANN model that need to be calculated are the corrosion rate and the remaining life of the pipelines. To calculate the corrosion rate, following API 570, the Point-to-Point analysis method is applied using Equation (2):

$$corrosion\ rate = \frac{initial\ thickness - minimum\ thickness}{years\ in\ service} \tag{2}$$

To estimate remaining life, equation (3) integrates initial thickness readings, corrosion rates, and guidance from a corrosion specialist, facilitating accurate predictions for pipeline longevity.

$$remaining\ life = corrosion\ rate * \frac{t\ actual - t\ required}{corrosion\ rate} \tag{3}$$

where “t_actual” represents the actual thickness measured during the inspection, and “t_required” denotes the essential thickness before adding the corrosion allowance and manufacturer's tolerance.

The 't_required' is calculated using the design formulas [18]. The required thickness “t_required” is determined using the equation (4):

$$\text{t\_required} = \frac{PD}{2S} \tag{4}$$

Where P represents the internal design pressure of gas pipelines, D is the external diameter of the pipe and S is the specified minimum yield strength.

Sirte Oil Company uses specific measurements and standards for its gas pipelines. The internal design pressure, denoted by 'P', is set at 600 psi. The external diameter of the pipe, represented as 'D', is 34 inches. The term 'S' signifies the specified minimum yield strength of the pipe material. For the material examined in this study, 'S' is measured to be 52,000 psi.

$$t_{required} = \frac{600*34}{2*52000} = 0.1962\ inch$$

This calculation, along with the corrosion rate, allows the estimation of the remaining life of a piping system based on its present state [19].

Data processing

The data processing for corrosion rate prediction involves cleaning and transforming the dataset. This includes removing duplicates, managing missing values, and normalization to rescale features within a specified range, such as [0, 1] or [-1, 1] [21, 22]. In MATLAB, 'mapminmax' from the Neural Network Toolbox normalizes data, typically within [-1,1] or [0,1], enhancing corrosion rate predictions for pipelines [20].

Table 1 shows encoding transforms input data into numerical formats, standardizing and capturing features. Decoding translates outputs into human-readable form, crucial for tasks like image recognition and natural language processing, bridging communication between network processing and real-world interpretation [21]

Table 1. *Encoding of corrosion types*

Corrosion Type	Encoding
Pinhole corrosion	1
Pitting corrosion	2
General corrosion	3
Circumferential grooving	4
Axial grooving	5
Circumferential slotting	6
Axial slotting	7

Corrosion types in gas pipelines are classified based on severity and metal loss characteristics. Key types include pinhole corrosion, which causes localized leaks; general corrosion, where material loss is uniform across a large surface; circumferential grooving, which weakens the pipe around its circumference; and axial grooving, where grooves run along the pipe's length. Classification is based on factors like pipe thickness (t), corrosion width (w), and length (l), helping assess the impact on pipeline integrity. Techniques like NDT and ILI aid in evaluating and managing these corrosion types to ensure safety.

Classification of corrosion types

Metal loss in gas pipelines is classified based on degradation severity, crucial for preventing failures [18]. Inline inspection (ILI) and nondestructive testing (NDT) methods like intelligent pigging (MFL testing) are commonly employed for assessment [22].

Distinct types of metal loss include pinhole corrosion, which compromises integrity and risks leaks; general corrosion, which erodes surfaces uniformly; circumferential grooving and pitting, which weaken structures; and axial grooving, which compromises integrity due to flow patterns or material variations. These anomalies can be understood by examining the distribution of their surface dimensions, such as bar graphs showing each type's surface dimensions. This classification is vital for comprehending the various metal loss anomalies and their impact on pipeline integrity [23-27].

The parameter (A), used to classify metal loss anomalies in pipeline inspections, is determined by the pipe's thickness, denoted as (t). These classifications stem from Non-Destructive Testing (NDT) methods, which evaluate the pipe's condition without inflicting any damage. If the thickness (t) is less than 10 mm, (A) is assigned a minimum value of 10 mm. This adjustment reflects the idea that a thinner pipe may not accurately capture certain characteristics. Conversely, if the thickness (t) is 10 mm or greater, the value of (A) matches that of (t). Within this framework, (w) signifies the width of the corrosion, while 'l' represents its length. These dimensions are crucial for comprehending and quantifying the pipeline's corrosion extent [28].

In the data processing phase, thorough preprocessing ensures that metal loss anomalies are accurately captured and classified. This involves not only cleaning and normalizing the dataset but also transforming key features such as pipe thickness (t), width (w), and length (l) of corrosion to better reflect the severity of degradation. By incorporating these dimensions and applying appropriate adjustments, such as setting a minimum value for thinner pipes, the model can more effectively analyze and predict the impact of various corrosion types on pipeline integrity, enhancing prediction accuracy.

ANN model development

The study employed a feedforward neural network with 10 inputs, 3 outputs, and 1-3 hidden layers of 5, 10, or 15 neurons, selecting architecture based on problem complexity, dataset size, and computational resources. Activation functions such as sigmoid, tanh, and ReLU introduced non-linearity for complex pattern learning [29]. Hyperparameter tuning, including adjusting learning rate, neuron per layer selection, and epochs, enhanced model performance, with data segmented into 70:15:15 or 60:20:20 ratios for training, validation, and testing. The training involved data collection, preprocessing, architecture definition, weight and bias initialization, forward and backpropagation, and loss minimization through optimization [30]. Validation and testing ensured model generalizability using performance metrics like MSE and R^2 on unseen data subsets [18].

Contribution Factor and P-value

The contribution factor is a measure of the impact of input factors on a model's output, identifying the most significant ones for further analysis. It's used to predict pipeline failures, and its significance is measured by the R-squared value in regression analysis. A p-value assesses significance in hypothesis testing. Small values (<0.05) imply the null hypothesis may be false, prompting consideration of alternative hypotheses [9].

RESULTS

The study provides a detailed analysis of statistical results, contribution factors, P-Value, and comparisons of ANN model experiments, focusing on optimal experiments and predicted results. Table 2 presents a statistical analysis of variables used in the ANN model, revealing corrosion distances, material thickness, wall thickness, pipe age, moisture content, soil resistivity, and chloride concentration. The corrosion rate averages 0.11 mm/year, and the remaining life spans range from 0.01 to 59.14 years. The data could help predict failures and plan infrastructure maintenance.

Table 2. *Corrosion variables and statistical analysis*

Variables	Corrosion distance (km)	Max corrosion depth (mm)	Current thickness (mm)	Corrosion length (mm)	Corrosion width (mm)	Wall thickness (mm)	Pipe age (year)	Moisture content (%)	Soil resistivity (ohm-cm)	Chloride concentration (ppm)	Corrosion rate (mm/year)	Remaining life (year)	Corrosion type
Min	5.47	1.75	4.98	9	14	8.74	29	3.18	538	37.43	0.06	0.01	2
Max	142.5	6.99	10.16	328	414	12.7	29	30.17	8469.47	271	0.24	59.14	6
Mean	40.03	3.1505	6.99	41.62	57.54	10.15	29	22.14	2707.32	199.3	0.11	20.39	3
Standard deviation	36.2	1.3	1.12	71	84	1.9	1	7.33	1917.37	60.77	0.04	10.52	5

Table 3 shows statistically significant correlations between variables and corrosion rate, with max corrosion depth showing a near-perfect positive correlation of 0.9963 with the rate, indicating robustness and high significance. Thickness has a weaker but still significant correlation (0.2334) with corrosion rate. Corrosion width (0.574) and wall thickness (0.8178) show strong correlations with p-values of 0, emphasizing their significance. Soil resistivity (-0.495) has a moderate negative correlation.

Table 3. *Correlation coefficient and P-values*

Variable	Correlation Coefficient	P-value
Corrosion distance	0.4964	$7.707*10^{-14}$
Max corrosion depth	0.9963	0
Current thickness	0.2334	$0.00087*10^{-14}$
Corrosion length	0.3809	$2.623*10^{-8}$
Corrosion width	0.574	0
Wall thickness	0.8178	0
Moisture content	0.509	$1.421*10^{-14}$
Soil resistivity	-0.495	$9.311*10^{-14}$
Chloride concentration	0.5065	$1.976*10^{-14}$

Figure 2 highlights the multifaceted factors influencing corrosion. The depth of corrosion depth (20%) and wall thickness (16%) are crucial factors in susceptibility. Other contributors include corrosion width (11%), distance, moisture, soil resistivity, and chloride concentration (each 10%). Understanding corrosion requires accounting for all factors' cumulative impact on material integrity.

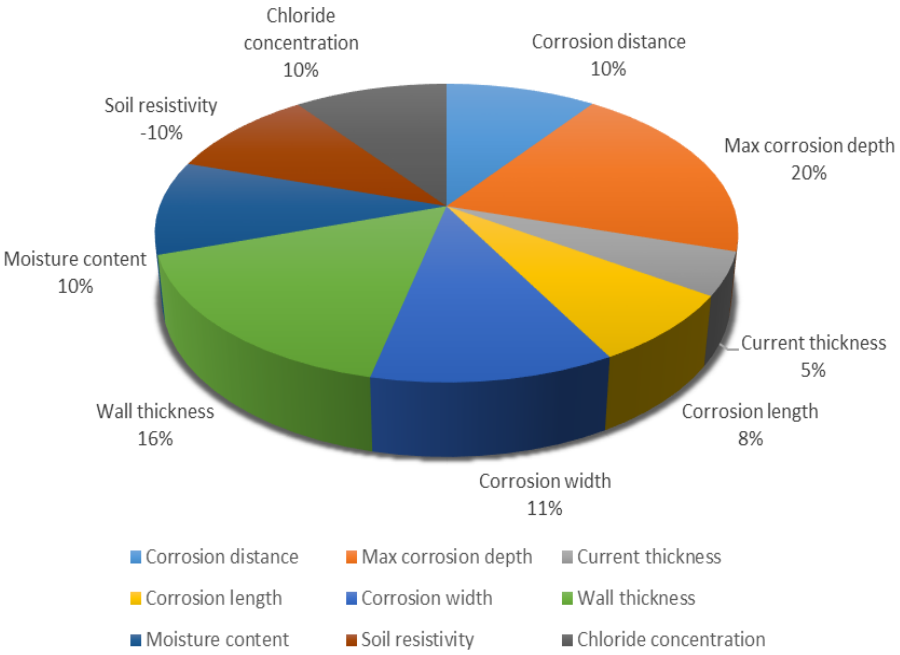


Fig. 2. Contribution factor

Comparative Analysis of ANN Model Experiments

Table 4 compares training algorithms based on MSE and R^2 values. Levenberg-Marquardt excels, especially with a single hidden layer. Experiment 3 achieves near-perfect accuracy (MSE: 0.03890, R^2: 0.9999) with 10 neurons. However, its efficiency decreases with more layers. Experiment 26 shows the gradient descent's inefficiency (MSE: 13.243). Data split of 70:15:15 yields better performance than 60:20:20. Accuracy improves with more training data. A layer with 10 neurons may be optimal, but increasing complexity doesn't always enhance performance. No direct correlation exists between epochs and results, necessitating individualised adjustments. Epoch selection must align with model needs and problem nature for optimal performance. Levenberg-Marquardt, especially with one hidden layer, shows promise in MSE and R^2 values, but optimal configuration depends on the use case and complexity-performance balance. Both metrics should guide algorithm selection and evaluation.

Experiment 3 in the Table 4 was identified as the optimal model after thoroughly investigating several ANN models. This section explores further the reasons for and strategies used by this exceptional performance of the model.

Figure 3 illustrates a feedforward ANN's structure with input-to-output data flow. It employs a hyperbolic tangent sigmoid transfer function in the hidden layer and a linear transfer function in the output layer.

Figure 4 in MATLAB displays the training state of an ANN, featuring gradient, learning rate, and validation performance plots. The x-axis denotes epochs, while the y-axis offers insights into gradient changes, adaptive learning rates, and validation set performance, aiding in real-time evaluations and algorithm adjustments.

Table 4. *Training experiments of the ANN model*

Exp	Training algorithm	Data division %	Hidden layers	Neron of hidden layers	Epochs	MSE	R^2
1	Levenberg-Marquardt	70:15:15	1	5	12	0.0659	0.9997
2	Levenberg-Marquardt	60:20:20	1	5	13	0.0994	0.9996
3	Levenberg-Marquardt	70:15:15	1	10	86	0.03890	0.9999
4	Levenberg-Marquardt	60:20:20	1	10	32	0.0866	0.9997
5	Levenberg-Marquardt	70:15:15	1	15	10	0.0938	0.9996
6	Levenberg-Marquardt	60:20:20	1	15	12	0.1708	0.9993
7	Levenberg-Marquardt	70:15:15	2	5-5	30	0.3837	0.9975
8	Levenberg-Marquardt	60:20:20	2	5-5	88	0.4599	0.9980
9	Levenberg-Marquardt	70:15:15	3	5-5-5	52	0.3118	0.9986
10	Levenberg-Marquardt	60:20:20	3	5-5-5	16	1.6388	0.9923
11	Scaled conjugate gradient	70:15:15	1	5	46	0.8655	0.9963
12	Scaled conjugate gradient	60:20:20	1	5	31	1.9582	0.9925
13	Scaled conjugate gradient	70:15:15	1	10	72	0.6030	0.9969
14	Scaled conjugate gradient	60:20:20	1	10	50	1.7552	0.9927
15	Scaled conjugate gradient	70:15:15	1	15	30	2.6087	0.9908
16	Scaled conjugate gradient	60:20:20	1	15	13	3.0167	0.9848
17	Scaled conjugate gradient	70:15:15	2	5-5	38	0.9734	0.9962
18	Scaled conjugate gradient	60:20:20	2	5-5	65	1.1118	0.9947
19	Scaled conjugate gradient	70:15:15	3	5-5-5	34	1.8842	0.9921
20	Scaled conjugate gradient	60:20:20	3	5-5-5	44	2.0798	0.9899
21	Quasi-Newton methods	70:15:15	1	10	52	0.1470	0.9994
22	Quasi-Newton methods	70:15:15	2	5-5	29	0.5150	0.9975
23	Quasi-Newton methods	70:15:15	3	5-5-5	46	1.9367	0.9927
24	Gradient descent-learning rate	70:15:15	1	10	79	3.3097	0.9884
25	Gradient descent-learning rate	70:15:15	2	5-5	80	2.0571	0.8475
26	Gradient descent-learning rate	70:15:15	3	5-5-5	102	13.243	0.9407
27	One step scant	70:15:15	1	10	28	1.3944	0.9934
28	One step scant	70:15:15	2	5-5	106	0.4795	0.9981
29	One step scant	70:15:15	3	5-5-5	98	0.4445	0.9984
30	Resillient propagation	70:15:15	1	10	68	1.0868	0.9968
31	Resillient propagation	70:15:15	2	5-5	60	1.0936	0.9952
32	Resillient propagation	70:15:15	3	5-5-5	34	1.7083	0.9931

At epoch 92, the gradient magnitude at 0.16419 suggests ongoing learning with room for error reduction. A low learning rate of 0.0001 promotes stability but may delay convergence. A validation check value of 6 indicates potential nearing of optimal state or overfitting, guiding training decisions. Figure 5 shows that epoch 86 records the highest validation performance at 0.03891, indicating successful handling of unseen data. The stability or decline in the performance of subsequent epochs may suggest overfitting or that the training is nearing completion.

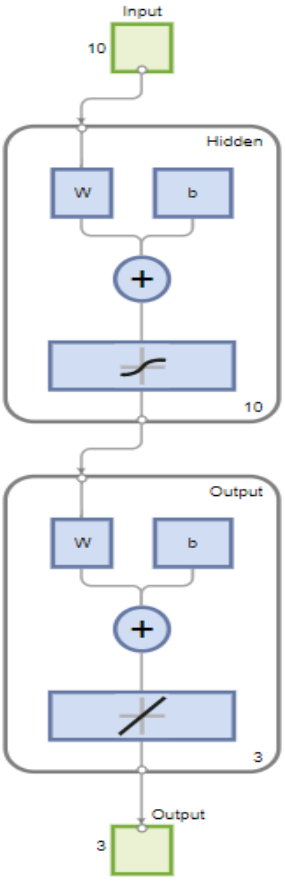


Fig. 3. Architecture of optimal model

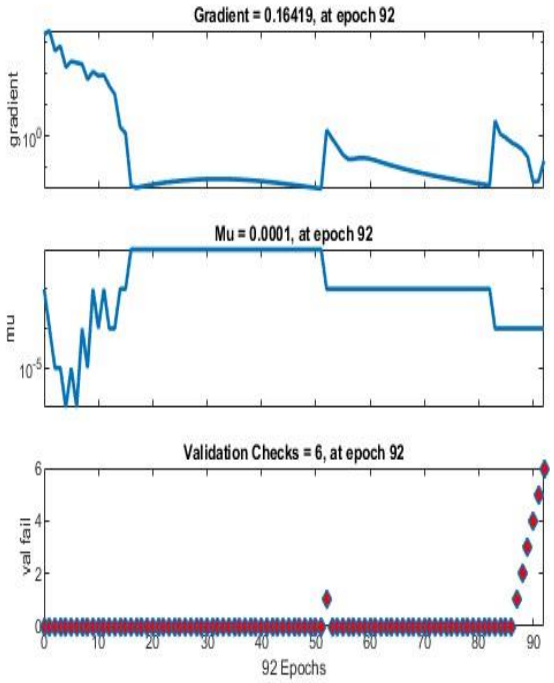


Fig. 4. Training state plot

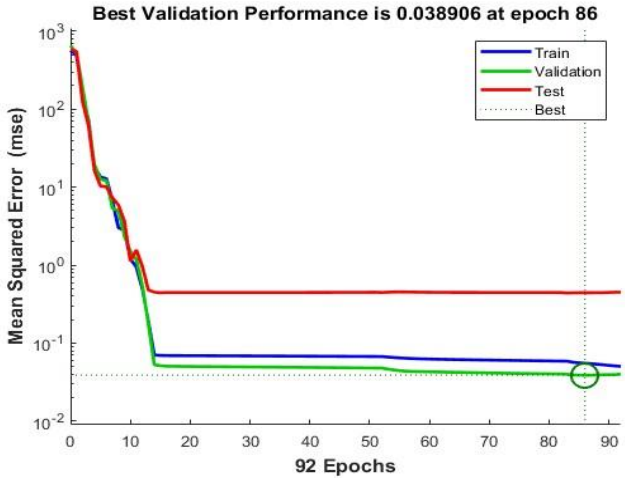


Fig. 5. Performance plot

Figure 6 illustrates the regression performance of the ANN by comparing the actual and predicted corrosion rates. The R^2 values are exceptionally high across the board: Training (R^2 = 0.99977) shows nearly perfect learning, Validation (R^2 = 0.99983) demonstrates excellent

predictions on unseen data, Testing (R^2 = 0.99804), though slightly lower, still confirms strong generalization to new data.

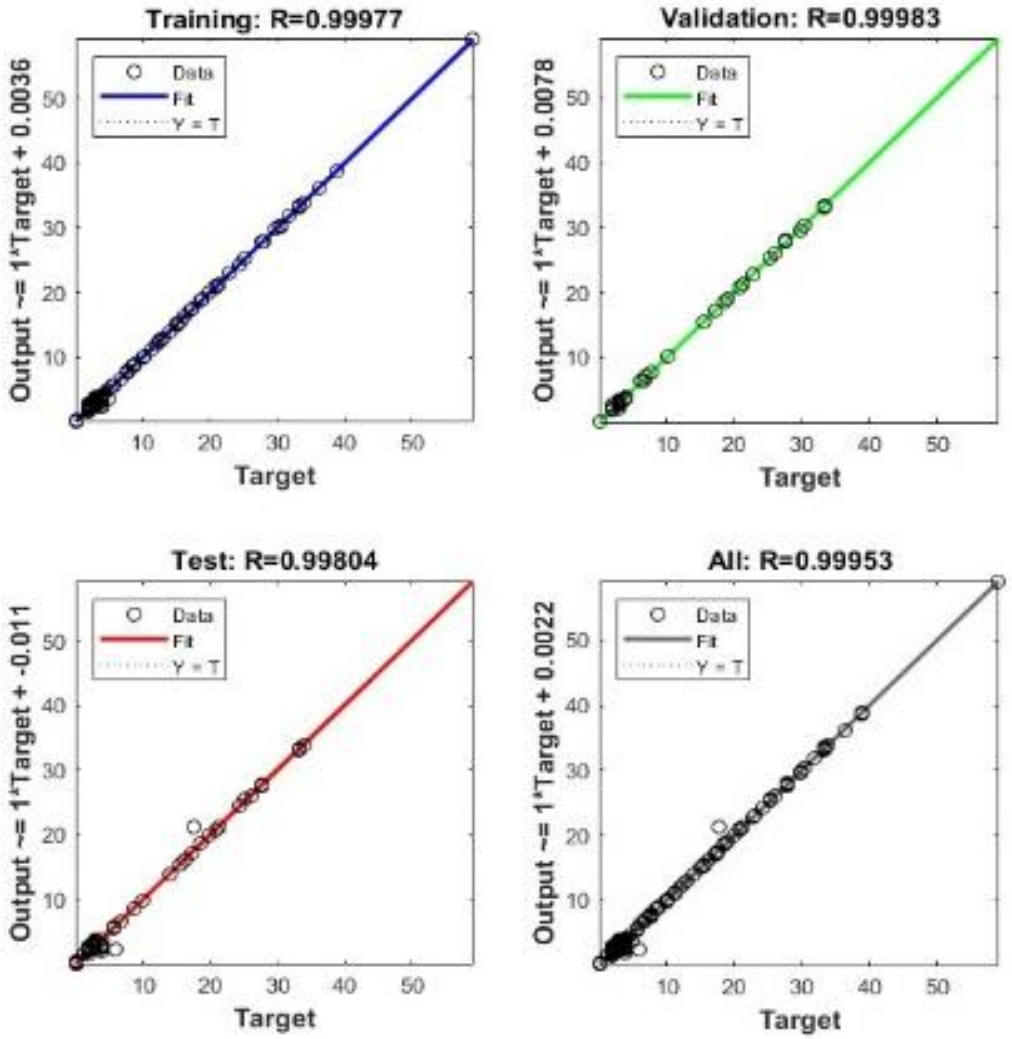


Fig. 6. ANN model regression plots

Table 5 compares actual and predicted corrosion rate, remaining life, and corrosion type for 24 unseen samples, with error percentages ranging from 0% to 25%. Most samples exhibit errors under 10%, indicating adequate accuracy. Remaining life errors are generally low (0.02% to 1.70%), with only a few exceeding 1%. Corrosion-type prediction errors are minimal, though some discrepancies occur in specific samples. The forecasting model generally performs well, especially in predicting remaining life, but occasional significant errors in corrosion rate and type require investigation for improvement.

Table 5. *Comparison between the actual values and predicted values of corrosion rate*

No	Actual values			Predicted values			Error percentage %		
	Corrosion rate	Remaining life	Corrosion type	Corrosion rate	Remaining life	Corrosion type	Corrosion rate	Remaining life	Corrosion type
1	0.1	51.13	3	0.1	51.32	3	0	0.37	0
2	0.13	29.76	3	0.14	29.64	3	7.69	0.4	0
3	0.12	36.29	3	0.13	36.25	3	8.33	0.11	0
4	0.11	38.8	2	0.11	38.78	2	0	0.05	0
5	0.21	6.98	3	0.19	7.07	3	9.52	1.29	0
6	0.13	29.76	3	0.14	29.74	3	7.69	0.07	0
7	0.11	41.51	2	0.1	41.35	3	9.09	0.39	0.5
8	0.2	10.17	2	0.17	10.16	3	15	0.09	0.5
9	0.11	44.45	3	0.11	44.39	3	0	0.13	0
10	0.12	36.29	2	0.11	36.22	2	8.33	0.19	0
11	0.12	33.96	3	0.13	33.84	3	8.33	0.35	0
12	0.11	38.8	2	0.11	38.78	2	0	0.05	0
13	0.12	33.96	3	0.12	33.85	3	0	0.32	0
14	0.13	31.79	3	0.14	31.8	3	7.69	0.03	0
15	0.12	36.29	3	0.13	36.16	3	8.33	0.36	0
16	0.11	38.8	3	0.12	38.77	3	9.09	0.08	0
17	0.09	59.14	4	0.09	59.13	4	0	0.02	0
18	0.14	24.42	3	0.15	24.4	3	7.14	0.08	0
19	0.1	9.99	3	0.11	9.82	3	10	1.7	0
20	0.08	15.56	2	0.1	15.44	2	25	0.77	0
21	0.06	33.38	2	0.07	32.94	3	16.67	1.32	0.5
22	0.08	18.98	2	0.09	19.02	2	12.5	0.21	0
23	0.1	9.99	3	0.12	9.83	3	20	1.6	0
24	0.07	25.24	2	0.08	25.48	2	14.29	0.95	0

Figures 7a, 7b and 7c compare actual and predicted corrosion rates, remaining life, and corrosion types across 24 samples. Close alignment between actual and predicted values is observed for corrosion rates and remaining life, with minor discrepancies. Although some deviations occur in corrosion-type predictions, e.g. samples 7, 8 and 21, the model's reliability remains high, (see Figure 7c).

The ANN model was developed using a Levenberg-Marquardt algorithm with one hidden layer of 10 neurons, achieving high accuracy with an R^2 of 0.9998 and an MSE of 0.038. However, discrepancies in corrosion-type predictions indicate the need for further improvements by incorporating more input variables or environmental factors. Figure 7b, showing actual vs. predicted remaining life, reflects strong accuracy but could be enhanced by adding residual analysis to identify and address minor prediction errors.

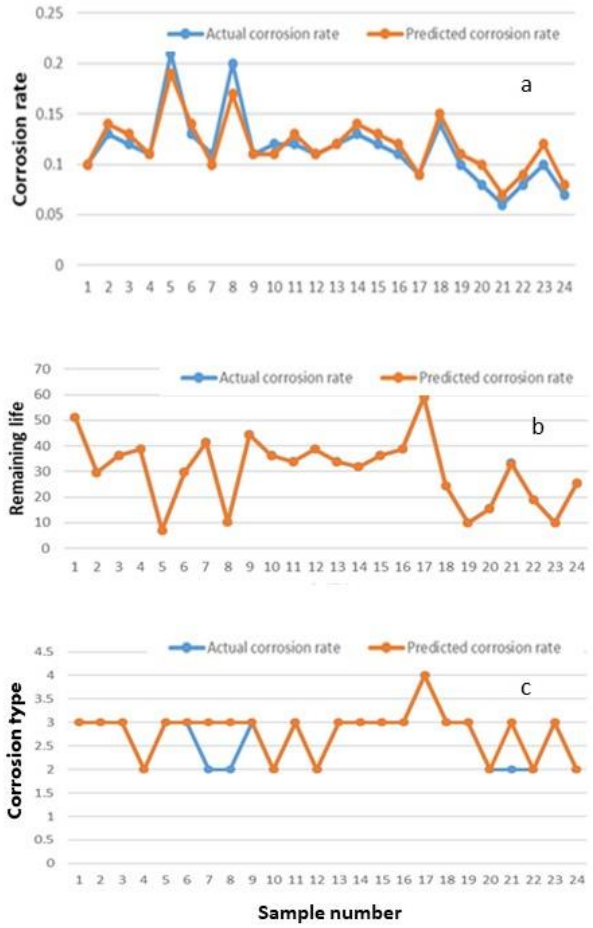


Fig. 7. Actual vs. predicted results for a) corrosion rate, b) remaining life, c) corrosion type

DISCUSSION

The results of this study demonstrate the significant impact of physical factors, such as max corrosion depth and wall thickness, on the progression of pipeline corrosion, with max corrosion depth showing an exceptionally high correlation (0.9963) to corrosion rate. These findings highlight the predominant role of material properties over environmental factors like moisture and soil resistivity, which, while moderately influential, exhibit weaker correlations. This reinforces the necessity of prioritizing physical characteristics in predictive models. The ANN model, particularly the one optimized with the Levenberg-Marquardt algorithm and a single hidden layer of 10 neurons, exhibited superior performance with an R^2 of 0.9999 and an MSE of 0.03890, underscoring its robustness and predictive accuracy. The comparison of training experiments revealed that increasing model complexity, such as adding more hidden layers, did not necessarily improve performance, emphasizing the importance of a well-balanced architecture. Furthermore, the ability of the ANN model to generalize effectively is evident from the high R^2 values across training, validation, and testing phases, with minimal errors observed in remaining life predictions.

Analysis of Table 3 reveals moderately corrosive environmental factors like moisture, soil resistivity, and chloride concentration. Despite falling within corrosive ranges, physical factors like corrosion distance, depth, and wall thickness exert a greater influence on corrosion progression, evident from their substantial variability and higher percentages in corrosion considerations. Wall thickness, for instance, represents 16% of corrosion considerations, surpassing most environmental factors which hover around 10%. Thus, while environmental conditions contribute moderately to corrosion, physical factors predominantly dictate corrosion development. The feedforward ANN model is distinguished by its discrete layers and unidirectional propagation of data, which embodies an ideal structure for achieving optimum performance. The hidden layer of the system uses hyperbolic tangent sigmoid transfer functions, whereas the output layer utilises linear transfer functions. The strategic organisation not only enables enhanced visualisation but also offers possible avenues for optimisation. In general, the design emphasises an architectural decision that is specifically designed to enhance the performance of the ANN model.

ANNs have demonstrated exceptional performance in predicting the corrosion rate of underground gas pipelines, providing crucial insights for oil and gas companies. By accurately forecasting corrosion rates, ANNs enable these companies to optimise maintenance schedules, minimise the risk of pipeline failures, and enhance overall safety. The impressive performance and regression metrics of the ANN model serve as strong evidence of its effectiveness in addressing this critical issue. Optimizing an ANN model for underground gas pipeline corrosion prediction demands a multifaceted approach, evident from diverse training experiments. Hyperparameter tuning, including selecting learning rates and layers, is crucial. Proper data preprocessing, addressing missing values and normalization, ensuring training on structured data, and bolstering accuracy.

The results in Table 5 indicate that the model demonstrates strong performance in predicting corrosion rates, remaining life, and corrosion types. The predicted corrosion rates closely align with the actual values, with error percentages generally under 10%, except for a few cases, such as 25% for sample 20 and 20% for sample 23. This consistency suggests the model's reliability in predicting corrosion rates. Additionally, the predicted remaining life values show minimal deviations from actual values, with most errors being less than 1%. This high level of accuracy is crucial for effective maintenance planning and safety assurance. The model also consistently predicts the correct corrosion type, with a 0% error rate in the majority of cases, indicating its robustness in classification.

Despite some discrepancies in corrosion-type predictions, the overall error rates were low, validating the model’s applicability in real-world scenarios. These results underscore the model’s potential for integration into pipeline maintenance and safety strategies, offering a powerful tool for predicting corrosion rates and enhancing decision-making processes in the oil and gas industry.

Utilizing ANN models for predicting corrosion rates in subterranean gas pipelines offers numerous technological benefits to oil firms. These models boast enhanced predictive accuracy compared to traditional deterministic and probabilistic models, potentially bolstering pipeline integrity evaluations and safety protocols. Additionally, ANN models can adapt and continuously learn from new data, making them adept at navigating dynamic settings. Furthermore, their ability to integrate seamlessly with other inspection instruments, like in-line inspection (ILI) tools, provides a comprehensive approach to corrosion detection.

CONCLUSIONS

1. The contribution factor emphasises that while the depth of corrosion (20%) and wall thickness (16%) are key determinants, the cumulative influence of factors like corrosion area width, moisture content, and soil resistivity crucially shapes material integrity. A holistic view of all these elements is essential for a comprehensive understanding of corrosion dynamics.
2. Through a systematic exploration of various ANN architectures, an optimal feedforward Levenberg-Marquardt configuration was determined that balances model complexity with computational efficiency. This architecture consists of one hidden layer equipped with 10 neurons. The hidden layer employs a hyperbolic tangent sigmoid transfer function, while the output layer utilises a linear transfer function.
3. The developed ANN model demonstrated a high level of accuracy in predicting the corrosion rate of underground gas pipelines. The model's performance was evaluated using various metrics, such as MSE of 0.9998 and R^2 of 0.038. The results indicated that the ANN model outperformed traditional statistical models and other machine-learning techniques in terms of prediction accuracy.
4. The ANN model proved to be a valuable tool for pipeline integrity management, given its ability to accurately predict the corrosion rate of underground gas pipelines across various corrosion types and environmental conditions. This performance not only showcases its good generalisability but also suggests its applicability to a broad spectrum of pipeline systems.
5. By accurately predicting corrosion rates, ANN models can significantly reduce maintenance costs for energy companies, enable targeted repairs, and extend pipeline lifetimes, ensuring consistent production and maximising return on investment.

The study presented promising outcomes but faced limitations in integrating factors like soil pH and microbial influences due to time constraints. Selecting optimal ANN architecture and hyperparameters is challenging and time-consuming, demanding extensive experimentation and validation.

ACKNOWLEDGEMENT

The authors want to acknowledge the funding and support provided for this study. This MSc course scholarship was funded by Sirte Oil Company, owned by the National Oil Corporation in Libya, and De Montfort University in the United Kingdom. Their generous contributions have been instrumental in the successful completion of this research.

REFERENCES

1. N.B. Shaik, W. Benjapolakul, S.R. Pedapati, K. Bingi, N.T. Le, W. Asdornwised, and S. Chaitusaney, "Recurrent neural network-based model for estimating the life condition of a gas pipeline," *Process Safety*, vol. 164, pp. 639-650, 2022.
2. M.E.A.B. Seghier, B. Keshtegar, M. Taleb-Berrouane, R. Abbassi, and N.T. Trung, "Advanced intelligence frameworks for predicting pitting corrosion depth in oil and gas pipelines," *Process Safety*, vol. 147, pp. 818-833, 2021.

3. H. Jiang, Y. Wang, J. Deng, and M. Xu, "A gas pipeline anti-damage early warning monitoring system," *Journal of Physics: Conference Series*, vol. 2033, no. 1, p. 012184, September 2021.
4. M.S. Liew, E.S. Lim, K.L. Na, and N.F. Mohd Sidek, "Parametric Study on the Factors of External Corrosion of Offshore Pipelines in Malaysia," 2012.
5. "API 5L X52 PSL1 LSAW Pipe, 34 Inch, SCH 40," Derbo Steel Pipe Co., Ltd. [Online]. Available: https://www.derbosteelpipe.com/api-5l-x52-psl1-lsaw-pipe-34-inch-sch-40.html. Accessed on: Jul. 2, 2023.
6. Y. Grosu, O. Bondarchuk, and A. Faik, "The effect of humidity, impurities and initial state on the corrosion of carbon and stainless steels in molten HitecXL salt for CSP application," *Solar Energy Materials and Solar Cells*, vol. 174, pp. 34-41, 2018.
7. R. Abubakirov, M. Yang, and N. Khakzad, "A risk-based approach to determination of optimal inspection intervals for buried oil pipelines," *Process Safety*, vol. 134, pp. 95-107, 2020.
8. C. Kim, L. Chen, H. Wang, and H. Castaneda, "Global and local parameters for characterizing and modeling external corrosion in underground coated steel pipelines: A review of critical factors," *Journal of Pipeline Science and Engineering*, vol. 1, no. 1, pp. 17-35, 2021.
9. A. Senouci, M. Elabbasy, E. Elwakil, B. Abdrabou, and T. Zayed, "A model for predicting failure of oil pipelines," *Structure and Infrastructure Engineering*, vol. 10, no. 3, pp. 375-387, 2014.
10. M. Obaseki, "Artificial Neural Network Simulation Model for Predicting Oil and Gas Pipeline Corrosion Rate in Nigerian Niger Delta," *FUPRE Journal of Scientific and Industrial Research (FJSIR)*, vol. 1, no. 1, pp. 86-94, 2017.
11. M. V. Biezma, D. Agudo, and G. Barron, "A Fuzzy Logic method: Predicting pipeline external corrosion rate," *International Journal of Pressure Vessels and Piping*, vol. 163, pp. 55-62, 2018.
12. N. B. Shaik, S. R. Pedapati, S. A. A. Taqvi, S. Ahsan and F. A. Abd Dzubir, "Classification of faults in oil and gas pipelines using support vector machines," *Pertanika Journal of Science and Technology*, vol. 28, no. S1, pp. 173-184, 2020.
13. N. B. Shaik, S. R. Pedapati, A. R. Othman, K. Bingi, and F. A. A. Dzubir, "An intelligent model to predict the life condition of crude oil pipelines using artificial neural networks," *Neural Computing and Applications*, vol. 33, no. 21, pp. 14771-14792, 2021.
14. E. Kaya and C. Başтemur Kaya, "A novel neural network training algorithm for the identification of nonlinear static systems: Artificial bee colony algorithm based on effective scout bee stage," *Symmetry*, vol. 13, no. 3, p. 419, 2021.
15. S.R. Pedapati, N.B. Shaik, , and F.A.B. Dzubir, "Remaining useful life prediction of a piping system using artificial neural networks: A case study," *Ain Shams Engineering Journal*, vol. 13, no. 2, p. 101535, 2022.
16. A. Le Nai, "*Neural Network SVG Visualizer*," Accessed: Sep. 7, 2023. [Online]. Available: https://alexlenail.me/NN-SVG/
17. H. Song, L. Yang, G. Liu, G. Tian, D. Ona, Y. Song, and S. Li, "Comparative analysis of in-line inspection equipments and technologies," IOP Conference Series: *Materials Science and Engineering*, vol. 382, no. 3, p. 032021, July 2018.
18. M. S. El-Abbasy, A. Senouci, T. Zayed, and F. Mosleh, "A condition assessment model for oil and gas pipelines using integrated simulation and analytic network process," *Structure and Infrastructure Engineering*, vol. 11, no. 3, pp. 263-281, 2015.

19. E.S. Menon, *Gas Pipeline Hydraulics*, CRC Press, 2005.

20. M.A.A. Ayoub. B.M.D. Demiral, 2011. Application of resilient back-propagation neural networks for generating a universal pressure drop model in pipelines. *University of Khartoum Engineering Journal (UOKEJ)*, 1(2), pp.9-21.

21. A. Essien and C. Giannetti, "A deep learning model for smart manufacturing using convolutional LSTM neural network autoencoders," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 9, pp. 6069-6078, 2020.

22. H. Woldesellasse and S. Tesfamariam, "Data augmentation using conditional generative adversarial network (cGAN): Application for prediction of corrosion pit depth and testing using neural network," *Journal of Pipeline Science and Engineering*, vol. 3, no. 1, p. 100091, 2023.

23. S.A. Khwaja and S. Paul, "Inspection of Coated Hydrogen Transportation Pipelines," *Applied Sciences*, vol. 12, no. 19, p. 9503, 2022.

24. E.S. Ameh, S.C. Ikpeseni, and L.S. Lawal, "A review of field corrosion control and monitoring techniques of the upstream oil and gas pipelines," *Nigerian Journal of Technological Development*, vol. 14, no. 2, pp. 67-73, 2017.

25. M. E. Orazem, "*Underground Pipeline Corrosion*," Elsevier Science, 2014.

26. H. Zhang, S. Sha, C. Willis, and P. Chen, "Feasibility Study of Pinhole Inspection via Magnetic Flux Leakage and Hydrostatic Testing in Oil & Gas Pipelines," in IOP Conference Series: *Materials Science and Engineering*, vol. 1043, no. 2, p. 022053, 2021.

27. Y. Guang, W. Wang, H. Song, H. Mi, J. Tang, Z. Zhao, Prediction of external corrosion rate for buried oil and gas pipelines: A novel deep learning method with DNN and attention mechanism, International Journal of Pressure Vessels and Piping, Vol. 209, 2024.

28. “Pipeline Operators Forum” POF 100 series, Mar. 30, 2021. Accessed: Jul. 9, 2023. [Online]. Available: https://pipelineoperators.org/documents.

29. N.B. Shaik, S.R. Pedapati, S.A.A. Taqvi, A.R. Othman, and F.A.A. Dzubir, "A feed-forward back propagation neural network approach to predict the life condition of crude oil pipeline," *Processes*, vol. 8, no. 6, p. 661, 2020.

30. Y. Wu, L. Liu, J. Bae, K.H. Chow, A. Iyengar, C. Pu, W. Wei, L. Yu, and Q. Zhang, "Demystifying learning rate policies for high accuracy training of deep neural networks," 2019 *IEEE International Conference on Big Data*, Dec. 2019, pp. 1971-1980.