Prediction of Equipment Failure Rates in Power Distribution Networks based on Machine-learning Method

Rita Jamasheva[1], Noor Hanoon Haroon[2], Ahmed Read Al-Tameemi[3], Israa Alhani[4], Ali Murad Khudadad [5], Bahira Abdulrazzaq Mohammed[6], Ali H. O. Al Mansor[7], Mustafa Asaad Hussein[8]
1- Department of Automation and Robotics, Almaty Technological University, Almaty, Kazakhstan.
Email: rita_2206@mail.ru (Corresponding author)
2- Information and Communication Technology Research Group, Scientific Research Center, Al-Ayen University, Thi-Qar, Iraq.
Email: noor.hh@gmail.com
3- Department of Medical Laboratory Technics, Al-Nisour University College, Baghdad, Iraq.
Email: ahmed.itm@nuc.edu.iq
4- Department of Medical Laboratory Technics, Mazaya University College, Dhi Qar, Iraq.
Email: israa.alhani@yahoo.com
5- Department of Medical Laboratory Technics, Al-Esraa University College, Baghdad, Iraq
Email: a.m.khudadad@esraa.edu.iq
6- Department of Medical Engineering, Al-Hadi University College, Baghdad 10011 Iraq.
Email: bahira.abdmoh@gmail.com
7- Department of Medical Laboratory Technics, Al-Zahrawi University College, Karbala, Iraq.
Email: alimansor@yahoo.com
8- National University of Science and Technology, Dhi Qar, Iraq.
Email: mustafa.assad1986@gmail.com

Received: 27 May 2023 Revised: 11 July 2023 Accepted: 4 August 2023

ABSTRACT:
This paper explores the application of a machine learning approach to predict equipment failure rates in power distribution networks, motivated by the significant impact of power outages on citizens' daily lives and the economy. In this research, data on equipment failure rates and maintenance records were collected from power distribution networks in Baghdad, Iraq. The collected data underwent preprocessing, and features were extracted to train Adaptive Neuro-Fuzzy Inference System (ANFIS) and Periodic Autoregressive Moving Average (PARMA) time series models. To initiate the project, information regarding blackouts that occurred between January 2018 and December 2021 was retrieved from the database. The RMSE index results for the PARMA time series and ANFIS model are 3.518 and 2.264, respectively, demonstrating the superior performance of the ANFIS model in predicting equipment failure rates and its potential for future predictions. This study highlights the ANFIS model's capacity to anticipate equipment failure rates, potentially enhancing maintenance efficiency and reducing power outages in Baghdad. The error mean square was employed to evaluate the proposed models' error rate.

KEYWORDS: Failure Rates; Machine Learning; Adaptive Neuro-Fuzzy Inference System model; Periodic Autoregressive Moving Average.

1. INTRODUCTION

Electricity distribution companies are required to provide reliable and stable electricity with minimum blackouts and standard voltage [1], [2]. However, power outages do occur due to various reasons, such as equipment failure, natural disasters, and human errors [3], [4]. These outages can cause inconvenience to customers and lead to economic losses. Therefore, it is crucial for electricity distribution companies to maintain and repair their equipment regularly and predict their failure rates accurately to minimize the outages [5]–[7].

Maintenance and repair planning is an essential task for electricity distribution companies to ensure that their equipment is functioning correctly and reliably [8], [9]. The identification of weak points in the distribution network and the discovery of unusual events can help companies develop better maintenance plans and prevent potential equipment failures [10], [11]. In this regard, tracing the outages that occurred in the network can provide valuable insights into the causes of

Paper type: Research paper
DOI: 10.30486/mjee.2023.1994835.1238
How to cite this paper: R. Jamasheva, N. Hanoon Haroon, A. R. Al-Tameemi, I. Alhani, A. Murad Khudadad, B. Abdulrazzaq Mohammed, A. H. O. Al Mansor, and M. A. Hussein8, **"Prediction of Equipment Failure Rates in Power Distribution Networks based on Machine-learning Method"**, *Majlesi Journal of Electrical Engineering*, Vol. 17, No. 3, pp. 77-82, 2023.

equipment failures and guide maintenance and repair planning [12].

Several studies have been conducted on equipment failure prediction in the power system [13]–[16]. Time series models, such as ARIMA and exponential smoothing, are widely used to predict equipment failure rates based on historical data [17]. For example, Jurgensen et al. [18] developed a time series model to predict the failure rate of power transformers. They used ARIMA and exponential smoothing methods and found that ARIMA performed better in predicting transformer failure rates.

In recent years, machine learning approaches, such as neural networks and fuzzy logic, have gained popularity in equipment failure prediction [19]. These models can handle nonlinear relationships between input and output variables and can capture complex patterns in the data. For instance, Lim et al. [20] used a neural network model to predict the remaining useful life of power transformers. They used oil samples from transformers as input variables and found that the neural network model outperformed the traditional time series models.

In another study, Idreeset al. [21] used a fuzzy logic-based model to predict the failure rate of distribution transformers. They used input variables such as ambient temperature, relative humidity, and load demand to train the model and found that the fuzzy logic model performed better than the time series models in predicting transformer failure rates.

However, these studies mainly focused on predicting the failure rate of power transformers and did not consider other equipment types. In addition, most studies used environmental variables such as temperature and humidity as input variables and did not consider other factors such as rainfall and wind speed.

In this study, we aim to predict the failure rate of a specific equipment type in one of the areas covered by the Baghdad Electricity Distribution Company, Iraq, using the Periodic Autoregressive Moving Average (PARMA) time series approach and the ANFIS model. We selected average air temperature, average rainfall, and average wind speed as input variables to the neural network. We also evaluated the performance of the proposed models using the mean square error. Our results show that the ANFIS model outperforms the time series models in predicting the failure rate of the equipment in question and can be used to predict future periods.

2. MATERIALS AND METHODS

Adaptive Neuro-Fuzzy Inference System (ANFIS) and Periodic Autoregressive Moving Average (PARMA) time series models were trained using features extracted from preprocessed collected data. The formulation of these models is evaluated in the following section.

2.1. Time Series Model

PARMA models are periodic or self-correlated moving average models [22]. They are a special case of ARMA models, with the only difference being that the number of periods is multiplied by the model coefficients for periodic models [23]. To model the monthly values, for instance, one model must be run for each month. Equation 1 represents the PARMA(p,q) model.

$$Y_{v,\tau} = \sum_{\mathrm{i}=1}^{\mathrm{p}} \varphi_{\mathrm{i},\tau} Y_{\mathrm{v},\tau-1} + \varepsilon_{\mathrm{v},\tau} - \sum_{\mathrm{i}=1}^{\mathrm{q}} \theta_{\mathrm{j},\tau} \varepsilon_{\mathrm{v},\tau-1} \qquad (1)$$

Where $Y_{v,\tau}$ is equal to the data values of the investigated case or an estimate in year V and season (period) τ which is normal and standard for each period. $\varepsilon_{\mathrm{v},\tau}$ is equal to the normal and standard time series and has φ and θ are the parameters of the model.

2.2. ANFIS Model

ANFIS is a type of machine learning algorithm that combines the adaptive capabilities of neural networks with the interpretability of fuzzy logic [24]. The ANFIS model consists of a set of fuzzy rules and a neural network structure that learns the parameters of these rules based on the input-output data.

The ANFIS model formulation can be divided into the following steps [25]:

- Fuzzification: In this step, the input variables are transformed into fuzzy sets using membership functions. The membership functions represent the degree of membership of each input variable to each fuzzy set. The fuzzy sets can be defined using different shapes such as triangular or Gaussian.
- Rule generation: The ANFIS model generates a set of fuzzy rules based on the input-output data. Each rule consists of a set of antecedents and a consequent. The antecedents are the fuzzy sets of the input variables, and the consequent is the output variable. The rules are generated using a clustering algorithm such as the subtractive clustering method.
- Rule aggregation: In this step, the fuzzy rules are combined to form a single fuzzy set that represents the output variable. The aggregation method can be a weighted average or a weighted product.
- Defuzzification: In the final step, the fuzzy output is transformed into a crisp output value. This is done by using a defuzzification method such as the centroid or the weighted average method.

3. MODELING AND ANALYSIS OF DATA

In this part, the failure rate of the equipment examined by the Baghdad Electricity Distribution Company has been estimated using the ANFIS model. There is a power failure. Owing to having the most electricity consumers in the nation, this corporation has a significant and weighty responsibility to provide dependable electricity. In order to initiate the project, the database was queried for information regarding blackouts that occurred between the beginning of January 2018 and the end of December 2021. According to this company's policies, the data of a critical area have been studied and the names of the equipment and the area have been omitted; these areas are indicated by R.

3.1. Data Analysis with Time Series Models in the Study Area

In this subsection, we will analyze and model the A region's time series data. With the aid of a time series graph, autocorrelation, and partial autocorrelation functions, the first step entails determining the stability of the mean and variance over time. Figs. 2 to 5 represent the first step of modeling implementation.

As seen in Fig. 2, the unreliability of the time series is illustrated by the downward trend over time. For a more precise investigation, however, the unreliability of the variance is analyzed first.

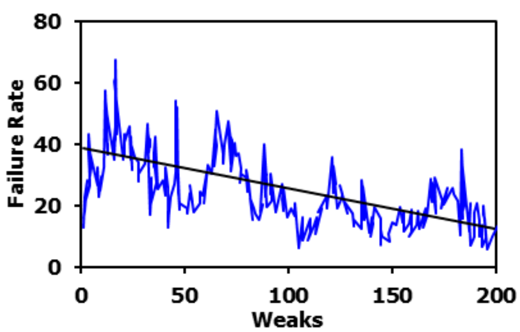

Fig. 1. The weekly rate of equipment failure.

The proposed conversion parameter for this test is depicted in Fig. 2, which indicates that the variance values do not remain constant over time. Therefore, the second root of the data must be computed to establish their reliability. After that, the autocorrelation function and partial autocorrelation of the data are drawn to determine the means' dependability (Fig.s 3). The results demonstrate that the autocorrelation function is gradually decreasing; consequently, it is evident that the mean is also unstable. To eliminate the instability in the mean, we must differentiate our data at least once; that is, we must subtract each data from the subsequent data and redraw its autocorrelation and partial autocorrelation function. If the autocorrelation function decreases rapidly, differentiation is terminated.

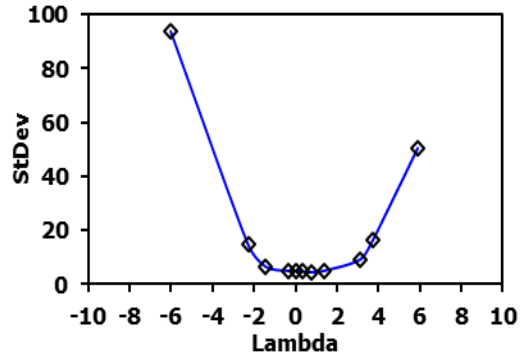

Fig. 2. The proposed transformation parameter for the stability of the variance of the area data.

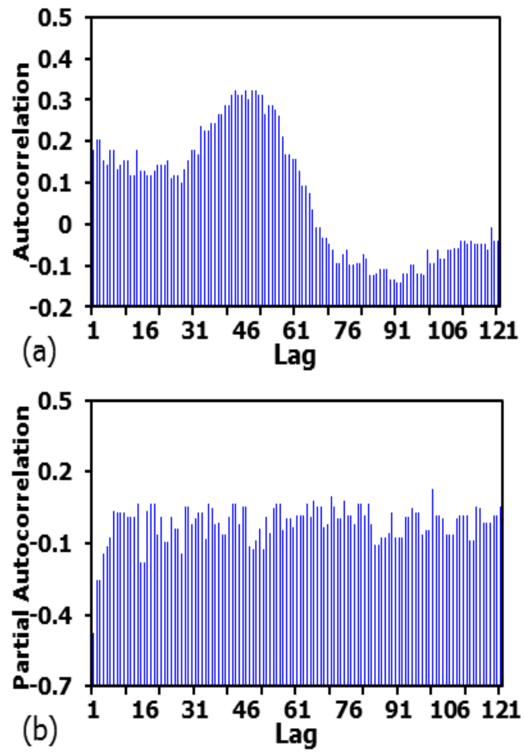

Fig. 3. Functions a) autocorrelation b) partial autocorrelation, after applying the transformation parameter on the area data.

Using normal and standardized data, the PARMA (1,0) model among other models was considered as the best model.

The Chi-Square test is utilized to determine the model's validity. This test is used to assess the independence of the model's remaining values. The outcome of this test is also reported alongside the P-Value. As the P-Values are greater than 0.05, the null hypothesis is accepted and the remaining residuals are independent, indicating that the fitted model is adequate.

Table 1. Chi-square test results for the fitted model of the study area

Lag	Chi-Square	DF	P-Value
12	2.5	7	0.87
24	18.8	19	0.52
36	25.1	31	0.69
48	39.7	48	0.71
60	48.5	51	0.62

3.2. Data analysis with ANFIS model

ANFIS is a type of machine learning model that combines fuzzy logic and neural network techniques. ANFIS was first introduced by Jang in 1993 and has been widely used in various fields, including engineering, finance, and medicine.

ANFIS works by using a set of rules based on fuzzy logic to make inferences from input data. Fuzzy logic is a mathematical system that deals with degrees of truth rather than binary true/false statements. It allows for more nuanced interpretations of data, which can be useful in complex systems.

The ANFIS model consists of five layers, including an input layer, a fuzzy layer, a normalization layer, a rule layer, and an output layer. The input layer receives data from the dataset, and the fuzzy layer applies fuzzy logic to the input data. The average air temperature, average rainfall, and average wind speed were selected as input variables, also because only one output is considered, one neuron is considered in the output layer. The normalization layer then normalizes the output (the failure rate) of the fuzzy layer, and the rule layer uses the normalized output to generate a set of rules. Finally, the output layer uses the rules to make predictions.

The ANFIS model can be trained using a backpropagation algorithm, which adjusts the parameters of the model to minimize the error between the predicted output and the actual output. The backpropagation algorithm uses gradient descent to find the optimal set of parameters. In the subsequent phase, model testing, the model created using experimental data is evaluated so that the inputs and weights used in the previous phase are maintained. Assuming that the previous outputs were merely information about the problem, we now calculate the output using the inputs and weights we have and compare the result to the previous outputs. Our level of success in creating the model is determined by the difference between the predicted and actual outputs. In fact, the effectiveness of models is also determined by the square root of the mean error.

One of the main advantages of ANFIS is its ability to handle complex and nonlinear systems. It can also be used for both regression and classification tasks. However, ANFIS requires a large amount of data for training and can be computationally expensive.

4. EQUIPMENT FAILURE RATE PREDICTION

In this section, the aim is to predict the failure rate of the considered equipment in the R region with the help of the PARMA time series model and ANFIS technique.

On the basis of the preliminary results and the time series and ANFIS models developed for the study area, the failure rate for January and February 2022 has been predicted; the results are presented in Table 5. As it is clear, the time series models fitted in the studied area of the Baghdad Electricity Distribution Company have performed well in predicting the failure rate of the equipment.

Table 2. Chi-square test results for the fitted model of the study area

RMSE	MSE	Predicted	Real
PARMA			
3.518	12.375	12	11
		11	9
		15	12
		10	10
		13	7
		18	12
		11	13
		17	14
ANFIS			
2.264	5.125	10	11
		9	9
		15	12
		8	10
		10	7
		16	12
		14	13
		15	14

The results indicate the better performance of time series models in comparison with the ANFIS model with regard to the MSE and RMSE index, meaning that the ANFIS model has a smaller mean square error index. Of course, this advantage may be due to the non-use of other time series models and the general non-optimality of the selection conditions.

5. CONCLUSION

Overall, the study demonstrates the potential of machine learning approaches to improve the efficiency of maintenance activities in power distribution networks. The findings of this study could have practical implications for power distribution companies in Baghdad and other regions facing similar challenges. By utilizing machine learning models, power distribution companies could anticipate equipment failure rates and take proactive measures to prevent power outages, thus enhancing the reliability and sustainability of the power supply.

REFERENCES

[1] H. Ghaedi, S. R. Kamel Tabbakh, and R. Ghaemi, "**Improving Electricity Theft Detection using Combination of Improved Crow Search Algorithm and Support Vector Machine**," *Majlesi Journal of Electrical Engineering*, vol. 15, no. 4, pp. 63–75, Dec. 2021, doi: 10.52547/mjee.15.4.63.

[2] D. L. Gerber, O. A. Ghatpande, M. Nazir, W. G. B. Heredia, W. Feng, and R. E. Brown, "**Energy and power quality measurement for electrical distribution in AC and DC microgrid buildings**," *Applied Energy*, vol. 308, p. 118308, Feb. 2022, doi: 10.1016/j.apenergy.2021.118308.

[3] A. Rafieyan, H. Sarvari, and D. W. M. Chan, "**Identifying and Evaluating the Essential Factors Affecting the Incidence of Site Accidents Caused by Human Errors in Industrial Parks Construction Projects**," *International Journal of Environmental Research and Public Health*, vol. 19, no. 16, p. 10209, Aug. 2022, doi: 10.3390/ijerph191610209.

[4] G. Elidolu, E. Akyuz, O. Arslan, and Y. Arslanoğlu, "**Quantitative failure analysis for static electricity-related explosion and fire accidents on tanker vessels under fuzzy bow-tie CREAM approach**," *Engineering Failure Analysis*, vol. 131, p. 105917, Jan. 2022, doi: 10.1016/j.engfailanal.2021.105917.

[5] C. G. Marcelino, V. Torres, L. Carvalho, M. Matos, and V. Miranda, "**Multi-objective identification of critical distribution network assets in large interruption datasets**," *International Journal of Electrical Power & Energy Systems*, vol. 137, p. 107747, May 2022, doi: 10.1016/j.ijepes.2021.107747.

[6] L. Fayazi Rad, R. Dashti, and R. Ahmadi, "**Dynamic Evaluation of Electricity Distribution Lines to Speed Up the Fault Repairs**," *Electric Power Components and Systems*, vol. 50, no. 11–12, pp. 649–661, Jul. 2022, doi: 10.1080/15325008.2022.2139434.

[7] P. P. Mawle, G. A. Dhomane, and P. G. Burade, "**Application of artificial intelligence in early fault detection of transmission line-a case study in India**," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 6, p. 5707, Dec. 2022, doi: 10.11591/ijece.v12i6.pp5707-5716.

[8] Y. N. Barykina and A. G. Chernykh, "**Ensuring of reliability and security of energy systems in the Russian Federation**," *IOP Conference Series: Earth and Environmental Science*, vol. 990, no. 1, p. 012001, Feb. 2022, doi: 10.1088/1755-1315/990/1/012001.

[9] X. Liang, M. A. Saaklayen, M. A. Igder, S. M. R. H. Shawon, S. O. Faried, and M. Janbakhsh, "**Planning and Service Restoration Through Microgrid Formation and Soft Open Points for Distribution Network Modernization: A Review**," *IEEE Transactions on Industry Applications*, vol. 58, no. 2, pp. 1843–1857, Mar. 2022, doi: 10.1109/TIA.2022.3146103.

[10] L. I. Alvarez Quiñones, C. A. Lozano-Moncada, and D. A. Bravo Montenegro, "**Machine learning for predictive maintenance scheduling of distribution transformers**," *Journal of Quality in Maintenance Engineering*, vol. 29, no. 1, pp. 188–202, Mar. 2023, doi: 10.1108/JQME-06-2021-0052.

[11] L. K. Mortensen, H. R. Shaker, and C. T. Veje, "**Relative fault vulnerability prediction for energy distribution networks**," *Applied Energy*, vol. 322, p. 119449, Sep. 2022, doi: 10.1016/j.apenergy.2022.119449.

[12] A. Marot *et al.*, "**Perspectives on Future Power System Control Centers for Energy Transition**," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 2, pp. 328–344, 2022, doi: 10.35833/MPCE.2021.000673.

[13] M. O. Mahgoub, S. M. Mazhari, C. Y. Chung, and S. O. Faried, "**A Prediction Interval Based Cascading Failure Prediction Model for Power Systems**," in *2020 IEEE Electric Power and Energy Conference (EPEC),* Nov. 2020, pp. 1–6. doi: 10.1109/EPEC48502.2020.9320125.

[14] M. Dong, "**Combining Unsupervised and Supervised Learning for Asset Class Failure Prediction in Power Systems**," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 5033–5043, Nov. 2019, doi: 10.1109/TPWRS.2019.2920915.

[15] N. N. Ravi, S. Mohd Drus, and P. S. Krishnan, "**Data Mining Techniques for Transformer Failure Prediction Model: A Systematic Literature Review**," in *2019 IEEE 9th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, Apr. 2019, pp. 305–309. doi: 10.1109/ISCAIE.2019.8743987.

[16] J. Fang, H. Wang, F. Yang, K. Yin, X. Lin, and M. Zhang, "**A failure prediction method of power distribution network based on PSO and XGBoost**," *Australian Journal of Electrical and Electronics Engineering*, vol. 19, no. 4, pp. 371–378, Oct. 2022, doi: 10.1080/1448837X.2022.2072447.

[17] Y. Wang, P. Guo, N. Ma, and G. Liu, "**Robust Wavelet Transform Neural-Network-Based Short-Term Load Forecasting for Power Distribution Networks**," *Sustainability*, vol. 15, no. 1, p. 296, Dec. 2022, doi: 10.3390/su15010296.

[18] J. H. Jurgensen, L. Nordstrom, and P. Hilber, "**Estimation of Individual Failure Rates for Power System Components Based on Risk Functions**," *IEEE Transactions on Power Delivery*, vol. 34, no. 4, pp. 1599–1607, Aug. 2019, doi: 10.1109/TPWRD.2019.2913777.

[19] Y. Lyu and Y. Jiang, "**Examination on avionics system fault prediction technology based on ashy neural network and fuzzy recognition**," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 4, pp. 3939–3947, Apr. 2020, doi: 10.3233/JIFS-179619.

[20] B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister, "**Temporal Fusion Transformers for interpretable multi-horizon time series forecasting**," *International Journal of Forecasting*, vol. 37, no. 4, pp. 1748–1764, Oct. 2021, doi: 10.1016/j.ijforecast.2021.03.012.

[21] M. Idrees *et al.*, "**Fuzzy Logic Based Calculation and Analysis of Health Index for Power Transformer**

Installed in Grid Stations," in *2019 International Symposium on Recent Advances in Electrical Engineering (RAEE)*, Aug. 2019, pp. 1–6. doi: 10.1109/RAEE.2019.8887016.

[22] M. Shimi, M. Najjarchi, K. Khalili, E. Hezavei, and S. M. Mirhoseyni, "**Investigation of the accuracy of linear and nonlinear time series models in modeling and forecasting of pan evaporation in IRAN**," *Arabian Journal of Geosciences*, vol. 13, no. 2, p. 59, Jan. 2020, doi: 10.1007/s12517-019-5031-7.

[23] Q. Shao and R. Lund, "**Computation and Characterization of Autocorrelations and Partial Autocorrelations in Periodic ARMA Models**," *Journal of Time Series Analysis*, vol. 25, no. 3, pp. 359–372, May 2004, doi: 10.1111/j.1467-9892.2004.00356.x.

[24] U. Çaydaş, A. Hasçalık, and S. Ekici, "**An adaptive neuro-fuzzy inference system (ANFIS) model for wire-EDM**," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6135–6139, Apr. 2009, doi: 10.1016/j.eswa.2008.07.019.

[25] V. V. Vasilevskij and M. O. Poliakov, "**Reproducing of the humidity curve of power transformers oil using adaptive neuro-fuzzy systems**," *Electrical Engineering & Electromechanics*, no. 1, pp. 10–14, Feb. 2021, doi: 10.20998/2074-272X.2021.1.02.