Heliyon 10 (2024) e26837

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Pipeline failure evaluation and prediction using failure probability and neural network based on measured data

H. Noroznia, M. Gandomkar *, J. Nikoukar

Department of Electrical Engineering, Islamic Azad University, Saveh Branch, Saveh, Iran

ARTICLE INFO

Keywords:
Stray currents
Remaining life
Condition monitoring
Cathodic protection
Failure probability
Neural networks

ABSTRACT

The chemical corrosion of metals in large industries such as oil and gas is a fundamental and costly problem. Gas transmission and distribution pipes and the other structures submerged in the soil and in an electrolyte, according to the existing conditions and according to the metallurgical structure, are corroded, and after a period of work, they disrupt an active system and process and lead to loss. The worst corrosion that occurs for metals embedded in the soil is where there are stray electric currents. Based on this, the cathodic protection of metal pipes is known as the most effective protection method to prevent the corrosion of structures buried in the ground, which is widely used to protect the corrosion distribution and transmission pipes of gas, oil, and water. In gas networks, current and voltage measurements for cathodic protection are carried out and recorded in specific periods according to the standards approved by the National Gas Company. The effect of stray currents on the obtained results is significant. The reason for this is that the available data is recorded as a time series, and as a result, the critical value of this time series will significantly impact the remaining life of the gas pipelines. Therefore, the purpose of this article is to investigate the stray currents effect on failure rate using normal probability distribution. In the following, the estimation of the remaining useful life of gas pipelines under cathodic protection is obtained using neural networks and compared with the results of the failure probability to check the accuracy of the results. According to the data history of the equipment, the amount of failure and the remaining useful life of the gas pipelines will be obtained.

1. Introduction

Knowing amount of damage and remaining life of apparatus helps managers and decision makers estimate plans, costs, and budgeting. These decisions will be very effective in annual budget planning, expenses, missions, purchase planning, and direct budgeting [1]. Considering that estimating the remaining useful life of apparatus is particularly important in making decisions in critical industrial fields, especially in oil and gas, and considering that premature replacement involves additional costs and late replacement, It also causes loss of life and money and increases the cost of maintenance and repairs, as a result, by knowing the useful life of the apparatus, it is possible to facilitate the management of goods and appliances.

Through the investigations, it has been determined that the worst corrosion that occurs for the metals in the soil is in places with stray electric currents. The specific resistance of soils is high even when they contain water, so electric currents in the ground will pass through metals embedded in the soil with little resistance. The stray current can cause pipe corrosion when it enters from one part of

* Corresponding author.
E-mail address: gandomkar_m@yahoo.com (M. Gandomkar).

https://doi.org/10.1016/j.heliyon.2024.e26837
Received 26 August 2023; Received in revised form 17 February 2024; Accepted 20 February 2024
Available online 22 February 2024
2405-8440/

the pipe and discharges from the other, and the flow circuit is complete.

Wandering currents are classified into the following three categories.

- Direct Current (DC),
- Alternative Current (AC) and
- Telluric Current

Direct pointed in drilling operations and welding operations using DC current. Underground train systems and the like, as well as the earth's magnetic field around the invaded pipe, also affect and cause disruption.

A system usually operates under various operating conditions, which may affect the degradation path of the system differently and thus reduce the accuracy of remaining practical life estimation. Reference [1] has proposed a model with a long time window to deal with this issue. First, a long time window is created in data preprocessing. Then, in model development, multiple degradation features are extracted by an improved various method, and these features are added to the raw data as additional features.

Machine learning-based algorithms have become promising in remaining useful life (RUL) estimation. Reference [2] presents the repeated architecture of estimating the RUL of turbofan engines. First, a deep long short-term memory (DLSTM) network with dropout in multiple layers is shown for RUL prediction. In the next step, the DLSTM model is improved to control the sequence in forward and backward direction using a bidirectional deep long short-term memory (BiDLSTM). Finally, an Attention-based deep LSTM (Attn-DLSTM) is presented, considering all RUL estimation time steps. Incorporating the attention mechanism helps improve the accuracy and interpretability of the deep LSTM network.

Reference [3] proposes a time-dependent survival neural network (NN) that incrementally estimates the latent failure risk and performs multiple binary classifications to produce RUL-specific probabilistic error predictions. A NN is trained with a new survival learning criterion in this reference.

New hybrid process for predicting the remaining useful life (RUL) of multi-functional spoiler (MFS) systems introduces by Ref. [4]. Also, in Ref. [5], a multi-reservoir echo mode network estimates fuel cell degradation and its remaining useful life. It is a method for predicting the evolution of output voltage of fuel cell with time.

Performance and maintenance data are reported in Ref. [6] from a list of cathodic protection (CP) systems installed on about 100 structures between 1987 and 2010. A large majority of them provide long-term corrosion protection. Failure of components and whole systems is determined as a function of age. Based on the statistical analysis of field data, the maintenance cost of a CP system is predicted using the life cycle cost model.

Reference [7] has discussed evaluating DC and AC stray corrosion phenomena on steel fibers and analyzing the main influencing parameters. Electrochemical instrumental methods, including Tofel polarization, CP, and electrochemical impedance spectroscopy (EIS), were used to evaluate the corrosion resistance of steel fibers, which have great potential to replace conventional steel reinforcements in construction. Reference [8] also creates a mathematical model of stray current distribution. References [9–13] have also done life estimation on railways. Reference [14] has investigated the functional life of pipeline using a general condition index and neural networks.

Reference [15] has introduced a new method called the self-evaluation decision-making algorithm (SEDMA) for prioritizing repairing and maintaining apparatus, including circuit breakers. In Ref. [16], By using normal probability distribution and condition monitoring of the equipment, it has been determined the failure and as a result scheduling the maintenance of the equipment.

Reference [17] has investigated the effect of pipes on the performance of the natural gas transmission network. Based on this, the performance of three groups of pipes, including new, 10 and 20-year-old pipes, has been evaluated and compared with each other.

Reference [18] has predicted the failure of oil and gas pipes using machine learning approach and reviewed the work of others in this field. In Ref. [19], Failure pressure estimation in oil and gas pipes has been done using a finite element study.

In reference [20], assessment of aging and maintenance of water-supply infrastructure has been done. Reference [21] has also evaluated the health of urban water-supply using the Bayesian method and triangular fuzzy number optimization.

Machine learning and artificial intelligence play pivotal roles across a spectrum of disciplines, revolutionizing industries and research endeavors. In healthcare, they aid in early disease diagnosis and drug discovery. Finance benefits from risk assessment and fraud detection algorithms, while autonomous vehicles leverage these technologies for real-time decision-making. Natural language processing enhances communication, and image recognition transforms various sectors, from e-commerce to manufacturing [22–24]. AI's influence extends to personalized education platforms and cybersecurity measures. With applications in climate prediction, these technologies contribute significantly to environmental sciences. Overall, machine learning and artificial intelligence have become indispensable tools, driving innovation, efficiency, and breakthroughs in diverse domains [25–27].

Therefore, failure prediction and health monitoring of gas pipelines, in addition to maintaining the security of the gas network, will lead to a reduction in the costs of the gas transmission network. So, the main goal of this article is to present a model for estimating the RUL of gas pipelines under cathodic protection in operational conditions, which is considered a suitable tool for operation management.

Now, the way to detect the failure of gas pipelines as well as the functional age of these pipelines in the gas transmission network should be correctly recognized so that their maintenance scheduling can be properly adjusted. Therefore, in the second part of this article, the information required to check the cathodic protection of the apparatus is examined. The third part will introduce the conditions' assessment methods, including the failure rate method based on normal probability distribution and neural network. Next, a case study conducted on an actual sample is given in the fourth section. In the end, the conclusion of this article is presented.

2. Data collection and monitoring

As mentioned, the chemical corrosion of metals is the most critical and costly problem in large industries, especially the oil and gas industry. Gas transmission and distribution pipes and other structures buried (or immersed) in an electrolyte are corroded according to the existing conditions and metallurgical structure. After work, they disrupt an active system and process and lead to unpredictable losses. Also, the worst corrosion for metals buried in the soil relates to places with stray electric currents. The monitored data in this article was measured within two years. These measurements have been carried out in the form of Table (1).

Important points that should be considered during these measurements include the following.

- The coating test has been done with a Holliday device or by installing a current source and insulating the damaged points.
- There must be a protected structure next to the protected gas pipeline (like water pipeline).
- To prevent corrosion in these pipes, two lines must be equal potential.
- The existence of 700 V DC voltage of trains in city next to these pipes is so harmful.
- The existence of 20 kV AC voltage of the subway next to these pipes is harmful.

It is also important to mention that the factors that lead to an increase in the resistance of the circuit have the lowering of the water level, the cross-section of the cable, the end of the anodes, the incomplete coil, the sulfate of the cable and the bus-bars in the bandbox.

AC and DC voltages and currents are measured according to the NACE-177 standard, and when a metal structure undergoes interference corrosion, 9.4 kg (20.7 lb) of steel to iron oxide per ampere of interference current for one year. Becomes. This current leaves the holes and gaps and quickly causes corrosion. If the AC current interferes with the DC, this causes the amount of DC current to decrease, and if it is less than the standard value, corrosion occurs, so that in the part where we have the highest amount of AC current, corrosion occurs faster and faster. If this value is 5–15 V AC, it will only cause corrosion, and if it is more, it will also destroy the insulation, and if the interference points are close to the transformer, it can also damage the transformer.

By measuring the AC voltage at different points of the pipeline and according to the NACE-177 standard, if the measured AC current of the line is higher than 5 V, in terms of corrosion, this standard can damage the pipeline. As a result, with the planned measurement of this voltage, it is possible to find stray current interference and comment on corrosion by checking and monitoring it.

3. Evaluation method

In this part of the paper, the evaluation method of the amount of failure and calculating the RUL of the equipment will be described. It should be noted that by failure rate detection of equipment, it is possible to prioritize its maintenance. Also, by evaluating the useful life of the equipment, it will be possible to maintenance visit schedule, the purchase of reservation equipment, the subsequent investments of the equipment, and the prevention of gas pipeline transmission network interruptions. Therefore, the two methods of failure detection and useful life determination are related to each other.

3.1. Normal probability distribution index

Probability distribution calculation is based on the normal function for apparatus condition parameters so that the historical information of the apparatus is received first. Then, according to the received data, the standard distribution model for each parameter is obtained as equation (1).

$$X \approx N\left(\mu, \sigma^2\right) \tag{1}$$

In the above relationship, μ is mean, and σ is the standard deviation of the data. The above normal distribution is calculated based

Table 1
The procedure of measuring values.

Type of measurement	Duration of the test	Description
Setting (Cathodic protection station) (CPS)	twice a month (once every 15 days)	It includes voltage before adjustment, current before adjustment, injection voltage before adjustment, shutdown voltage before adjustment, transformer voltage and current after adjustment, and silica gel color.
Measurement (test point) (TP)	Every 4 months	If these measurements are in the form of a pond, the cleaning of the pond should also be taken into account, and if it is in the form of a funnel, the measurement is made from the three parts of the funnel, sheath, and lightning arrester.
Measurement of flange insulation (Flange insulation) of TBS stations	Every 6 months	…
Anodic substrate control	Every 6 months	The measurements are related to the anode current
AC line voltage measurement	Every 4 months	This AC voltage is caused by stray currents, which, according to the American NACL standard, can be ignored if it is less than 15 V. This voltage enters the line from one point and exits from the same point, causing pipe corrosion in the same place.

Therefore, according to the things mentioned above, the desired data for failure analysis will be in the form of Table (2). As can be seen, the measured data includes nine items.

on equation (2).

$$f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\quad x=X_1,X_2,\ldots,X_9 \tag{2}$$

The normal distribution is updated in each step by receiving new data, and the new μ and σ values are calculated. Then, in each stage and according to the specified limits of each data, the probability of error, partial failure, total failure, and the health of each part of the apparatus will be obtained.

A normal distribution has been used to model the dispersion of apparatus parameters. Fig. (1) shows an example of a normal distribution for a type of apparatus parameter. For all these parameters, according to Fig. (1), four ranges are defined, which include apparatus health range, partial damage range, total damage range, and apparatus fault range. The reason for using these four ranges is that the middle half of it can be considered entirely healthy. As we get closer to the number of apparatuses failures, the range of partial and total failure becomes more intense. Therefore, anything outside these four ranges will be an apparatus error. If a sample of the value of x falls within the healthy range of the apparatus, it indicates the proper performance of the apparatus in this sample of the value of x. If the sample of x is in the second interval, it shows that some parts are associated with a little delay and may have a minor failure. If sample x is in the third interval, the relevant parts cannot work correctly, and the apparatus has a general breakdown. Finally, if sample x is outside the three limits, the apparatus has far more difficult problems in its correct operation. Of course, these values for each apparatus and parameter have their weight and can be obtained for a device sample once and forever.

In general, the apparatus probabilities in parameter x as $P(x^1) = Pr(l_i < x < u_i^1)$, $P(x^2) = Pr(u_i^1 < x < u_i^2)$ and $P(x^3) = Pr(u_i^2 < x < u_i^3)$ will be; So that l_i is the lower limits of each interval, u_i is the upper limits of each interval, and the upper bounds of 1, 2, and 3 express the state of health of the apparatus, minor deterioration of the apparatus, and major deterioration of the apparatus, respectively. These possibilities can be defined for all other apparatus parameters as well.

According to the distribution of the mentioned probabilities, the proposed model will be built to calculate the probability of health, minor deterioration, major deterioration, and fault in apparatus according to the following assumptions.

1. Only if all the desired parameters are within the health range of the normal distribution, then the apparatus is in a healthy state.
2. If even one parameter is in the minor deterioration range, the apparatus is in a minor deterioration state.
3. If even one parameter is in the major deterioration range, the apparatus is in a major deterioration state.
4. The apparatus is in fault state if even one parameter is in the fault range. The apparatus may work slowly or even stop working properly in this situation.

According to the abovementioned situations, we will calculate the probability of each situation. Here is an example that includes the overall status of the entire apparatus (i.e., parameters X_1 to X_9 according to Table (2)). The probability of fault is calculated according to equation (3):

$$\begin{aligned}P_f(Eq^f) = 1-\sum_{j=1}^{3}P(X_1^j)\cdot\sum_{j=1}^{3}P(X_2^j)\cdot\sum_{j=1}^{3}P(X_3^j)\cdot\sum_{j=1}^{3}P(X_4^j)\cdot\\ \cdot\sum_{j=1}^{3}P(X_5^j)\cdot\sum_{j=1}^{3}P(X_6^j)\cdot\sum_{j=1}^{3}P(X_7^j)\cdot\sum_{j=1}^{3}P(X_8^j)\cdot\sum_{j=1}^{3}P(X_9^j)\end{aligned} \tag{3}$$

The probability of major deterioration of the apparatus is calculated according to equation (4):

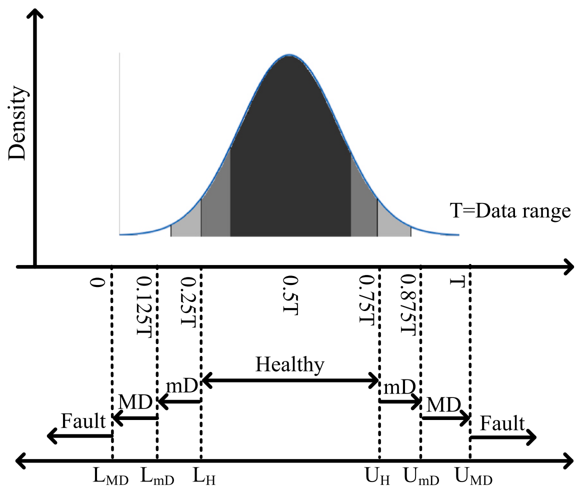

Fig. 1. The probability density of the parameter x.

Table 2
Measured data type and range.

No.	Data Title	Measured data type	Date Range
1	X_1	DC voltage of measurement points in gas networks	The normal value: between 0.85 and 1.2 V
2	X_2	AC voltage of measurement points in gas networks	The maximum acceptable voltage is 15 V
3	X_3	DC voltage of measurement points before adjustment	The normal value: between 0.85 and 1.2 V
4	X_4	Circuit resistance in the GPS stations	The normal value: less than 3 Ω
5	X_5	Transformer output voltage	A normal value: between 0.6 and 50 V
6	X_6	The output current of the transformer	A normal value: between 5 and 20 amps
7	X_7	The amount of anode current	MMO in water: 8 amps and Silicon in water: 4 amps
8	X_8	The voltage of the measurement points in the transformer off state	The normal value: between 0 and 75 V
9	X_9	water column	At the beginning of the good drilling: about 10 m above the Andes

$$P_f(Eq^3) = (1 - P_f(Eq^f)) - - \left[\sum_{j=1}^{2} P(X_1^j) \cdot \sum_{j=1}^{2} P(X_2^j) \cdot \sum_{j=1}^{2} P(X_3^j) \cdot \sum_{j=1}^{2} P(X_4^j) \cdot \cdot \sum_{j=1}^{2} P(X_5^j) \cdot \sum_{j=1}^{2} P(X_6^j) \cdot \sum_{j=1}^{2} P(X_7^j) \cdot \sum_{j=1}^{2} P(X_8^j) \cdot \sum_{j=1}^{2} P(X_9^j) \right] \tag{4}$$

The probability of minor deterioration of the apparatus is calculated according to equation (5):

$$P_f(Eq^2) = (1 - P_f(Eq^f) - P_f(Eq^3)) - - [P(X_1^1) \cdot P(X_2^1) \cdot P(X_3^1) \cdot P(X_4^1) \cdot P(X_5^1) \cdot \cdot P(X_6^1) \cdot P(X_7^1) \cdot P(X_8^1) \cdot P(X_9^1)] \tag{5}$$

And finally, the probability of health of the apparatus is calculated from equation (6):

$$P_f(Eq^1) = P(X_1^1) \cdot P(X_2^1) \cdot P(X_3^1) \cdot P(X_4^1) \cdot \cdot P(X_5^1) \cdot P(X_6^1) \cdot P(X_7^1) \cdot P(X_8^1) \cdot P(X_9^1) \tag{6}$$

So that the superscripts f and 1, 2, and 3, respectively, show the states of fault, health, minor deterioration, and major deterioration of the apparatus. This situation also includes all the first to ninth parameters for the general conditions of the apparatus. Now, if any data is unavailable, it is necessary to make fewer calculations according to the available data.

3.2. Neural network

As we know, neural networks have different types and different applications. Therefore, according to the available data, this article has used the NARX dynamic NN. Because the NARX neural network structure is more accurate in assessment than other existing models, this article attempts to estimate the useful life of three gas systems under cathodic protection and the RUL of the pipes. For this aim, this problem has been solved by using the dynamic NN and considering the data in the form of continuous time data, by time series tool.

NARX dynamic neural network has a hidden layer. The number of neurons in this layer are 10 neurons by fault and trial. The number of signals used in the model has 4 inputs and 5 outputs to have the best matching, which has been selected experimentally. The driving function of hidden layer neurons is considered a sigmoid function, and the output layer's driving function is considered nonlinear.

To train the NN, 70 percent of the data have been applied, and the expert people and experts of the gas department have sampled these data. So, the validation and test data sets are each 15 percent of the primary data. By these data and using the NN-toolbox in MATLAB software, the NN was trained, and the nonlinear f (the nonlinear function of the input and output of the system) was defined and determined to be the values after training.

The impacts and indicators are determined using documents, sources, and experts' opinions according to the approved standards B31G (ANSI/ASME B31G) and B31.8 (ANSI/ASME B31.8). Operating life is calculated by multiplying the actual life by the effect of each effective factor on the life of the relevant apparatus and the coefficients related to the system conditions according to equation (7).

$$t_f = \sum_{i=1}^{n} W_i.T \tag{7}$$

Thus, t_f is the functional age of the apparatus, T is the calendar (actual) life, i is the counter of the number of factors affecting the functional life of the apparatus, and W is the influence of the i-th index.

The remaining operating life for each apparatus, according to the multiplication of "remaining standard life" by the effect of each of the effective indicators on the life for the conditions in which the apparatus is supposed to operate in the future, equation (8) will be counted.

$$Rt_f = \sum_{i=1}^{n} \left(\frac{1}{W_i} \right) \times t_R \tag{8}$$

So that Rt_f is the remaining functional life, and t_R is the remaining standard life. The remaining standard life is obtained from the difference between the standard life and the operating life according to equation (9).

$$t_R = t_s - t_f \tag{9}$$

So that t_S is the standard life.

4. Case study

In this section, the results of the proposed procedure based on normal probability distribution and neural network for calculating the failure rate of the apparatus and estimating its real life will be expressed. The relevant data have been obtained considering the conditions in Table (1). Also, the measured data includes the items stated in Table (2), and the range of each data is also given in this table.

4.1. Results of normal probability distribution

In this part of the article, the results of normal probability distribution will be examined. According to Fig. (2-a) to Fig. (2-f), the fault probabilities of apparatuses 1 to 6 have been compared, respectively. These outputs are obtained using the method described in sections 1-3. According to this figure, the amount of apparatus failure is calculated based on equations (3)–(6) and measured with a scale between 0 and 1. The closer the failure rate of the apparatus to 1, the worse the condition of the apparatus, and the closer the failure rate to 0, the better the condition of the apparatus. According to this figure, it can be seen that in all the apparatuses, the state of failure of the apparatus has been increasing.

4.2. Prediction of cathodic protection failure time by neural network

As mentioned, this article uses the NARX neural network model and Levenberg-Marquardt train model. This study will use the

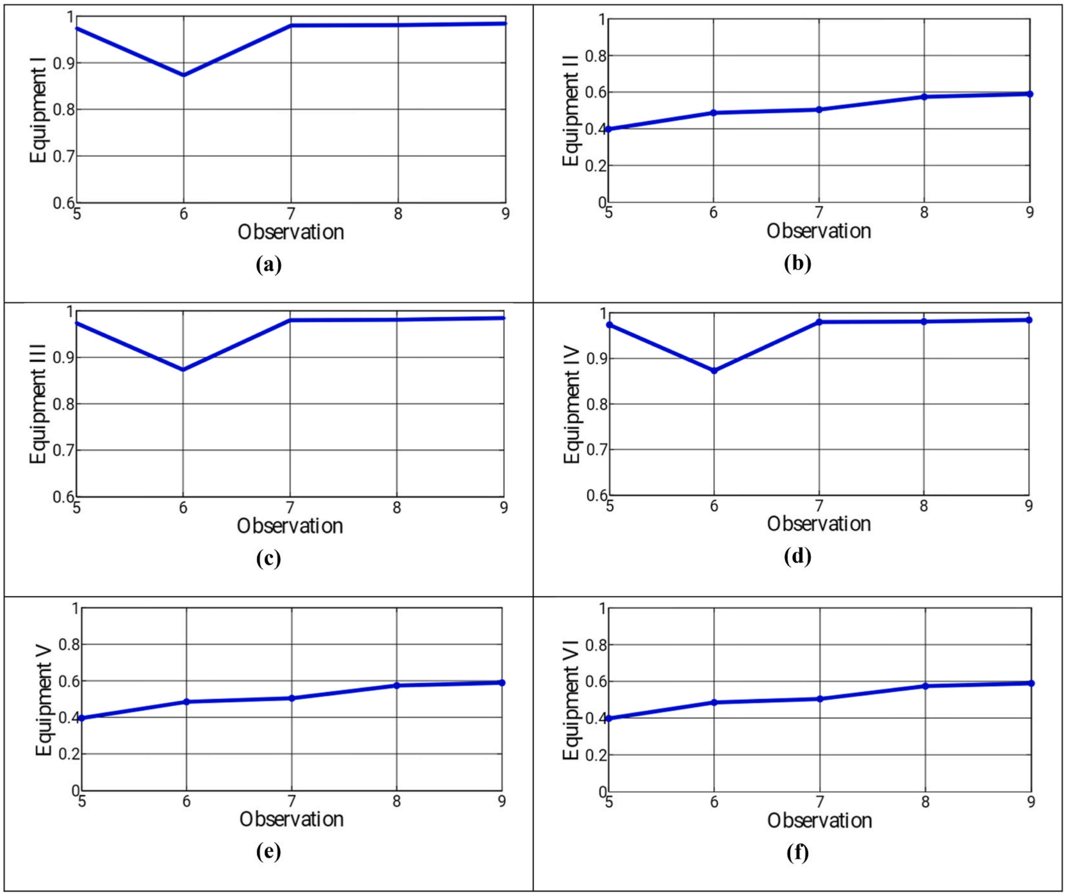

Fig. 2. Failure rate based on normal probability distribution for a) Equipment 1 b) Equipment 2 c) Equipment 3 d) Equipment 4 e) Equipment 5 f) Equipment 6.

multi-layer perceptron NN to estimate the remaining useful life (functional life). After performing the simulation with MATLAB software, details of the output will be checked. According to Fig. (3), the failure time of cathodic protection is predicted. The blue, green and red graphs are the training, validation and test results, respectively. Due to the lack of information time, the training data chart does not match the test data chart well. It is possible to validate the correctness of network training by comparing the validation and test graphs. When the graph of the training result has the biggest difference from the test result, and there are rapid changes in the test result, it is time to pay more attention to the cathodic protection system and survey the pipe's status.

Fig. (4-a) to Fig. (4-c) show data distribution in three situations, respectively: training, validation, and testing. Also, all these three situations are shown in Fig. (4-d). According to this figure, it can be seen that when the cathodic protection is done optimally (in the training state), the dispersion of the data will be less compared to the regression line.

The following points are considered in connection with the obtained results.

- The period is three months, and the value of regular is in cycles. Regulat changes are from zero to one hundred and fifty. If the regulation value is more than 100, the short visit period and the network should be observed and the necessary checks should be done on the pipe.
- Data is available for up to twenty time periods. With the passage of time and recording more data, it will be possible to update the data and improve the output process. In this study, 120 cycles are given, for example, which can be drawn from the beginning to the end until the time of changing the tube. Also, the regulator is the line voltage regulator; therefore, changes in it should be considered.
- The potential difference is measured and recorded before and after station transformer adjustment. If this difference is large, the visit periods are shortened, and the network is monitored.
- The potential difference is measured after adjustment. The circuit should be shorted and monitored if this value is more than 0.5V.
- The potential difference is measured before adjustment. If the value of the potential difference exceeds 0.5 V, the visit period should be short, and the network should be monitored.

The results obtained from this method are shown in Table (3). In this table, the results of the previous and the new methods are compared. In the previous model, the remaining life is calculated using the organizational standard, operational, and calendar life. The basis of calculations is the indicators determined in mathematical formulas. While in the new method, neural networks, which have more accuracy in estimation, have been used. It is worth mentioning that considering that the results of the two methods are close to each other, the correctness of the implementation steps of the new method is confirmed.

5. Discussion

In this article, two methods were used to analyze the failure and lifespan of the equipment and the results of these two methods were compared.

- According to the normal probability distribution of the apparatus, it was observed that the failure rate of the apparatus was mostly increasing, and this procedure was seen differently for different apparatuses.
- This difference was also dependent on the natural and functional life of the apparatus.
- In all these apparatuses, the index trend showed an improvement at first and then went to failure.
- Then, a model of neural networks was proposed. So MATLAB software and multilayer perceptron NN was used to estimate the RUL.
- By using this method, first of all, the equipment remaining useful life in the industries will be estimated, the commercial risks caused by the non-functioning and failure of the devices will be managed.
- Also, this analysis has many positive results, including investigating the causes of destruction of parts and equipment related to industries, economic aspects, maintaining production safety, increasing the burden of preventing financial and life losses, applying coefficients of high reliability, experiencing and analyzing unwanted stresses, unforeseen operational and environmental factors, use outside the design range (such as load of high cyclic and temperature), reducing the properties of materials in service in sensitive areas, and experiencing and analyzing damage identification.

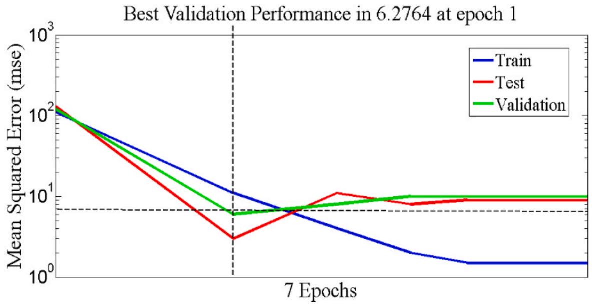

Fig. 3. Neural network train, test, and validation according to epoch.

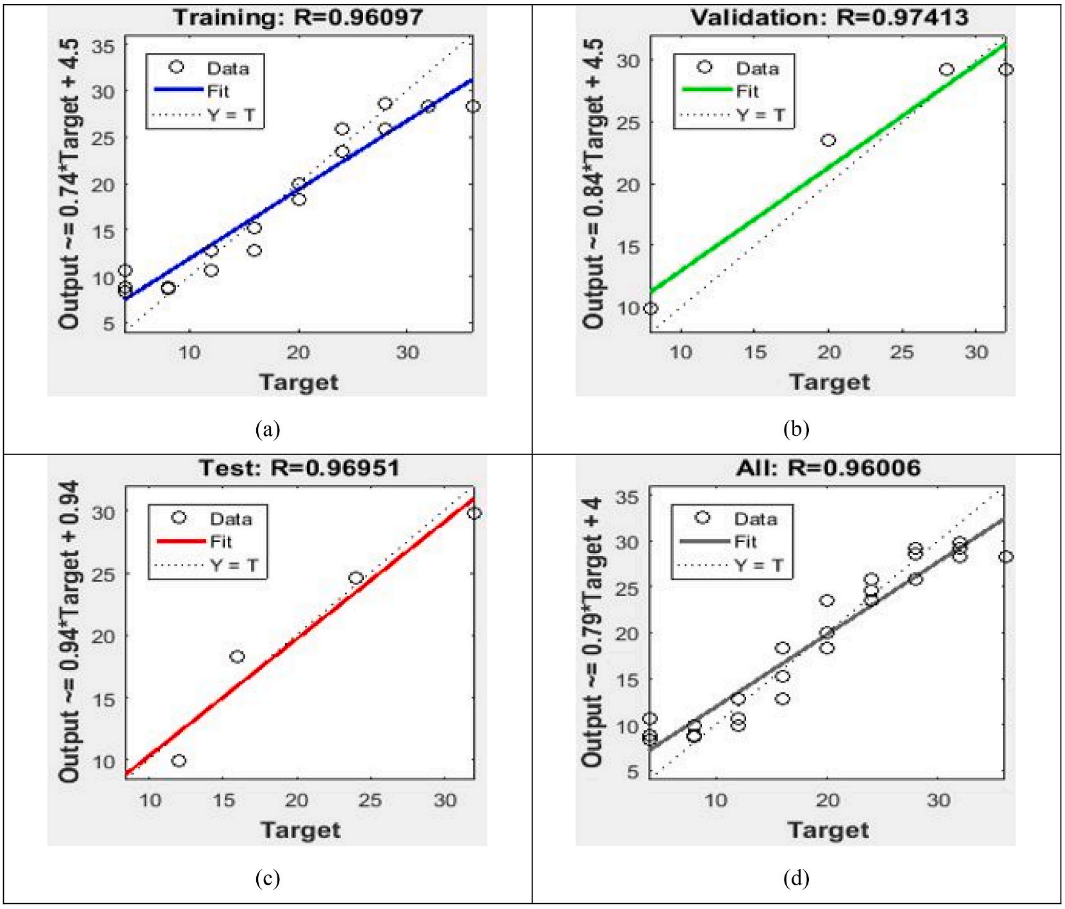

Fig. 4. Scattering of data relative to the regression line (a) training (b) validation (c) test (d) all.

Table 3
RUL of gas pipes under cathodic protection.

Method	Nazarabad		Eshthard		Karaj	
	years	months	years	months	years	months
Previous	17	2	19	9	17	7
New	16	1	18	3	16	3

6. Conclusion

This article investigated the stray currents effects on the failure rate and estimation of the RUL of gas pipes under cathodic protection by using the apparatus failure model using normal probability distribution and neural networks. In addition to stating the amount of failure and the RUL, this article also plans and announces the time of the next visit according to the pipes' condition and failure level. According to the comparison of the results of the two processes proposed in this article, the accuracy of the input data as well as the models proposed to measure the failure and the RUL of the equipment have been validated. The results of this article are sufficient and can be revised with new data.

Data availability statement

The data are available on request.

CRediT authorship contribution statement

H. Noroznia: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **M. Gandomkar:** Supervision. **J. Nikoukar:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

While appreciating Alborz Gas Company, this article has been formed with the help of this organizations.

References

[1] X. Zhang, et al., Remaining useful life estimation using CNN-XGB with extended time window, IEEE Access 7 (2019) 154386–154397, https://doi.org/10.1109/ACCESS.2019.2942991.

[2] A. Das, S. Hussain, F. Yang, M.S. Habibullah, A. Kumar, Deep recurrent architecture with attention for remaining useful life estimation, TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON) (2019) 2093–2098, https://doi.org/10.1109/TENCON.2019.8929267.

[3] J. Zhang, S. Wang, L. Chen, G. Guo, R. Chen, A. Vanasse, Time-dependent survival neural network for remaining useful life prediction, in: Q. Yang, Z.H. Zhou, Z. Gong, M.L. Zhang, S.J. Huang (Eds.), Advances in Knowledge Discovery and Data Mining. PAKDD, Lecture Notes in Computer Science, vol. 11439, 2019, https://doi.org/10.1007/978-3-030-16148-4_34, 2019.

[4] M. Kordestani, M.F. Samadi, M. Saif, A new hybrid fault prognosis method for MFS systems based on distributed neural networks and recursive bayesian algorithm, IEEE Syst. J. 14 (4) (2020) 5407–5416, https://doi.org/10.1109/JSYST.2020.2986162.

[5] R. Mezzi, S. Morando, N.Y. Steiner, M.C. Péra, D. Hissel, L. Larger, Multi-reservoir echo state network for proton exchange membrane fuel cell remaining useful life prediction, IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society (2018) 1872–1877, https://doi.org/10.1109/IECON.2018.8591345.

[6] Rob B. Polder, Greet Leegwater, Daniël Worm, Wim Courage, Service life and life cycle cost modelling of cathodic protection systems for concrete structures, Cement Concr. Compos. 47 (2014) 69–74.

[7] K. Tang, Stray alternating current (AC) induced corrosion of steel fibre reinforced concrete, Corrosion Sci. 152 (2019) 153–171.

[8] Y. Wang, Z. Li, H. He, W. Lu, Y. Xu, Y. Mao, An Inductance Forcedly Absorbing Current Circuit to Reduce the Stray Current in the DC Power Supply System, 2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2019, pp. 2303–2308, https://doi.org/10.1109/ICIEA.2019.8834227.

[9] K.V. Avdeeva, Experimental research of stray currents influence of DC railway transport to grounding grid, 2019 Dynamics of Systems, Mechanisms and Machines (Dynamics) (2019) 1–4, https://doi.org/10.1109/Dynamics47113.2019.8944418.

[10] L. Zhao, J. Li, M. Liu, Simulation and analysis of metro stray current based on multi-locomotives condition, 2016 35th Chinese Control Conference (CCC) (2016) 9252–9258, https://doi.org/10.1109/ChiCC.2016.7554829.

[11] W. Liu, T. Li, J. Zheng, W. Pan, Y. Yin, Evaluation of the effect of stray current collection system in DC-electrified railway system, IEEE Trans. Veh. Technol. 70 (7) (July 2021) 6542–6553, https://doi.org/10.1109/TVT.2021.3084340.

[12] P. Yuan, W. Mao, H. Ye, Y. Liu, Model Construction and Analysis of Transformer DC Magnetic Bias Induced by Rail Transit Stray Current, 2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2), 2019, pp. 1710–1713, https://doi.org/10.1109/EI247390.2019.9061729.

[13] Y. Lin, K. Li, M. Su, Y. Meng, Research on stray current distribution of metro based on numerical simulation, in: 2018 IEEE International Symposium on Electromagnetic Compatibility and 2018 IEEE Asia-Pacific Symposium on Electromagnetic Compatibility, EMC/APEMC), 2018, pp. 36–40, https://doi.org/10.1109/ISEMC.2018.8393734.

[14] H. Noroznia, M. Gandomkar, J. Nikoukar, A. Aranizadeh, M. Mirmozaffari, A novel pipeline age evaluation: considering overall condition index and neural network based on measured data, Machine Learn. Knowledge Extract. 5 (2023) 252–268.

[15] A. Aranizadeh, S.M. Shahrtash, A. Gholami, Prioritizing CBs maintenance and identifying mandatory maintenance at higher priorities, Int. Trans. Elect. Energy Sys. (2022).

[16] A. Aranizadeh, S.M. Shahrtash, A. Gholami, "Comprehensive Condition Assessment of Circuit Breakers in a Power Network for Maintenance Scheduling", IET Generation, Transmission & Distribution, 2023.

[17] A.D. Woldeyohan, M.A.A. Majid, Effect of age of pipes on performance of natural gas transmission pipeline network system, J. Appl. Sci. 11 (No. 9) (2011), https://doi.org/10.3923/jas.2011.1612.1617.

[18] A.M.A. Sabaeei, H. Alsussian, S.J. Abdulkadir, A. Jagadeesh, Prediction of oil and gas pipeline failures through machine learning approaches: a systematic review, Energy Rep. 10 (2023) 1313–1338.

[19] J.C. Velazquez, N.E. Gonzalez-Arevalo, M. Diaz-Cruz, A. Cervantes-Tobon, H. Herrera-Hernandez, E. Hernandez-Sanchez, Failure pressure estimation for an aged and corroded oil and gas pipeline: a finite element study, J. Nat. Gas Sci. Eng. 101 (2022).

[20] H. Sakai, M. Satake, Y. Arai, S. Takizawa, Report cards for aging and maintenance assessment of water-supply infrastructure, J. Water Supply Res. Technol. Aqua 69 (No. 4) (2020) 355–364.

[21] X. Ding, S. Liu, X. Shi, J. Chu, X. Guo, Health evaluation of urban water supply pipe network using the bayesian method based on triangular fuzzy number optimization, IOP Conf. Ser. Earth Environ. Sci. 467 (2020), https://doi.org/10.1088/1755-1315/467/1/012124. EEEP.

[22] M. Mirmozaffari, M. Yazdani, A. Boskabadi, H. Ahady Dolatsara, K. Kabirifar, N. Amiri Golilarz, A novel machine learning approach combined with optimization models for eco-efficiency evaluation, Appl. Sci. 28 (15) (2020) 5210, 10.

[23] M. Mirmozaffari, E. Shadkam, S.M. Khalili, K. Kabirifar, R. Yazdani, T. Asgari Gashteroodkhani, A novel artificial intelligent approach: comparison of machine learning tools and algorithms based on optimization DEA Malmquist productivity index for eco-efficiency evaluation, Int. J. Energy Sect. Manag. 15 (3) (2021) 523–550.

[24] M. Mirmozaffari, E. Shadkam, S.M. Khalili, M. Yazdani, Developing a novel integrated generalised data envelopment analysis (DEA) to evaluate hospitals providing stroke care services, Bioengineering 8 (12) (2021) 207.

[25] M. Mirmozaffari, R. Yazdani, E. Shadkam, S.M. Khalili, L.S. Tavassoli, A. Boskabadi, A novel hybrid parametric and non-parametric optimisation model for average technical efficiency assessment in public hospitals during and post-COVID-19 pandemic, Bioengineering 9 (1) (2021) 7.

[26] M. Mirmozaffari, R. Yazdani, E. Shadkam, L.S. Tavassoli, R. Massah, VCS and CVS: new combined parametric and non-parametric operation research models, Sustainable operations and computers" 2 (2021) 36–56.

[27] M. Mirmozaffari, N. Kamal, The application of data envelopment analysis to emergency departments and management of emergency conditions: a narrative review, Healthcare 11 (No. 18) (2023) 2541. MDPI.