Corrosion Rate Prediction for Underground Gas Pipelines Using Artificial Neural Network (ANN)

[1]*A. Ahmaid*[1]*, and [2]*F. Khoshnaw*

[1]*Sirte Oil company, Brega, Libya, AShwigi@sirteoil.com.ly*
[2] *School of Engineering and Sustainable Development, De Montfort University, UK*
ORCID No. (Author) :0000-0003-4344-0703 scopus Author ID: 55351723700

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 0.9999. The study addresses the challenge of accurately predicting corrosion rates and remaining life in underground gas pipelines by developing an optimized 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 rate, ANN

التنبؤ بمعدلات التآكل لخطوط أنابيب الغاز المدفونة تحت الأرض باستخدام الشبكات العصبية الاصطناعية (ANN)

اشرف احميد – فؤاد خوشناو
المؤلف: اشرف احميد – شركة سرت لانتاج النفط والغاز
مساعد المؤلف : فؤاد خوشناو- جامعة دي منتفورت بريطانيا

الملخص

تتناول هذه الدراسة تحدي التنبؤ الدقيق بمعدلات التآكل وتقدير العمر المتبقي لخطوط أنابيب الغاز المدفونة تحت الأرض، حيث تتأثر هذه العملية بالتفاعل المعقد بين العوامل الفيزيائية والظروف البيئية. وتُعد النماذج التقليدية غير كافية لتمثيل هذه المتغيرات بصورة دقيقة، مما يؤدي إلى انخفاض موثوقية التنبؤات. وللتغلب على هذه المشكلة، تهدف الدراسة إلى تطوير نموذج أكثر دقة ومحسّناً باستخدام الشبكات العصبية الاصطناعية (ANN)
تركز الدراسة على التنبؤ بمعدل التآكل وتقدير العمر المتبقي لخطوط أنابيب الغاز المدفونة تحت الأرض باستخدام شبكات عصبية اصطناعية تم تنفيذها باستخدام برنامج MATLAB وقد أُخذت في الاعتبار كل من العوامل الفيزيائية، مثل أقصى عمق للتآكل وسُمك الأنبوب، والمتغيرات البيئية، مثل الرطوبة والمقاومية النوعية للتربة وتركيز الكلوريد. وأظهر التحليل أن عمق التآكل وسُمك جدار الأنبوب يمثلان من أهم العوامل المؤثرة، حيث بلغت مساهمتهما في التأثير على سلامة المادة 20%و16%، على التوالي.

تحقق النموذج الآمثل للشبكة العصبية الاصطناعية، باستخدام خوارزمية Levenberg–Marquardt وطبقة مخفية واحدة تضم 10خلايا عصبية، دقة تنبؤية عالية، حيث بلغ متوسط مربع الخطأ (MSE) = 0.038، ومعامل التحديد $R^2 = 0.9999$. وتتمثل مساهمة هذه الدراسة في تطوير أداة تنبؤية عالية الدقة للتنبؤ بمعدل التآكل والعمر المتبقي لخطوط أنابيب الغاز المدفونة تحت الأرض، بما يتفوق على النماذج التقليدية، ويدعم اتخاذ قرارات أكثر فاعلية فيما يتعلق بصيانة خطوط الأنابيب وسلامتها.

الكلمات المفتاحية :سلامة خطوط الأنابيب؛ معدل التآكل؛ الشبكات العصبية الاصطناعية (ANN)

1. Introduction

Pipelines are essential for the safe and efficient transport of petrochemical products, connecting production sites with industrial and consumer centers while supporting economic growth and reducing dependence on higher-emission transportation methods M. E. A. B. Seghier; etal (2021). Underground gas pipelines provide additional safety and environmental benefits through continuous monitoring and early fault detection. However, external corrosion resulting from soil–pipe interactions remain a major threat to pipeline integrity, causing safety hazards, environmental damage, costly maintenance, and potential loss of public confidence M. S. Liew; etal (2012).

Corrosion rate, commonly measured in millimeters or inches per year, indicates the speed of metal deterioration caused by environmental chemical reactions. It is influenced by factors such as material properties, chemical composition, temperature, and corrosive agents. Remaining life estimates the safe operational period of a structure based on corrosion rate and material loss, enabling effective maintenance planning and safe operation (Derbo Steel Pipe Co., Ltd. [Online],(2023).

Maintaining pipeline integrity requires understanding corrosion mechanisms and implementing effective prevention and control strategies. This study aims to enhance the predictive capability of Artificial Neural Networks (ANNs) for corrosion forecasting. A reliable ANN-based approach can improve asset management, reduce failure risks, and support safer pipeline operations. ANN models offer advantages including high prediction accuracy, adaptability, and integration with inspection systems, helping reduce maintenance costs, extend pipeline lifespan, improve operational reliability, and optimize investment returns R. M. Mahmoodian and C. Q. Li, (2020).

Biezma et al. (2018), developed a fuzzy logic model to predict external pipeline corrosion rates and support inspection planning. Shaik etal. (2020), used Support Vector Machines (SVM) to predict metal loss defects in Sudanese oil pipelines, achieving 93% accuracy. In another study, Shaik etal, (2021), proposed an ANN-based model for predicting pipeline lifespan and classifying metal loss defects.

1.1 Research gap: Although previous studies advanced corrosion prediction, most focused on limited corrosion types and input variables, restricting practical applicability. This study addresses these limitations by developing an optimized ANN model using the Levenberg–Marquardt algorithm to analyze seven corrosion types, integrate physical and environmental factors, and predict corrosion rate, remaining life, and corrosion type.

2. Methodology

Study Area: This study analyzes a 34-inch underground gas pipeline operated by Sirte Oil Company in Libya using 200 samples from intelligent pigging data collected in

2009 and soil properties including resistivity, moisture, and chloride content.

2.1 Data Calculations: Two key ANN outputs, corrosion rate and remaining life, were calculated from the input data. Corrosion rate was determined using the Point-to-Point analysis method in accordance with API 570 standards, while remaining life was estimated based on corrosion progression and wall thickness loss E.S. Menon (2005), and S. R. Pedapati,etal (2022).

2.2 Meteorological Data: The ANN model uses corrosion distance, maximum depth, current thickness, pipe age, moisture, soil resistivity, and chloride content as inputs, while predicting corrosion rate, remaining life, and corrosion type (see Figure 1).

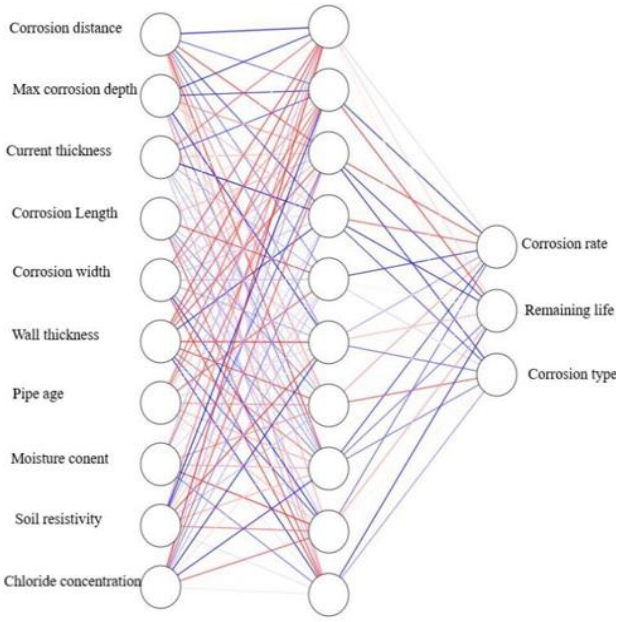


Figure 1. The ANN architecture

2.3 Data processing: Data preprocessing involved removing duplicates, handling missing values, and normalizing variables to improve ANN performance. Corrosion types were encoded numerically from 1 to 7, as shown in table 1, facilitating ANN training and subsequent decoding into interpretable classifications B. Akhlaghi, etal; (2023).

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

2.4 Classification of corrosion types: Metal loss anomalies were classified using intelligent pigging and nondestructive testing data figure 2. The parameter A was derived from pipe thickness (t), where A = 10 mm when $t < 10$ mm, and A = t when $t \geq 10$ mm, H. Zhang etal; (2021).

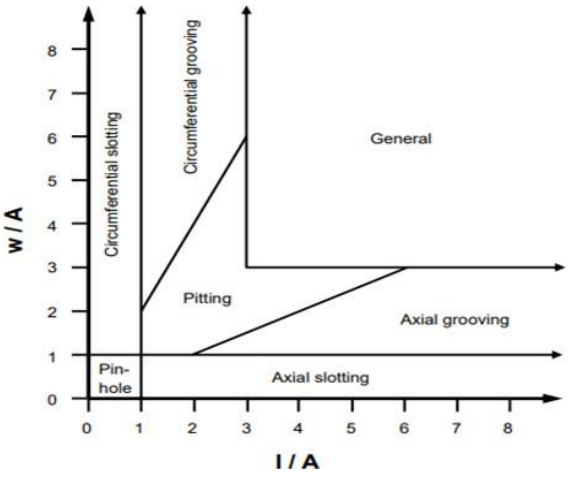


Figure 2. Classification of metal loss anomalies per dimension class

2.5 ANN model development: A The study used a feedforward neural network with 10 input variables, 3 output variables, and 1–3 hidden layers containing 5–15 neurons. Different network architectures, layer sizes, and data divisions were evaluated iteratively to identify the optimal configuration. The MATLAB Data Division function was used to randomly divide the data into training, validation, and testing sets using a 70:15:15 ratio. Sigmoid, tanh, and ReLU activation functions were considered to capture nonlinear relationships. The model was trained through data preprocessing, forward propagation, backpropagation, and optimization, and its performance was evaluated using MSE and R^2.

3. Results

3.1 Meteorological Analysis: Statistical analysis was conducted on variables used in the ANN model, including corrosion distance, corrosion depth, wall thickness, pipe age, moisture content, soil resistivity, and chloride concentration. Results showed an average corrosion rate of 0.11 mm/year and remaining life ranging from 0.01 to 59.14 years. These findings support failure prediction and maintenance planning while evaluating optimal ANN configurations.

Table 2 shows significant correlations with corrosion rate. Maximum corrosion depth has a near-perfect positive correlation (0.9963). Thickness is weak (0.2334). Corrosion width (0.574) and wall thickness (0.8178) are strong. Soil resistivity shows a moderate negative correlation (-0.495).

Table 2. 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 3 highlights the multifaceted factors influencing corrosion. The depth of corrosion depth (20%) and wall thickness (16%) are crucial factors in susceptibility.

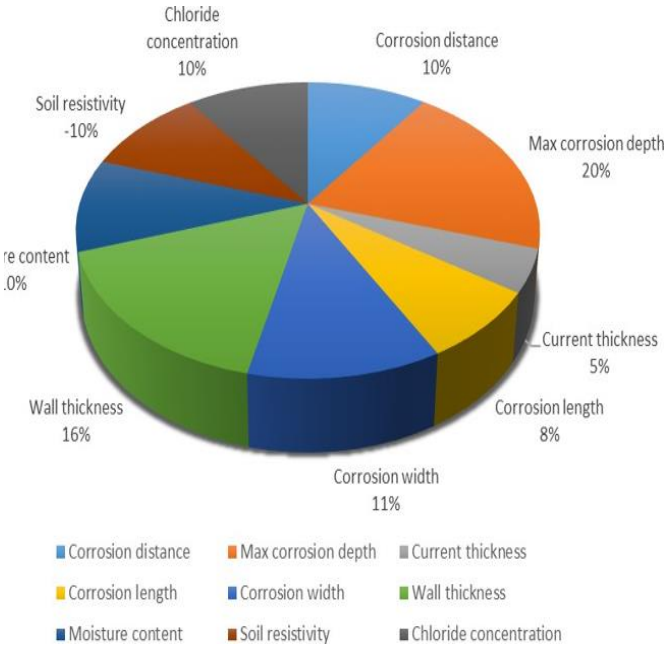


Figure 3. Contribution factor

3.2 Comparative Analysis of ANN Model Experiments: The ANN experiments compared training algorithms using MSE

and R^2 as performance metrics. The Levenberg–Marquardt algorithm achieved the best results, particularly with a single hidden layer, reaching an MSE of 0.03890 and an R^2 of 0.9999 using 10 neurons. Results indicate that moderate network complexity yields optimal performance, whereas additional layers do not necessarily improve accuracy. Figure 4 illustrates a feedforward ANN's structure with input-to-output data flow.

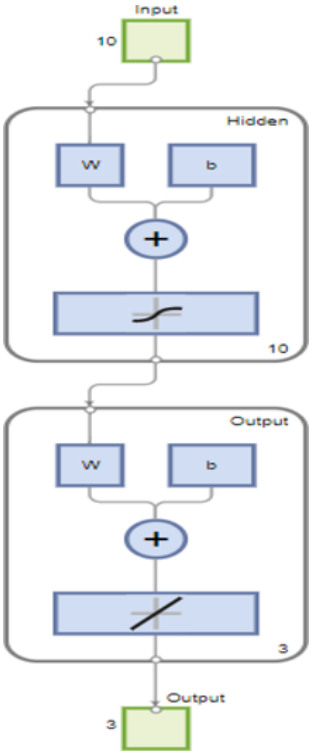


Figure 4. Architecture of optimal model

Figure 5 in MATLAB shows the ANN training state, including gradient, learning rate, and validation performance across epochs. At epoch 92, the gradient (0.16419) indicates ongoing learning. A low learning rate (0.0001) improves stability but slows convergence. A validation check value of 6 suggests the model may be nearing optimal performance or potential overfitting, guiding training adjustments.

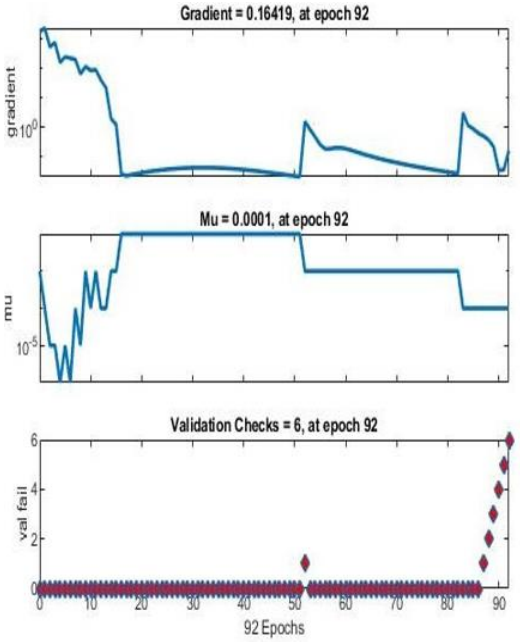


Figure 5. Training state plot

Figure 6 shows that epoch 86 records the highest validation performance at 0.03891, indicating successful handling of unseen data.

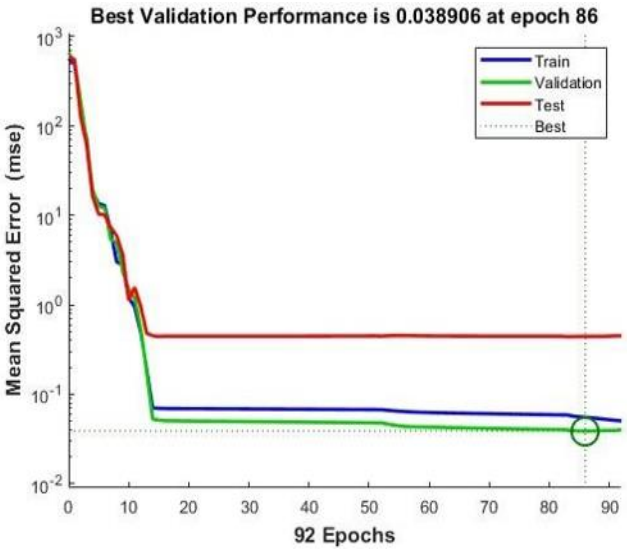


Figure 6. Performance plot

Figure 7 shows ANN regression performance comparing actual and predicted corrosion rates.

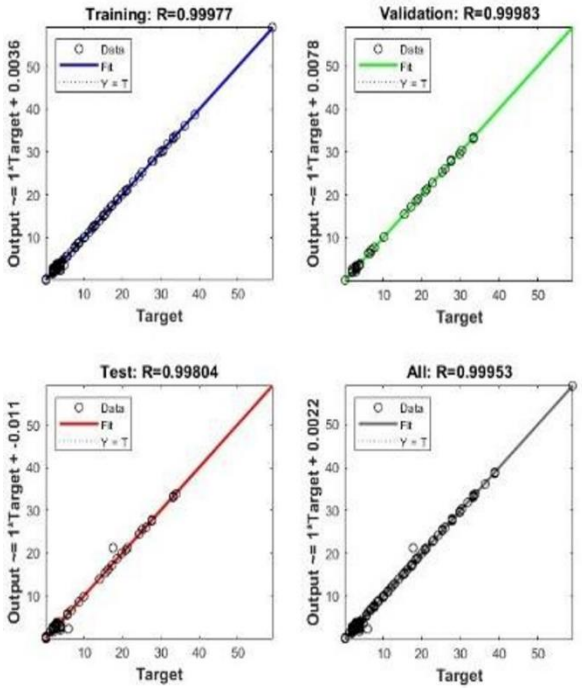


Figure 7. ANN model regression plots

Table 3 compares actual and predicted corrosion rate, remaining life, and corrosion type for 24 unseen samples. Most errors are below 10%, indicating good accuracy. Remaining life predictions are highly accurate (0.02%–1.70%).

Table 3. 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 8, 9 and 10 compare actual and predicted corrosion rates, remaining life, and corrosion types across 24 samples.

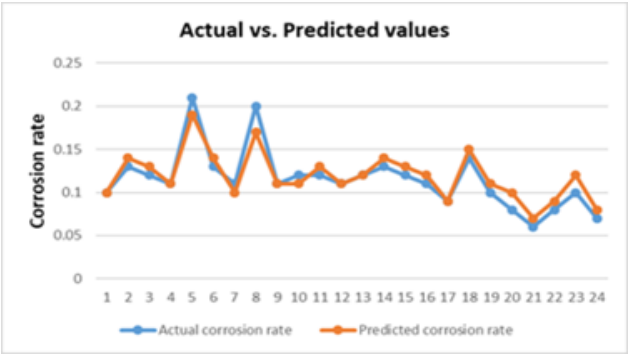


Figure 8. actual vs. predicted values for corrosion rate

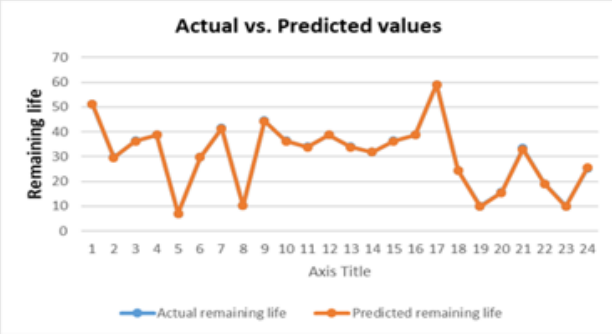


Figure 9. actual vs. predicted values for remaining life

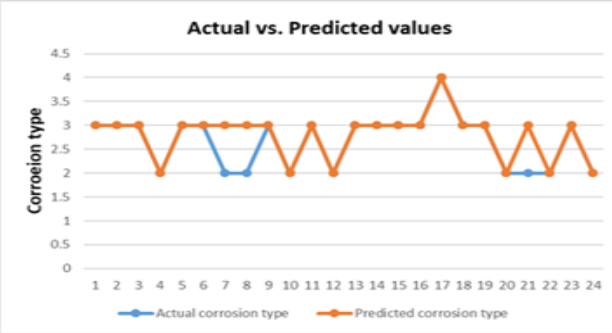


Figure 10. actual vs. predicted values for corrosion types

4. Results and Discussion

The results demonstrate that physical factors, particularly maximum corrosion depth and wall thickness, have the greatest influence on pipeline corrosion progression. Maximum corrosion depth exhibits an exceptionally high correlation (0.9963) with corrosion rate. Environmental factors show weaker correlations.

The ANN model optimized using the Levenberg–Marquardt algorithm and a single hidden layer with 10 neurons achieved the best performance, with an R^2 of 0.9999 and an MSE of 0.03890. Increasing network complexity by adding hidden layers did not improve accuracy, highlighting the importance of selecting a balanced architecture. High R^2 values across training, validation, and testing datasets indicate strong generalization and reliable remaining-life predictions.

Evaluation on unseen data showed that most predictions were within a 10% error range. Remaining life predictions were particularly accurate, with errors ranging from 0.02% to 1.70%. Although minor discrepancies occurred in corrosion-type classification, the model consistently provided reliable predictions of corrosion rate, remaining life, and corrosion type, supporting its application in pipeline integrity management.

The feedforward ANN architecture, employing a hyperbolic tangent sigmoid activation function in the hidden layer and a linear transfer function in the output layer, enabled stable and effective learning. Model performance was further enhanced through proper data preprocessing and careful hyperparameter tuning, including learning rate and layer configuration.

ANNs offer significant advantages for corrosion prediction, including high accuracy, adaptability to new data, and compatibility with inspection technologies such as in-line inspection systems.

4.1 Limitations: The study did not incorporate additional factors such as soil pH and microbial activity due to time constraints. Furthermore, optimizing ANN architecture and hyperparameters requires extensive experimentation and validation.

4.2 Implications: The proposed ANN model accurately predicts corrosion rate and remaining life, enabling targeted maintenance, reducing failure risks and costs, and extending pipeline lifespan. Future improvements may be achieved by incorporating additional environmental factors and integrating the model with existing inspection technologies.

5. Conclusion

- The contribution factor emphasizes 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.
- Through a systematic exploration of various ANN architectures, an optimal feedforward Levenberg-Marquardt configuration was determined that balances model complexity with computational efficiency.
- The developed ANN model demonstrated a high level of accuracy in predicting the corrosion rate of underground gas pipelines. The results indicated that the ANN model outperformed traditional statistical models and other machine-learning techniques in terms of prediction accuracy.
- 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 maximizing return on investment.

6. Recommendations:

- Include additional environmental factors like soil pH and microbes for better accuracy.
- Continue tuning ANN architecture to improve performance.
- Combine ANN with inspection tools for comprehensive monitoring.
- Collect more varied data to enhance model generalization.
- Develop user-friendly tools for maintenance planning.
- Regularly validate and update the model with new data.
- Evaluate cost benefits to support ANN adoption

Acknowledgements

The authors gratefully acknowledge Sirte Oil Company, Libya, for providing pipeline inspection data and operational support essential to this study.

I also would like to sincerely thank Dr. Fuad Khoshnaw of the School of Engineering and Sustainable Development, De Montfort University, UK, for his invaluable guidance, supervision, and support throughout this research.

Finally, I appreciate the constructive feedback from colleagues and reviewers, which improved the quality of this research.

References

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 and Environmental Protection*, vol. 147, pp. 818–833, 2021.

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.

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.

R. M. Mahmoodian and C. Q. Li, "Failure assessment and safe life prediction of corroded oil and gas pipelines," J.S.E, vol. 151, pp. 434-438, 2017. Abubakirov, M. Yang, and N. Khakzad, “A risk-based approach to determination of optimal inspection intervals for buried oil pipelines,” *Process Safety and Environmental Protection*, vol. 134, pp. 95–107, 2020.

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.

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.

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.

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

B. Akhlaghi, H. Mesghali, M. Ehteshami, J. Mohammadpour, F. Salehi, and R. Abbassi, "Predictive deep learning for pitting corrosion modeling in buried transmission pipelines," Process Safety, vol. 174, pp. 320-327, 2023.

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.