WEP Warwick Evans Publishing

International Journal of Natural Resources and Environmental Studies

ISSN: 3006-2012 (Print), ISSN: 3006-0834 (Online) | Volume 7, Number 2, Year 2025
DOI: https://doi.org/10.62051/ijnres.v7n2.04
Journal homepage: https://wepub.org/index.php/IJNRES/index

An Integrated Study on Data-Driven Corrosion Prediction and Maintenance Decision-Making for Buried Pipelines

Tai Zhang, Jianchang Zhang, Ni Zhang and Dan Yu

The Third Oil Transportation Department of Changqing Oilfield Company, PetroChina, Yinchuan 750001, China.

ABSTRACT

This paper addresses the corrosion issues of buried pipelines by constructing a corrosion rate prediction model based on neural network algorithms, integrating multi-source data such as in-situ soil physicochemical parameters. Key input parameters of the model were identified through feature analysis, and the model was trained and validated using historical monitoring data. The results demonstrate that the model effectively captures the nonlinear relationships between multiple factors and corrosion rate with high prediction accuracy, particularly in the low corrosion rate range. Furthermore, the paper systematically reviews pipeline anti-corrosion maintenance decision-making methods based on maintenance timing, strategies, and evaluation metrics, comparing the applicability and limitations of reactive, periodic, and predictive maintenance strategies. It emphasizes that an intelligent maintenance system integrating multidisciplinary approaches such as machine learning, risk assessment, and reliability analysis can achieve closed-loop management of pipeline corrosion protection. This approach enhances the economy, safety, and sustainability of maintenance practices, providing theoretical support and decision-making guidance for the intelligent operation and maintenance of oil and gas station pipelines.

KEYWORDS

Buried Pipeline; Corrosion Prediction; Neural Network; Corrosion Protection Decision-Making.

1. INTRODUCTION

Oil and gas pipelines serve as critical infrastructure for energy transportation. In China, most pipelines are buried underground, offering advantages in safety, economic efficiency, space saving, and pollution reduction. However, buried pipelines are subjected to external corrosion risks—such as soil conditions and coating degradation—as well as internal corrosion caused by the properties of conveyed fluids, temperature, and pressure fluctuations [1, 2]. These factors significantly increase the likelihood of pipeline corrosion and leakage, posing serious threats to personnel safety and the ecological environment [3, 4].

Corrosion of buried pipelines involves complex mechanisms, which can be broadly categorized into three types: chemical, electrochemical, and physical corrosion. Among these, soil-induced corrosion is typically electrochemical in nature. Soil properties—such as moisture content, resistivity, pH, redox potential, and anion concentration—significantly influence the corrosion rate and remaining service life of pipelines, thereby maintaining them under high-risk conditions for extended periods [5, 6]. Furthermore, stray currents generated by electrical equipment within the station can accelerate corrosion, shifting the process from natural to aggressive degradation, considerably shortening the pipeline's lifespan, and increasing the risk of leaks. Therefore, analyzing and predicting corrosion factors for station pipelines is essential to ensuring operational safety, as it enables early detection of potential corrosion, extends service life, and helps control maintenance costs [7, 8].

Fig. 1 Exterior View of Corroded Buried Pipelines in an Oil Transfer Station

Based on historical operational data and external corrosion monitoring results of buried pipelines, this study integrates soil physicochemical parameters—such as moisture content, resistivity, pH, redox potential, and anion concentration—as well as electrical interference factors. Using corrosion rate as the core evaluation metric, machine learning algorithms and optimization methods are employed to develop a comprehensive corrosion assessment model for buried pipelines within the station. The model aims to provide a theoretical basis and decision support for corrosion prevention strategies and lifecycle management of pipelines [9, 10].

2. CORROSION RATE PREDICTION MODEL

Corrosion prediction for buried pipelines is a process that estimates the likelihood and severity of future corrosion by integrating historical and real-time monitoring data, operational conditions, and relevant environmental factors, combined with expert knowledge. It aims to formulate proactive maintenance strategies for corrosion prevention.

2.1. Methodology Overview.

With the increasing intelligence and digitalization of oil and gas pipeline systems, machine learning has gradually become a mainstream method for corrosion prediction, owing to its powerful nonlinear processing capability, high computational efficiency, and low dependency on physical mechanisms. Compared to traditional mechanistic models and statistical methods, it demonstrates superior performance in terms of accuracy and stability. Commonly used machine learning algorithms include basic classification algorithms—such as decision trees, support vector machines, neural networks, and Bayesian networks—as well as ensemble algorithms like random forests and clustering algorithms. Among these, neural networks, support vector machines, Bayesian networks, and random forests are widely applied in practice. The characteristics of these methods are summarized in the table below.

Table 1 Performance Comparison of Selected Machine Learning Algorithms

Method	Characteristics	Advantages
Neural Network	Self-learning, associative storage; high-speed optimization seeking	Autonomous adjustment of multi-factor weights; excellent performance in image recognition and predictive analysis
Support Vector Machine	Concise model; effective learning with high-dimensional data	Fast prediction speed; strong generalization ability and robustness; capable of handling nonlinear problems
Bayesian Network	Handles complex relationships among large-scale variables; modifiable probability distributions	Effectively describes dependencies between variables; enables construction of dynamic models with multiple variables
Random Forest	High accuracy and robustness; handles large-scale and high-dimensional data	Effectively reduces overfitting risk; parallel training of multiple decision trees; evaluates feature importance efficiently

This study primarily employs a neural network to construct a corrosion prediction model for buried pipelines, leveraging its powerful capability for nonlinear fitting and feature learning. The model effectively integrates multi-source data, including soil environmental parameters such as moisture content, resistivity, and pH value, to estimate the pipeline corrosion rate. By optimizing the network structure and parameters through training, the model can deeply reveal the complex influences of various factors on the corrosion process, providing a reliable basis for pipeline safety and intelligent maintenance.

2.2. Neural Network.

A neural network is a computational model inspired by the structure of the human brain. Its core consists of multiple layers of interconnected artificial neurons. Each neuron processes input data through weighted summation and nonlinear activation functions, progressively extracting and transforming features across layers. By optimizing parameters via backpropagation and gradient descent, the network autonomously learns complex nonlinear mappings from large-scale data without relying on explicit physical mechanisms. Owing to its powerful feature learning capability and high predictive accuracy, this algorithm is widely applied in scenarios such as corrosion prediction.

(1) The linear transformation of a single neuron is expressed as follows, achieving a linear combination of the input signals:

$$z = (w_1 \cdot x_1 + w_2 \cdot x_2 + \cdots + w_n \cdot x_n) + b = \sum_{i=1}^{n} w_i x_i + b \tag{1}$$

(2) The activation function serves as a critical step for introducing nonlinearity, enabling the network to learn complex patterns.

$$a = \sigma(z) \tag{2}$$

(3) The loss function quantitatively measures the discrepancy between the model's predictions and the true values, thereby enhancing and stabilizing the prediction accuracy of the model.

$$\mathrm{L} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{K} y_{i,j} \,\mathrm{lo} \tag{3}$$

The neural network computational framework of this study is illustrated in the figure below, which systematically demonstrates the complete workflow from data input to model application.

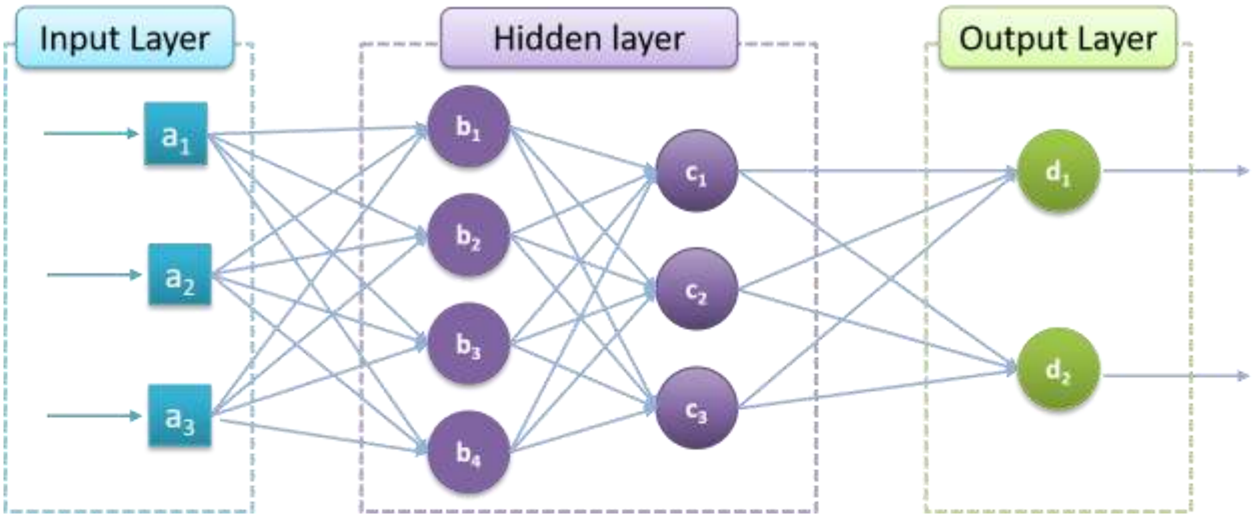

Fig. 2 Architecture of the Neural Network Computational Framework

2.3. Prediction Model Construction.

This study collected 50 sets of corrosion point data from buried pipelines, including various soil physicochemical parameters and actual pipeline corrosion rate values. Through significance analysis and feature selection, key input features—such as soil moisture content, resistivity, pH, and other indicators (A–G)—were identified, with the pipeline corrosion rate set as the output target. The dataset was randomly divided into training and testing sets in a 7:3 ratio and standardized to eliminate scale effects. A backpropagation (BP) neural network with two hidden layers was constructed to learn the complex nonlinear mapping between the selected features and the corrosion rate using the training set. The generalization ability of the model was evaluated using the test set. The results demonstrate that the model achieves high prediction accuracy and can serve as an effective tool for corrosion protection of pipelines.

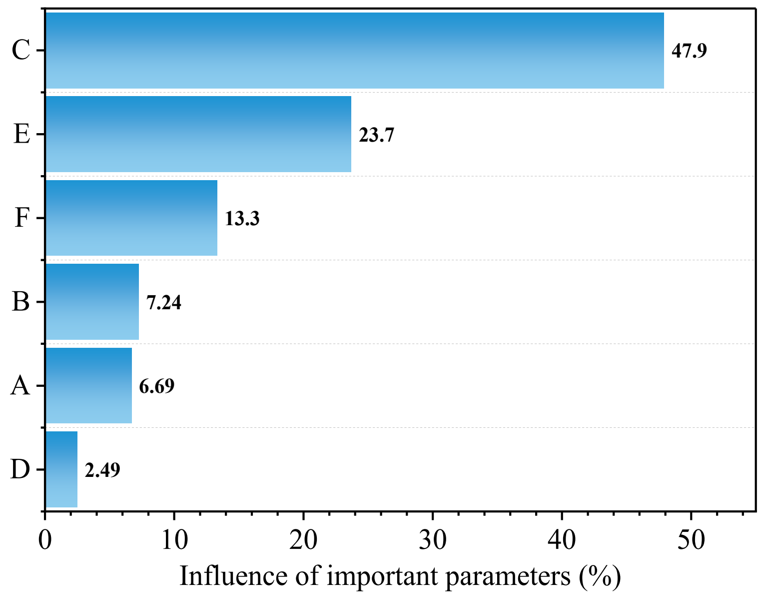

Fig. 3 Significance Screening of Influencing Factors for Buried Pipeline Corrosion

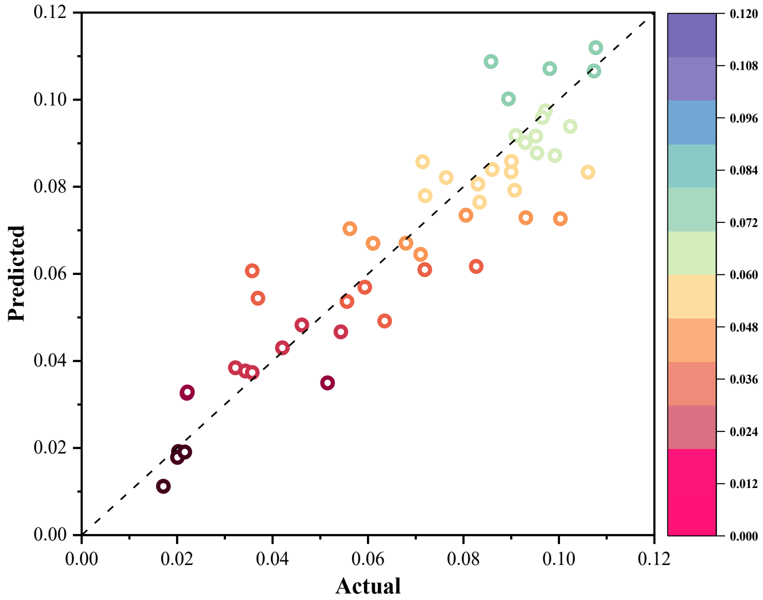

Fig. 4 Comparison Between Predicted and Actual Results of Pipeline Corrosion Rate Model

Analysis of the scatter plot between predicted and true values indicates that the pipeline corrosion rate prediction model demonstrates favorable performance. Most data points are evenly distributed on both sides of the y=x reference line, suggesting close agreement between predictions and actual measurements, with an overall small prediction error. Particularly in the low corrosion rate range (0.00–0.04), the model shows high predictive accuracy, as data points are tightly clustered around the reference line. In medium and higher corrosion rate intervals, although minor deviations are observed in some points, the vast majority remain within an acceptable error margin. The model effectively captures the nonlinear relationship between soil physicochemical parameters and corrosion rates, validating the selected features and the efficacy of the neural network architecture. It thus provides reliable data support for pipeline safety management and maintenance decision-making.

3. DECISION-MAKING FOR PIPELINE CORROSION PROTECTION AND MAINTENANCE

Upon completion of pipeline corrosion rate prediction, maintenance decisions must be further developed based on the predictive results. This process takes the corrosion rate assessment as the core basis, comprehensively considers factors such as safety risks, maintenance costs, and benefits, and employs maintenance decision-making theories and technologies to determine recommended solutions. Thereby, predictive data are transformed into executable maintenance strategies, achieving closed-loop management from condition early-warning to precision operation and maintenance. Corrosion of buried oil and gas pipelines results from the coupling of multiple factors such as soil properties and medium composition [11, 12]. Traditional online inspection and periodic maintenance exhibit limitations and lagging responses, while existing passive maintenance strategies often lead to resource waste and cost increases, failing to adapt to the dynamic evolution of corrosion. With the advancement of technologies such as machine learning and fault detection, these methods have promoted the diversification and refinement of maintenance decision-making. They provide new pathways for dynamic adjustment and optimized resource allocation in pipeline maintenance, contributing to the establishment of a more efficient and sustainable intelligent maintenance system [13, 14]. The disciplines related to pipeline anti-corrosion maintenance decision-making are shown in the figure below.

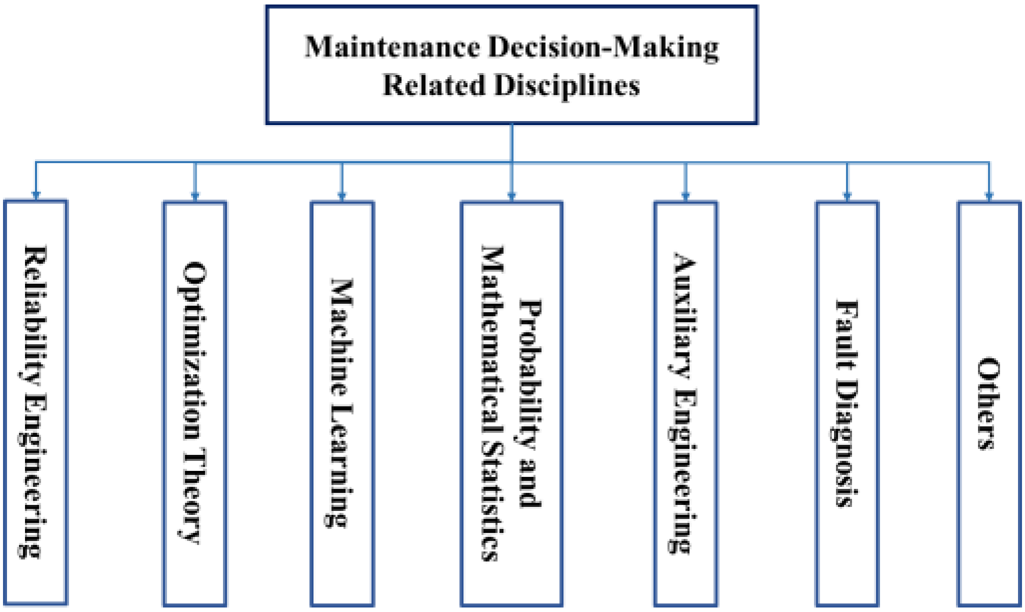

Fig. 5 Theories and Methods for Decision-Making in Pipeline Corrosion Protection and Maintenance

3.1. Maintenance Strategies Based on Timing.

Based on pipeline corrosion status and failure risk, maintenance strategies can be categorized into three types: reactive, periodic, and predictive maintenance. Reactive maintenance is performed after a failure occurs, characterized by passive response and a high likelihood of production interruptions, safety incidents, and substantial costs. Periodic maintenance involves scheduled inspections and interventions, which can prevent certain risks but relies on fixed intervals, potentially resulting in unnecessary repairs and insufficient adaptability. In contrast, predictive maintenance utilizes online monitoring and artificial intelligence technologies to dynamically formulate proactive maintenance plans through real-time data analysis and corrosion prediction. This approach combines high efficiency, cost-effectiveness, and safety, significantly enhancing the reliability and sustainability of pipeline operations [15, 16].

3.2. Maintenance Strategies Based on Methods.

Based on the severity of pipeline defects and repair techniques, maintenance strategies for oil and gas pipelines can be classified into corrective, replacement, and preventive maintenance. Corrective maintenance (e.g., reinforcement, sleeve repair, welding) targets local defects with low cost and rapid restoration of integrity, yet it fails to address systemic issues. Replacement maintenance is suitable for severe defects (such as corrosion depth exceeding 80% of wall thickness), eliminating risks thoroughly through pipe replacement, but it involves high costs and production disruptions, implemented via trenching or trenchless methods. Preventive maintenance, meanwhile, focuses on corrosion prevention and retardation, reducing risks and lifecycle costs through regular inspections and interventions. However, its effectiveness relies on a robust integrity management data system. Future efforts should integrate big data and machine learning to enable more intelligent decision-making [17, 18].

3.3. Maintenance Strategies Based on Evaluation Metrics.

With the integration of multidisciplinary knowledge, maintenance strategies for pipeline corrosion have become increasingly diverse and can be categorized into decision-making methods based on evaluation metrics such as reliability, remaining useful life, risk assessment, and fault tree analysis.

Among these, reliability-based methods estimate the probability of failure by analyzing defect data, operational records, and maintenance history, thereby formulating maintenance strategies aimed at loss minimization. In contrast, risk-based methods further quantify both the likelihood of failure and its potential consequences, guiding precise maintenance actions oriented towards risk-cost optimization, which enhances both targeting accuracy and operational effectiveness [19, 20].

3.4. Maintenance Decision-Making Based on Remaining Useful Life.

Maintenance decision-making based on remaining useful life (RUL) prioritizes repair actions by assessing pipeline performance degradation stages and quantifying damage trends incorporating material properties, operational conditions, and environmental factors. Although this approach provides an intuitive reflection of equipment status, prediction deviations may occur due to the complexity of corrosion processes. In contrast, fault tree analysis (FTA) starts from a top event and identifies systemic vulnerabilities through causal logic, enabling the development of cost-optimal maintenance strategies based on criticality ranking.

4. SUMMARY

(1) The machine learning-based corrosion prediction model demonstrates significant applicability and reliability. The neural network prediction model developed in this study effectively integrates multi-source environmental and operational data, accurately capturing the complex nonlinear relationships between soil physicochemical parameters and corrosion rates. It exhibits high predictive accuracy, particularly in the low to medium corrosion rate range, thereby providing a scientific basis for pipeline condition assessment and risk early warning.

(2) Corrosion maintenance decision-making is critical to ensuring the long-term safe operation of pipelines. Traditional passive and time-based maintenance strategies are often inadequate in addressing complex corrosion issues due to their reactive nature and inherent delays. In contrast, proactive and dynamic decision-making methods—based on real-time monitoring, corrosion prediction, and optimization—significantly enhance the targeting, cost-effectiveness, and systemic safety of maintenance activities. This approach represents the future direction of intelligent pipeline operation and maintenance.

(3) Multidisciplinary integration is driving the evolution of pipeline maintenance toward greater refinement and intelligence. Effective maintenance decision-making requires the synthesis of multiple indicators, including reliability, risk, corrosion rate, and remaining service life. By integrating big data analytics, machine learning, and modern inspection technologies, a closed-loop management system unifying “prediction–assessment–decision” can be established. This integrated framework not only enhances pipeline safety but also optimizes life-cycle costs and improves operational sustainability.

REFERENCES

[1] X. Wu, H.Y. Wu, B. Peng, et al., Soil Corrosion Rate Prediction along Sichuan to Eastern China Gas Transmission Pipeline Based on BP Neural Network, Pipeline Technique and Equipment. (01) (2015) 7-9.

[2] Senouci Ahmed, Elabbasy Mohamed, Elwakil Emad, Abdrabou Bassem, Zayed Tarek. A model for predicting failure of oil pipelines [J]. Structure and Infrastructure Engineering, 2014, 10(3): 375-387.

[3] Ma Haonan, Geng Mengying, Wang Fan, Zheng Wenyue, Ai Yibo, Zhang Weidong. Data augmentation of a corrosion dataset for defect growth prediction of pipelines using conditional tabular generative adversarial networks [J]. Materials, 2024, 17(5): 1142.

[4] Li Yun-Tao, He Xiao-Ning, Shuai Jian. Risk analysis and maintenance decision making of natural gas pipelines with external corrosion based on Bayesian network [J]. Petroleum Science, 2022, 19(3): 1250-1261.

[5] Gong Changqing, Zhou Wenxing. Multi-objective maintenance strategy for in-service corroding pipelines using genetic algorithms [J]. Structure and Infrastructure Engineering, 2018, 14(11): 1561-1571.

[6] Chen Yunxiang, Wang Zezhou, Cai Zhongyi. Optimal maintenance decision based on remaining useful lifetime prediction for the equipment subject to imperfect maintenance [J]. Ieee Access, 2020, 8: 6704-6716.

[7] He Chu, Tan Zijing, Chen Qing, Sha Chaofeng. Repair diversification: A new approach for data repairing [J]. Information Sciences, 2016, 346: 90-105.

[8] Xie Mingjiang, Zhao Jianli, Zuo Ming J, Tian Zhigang, Liu Libin, Wu Jinming. Multi-objective maintenance decision-making of corroded parallel pipeline systems [J]. Applied Energy, 2023, 351: 121822.

[9] Jia Xuefei, Fu Chao, Chang Wenjun. A transfer-based decision-making method based on expert risk attitude and reliability [J]. Applied Intelligence, 2025, 55(7): 1-14.

[10] Tulabandhula Theja, Rudin Cynthia. On combining machine learning with decision making [J]. Machine learning, 2014, 97: 33-64.

[11] Z.B. Yin, S.S. Wang, Z.H. Zhu, et al., Key parameters of soil corrosivity and a model for predicting the corrosion rate of Q235steel in Beijing, Chinese Journal of Engineering. 45 (11) (2023) 1939-1947.

[12] Y. Guang, W.H. Wang, H.W. Song, et al., 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. 209 (2024) 105218, https://doi.org/10.1016/j.ijpvp.2024.105218.

[13] Tee Kong Fah, Pesinis Konstantinos. Reliability prediction for corroding natural gas pipelines [J]. Tunnelling and Underground Space Technology, 2017, 65: 91-105.

[14] Wang Wenbin, Jiang Shibin, Liu Jianguo, Cui Gan. An evaluation method for pipeline corrosion risk index weighting in beach and sea oil fields based on combined weighting with improved hierarchical analysis and Bayesian networks [J]. Applied Ocean Research, 2025, 158: 104522.

[15] Ossai Chinedu I. Corrosion defect modelling of aged pipelines with a feed-forward multi-layer neural network for leak and burst failure estimation [J]. Engineering Failure Analysis, 2020, 110: 104397.

[16] Elshaboury Nehal, Al-Sakkaf Abobakr, Alfalah Ghasan, Abdelkader Eslam Mohammed. Data-driven models for forecasting failure modes in oil and gas pipes [J]. Processes, 2022, 10(2): 400.

[17] Jiang Fengyuan, Dong Sheng, Zhao Enjin. A study on burst failure mechanism analysis and quantitative risk assessment of corroded pipelines with random pitting clusters [J]. Ocean Engineering, 2023, 284: 115258.

[18] Zhou Wenxing. System reliability of corroding pipelines [J]. International Journal of Pressure Vessels and Piping, 2010, 87(10): 587-595.

[19] Gong Changqing, Zhou Wenxing. First-order reliability method-based system reliability analyses of corroding pipelines considering multiple defects and failure modes [J]. Structure and Infrastructure Engineering, 2017, 13(11): 1451-1461.

[20] C.T. Wang, M.F. Hassanein, M.M. Li, Numerical simulation of oil and gas pipeline corrosion based on single- or coupled-factor modeling: A critical review, Natural Gas Industry B. 10 (5) (2023) 445-465, https://doi.org/10.1016/j.ngib.2023.08.001.