Thermal Science and Engineering (2023) Volume 6 Issue 2
doi: 10.24294/tse.v6i2.2087

ORIGINAL RESEARCH ARTICLE

Proposing a new optimized forecasting model for the failure rate of power distribution network thermal equipment for educational centers

Aidin Shaghaghi[1], Reza Omidifar[2], Rahim Zahedi[3,*], Ali Asghar Pourezzat[4], Mansour Keshavarzzadeh[5]

[1] *Department of Energy Systems Engineering, Iran University of Science and Technology, Tehran 11336, Iran*

[2] *Faculty of Governance, University of Tehran, Tehran 11366, Iran*

[3] *Department of Renewable Energy and Environmental Engineering, University of Tehran, Tehran 11366, Iran*

[4] *Department of Public Administration, Faculty of Management, University of Tehran, Tehran 11336, Iran*

[5] *Department of Mechanical Engineering Science, University of Johannesburg, Johannesburg 1560, South Africa*

*** Corresponding author:** Rahim Zahedi, rahimzahedi@ut.ac.ir

ABSTRACT

To gain a deep understanding of maintenance and repair planning, investigate the weak points of the distribution network, and discover unusual events, it is necessary to trace the shutdowns that occurred in the network. Many incidents happened due to the failure of thermal equipment in schools. On the other hand, the most important task of electricity distribution companies is to provide reliable and stable electricity, which minimal blackouts and standard voltage should accompany. This research uses seasonal time series and artificial neural network approaches to provide models to predict the failure rate of one of the equipment used in two areas covered by the greater Tehran electricity distribution company. These data were extracted weekly from April 2019 to March 2021 from the ENOX incident registration software. For this purpose, after pre-processing the data, the appropriate final model was presented with the help of Minitab and MATLAB software. Also, average air temperature, rainfall, and wind speed were selected as input variables for the neural network. The mean square error has been used to evaluate the proposed models' error rate. The results show that the time series models performed better than the multi-layer perceptron neural network in predicting the failure rate of the target equipment and can be used to predict future periods.

Keywords: failure rate prediction; seasonal time series models; artificial neural network; electricity distribution company

ARTICLE INFO

Received: 29 April 2023
Accepted: 19 July 2023
Available online: 19 October 2023

1. Introduction

To optimize operation and planning in distribution networks, including load forecasting, event management, rearrangement, allocation of losses, etc., state estimation is used in distribution management systems, using advanced monitoring, control, and technology systems[1]. Self-healing in intelligent distribution networks has been an important factor in guiding these networks using distribution mode estimation[2]. By estimating the distribution mode, the electrical parameters of the network can be calculated and estimated in a short period, thus providing the necessary information for planning, operation, and other controls.

Tehran Electricity Distribution Company is one of the service companies in the Tehran metropolitan area. As a provider of one of the vital services of urban services, this company plays an essential role in the city's ecosystem, and many of the requirements and functional needs of the company arise from this position[3]. The set of electricity distribution companies with 39 independent companies is managed in a large holding under the supervision of the specialized mother company, Tavanir. According to the comprehensive plan of information and communication technology, the services of this sector because it includes all the layers of society and all the different resources and jobs, and on the other hand, the need for the members of society to have reliable and continuous electricity to the customers, who are the economic wheels of the industries and the whole organization[4]. Depending on it, it is very important. The most important task of the distribution company is to provide reliable and stable electricity, which should reach consumers with minimal outages and a standard voltage. The results of the investigations show that blackouts in the distribution area are one of the main reasons for customers' electricity blackouts[5]. Outages in the distribution area are divided into two categories: unplanned (accidental) outages and planned outages; the first category occurs due to technical and non-technical events in the electricity distribution networks and is cut off without the will of the company and the electricity personnel[6]. It is very critical. In the last few years, the big Tehran electric power distribution company has provided a tool for maintaining the basic information of the network, the events that have occurred, and so on, with the establishment of the event registration software (ENOX), as a result of which it is possible to obtain comprehensive information about the errors that have occurred[7]. (Including time of occurrence, resolution duration, type, cause, etc.) Which is stored in detail in the database of that system. In this regard, various forecasting techniques can be used to determine the number of equipment failures in the future[8]. Among these are the time series analysis technique, the neural network, and the decision tree[9].

In the study of Plieva et al.[10], a statistical analysis is presented regarding the leading causes of damage in the electricity distribution network. Based on the monitoring information of a distribution company located in Spain between 2014 and 2019, they discovered the leading causes of accidents and malfunctions in electricity distribution networks. Also, a comprehensive model for predicting the reliability of the power distribution system has been presented, consisting of two parts: power distribution system failure models and planned shutdowns[11]. This comprehensive model consists of a three-layer neural network and gray analysis. In the research of Richter et al.[12], an accurate reliability forecasting model is presented using a combination of seasonal integrated self-correlated moving average models and artificial neural networks to increase productivity and reduce costs. Rojek et al.[13] have proposed a model for predicting the failure rate of a Polish water supply company using a perceptron multi-layer artificial neural network. Shi et al.[14] have predicted the price of electricity by presenting a combined model of support vector regression and a single auto-correlated moving average. Also, during research, the adaptive neural-fuzzy inference system model has been used in simulating and forecasting the electric energy demand of the G8 countries until 2020[15]. Cherif et al.[16] have used the combined neural-fuzzy and particle mass algorithms in their studies to predict the long-term demand for electric energy in Algeria until 2025. The results of his study show the high power of the combinational algorithm of particle masses and the adaptive neural-fuzzy inference system in forecasting. Estelaji et al.[17], in their research, extracted blackout data recorded in the event registration software of the big Tehran electricity distribution company (ENOX) from 2012 to 2020 using the k-means algorithm as a clustering tool in the cause data mining technique. The errors that led to the blackout were classified into clusters based on behavioral similarity and finally prioritized based on the degree of impact on the blackout. In the article by Khah et al.[18], a decision support system has been designed using data mining techniques to find meaningful relationships between the features that sensors archive every day in a hydroelectric power plant. These relationships help the experts make accurate and quick decisions and prevent damage to the equipment due to wrong or late decisions. Fallahi et al.[19] have studied the prediction of the market settlement price for each competitive cluster of the Iranian market by using the genetic algorithm to perform the neural network training process. Finally, Chen et al.[20] have modeled and forecasted the hourly demand of electricity consumption with the help of integrated self-correlated moving average models, taking into account the main and mutual effects. As it can be seen, despite the widespread use of artificial neural networks and time series models in many sciences, until now, using these efficient tools, no operational model has been used to predict the failure rate of those types of power distribution network equipment that leads to blackouts.

The fire in the school of Shin Abad village (of Piranshahr) is one of a series of fires due to the failure of heating equipment in Iranian schools on Wednesday, 15 December 2018. In this incident, 39 students suffered burns, and four of them died due to the severity of their injuries. Also, the fingers of six students were amputated due to severe burns and rejection of transplants. Complaints indicate that this incident, which resulted in severe injuries to students, has caused serious mental and social consequences for the victims, their family members, and the residents of the village. Similar cases that happened due to the failure of thermal equipment in schools are seen in delegations. To solve this problem and reduce such incidents, forecasting the failure rate of power distribution network thermal equipment, especially for educational centers, is a must.

Therefore, in this article, we intend to obtain the weekly breakdown rate of the equipment in question from the incident registration software that caused accidental blackouts (in the period from the beginning of April 2018 to the end of March 2021, related to two regions belonging to the deputy coordination and supervision of the southeast) and using secret patterns. The article will consider the strong coupling of the thermal equipment operating state, the complexity of the failure mechanism, and the diversity of influencing factors; study the failure rate prediction method of the thermal equipment; and propose a combined model to improve the performance of the thermal equipment. In the combined model developed in this research, first, consider the parameters that affect the failure of the thermal equipment. Secondly, without increasing the complexity, based on the optimally weighted combination modeling theory, the minimum error square sum of the combination model is used as the objective function to solve the optimal weighting coefficients and the optimal weighted geometric average. Utilizing time and an artificial neural network, we will provide models to predict the failure rate of this equipment by separating each region so that each region has the necessary preparations to face these failures. In addition, the performance of the two proposed techniques is compared, and the superior approach is selected to predict the number of failures.

2. Prediction techniques

2.1. Time series models

A time series is an ordered sequence of observations[21]. Although it is usually arranged according to time, especially in equal intervals, sorting may be done according to other dimensions, such as distance[22]. Since past events can affect future events and the existence of interruptions in behavior is a common issue in various sciences, the time factor is a very important determining factor in time series data sets. A key feature of time series data that makes it more challenging to analyze than cross-sectional data is that observations over time are usually correlated. The meaning of this discussion is that most of the time series data have a close relationship with their observations in the recent past. Another important feature of time series data that follow a particular frequency is that they are likely to show strong seasonal patterns. This feature mainly occurs in weekly, monthly, or seasonal time series data. Box-Jenkins[23] models are widely used in time series modeling. In autoregressive models of degree p or AR (p), the autocorrelation function combines exponential and sinusoidal functions with a decreasing amplitude of fluctuations. The partial autocorrelation function for the process AR (p) becomes zero after the interval p.

$$\left(1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p\right)\mathrm{X}_t = a_t \tag{1}$$

In moving average degree q or MA (q) models, the autocorrelation function becomes zero after the interval q. The partial autocorrelation function combines exponential and sinusoidal functions with a decreasing amplitude of fluctuations.

$$\mathrm{X}_t = \left(1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_q B^q\right)a_t \tag{2}$$

In autoregressive-moving average processes, the autocorrelation and partial autocorrelation functions both approach zero. Subsequently, a general model that can model a broad class of unstable time series is the combined autoregressive moving average model with degree (p, d, q), summarized as ARIMA (p, d, q).

$$\emptyset_p(B)(1-B)^d X_t = \theta_q(B)_{a_t} \tag{3}$$

Finally, by considering the seasonal relationship between the data, the final seasonal time series model is $\mathrm{ARIMA}(p, d, q) \times \mathrm{ARIMA}(P, D, Q)_S$ or $\mathrm{SARIMA}(p, d, q) \times (P, D, Q)_S$. It is as follows:

$$\Phi_P(B^s)\emptyset_p(B)(1-B)^d(1-B^s)^D X_t = \theta_q(B)\Theta_Q(B^s)a_t \tag{4}$$

Time series modeling is an iterative method that starts with pattern recognition and parameter estimation. After estimating the parameters, we must evaluate the adequacy of the model by checking whether the assumptions are valid.

2.2. Neural networks

Neural networks are one of the most dynamic fields of research in the contemporary era, which has attracted many people from various scientific fields[24]. Artificial neural networks are new computing systems and methods for machine learning, displaying knowledge, and applying the obtained knowledge to optimize the output responses from complex systems[25]. An artificial neural network consists of three input, output, and processing layers. Each layer contains a group of nerve cells (neurons) that are generally connected with all the neurons of other layers. Perceptron networks are very worthy of attention because they have a suitable ability to evolve using input vectors. This type of neural network is very fast and reliable in solving problems[26]. One of the most important types of neural networks is the forward learning neural network. This paper uses a multilayer perceptron neural network with forward learning. Forward learning means that the value of the output parameter is determined based on the input parameters and a series of initial weights. The input values are combined and used in the hidden layers, and the values of these hidden layers are also combined to calculate the output values. Each time this algorithm is executed for all the data in the bank, it is called a period[27]. These periods continue until the error value does not change. Since the number of parameters in neural networks is large, the calculations of these networks can be time-consuming. However, if these networks are implemented for a long time, they will usually be successful. **Figure 1** shows an overview of a perceptron neural network with 3 inputs and 1 output for the given problem.

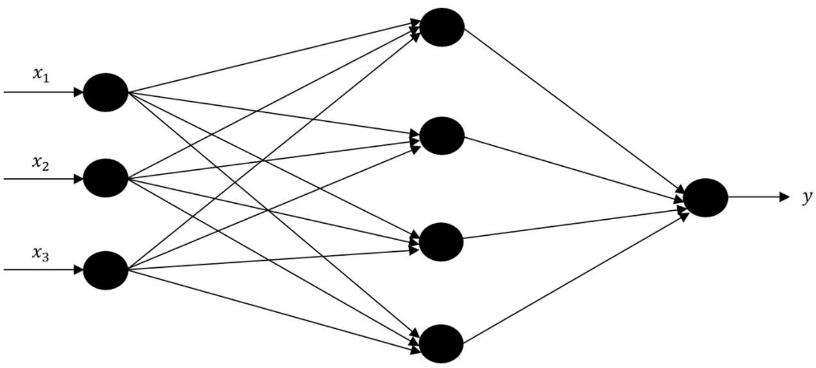

Figure 1. Multilayer perceptron neural network.

3. Data modeling and analysis

In this section, by using Box-Jenkins models and a multi-layer perceptron neural network, we model the failure rate of the desired equipment of Tehran's electric power distribution company. Greater Tehran Electric Power Distribution Company has 4,099,000 subscribers, four vice-chancellors, and 22 electricity regions. Due to the country's most significant number of electricity consumers, this company has an important and heavy duty in providing reliable electricity. The basis of this research is the information recorded in the power outage management software of the big Tehran electric power distribution company called ENOX. To start the work first, according to the need, the data related to the shutdowns that occurred from the beginning of April 2018 to the end of March 2021 were extracted from the database. According to this company's policies, the data of two key areas has been analyzed, and the names of the equipment and the areas have been avoided. These areas have been shown as A and B.

3.1. Data analysis with time series models

3.1.1. Area A

In this subsection, we will examine and model the time series data of the Ami region. The first step is to investigate the stability of mean and variance over time, which is done with the help of a time series graph, Box-Cox graph, autocorrelation, and partial autocorrelation functions. **Figures 2–5** represent the first step of implementation for modeling.

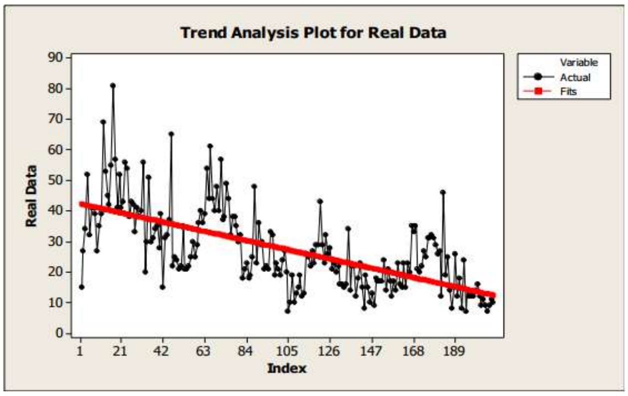

Figure 2. Weekly equipment breakdown rate chart in A region.

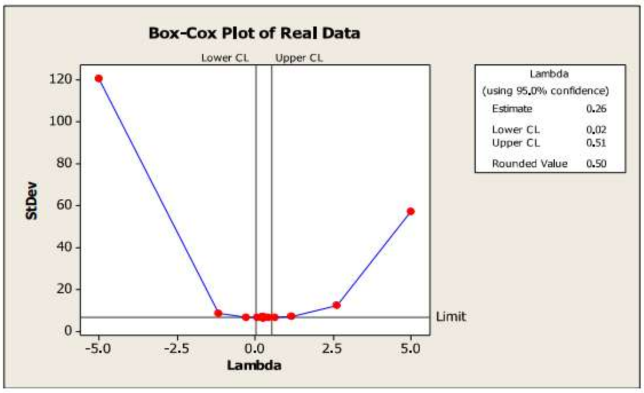

Figure 3. The proposed transformation parameter of the Box-Cox test for the stability of the variance of the data of region A.

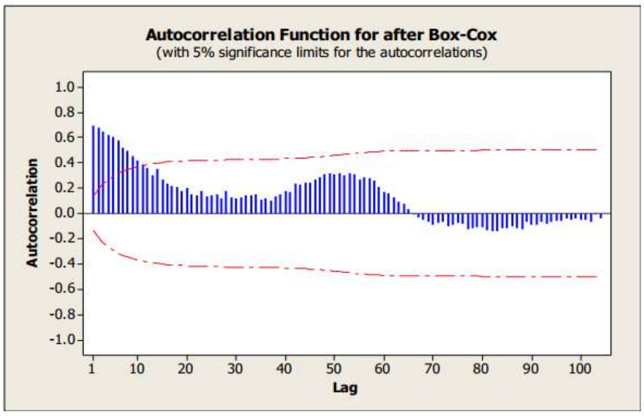

Figure 4. The graph of autocorrelation function after applying the transformation parameter on the data of region A.

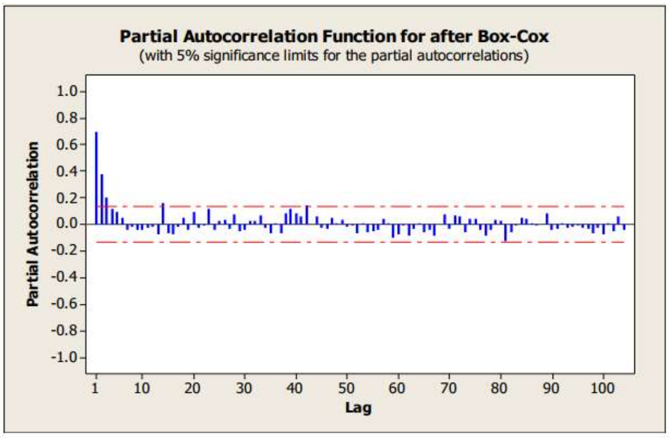

Figure 5. Partial autocorrelation function diagram after applying the transformation parameter on the data of region A.

As can be seen, **Figure 2** shows the unreliability of the time series because there is a downward trend over time. However, for a more detailed investigation, first, the unreliability in the variance is analyzed with the help of the Bucks-Cox test. The proposed conversion parameter of this test in the Minitab software in **Figure 3** is 0.5, which means that the values did not have constant variance over time. Therefore, the data's second root should be calculated to establish reliability. After that, the autocorrelation function and the partial autocorrelation of the data subject to the operation of the Bax-Cox test result have been drawn to test the reliability or unreliability of the mean (**Figures 4** and **5**). The results show that the autocorrelation function is slowly decreasing; therefore, it can be seen that the mean is also infinite. With these descriptions, AR, MA, and ARMA models are discarded for predicting this region's deterioration rate because these models are used only when the time series is stable. To fix the instability in the average, the differentiation operation must be done at least once from our data. In other words, subtract each data point from the next data point and draw its autocorrelation and partial autocorrelation functions again. If the autocorrelation function decreases rapidly, we stop the differentiation process. These operations were performed using Minitab software; the results are reported in **Figures 6** and **7**. The graphs show that the number of breakdowns varies from week to week in a month and from week to week in the same month over many years. Therefore, for example, to predict the number of breakdowns for a week in August, we must not only examine the number of breakdowns in the adjacent weeks but also test the data for August in previous years as well.

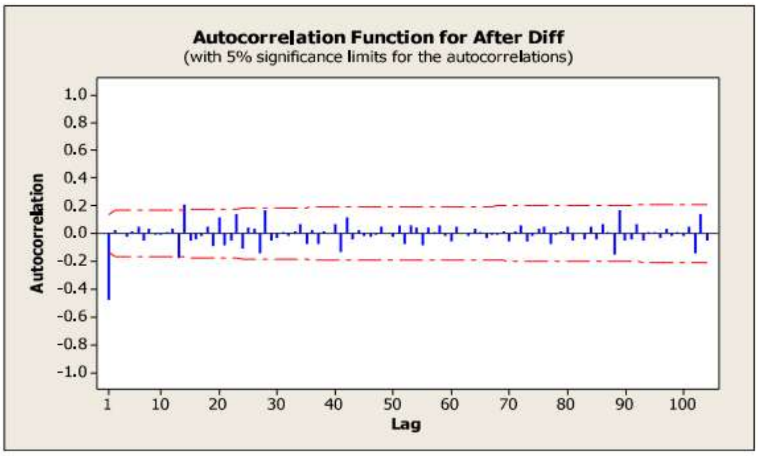

Figure 6. The graph of autocorrelation function after applying differentiation on the data of region A.

Based on the trend in **Figures 6** and **7**, the best guess can be in the form of a model $\mathrm{SARIMA}(0,1,1) \times (0,0,1)_{51}$. The Minitab software outputs resulting from this guess are described in **Table 1**.

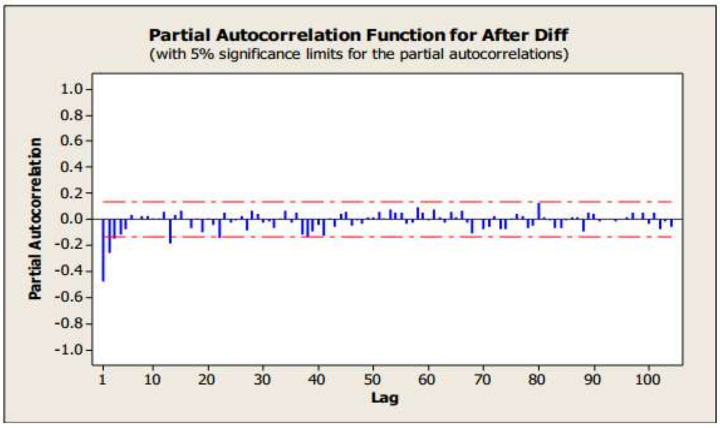

Figure 7. Diagram of the partial autocorrelation function after applying differentiation on the data of region A.

Table 1. Estimation of model coefficient $\mathrm{ARIMA}(0,1,1) \times \mathrm{ARIMA}(0,0,1)_{51}$.

Type	Coef	SE Coef	T	P
MA 1	0.6934	0.0501	13.84	0.000
SMA 51	−0.2320	0.0782	−2.97	0.003
Constant	−0.01028	0.02066	−0.50	0.619
Differencing: 1 regular difference				
Number of observations: Original series 208, after differencing 207				
Residuals: SS = 125.315 (back forecasts excluded) MS = 0.614 DF = 204				

Considering the significance level of 0.05 and the fact that the reported P-values for the estimated coefficients are less than this value, both coefficients are significant, and the model is as follows:

$$(1-B)\mathrm{X}_t = (1-0.6934B)(1+0.2320B^{51})a_t \tag{5}$$

In the following, a test called Lee Jang-Box is used to check the adequacy of the model. This test is used to evaluate the independence of the residual values of the model. The result of this test is also stated, along with the P-value. Since the P-value is more significant than 0.05, the null hypothesis is accepted, and the residuals are independent, showing the fitted model's adequacy. **Table 2** shows the Lee Jang-Box test results for the fitted model of region A.

Table 2. Lee Jang-Box test results for the fitted model of region A.

Lag	12	24	36	48
Chi-square	2.1	19.4	26.6	40.1
DF	9	21	33	45
***P*-value**	0.989	0.556	0.778	0.681

3.1.2. Region B

In this subsection, we examine and model the time series data of region B. Again, the first step is to check the stability of the mean and variance over time, which is reported in **Figures 8–11**.

The evidence in this area also indicates the unreliability of the time series. Still, for accurate investigation, first, the unreliability in the variance is analyzed with the help of the Box-Cox test, and the proposed conversion parameter is zero with respect to form 9. Therefore, the natural logarithm of the data should be calculated. To be able to establish reliability in them. Then, the examination of the autocorrelation and partial autocorrelation functions indicates the unreliability of the average. According to **Figures 10** and **11**, one time of normal differentiation and one time of seasonal differentiation are needed to solve the instability in the mean. After that, the autocorrelation and partial autocorrelation functions are drawn again (**Figures 12** and **13**).

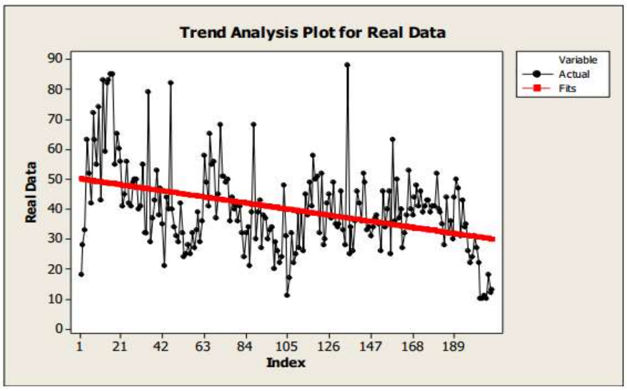

Figure 8. Weekly breakdown rate chart in area B.

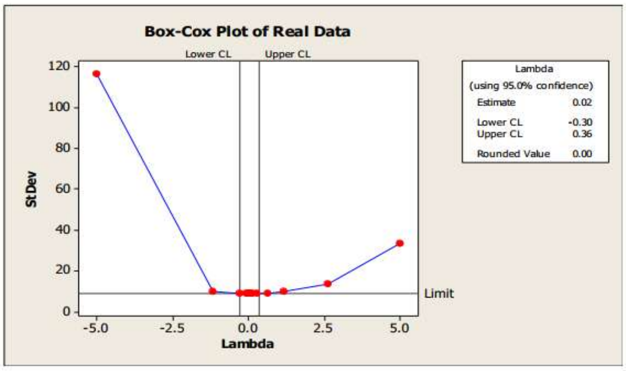

Figure 9. The suggested transformation parameter of the Box-Cox test for the stability of the variance of region B data.

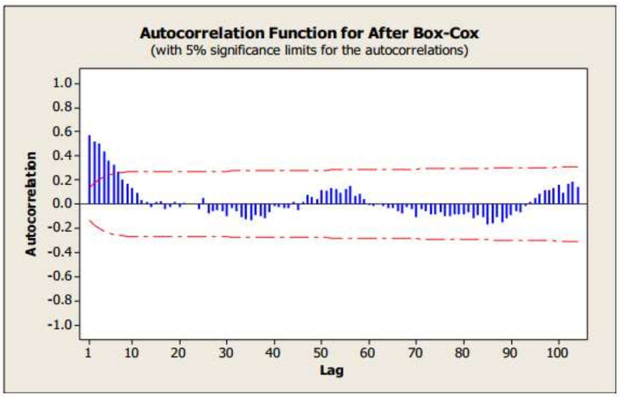

Figure 10. Graph of the autocorrelation function after applying the transform parameter on the data of region B.

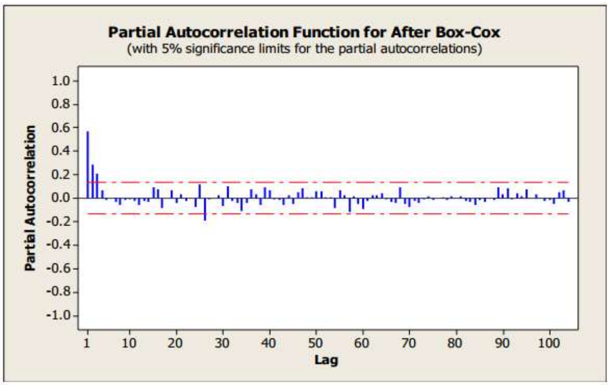

Figure 11. Partial autocorrelation function diagram after applying the transformation parameter on the data of region B.

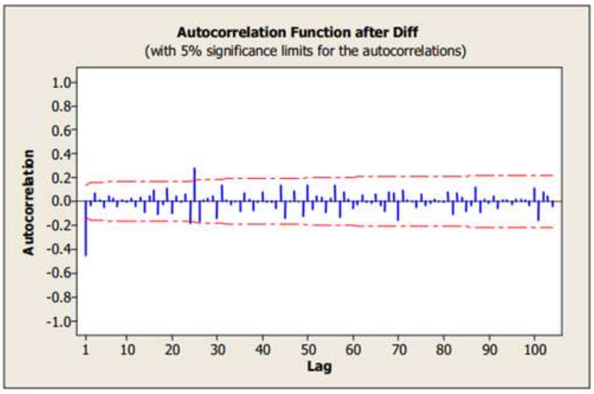

Figure 12. The graph of autocorrelation function after differentiating the data of region B.

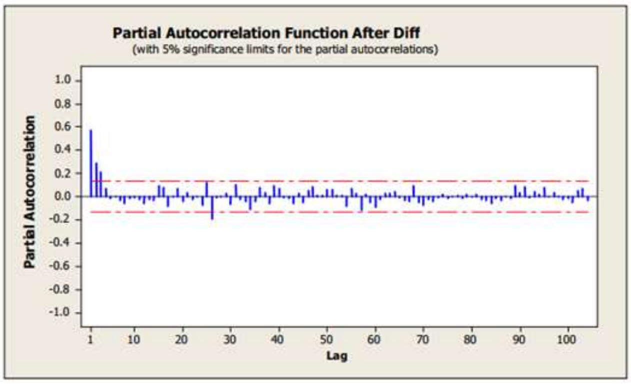

Figure 13. Partial autocorrelation function diagram after applying differentiation on the data of region B.

Based on the trend in **Figures 12** and **13** and after differentiating twice, the best possible option can be the model $\text{SARIMA}(0,1,1) \times (1,1,0)_{51}$. The outputs of the resulting Minitab software are described in **Table 3**.

Table 3. Estimation of model coefficients $\text{ARIMA}(0,1,1) \times \text{ARIMA}(1,1,0)_{51}$.

Type	Coef	SE Coef	T	P
SAR 50	−0.8120	0.0659	−12.33	0.000
MA 1	0.6886	0.0598	11.52	0.000
Differencing: 1 regular, 1 seasonal of order 50				
Number of observations: Original series 208, after differencing 157				
Residuals: SS = 15.0676 (back forecasts excluded) MS = 0.0972				
DF = 155				

According to the significance level of 0.05, both coefficients are significant, and the model is as follows:

$$(1 + 0.8120B^{50})(1 - B)(1 - B^{51})\text{X}_t = (1 - 0.6886B)a_t$$

Subsequently, the evaluation of the independence of the residuals of the model with the Lee Jung-Baks test in **Table 4** indicates that the residuals are independent and, therefore, the fitted model is sufficient.

Table 4. Lee Jung-Box test results for the fitted model of region B.

Lag	12	24	36	48
Chi-square	11.5	21.4	33.7	59.1
DF	10	22	34	46
***P*-value**	0.319	0.493	0.482	0.092

3.2. Data analysis with artificial neural network

Due to the complexity of certain phenomena and the diversity of various factors, the performance of single-term models in forecasting is always limited, so a combined forecasting model is established. The combination forecasting model uses appropriate methods to combine multiple single forecasting models, comprehensively process the forecast results of various single forecasting models, and generate a total forecasting model containing the forecasting information of multiple forecasting models. Due to the complexity of the failure rate of thermal equipment and the influence of various factors on this phenomenon, the use of a single model has certain limitations in predicting the failure rate, which reduces the accuracy of the single prediction model. For this reason, five different single models are used in this study, including the MLR model, GM (1, N) model, PLS model, BP model, and SVM model. On the basis of a single prediction, a combined prediction method with the sum of squared errors as the objective function is proposed for combined prediction. At the same time, the following methods are used to solve the weights, using the optimal combination model prediction method to select an objective function that can describe the prediction error, and then minimize the objective function to determine the optimal weight. Combining the failure rate prediction error of the five algorithms of the MLR model, GM (1, N) model, PLS model, BP model, and SVM model, the error function can be determined. +e sum of squares of the prediction errors is the objective function to construct the optimal combination model. Calculate the corresponding optimal weight. $xi(t)$ are the prediction values of MLR, GM (1, N), PLS, BP, and SVM algorithms at time t, respectively; the prediction error at time t can be obtained by $eit(t) = x(t) - xi(t)$; the forecast error information matrix is $E = [(ei)n \times m]$ $[(eit)\ n \times m]\ T$; five kinds of prediction results are analyzed and calculated with unequal weight combination; $W = [w1, w2, \ldots, wn]\ T$ are the corresponding output weight values of the two prediction methods. In order to do so, a set of 26,280 data was used, of which 75% were used for training and the rest 25% for testing. The execution steps of the combined model are shown in **Figure 14**.

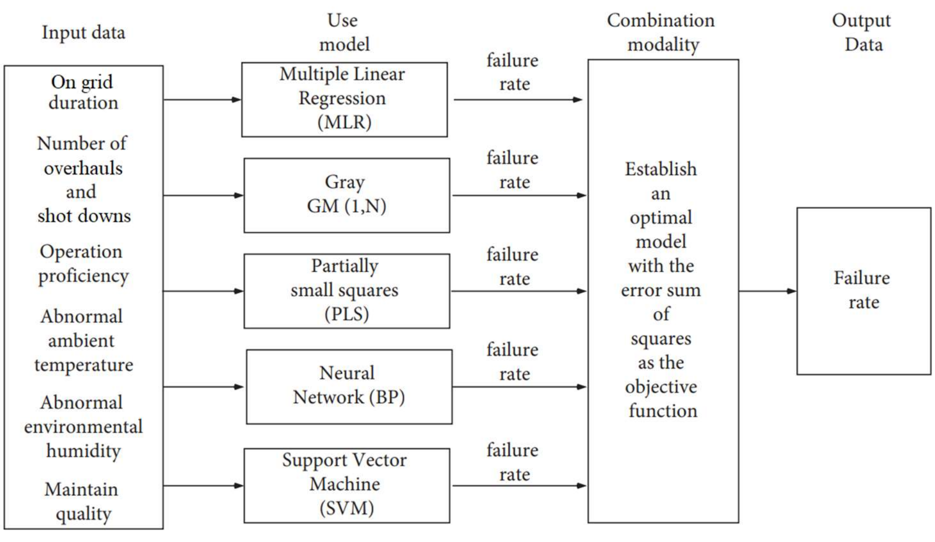

Figure 14. Execution steps of the combined model.

We use a multilayer perceptron neural network with forward learning to predict the desired equipment's failure rate. A neural network with one or two hidden layers is made for this problem. The number of input neurons was determined according to the number of parameters under study. As mentioned earlier, average air temperature, average rainfall, and average wind speed were selected as input variables for the neural network. Also, because only one output is considered, one neuron is considered in the output layer. The number of neurons in the hidden layer is determined according to the desired database. The neural network model is first built using the training data. In the training phase, the inputs and outputs are known in advance, the known data of the problem is calculated, and the weights between the neurons, which are part of the unknowns of the problem, are obtained. In the next stage, which is the model testing phase, the created neural network model is evaluated using experimental data in such a way that the inputs and weights obtained in the previous stage are only the information about the problem, and this time Using the inputs and the weights we have, we get an output and compare the obtained answer with the outputs we had before. The difference between the obtained outputs and the actual outputs determines the degree of our success in creating the model. The mean square of errors also measures the efficiency of neural network models.

4. Prediction of equipment failure rate

This section aims to predict the failure rate of the desired equipment in two regions, A and B, with the help of fitted time series models and artificial neural network techniques.

4.1. Forecasting with time series models

Based on the preliminary results and the $\mathrm{SARIMA}(0,1,1) \times (0,0,1)_{51}$ the model obtained in the region, the deterioration rate for the first six weeks of 2021 was predicted, the results of which are reported in **Table 5**.

Table 5. Predicted values for the first 6 weeks of 2021, the deterioration rate in region A based on the time series model.

Mean square error (MSE)	Predicted deterioration rate	Deterioration rate in real conditions
4.83	10	11
	9	6
	11	8
	10	10
	9	10
	11	14

Also, the deterioration rate forecast for the first six weeks of 2021 is based on $\mathrm{SARIMA}(0,1,1) \times (1,1,0)_{51}$, the model fitted in region B, as shown in **Table 6**.

Table 6. Predicted values for the first 6 weeks of 2021, the rate of deterioration in B region based on the time series model.

Mean square error (MSE)	Predicted deterioration rate	Deterioration rate in real conditions
7.66	13	16
	18	20
	19	21
	21	25
	15	18
	24	26

As it is clear, the time series models fitted in both regions A and B of the Tehran electricity distribution company have adequate performance in predicting equipment failure rates. They have acceptable and minor mean square errors.

4.2. Prediction with artificial neural network

In this study, in order to measure the accuracy of different models and prevent overfitting, the collected data are divided into two subsets: the test input data set (including 75% of the data) and the inspection data set (including 25% of the data). Use the sample 1–25 group working time distribution data of 1000–2000 h as an example for analysis. The first 18 sets of data from samples 1–18 are selected as modeling input, and the last 7 sets of data from samples 19–25 are data whose working hours are distributed between 1500 h and 2000 h for predictive model testing, and the MLR model and GM (1, N) are established, respectively. The model, PLS model, BP model, and SVM model, taking operation time, number of overhauls and shot downs, manipulation proficiency, abnormal environmental temperature, abnormal environmental humidity, and maintenance quality as independent variables, and the thousand-hour failure rate as the dependent variable to obtain sample data is shown in **Figure 15**. Related models have been established and studied.

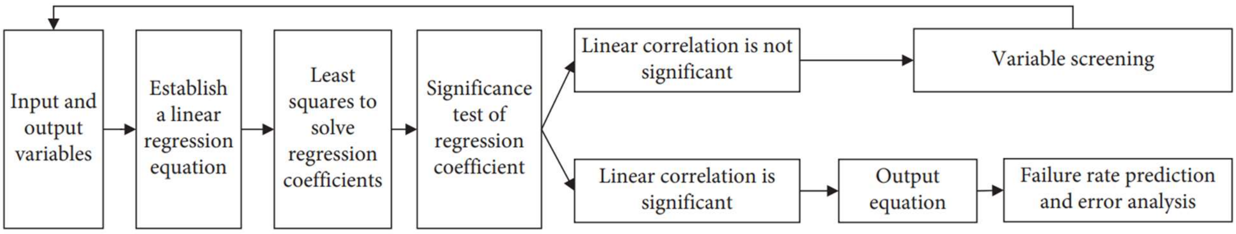

Figure 15. The MLR model process.

In this sub-section, the prediction of equipment failure rates in two regions, A and B, in the first six weeks of 2021 is done with an artificial neural network technique. As before, by using the average of the squared errors, the deviation of the forecast is measured against the actual values. **Tables 7** and **8** show the results.

Table 7. Predicted values for the first 6 weeks of 2021, the deterioration rate in area A based on the artificial neural network technique.

Mean square error (MSE)	Predicted deterioration rate	Deterioration rate in real conditions
9.16	14	11
	11	6
	10	8
	12	10
	8	10
	17	14

Table 8. Predicted values for the first 6 weeks of 2021, the deterioration rate in area B based on the artificial neural network technique.

Mean square error (MSE)	Predicted deterioration rate	Deterioration rate in real conditions
13.83	20	16
	23	20
	19	21
	30	25
	23	18
	28	26

In the following, we should compare the two approaches used. The results show the better performance of time series models in comparison with the neural network with regard to the mean square error index, meaning that time series models have a smaller mean square error index. Of course, this superiority may be due to not using appropriate training functions, not choosing the number of layers used, and, in general, the non-optimal stopping conditions in the selective neural network. Also, among the time series models, the model fitted to the region A data has the least error.

5. Conclusion

Electricity distribution companies are at the forefront of the electricity industry, on the front lines of communication with the people. They are particularly important as they are the external mirror of the electricity industry's performance and will be effective in social reflection. The results of the investigations show that blackouts in the distribution area are one of the main reasons for customer blackouts, which are affected by technical and non-technical events in electricity distribution networks, including equipment breakdowns. In this regard, various techniques can be used to model and predict the number of equipment failures in the future. This article used time series analysis and artificial neural network techniques to provide the optimal model and discover the latent pattern in the number of breakdowns of specific equipment in two key and sensitive areas covered by the Greater Tehran Electric Power Distribution Company. In order to provide good maintenance and maintenance decision-making, establish the health management mechanism of thermal equipment, and improve the prediction accuracy of the failure rate of airborne equipment, two single models based on the optimal combination prediction model were used. The accuracy, performance, and predictive ability of the combined model are studied and compared with those of different individual models. The use of the combined model increases the R2 index rate from 1.4% to 10%. It also reduced the MAPE index rate from 59.7% to 88.4%.

The results of the investigations based on the mean square error index indicated that the seasonal time series models $\mathrm{SARIMA}(0,1,1)\times(0,0,1)_{51}$ and $\mathrm{SARIMA}(0,1,1)\times(1,1,0)_{51}$ had a very good and appropriate performance in predicting the failure rate of the desired equipment in the two studied areas. Also, this category of models can be used in other areas of the electricity distribution company. Therefore, the results of this research greatly help predict and obtain the necessary preparation of operational crews to face unplanned blackouts and therefore make timely decisions for the authorities. The Tehran electricity distribution company will expand. Finally, as suggestions for future studies and research, it is possible to optimize the number of hidden layers and the number of neurons with the help of the response surface methodology approach and compare its performance with time series models to predict the equipment failure rate. It provides a new method of failure rate prediction for thermal equipment. This method can be used to predict the remaining life, failure time, and other aspects of thermal equipment. In addition, the prediction method has certain reference significance and engineering application value in the field of complex equipment fault diagnosis and prediction. In the future, new applications of combination models will be further explored, and complex multi-combination models will be developed to achieve more accurate prediction results.

Author contributions

Conceptualization, RO and AAP; methodology, AS; software, MK; validation, AS, RZ and MK; formal analysis, RO; investigation, RZ; resources, RO; data curation, MK; writing—original draft preparation, RZ; writing—review and editing, AAP; visualization, RZ; supervision, AAP; project administration, RO; funding acquisition, RO. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

1. Keshavarzzadeh M, Zahedi R, Eskandarpanahe R, et al. Estimation of NO_x pollutants in a spark engine fueled by mixed methane and hydrogen using neural networks and genetic algorithm. *Heliyon* 2023; 9(4): e15304. doi: 10.1016/j.heliyon.2023.e15304
2. Romanov OI, Nesterenko M, Mankivskyi V. The method of redistributing traffic in mobile network. In: *Data-Centric Business and Applications*. Springer; 2021. pp. 159–182.

3. Farshchian G, Darestani SA, Hamidi N. Developing a decision-making dashboard for power losses attributes of Iran's electricity distribution network. *Energy* 2021; 216: 119248. doi: 10.1016/j.energy.2020.119248
4. Zare P, Ghadimi H, Zare R, et al. The study impact of restructuring on efficiency of Iran's electricity distribution and transmission network. In: Proceedings of the 2023 8th International Conference on Technology and Energy Management (ICTEM); 8–9 February 2023; Iran.
5. Brown RE. In: *Electric Power Distribution Reliability*, 2nd ed. CRC Press; 2017.
6. Fuchs EF, Masoum MAS. *Power Quality In Power Systems And Electrical Machines*, 1st ed. Academic Press; 2011.
7. Safari A, Das N, Langhelle O, et al. Natural gas: A transition fuel for sustainable energy system transformation? *Energy Science & Engineering* 2019; 7(4): 1075–1094. doi: 10.1002/ese3.380
8. Estelaji F, Naseri A, Zahedi R. Evaluation of the performance of vital services in urban crisis management. *Advances in Environmental and Engineering Research* 2022; 3(4): 1–19. doi: 10.21926/aeer.2204057
9. Zahedi R, Ghodusinejad MH, Gitifar S. Threats evaluation of border power plants from the perspective of fuel type and providing solutions to deal with them: a case study of Iran. *Transactions of the Indian National Academy of Engineering* 2022; 8(1): 1–13. doi: 10.1007/s41403-022-00377-5
10. Plieva M, Gavrina O, Kabisov A. Analysis of technological damage at 110 kV substations in JSC IDGC of the north caucasus-«Sevkavkazenergo». In: Proccedings of the 2019 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon); 1–4 October 2019; Vladivostok, Russian.
11. Dehghanian P, Fotuhi-Firuzabad M, Aminifar F, Billinton R. A comprehensive scheme for reliability centered maintenance in power distribution systems—Part I: Methodology. *IEEE Transactions on Power Delivery* 2013; 28(2): 761–770. doi: 10.1109/TPWRD.2012.2227832
12. Richter L, Lehan M, Marchand S, et al. Artificial intelligence for electricity supply chain automation. *Renewable and Sustainable Energy Reviews* 2022; 163: 112459. doi: 10.1016/j.rser.2022.112459
13. Rojek I, Studzinski J. Detection and localization of water leaks in water nets supported by an ICT system with artificial intelligence methods as a way forward for smart cities. *Sustainability* 2019; 11(2): 518. doi: 10.3390/su11020518
14. Shi J, Guo J, Zheng S. Evaluation of hybrid forecasting approaches for wind speed and power generation time series. *Renewable and Sustainable Energy Reviews* 2012; 16(5): 3471–3480. doi: 10.1016/j.rser.2012.02.044
15. Samui P, Sekhar S, Balas VE. *Handbook of neural computation*, 1st ed. CRC Press; 2017.
16. Cherif K, Yahia N, Bilal B, Bilal B. Erosion potential model-based ANN-MLP for the spatiotemporal modeling of soil erosion in wadi saida watershed. *Modeling Earth Systems and Environment* 2023; pp. 1–23. doi: 10.1007/s40808-022-01657-3
17. Estelaji F, Aghajari AA, Zahedi R. Flood zoning and developing strategies to increase resilience against floods with a crisis management approach. *Asian Review of Environmental and Earth Sciences* 2023; 10(1): 14–27. doi: 10.20448/arees.v10i1.4439
18. Khah MV, Zahedi R, Eskandarpanah R, et al. Optimal sizing of residential photovoltaic and battery system connected to the power grid based on the cost of energy and peak load. *Heliyon* 2023; 9(3): e14414. doi: 10.1016/j.heliyon.2023.e14414
19. Fallahi S, Shaverdi M, Bashiri V. Applying GMDH-type neural network and genetic algorithm for stock price prediction of Iranian cement sector. *Applications and Applied Mathematics: An International Journal (AAM)* 2011; 6(2): 572–591.
20. Chen Q, Guo H, Zheng K, Wang Y. *Data Analytics in Power Markets*. Springer; 2021.
21. Zahedi R, Eskandarpanah R, Akbari M, et al. Development of a new simulation model for the reservoir hydropower generation. *Water Resources Management* 2022; 36(7): 1–16. doi: 10.1007/s11269-022-03138-9
22. Madsen H. *Time Series Analysis*. CRC Press; 2007.
23. Harvey AC, Todd PHJ. Forecasting economic time series with structural and Box-Jenkins models: A case study. *Journal of Business & Economic Statistics* 1983; 1(4): 299–307. doi: 10.2307/1391661
24. Zahedi R, Daneshgar S, Golivari S. Simulation and optimization of electricity generation by waste to energy unit in Tehran. *Sustainable Energy Technologies and Assessments* 2022; 53: 102338. doi: 10.1016/j.seta.2022.102338
25. Gupta M, Jin L, Homma N. *Static And Dynamic Neural Networks: From Fundamentals To Advanced Theory*, 1st ed. Wiley-IEEE Press; 2003.
26. Guo R, Lin Z, Shan T, et al. Physics embedded deep neural network for solving full-wave inverse scattering problems. *IEEE Transactions on Antennas and Propagation* 2021; 70(8): 6148–6159. doi: 10.1109/TAP.2021.3102135
27. Elakkiya E, Selvakumar S. Stratified hyperparameters optimization of feed-forward neural network for social network spam detection (SON2S). *Soft Computing* 2022; 26(2): 11915–11934. doi: 10.1007/s00500-022-07020-z