7th International Conference on
Reliability and Safety Engineering
(ICRSE2023)

7th International Conference on Reliability and Safety Engineering

Aerospace Research Institute
Ministry of Science Research and Technology

Forecasting of unsold energy in electricity distribution networks with the aim of improving network reliability

Elham Fallah Baghemoortini[1], Davood Shishebori[2*]

1-M.Sc., Department of Industrial Engineering, Yazd University, Yazd, Iran
2-Associate Professor, Department of Industrial Engineering, Yazd University, Yazd, Iran

Corresponding author, shishebori@yazd.ac.ir

Abstract

The finiteness of electricity sources, the high consumption of electricity in Iran and the high price of electricity production mean that an accident in the electricity distribution network will cause damage to the network and subscribers. When an incident occurs in the network, in addition to the cost of repairs to the network, the subscribers are without electricity for a period of time and the network is stopped from selling electricity. From an economic point of view, these incidents cause losses and loss of economic resources and reduce competitiveness in this area. Therefore, investigating the events that have occurred and predicting potential events that may occur in the future and cause energy not to be sold; It helps network management to improve reliability. The purpose of this article is to predict the unsold energy based on the events that have occurred. For this purpose, the ARIMAX time series model was implemented on the data of the electricity distribution network of Yazd province. The results showed that in June and July, which will be the peak load of the network, we will have an average of 6.42 MWH of unsold energy. Managers can reduce this amount of unsold energy by improving the reliability of power distribution networks.

Keyword: Electrical distribution network; Forecast; Time series; Unsold energy; Reliability

Introduction

One of the infrastructures of urban development in the world is the electric power distribution network, which exists more or less with differences in all countries of the world. Today, with the expansion of the industry and the increase in the sensitivity of electrical devices, electricity distribution networks are facing an increase in the demand for stable electricity supply; So that both the number and duration of blackouts should be minimized. During the years of creating urban and industrial distribution networks in more than a century, a lot of progress has been achieved, but despite the advancement of knowledge and the reduction of design challenges and limitations, it is still facing problems and complications.
In the designs of engineers in the electricity distribution network, the aerial network is still preferred in Iran and the world due to some simplicity and low costs compared to the ground network. The aerial network, like other networks, needs access (heavy vehicles and staff access) for maintenance and periodic inspections, and on the other hand, there is a dependency on the aerial network for lighting roads such as highways and roads. Therefore, the aerial network is placed next to roads and highways in most cases. This neighborhood is also present in the urban distribution network, with buildings and people's residences. The proximity of the air distribution network to the two main centers of traffic, work, construction and dozens of other issues will also cause damage to this network.
Today, one of the most important sources of managers' power is the information and perspective on the future of the system, in fact, correct statistics are the eyes of managers in micro and macro decisions. Every manager deals with forecasting in some way in his decisions; Of course, the prediction never exactly matches the reality, and you should try to reduce the prediction error to the minimum possible. Prediction means estimating something that will happen in the future.

The most important task of distribution companies is to provide reliable and stable electricity, which should reach consumers with minimum blackouts or standard voltage. The results of the investigations show that blackouts in the distribution field are divided into two categories: unplanned (accidental) blackouts and planned blackouts, the first category of which occurs as a result of technical and non-technical events in electricity distribution networks and is cut off without the will of the company

and electricity personnel and is very critical [1]. Due to the low voltage level and high current, distribution networks have very high losses and high voltage drop, which has always contributed the most to losses and reduced reliability of the electricity distribution network [2]. Therefore, the prediction of incidents and a comprehensive approach to the timing of unplanned outages are very valuable in order to prevent any power losses.
Every organization needs to predict the future in order to survive and ensure its success in the society. On the one hand, forecasting is an uncertain process, because it is impossible to speak with certainty about the future, and on the other hand, the wider the horizon of forecasting, the greater the uncertainty. Investigation of incidents leading to breakdowns and disruptions in electricity distribution networks has been investigated in many researches in the world [3].
In a research conducted on electricity demand data in one of Japan's regions, the amount of load consumed at different times was predicted by the SARIMAX method. This issue expresses the importance of examining the data available in electricity distribution companies [4]. A comprehensive model for predicting the reliability of the power distribution system has been presented, which is made of two separate parts of power distribution system failure models and planned outages [5].
A study was conducted in India [6] that provides an overview of recent electricity consumption forecasting approaches in the context of various applications. This study presents a review of energy forecasting works published between 2017 and 2020.
In a research, the annual electricity demand of Bangladesh was predicted using a multivariate time series model. Since the single-variable time series cannot include external factors, two exogenous variables, including population and GDP per capita, were introduced as exogenous variables to achieve better performance [7].
Researchers presented a technique to examine the importance of features in the data to predict the next hour demand of medium voltage electrical loads using artificial intelligence [8]. In 2018, a research was conducted by Arjamand et al., emphasizing the pre-processing of data before entering electric load prediction models [9].
According to the above research, it is clear that it is important to investigate the incidents and the time taken to repair the network. Also, the amount of consumption has often been predicted, while predicting the time of unplanned blackouts can give managers a good view so that they can plan to reduce them. In the last few years, Yazd Province Electricity Distribution Company has provided a tool to record events by deploying incident software (System 121), as a result of which it is possible to obtain comprehensive information about errors that occurred (including time of occurrence, duration of occurrence, duration of resolution, type , reason, etc.) which is stored in detail in the database of that system. In this regard, by examining the data recorded between November 1389 and Shahrivar 1400, we will predict the main causes of unplanned blackouts.
The purpose of this study is to predict unsold energy based on events. In the following, after examining the background of the research and the research method, the obtained results will be explained.
This article uses the time series model (ARIMAX) and R software to do the forecasting. The results of this study can help managers to make decisions for planning to reduce and face accidents (caused by unplanned shutdowns). Naturally, the right decision, taking into account the results of this study, will lead to the reduction and elimination of breakdowns, stoppages, sudden and unexpected events in the distribution network, increasing the credibility, reputation and goodwill of electricity distribution companies among their subscribers and significantly reducing costs. became. Also, it can be attractive for the electricity distribution organization as a seller of electricity to compensate for this lack of sale by adjusting the price of electricity.

research method

Today, with the expansion of data mining (data analysis), various methods for forecasting have been developed. Time series is one of the common methods for forecasting. In time series models, the information related to the past of the product is evaluated in their chronological order and based on that, a forecast is made for the future of the product. Time series of statistical estimation methods are used in forecasting. In this research, forecasting is analyzed based on data analysis during the time of the month, and according to seasonal changes and random changes, the researcher will find out the trend of incidents in five years. .
Regression modeling is the most widely used method for forecasting time series and is a statistical technique (linear or non-linear) that derives the most appropriate coefficients for the independent variables. These coefficients (or weights) explain the variance of the dependent variable as a function of a certain independent variable and thus predict the future values of the dependent variable. This relationship is controlled by the special order and structure of the model.
Model structure describes the type (or classification) of different polynomial terms that are included in a given model. The most common linear black box models for time series data incorporate the use of both AutoRegressive (AR) and Exogenous (X) terms. AR entries refer to the use of lagged (prior) values of the dependent variable; while X refers to the inputs of independent external variables. More complex models may include moving average (MA) terms that refer to the composition of the residuals (ie, the within-sample errors resulting from the training period). And these polynomial terms are usually combined to create specific model structures, such as the automatic moving average with exogenous inputs (ARMAX) model, with the aim of

providing varying degrees of flexibility, accuracy, and independence with respect to modeling system dynamics and error characteristics [10].

In the equations of these models, B is a regression function, the function related to auto regression, related to moving average and random error. Exogenous variables are characterized by β coefficient. The effect of seasonal coefficients is also determined by parameter d. The autoregressive (AR) model is used when the value of the series depends on the value immediately before it plus a random error, and includes everything new at time t that is not expressed by past values in the series. The process is reversible and the series is stationary if it is. ARX model is an explanatory auto-regressive model considering exogenous variables. A moving average (MA) is used when events produce an instantaneous effect that remains for short periods of time. This process is always static. Moving average with explanatory exogenous variables is called MAX. A moving average ARMA autoregressive process is obtained from the combination of different features of two autoregressive processes AR and moving average MA. The ARMAX model is one of the most flexible types of time series models. have explanatory and lag-dependent variables, for each finite variable, accommodate more than one explanatory variable. Here we assume that the errors are independent and uniformly distributed with zero mean and constant variance and zero covariance.

Table 1. Types of time series patterns

Row	Model	Eq
1	AR	$y_t = \varphi_1 y_{t-1} + \ldots\ldots\ \varphi_2 y_{t-2} + \varphi_3 y_{t-3}$ $\varphi_p y_{t-p} + \varepsilon_t$ $y_t = \varphi(B) y_t + \varepsilon_t$
2	ARX	$y_t = \varphi(B) y_t + \beta x_t + \varepsilon_t$
3	MA	$y_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} \cdots\cdots\cdots$ $\theta_p \varepsilon_{t-q} + \varepsilon_t$ $y_t = \theta(B) \varepsilon_t$
4	MAX	$y_t = \theta(B) \varepsilon_t + \beta x_t$
5	ARMA	$\varphi(B) y_t = \theta(B) \varepsilon_t$
6	ARMAX	$\varphi(B) y_t = \theta(B) \varepsilon_t + \beta x_t$
7	ARIMA	$\varphi(B)(1 - B)^d y_t = \theta(B) \varepsilon_t$
8	ARIMAX	$\varphi(B)(1 - B)^d y_t = \theta(B) \varepsilon_t + \beta x_t$

A general model that has the ability to represent a wide class of non-stationary time series is called ARIMA autoregressive moving average. By adding the exogenous variable to the ARIMA model, it is possible to evaluate the impact of explanatory exogenous variables on the dependent variable by the ARIMAX model.

Numerical findings

One of the main reasons for customers' power outages is the blackouts in the distribution field, which are affected by technical and non-technical events in the electricity distribution networks. According to the characteristics of electricity distribution airlines in Iran, distribution networks are dominant in the form of airlines. Categorization of aerial electricity distribution networks in the form of; The network is medium pressure, weak pressure and lighting. Various equipment are placed together in the power distribution network to distribute and flow electrical energy. According to the type of network arrangement, the equipment is different. The overhead power distribution network is exposed to many types of accidents that lead to power outages. The 121 system of the electricity distribution company of Yazd province records unexpected incidents along with the cause of the incident and the duration of the network interruption. These incidents are categorized and recorded by the operator of the dispatching unit of the electricity distribution company. Unplanned outages are recorded in the system over time and its database is available. Examining the trends and fluctuations of these times can provide an effective management perspective. Now, the goal is to predict the planned blackout time by considering the number of incidents as an exogenous variable, and also by considering the number of incidents and blackout times, the amount of unsold energy can be predicted.

With the cooperation of the Dispatching Unit of the Central Headquarters of the Yazd Province Electricity Distribution Company, a daily report of medium pressure feeder incidents and outages was prepared from the 121 system. This report is from November 1389 to Shahrivar 1400 on a monthly basis and includes the incidents that occurred, the time of network disconnection, the time of network connection, the load on the network, the location of the incident and its cause. The spatial scope of the three cities of Yazd is considered (from Janbaz Blvd.: Abolfazl Square to Imam Ali Square, to Fahraj). It should be noted that only urban subscribers are connected to this network. Figure 1 shows the initial form of the data.

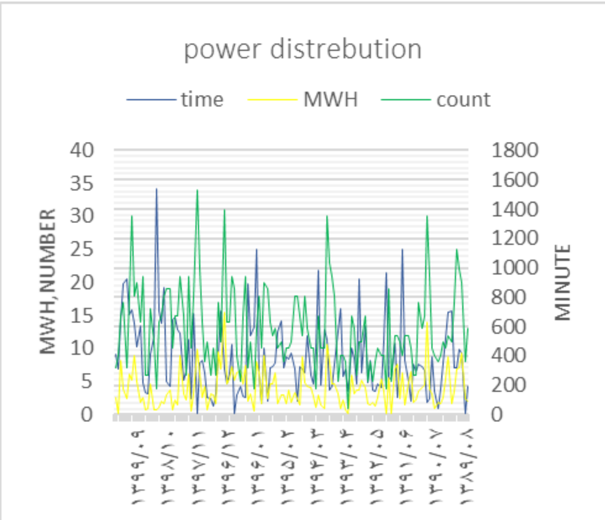

Fig. 1. Graph of variables of outage duration, number of incidents and unsold energy by time

According to the data of the above report, the time of blackouts, the number of incidents and the amount of unsold energy were extracted by month (130 records). To validate the prediction, 30% of the data was removed and only 92 records were entered into the model for prediction. And the forecast has been made for three periods.
To use different types of time series patterns and fit the model, first the instability of the series is checked. Instability in the mean means the existence of an upward or downward trend line in the data, but instability in the variance means the existence of uneven lows and highs in the data. ADF and KPSS tests investigate non-stationarity. The null hypothesis for the ADF test is that the data are non-stationary. Therefore, a p-value greater than 0.05 indicates instability, and a p-value less than 0.05 rejects the instability of the series. In the KPSS test, the null hypothesis is that the data are stationary. In this case, a p-value less than 0.05 indicates unstable series and a p-value greater than 0.05 does not reject the stationarity of the series. Many researchers determine the instability in the series by the result of a test or by examining the autocorrelation and partial autocorrelation diagrams. In this research, the instability has been investigated by considering the results of both tests and examining the diagrams. The results of these tests for the blackout time series in the software in Table 2 show that the blackout time series is an unstable series.

Table 2. Results of tests to check instability

Row	Variable	P-value	
		KPSS	ADF
1	unsold energy	0.02	0.11

In time series, the correlation between two terms (Zt, Zt+k) is determined by the time lag of the autocorrelation graph (ACF) (Figure 2). To check autocorrelation and partial autocorrelation, the maximum lag (lag max) is first determined. This value is usually equal to n/4, (n number of tens of series).

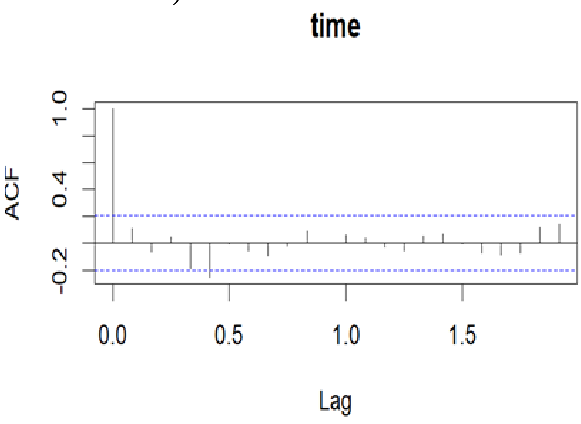

Fig. 2. Autocorrelation diagram

Autocorrelation of data, how much each data depends on the data before it. As you can see, from the second delay onwards, the autocorrelation has not exceeded the acceptance limit. Figure 3 shows the partial autocorrelation (PACF).

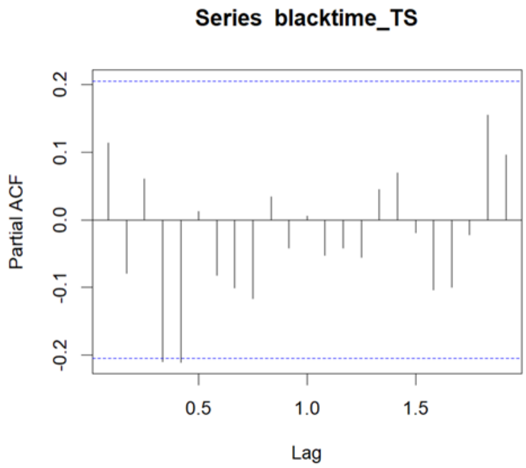

Fig. 3. Autocorrelation diagram

After checking autocorrelations and preliminary analysis, estimating the type of model and parameter values (coefficients) with the help of software, the models are fitted and considering the lowest Akaike value (AIC), the lowest error and checking the residuals, the best model is selected.
The equations of forecasting models for unsold energy are as follows:

(1) $(1-0.3679B)(1-B)^0 Z_t = (1+0.2471B-0.1459B^2)+\varepsilon_t$

(2) $(1-0.8872B)(1-B)^0\, y_t = (1+0.9434B)\varepsilon_t + 0.09078x_t + 0.0074k_t$

It has also been implemented for the amount of energy that has not reached the customer, and the equation of these models is formula 1 and 2. In Table 3, it is clear that the value of Akaike's criterion and the average absolute error have decreased. Therefore, the obtained model is more accurate. By considering the number of accidents and the duration of blackout as exogenous variables, the accuracy of the model has increased. The proposed model for predicting unsold energy is ARIMAX(1,0,1)(0,0,0) considering the number of incidents and time of unsold energy as exogenous variables.

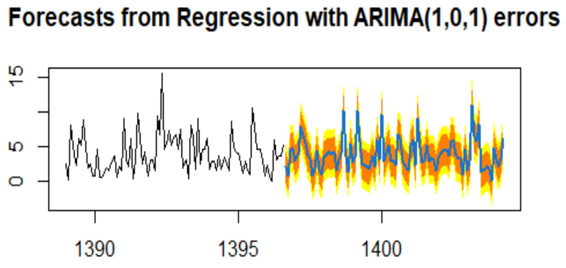

Fig. 4. ARIMAX unsold energy prediction chart (0,0,0)(1,0,1)

With the help of R software, the forecast for the next three periods (years) has been made (Figure 4) and since 30% of the data have been used for validation, in addition to the average absolute error predicted by the software, the error has been recalculated. It was found that ARIMAX models performed better than ARIMA in this type of error. In time series methods, after extracting trends and seasonal factors from the data, the residuals should follow the normal distribution function and no trend should be observed in them. Figure 5 shows the state of the residuals that change around zero, the normal distribution function's goodness-of-fit diagram is also drawn. The box test (Box.test) is performed to check the assumption of normality of the residuals. Chi-square statistic and p-value indicate that the values of the residuals follow the normal distribution function. Also, the number of protruding lags in the ACF chart is a case that shows that the residuals are not autocorrelated.

Table 3. Comparison of the results of the models

*	Model	Exogenous variable	Absolute error	Akaike's Criterion (AIC)
Forecasting unsold energy	ARIMA (1,0,1) (0,0,0)	-	2.122	457.84
	ARIMAX (1,0,1) (0,0,0) , xreg = count & black time	Duration of outages, number of incidents	1.65	5.363

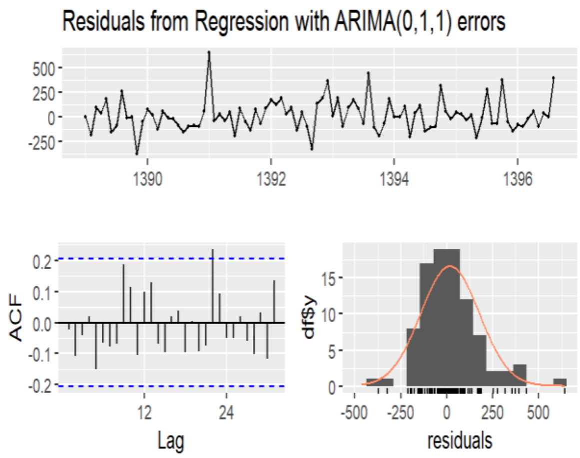

Fig. 5. The graph of the residuals of the ARIMAX model (0,1,1)(0,0,0)

Conclusion

Recording vital events and incidents is one of the long-standing human activities that plays an important and vital role in having a healthier and safer life. In most of the country's plans for the future, the preservation of human life and the continuation of a safe life for humans is considered an urgent need, and the need for statistical information and forecasting of such information is felt. Therefore, it is necessary to be able to predict the amount of events and consequently its optimal control. Forecasting for short-term, medium-term, and long-term planning and for building existing products with existing facilities and achieving the goals of the fire department and safety services is the main goal of predicting incidents, and the aforementioned applied time series models are the answer to these issues. In the research that was conducted, the number of incidents and the duration and amount of unsold energy (not reached to the customer) were predicted with an acceptable error. According to the obtained results (Table 4), we can say that we will have 6.42 MWH unsold (with an absolute error of 1.65) in June and July next year. On the other hand, the community is without electricity and dissatisfaction has arisen, which lies in economic and social losses. Therefore, with this warning, managers in their planning this month of the year should re-examine the factors of disruption and lack of electricity supply and think of measures to reduce these blackouts.

Table 4. Forecast of unsold energy for the first six months of 1401

Row	Date	Energy (MWH)
1	1401.01	3.14
2	1401.02	2.3
3	1401.03	6.72
4	1401.04	6.12
5	1401.05	3.20
6	1401.06	2.42

Smart electrical energy distribution networks are one of the latest technologies in the world. The main goal of these networks is to provide reliable electricity, increase the reliability factor, network stability and respond to the growing needs of customers with minimal damage to the environment, profit and high efficiency. Predicting the future state of the network with the least error brings us closer to the smart network. Therefore, in the future, researchers can increase the prediction accuracy by implementing immune intelligence in the network and considering more variables on the network.

References

[1] Asadzadeh, S., 2019. Optimal Modeling and Forecasting of Equipment Failure Rate for the Electricity Distribution Network. Iranian Electric Industry Journal of Quality and Productivity, 8(1), pp.53-61.

[2] Akbari, 2020. Optimum rearrangement of distribution networks with the aim of reducing losses, increasing reliability and improving voltage profile using wild mouse colony algorithm. New research in electricity, 9(1), pp.35-45.
[3] Karim Abadi, Haji Abadi, Ghayab and Ebadaleh, 2017. Review of repairs and breakdowns of transmission and super distribution substation equipment. New research in electricity, 5(2), pp.20-31
[4] Elamin, N. and Fukushige, M., 2018. Modeling and forecasting hourly electricity demand by SARIMAX with interactions. Energy, 165, pp.257-268.
[5] Xie, K., Zhang, H. and Singh, C., 2016. Reliability forecasting models for electrical distribution systems considering component failures and planned outages. International journal of electrical power & energy systems, 79, pp.228-234.
[6] Gupta, A., Chawla, M. and Tiwari, N., 2022. Electricity Power Consumption Forecasting Techniques: A survey. Available at SSRN 4019431.
[7] Rabbi, F., Tareq, S.U., Islam, M.M., Chowdhury, M.A. and Kashem, M.A., 2020, December. A Multivariate Time Series Approach for Forecasting of Electricity Demand in Bangladesh Using ARIMAX Model. In 2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI) (pp. 1-5). IEEE.
[8] Armano, G. and Pegoraro, P.A., 2022. Assessing Feature Importance for Short-Term Prediction of Electricity Demand in Medium-Voltage Loads. Energies, 15(2), p.549.
[9] Arjmand, A., Samizadeh, R. and Dehghani Saryazdi, M., 2019. Improved forecasting of short term electricity demand by using of integrated data preparation and input selection methods. Journal of Energy Management and Technology, 3(1), pp.48-57.
[10] Ahmed, A.I., McLeod, R.S. and Gustin, M., 2021. Forecasting underheating in dwellings to detect excess winter mortality risks using time series models. Applied Energy, 286, p.116517.