

RESEARCH PAPER

Midterm Load Forecasting Analysis For Erbil Governorate Based On Predictive Models

Warda Hussein Ali

Department of Electrical Power ,Engineering Technical College of Engineering, Sulaimani Polytechnic University, Sulaimani,
Iraq

ABSTRACT:

Electrical power supply is becoming more and more complex as a result of expansion, growing population, and unsuitable planning of administration and peoples. Electrical power load forecasting may be defined as the process of predicting electrical load values for future of the system with respect to current demands. This analysis is an important procedure for the power system planners and the demand controllers to ensure that the system can generate sufficient of electricity for different kinds of terms such as short, medium and long term load forecasting. The forecasting analysis allows us to manage the electrical loads with the increasing demand. For that purpose, we have used some predictive models to analyze of electrical load forecasting for Erbil Governorate in Iraq. This analysis helps us to manage our planning better, arrange system maintenance plan and enhance fuel control. This study raises an attempt for forecasting the peak (upper limit) monthly demand of electric power for one year ahead. Simple linear regression model and Auto Regressive Integrated Moving Average model were applied as forecasting models for a power consumer's dataset for the purpose of predicting forthcoming year electricity load demand. , also Forecasting models are then validated using some indicators, indicator used is Root Mean Square Error (RMSE) ,which is conceded a statistic metric that is commonly used for accuracy evaluation of LF methods, and Mean Absolute Error (MAE) both used as a forecasting accuracy criteria.

KEY WORDS: Load Forecasting Analysis, Linear Regression, Root Mean Square, ARIMA Modelling

DOI: <http://dx.doi.org/10.21271/ZJPAS.32.3.3>

ZJPAS (2020) , 32(3);20-29 .

1.INTRODUCTION :

Electrical loads are varying from day to day as a consequence of modern civilization and development in technology. Industrial loads, residential loads, and commercial loads are not steady, resulting mostly in the over loading of power systems. The same matter applies to Erbil governorate electric power system.

As a result, of this, there is sever need to forecast the future electrical load demands in the governorate therefore this will serve to find and evaluate the sum of the required load and prepare for the capacity of generation that would meet the electrical load demand.

The forthcoming events and situations based on earlier existing data can be defined as forecasting, the proses of making similar estimate can be called forecasting. Forecasting is important to make decisions. Load forecasting usually confused with the load prediction, however to study load forecasting, one should depend mainly on previous data that recorded. It can be said that

* Corresponding Author:

Warda Hussein Ali

E-mail: warda.ali@spu.edu.iq

Article History:

Received: 11/12/2019

Accepted: 23/01/2020

Published: 15/06 /2020

forecasting is more particular and it is able to cover a wide range of possible results. For now, a day's electrical load forecasting is a very important research area for electricity suppliers, financial organization, and transmission distribution, and other participants in electric power generation. It is the requirement estimation of electrical loads for a particular location depending on past-recorded data of electrical load demand (Bowerman et al., 2003).

With respect to the time period, there are three kinds of load forecasting, first long term load forecasting (LT), predicts electrical load from one year to 10 years, second if forecasting range is from one week up to a year, then it is said to be midterm load forecasting (MT), third, short term load forecasting (ST) relates to time period from minutes up to one day predictions.

Short-term electrical load forecasting relates to forecasting of loads from several minutes to one week ahead. A dependable (ST) forecasting supports energy suppliers and utilities to deal with the problems presented by the growth of electricity markets, and (ST) is very important as it affects strongly power system operation including estimation of variable transfer capability, power system stability margins, load elimination decisions, etc... As a result, proper load forecasting ensures more reliability in electrical power system operation while it improves the reduction of its operation cost by offering correct input day ahead scheduling.

(MT) is the second type of load forecasting, Mid-term load forecasting has a period time in a week to a year and this type of forecasting relies mainly on expansion factors such as main events increment of extra loads maintenance of large consumers and seasonal change, in this term of forecasting hourly loads are used for predictions of the day peak load or weeks peak load ahead (Feinberg and Genethlion, 2005; Ismael, 2019). (LT) is third type of load forecasting which plays a basic role for both of planners and utilities in term of progression of the grid and growth planning and it relates to time frame of one to ten years and sometime up to several decades.

A number of studies and wide range of model methodologies are presented in the study for various kind of electrical load forecasting: In

(Bruhns et al., 2005), researchers worked on model improvement of seasonality, and they studied midterm electrical load forecasting using nonlinear regression method.

Felly Njoku CF , Adewale A , Samuel IA . Carried out a study to calculate the medium-term electrical load forecasting using three regression models (Samuel et al., 2014).

In (Amjady, 2002), the rise of using intelligence techniques were shown, and many works have been performed using an approach of artificial neural network in long and short term electrical load forecasting. In addition, in (Nur et al., 2013), they proposed a method of exponential smoothing for forecasting of electrical load utilized of Malaysia.

A study on load forecasting in Karnataka, India has been performed using time series analysis, Three types of ARIMA models were developed which are Auto regressive model, ARIMA, and Auto regressive moving average. Result indicated that ARIMA model is the most reliable model (N. Amral, 2007).

The main work carried out in this study is indicating results of midterm electricity load forecasting using average the load data collected from substations related Electricity Control Center (ECC) of Kurdistan region in Iraq for Erbil Governorate. Testing and proven of the accuracy for the load forecasting has been done using mean absolute percentage error (MAPE) which measures absolute variance between the actual load demand and forecasted load values and then calculates the mean and the root mean square error (RMSE) which finds the variance between the actual load demand and the forecast load, squares the variance, calculates the mean of squares and finally calculates the square root.

The time series models employed in this study involves LR and ARIMA. The research question come to mind is; How to implement the predictive models to find the forecast load? The answer of this research question is important because it helps the practitioners to make a decision in advance about increasing or decreasing the electrical load for the next year. At the same time, it helps to know which part should be increased for the next year like adding some more electrical sub-stations

or distributing the loads to the electrical network. For that reason, we have to choose the optimum model to analyze the collected data. The organization of this study is abbreviated as following; section two introduces the electrical load profile of Erbil governorate. Section three represents methodology to analyze the data obtained experimental results. Section four describes the obtained experimental results. Finally, section five gives the conclusions.

2. ERBIL GOVERNORATE LOAD PROFILE

The main resources of electricity power supply is the governed is Erbil combined gas power plant which consists of 10 units (8 simple cycle 125

MW per unit and 2 combined cycle 250 MW per unit), which is also supply Sulaimani and Duhok Governorates. In the past years, a severe shortage of electric power supply in the Erbil Governorate has been recorded. Because the available electric power supply sources mentioned above dose not meets the power consumed in the governorate. The consumers provided by electricity for a very short period time, usually ten hours per each day and that is depending on the capacity of generation. The monthly energy consuming request (Unit in MWh) data has been taking form (ECC) Kurdistan region from 2016 to September 2019. Figure 1 captures growth of energy consumption and time periods.

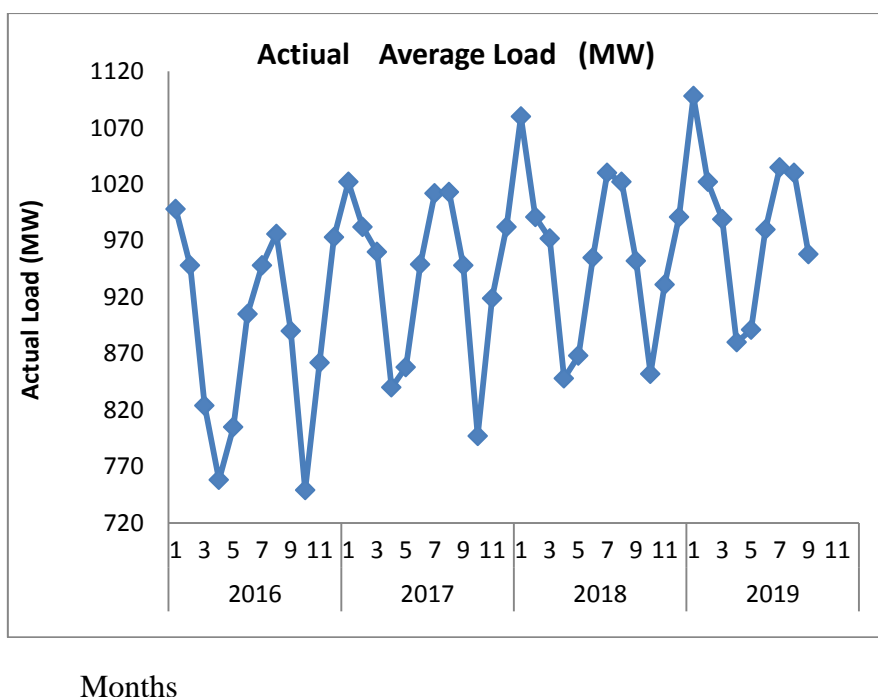


Fig.1 Monthly mean maximum demand from January 2016 to September of 2019

3.METHODOLOGY

Section bellow presents a procedure for creating an accurate model of time series that used for model electrical load forecasting in Erbil Governorate. The procedures contain selecting proper model, plotting various data, estimation of different parameter, and electrical load forecasting. The analysis is performed by using ARIMA time series modeling and linear

regression method. In this research, NCSS software is used which is a Statistical Analysis Software contains a variety of tools used for statistical tasks required in research. Here, we described the necessary background and predictive models to understand what we have done.

A. The data set

To demonstrate the high and low frequency demand features using the historical electricity load demand data measured in Megawatts (MW), which were recorded as monthly from Electricity Control Center (ECC) of the Kurdistan region of Iraq in 48 months from January 2016 to September 2019 are used.

The original behavior of monthly energy consumption of the mean load can be seen in Figure1. Numbers of factors are behind the nonlinear behavior of load demand, such as country growth, weather condition during the year and ect..It is clear from the graph it grew the upward demand every year. The variation of the series is frequently stable therefore no need for logarithmic or any other transformation. Values indicated from the plot that maximum indices are reached during January, February and December, while minimum indices are reached during April, May, September and October, and this is due to various changes in temperature.

B. Forecast Modeling

In this study, the load forecasting models used are Linear Regression Analysis and ARIMA in time series model.

Time series is a numerical analysis that deals with observation of data points, or trend analysis. To yield correct statistical inferences, these data must be repeatedly measured, often over a four to five time period.

It is a set of observations x_t , each value has been noticed at a particular time t , marked by $\{X_t\}$. It may show an achievement for the procedure that represented as follows:

$$X_t = m t + S_t Y_t \dots \dots \dots (1)$$

$$t = 1, 2 \dots n,$$

Here mt represents trend component, a trend is a stable directional change in the series.

Seasonal variations are represented in time series its marked as S_t which is a seasonal component; and it is particularly right in series that represent climate changes. and Y_t this component represents

random noise which is stationary (Brockwell and Davis, 2002).

The main purpose for modeling a time series is predicting series set data that are not deterministic in normal but random component are existed in it. mt and S_t components must be calculated and minimized as to make time series Y_t be stationary. The time series $\{X_t\}$ will consider as a stationary or non-varying if the auto covariance function and mean value of $\{X_t\}$ does not depend on time. A varying time series have to transfer to a stationary one. At that time only a suitable probabilistic time series model may be fined for Y_t to study its characteristics and to use it for forecasting objectives. In section bellow a brief summary about each of the used models are given in this sub- section.

2.Linear Regression Model

In Linear Regression method the relation between variables are found. It could be between two variables. In this case, it is named simple linear regression. If it is found between more variables, it is named multiple linear regressions.

After the relation between variables has been found, it is assumed that parameters changing with the similar relation; therefore, the similar relationship applied to the same upcoming parameters, which will give us the value of dependent variable for the matching upcoming independent variable, Linear Regressions is very simple to match the curve and calculate the coefficients. The model takes the shape of $y=mx+c$, where m is the curves slope, and the intercept point is c , and the variable x is independent variable and $f(x)$ is the dependent variable. The work is to calculate the parameters m and c with the help of variable data of x and y (Dara et al., 2013).

3.ARIMA Model

ARIMA represent auto regressive integrated moving average and is defined by three variables: (p, d, q) .

Using the past values in the linear regression equation in the time series Y , refers to auto

regressive (AR(p)) component . The parameter p indicates the number of delays carried out by the model.

where ϕ_1, ϕ_2 are parameters for the model.

The rate of deviation in the integrated (I(d)) component represented by d. Subtracting series current values and series previous values d times is represent differencing of a series. Often, to make a series stable whenever a stationary assumption dose not met, differencing

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q} \dots \dots \dots (2)$$

Where {Zt} is defined as a series of uncorrelated random variables with zero average and unchanged variance, and the polynomials $(1-\phi_1 z - \dots - \phi_p z^p)$ and $(1+\theta_1 z + \dots + \theta_q z^q)$ having no shared factors between them .

The process {Xt} may called an ARMA (p, q) process that has mean μ if {Xt- μ } is an ARMA (p, q) process and easily written in the briefer form of $\phi(B)X_t = \theta(B) Z_t \dots \dots \dots (3)$

Where $\theta(\cdot)$, $\phi(\cdot)$ are respectively the qth and pth degrees of the polynomials,

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \dots \dots \dots (4)$$

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q \dots \dots \dots (5)$$

B represents backward shift carrier ($B^j X_t = X_{t-j}$, $B^j Z_t = Z_{t-j}$, $j=0, \pm 1, \dots$)

The series {Xt} is said to be an autoregressive procedure taking the degree p if $\phi(z) = 1$ and having moving average procedure of degree q if $\theta(z) = 1$ (Brockwell and Davis, 2002).

After the power load data has been pre-treated and numerically tested, the calculation of the model order and parameters are also necessary. Akaike Information Criterion (AIC) method is performed to evaluate the order of time series models that is shown in equation 6: [3]

$$AIC = Ln (\sigma^2) + \frac{2(P + q)}{T} \dots \dots \dots (6)$$

Where:

- T present numbers of non-missing values in the series.

is used. Error of the time series model as a combination of prior error terms represented by moving average (MA(q)) component. The number of terms to involve in the model represents the order q.

A non-varying ARMA (p, q) time series model is shown as a sequence for random variables {Xt}, represented by:[5].

- P represents the degree of the AR component model.
- q represents the degree MA component model
- σ represents the standard deviation of the residuals.

It is well established, at least among statisticians of some higher caliber, that models with the values of the AIC statistic within a certain threshold of the minimum value should be considered as appropriate as the model minimizing the AIC statistic. (Li et al., 2014; Charles et al, 1999).

C. Evaluation of Performance using MAE and RMSE

Time series forecasting performance measures; provide a summary of the ability of the forecast model that performs the forecasting. Two measures are used in this work to measure the performance of electrical load forecasting, they are: first is the root mean square error (RMSE) and second is the mean absolute error (MAE) (Okolobah and Ismail, 2013).

MAE is defined in equation (7) as:

$$E = (La - Lf) \dots \dots \dots (7)$$

Where:

- E represents the error
- Lf represents the forecast load
- La represents the actual demand load

N represents the number of values
 The RMSE is shown in equation (8) as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n E_i^2} \dots\dots\dots (8)$$

1. THE RESULTS

In this section, we have shown our results with respect to the two predictive models:

1) Results Using the Linear Regression Model

Table I indicates the forecasted load values and actual load values, the values applying the linear regression model from 2016 to 2109 and forecasting load for 2020, in the meanwhile Figure 2 represents a comparison of the forecast and actual electric load values from January 2016 to September 2019 and forecasted load for 2020. The MAE and RMSE values are evaluated using linear regression models and the calculated values of MAE were 14.84 and RMSE value is 20.22.

Table I: forecasted and Actual load values applying regression method

Year	Month	Actual Load in MW	Forecasting load in MW
2016	1	998	993.90
	2	948	931.11
	3	824	909.00
	4	758	798.96
	5	805	813.93
	6	905	897.96
	7	948	961.46
	8	976	968.81
	9	890	896.54
	10	749	768.88
	11	862	869.60
	12	973	945.46
2017	1	1022	1026.71
	2	982	961.76
	3	960	938.84
	4	840	825.12
	5	858	840.51
	6	949	927.20
	7	1012	992.68
	8	1013	1000.19
	9	948	925.50

	10	797	793.65
	11	919	897.54
	12	982	975.76
2018	1	1080	1059.52
	2	991	992.41
	3	972	968.69
	4	848	851.28
	5	868	867.09
	6	955	956.44
	7	1030	1023.91
	8	1022	1031.56
	9	952	954.46
	10	852	818.42
	11	931	925.47
	12	991	1006.05
2019	1	1098	1092.33
	2	1022	1023.06
	3	989	998.53
	4	880	877.44
	5	891	893.66
	6	980	985.68
	7	1035	1055.13
	8	1030	1062.94
	9	958	983.42
2020	1	Predictive Value	1125.14
	2	Predictive Value	1053.72
	3	Predictive Value	1028.37
	4	Predictive Value	903.60
	5	Predictive Value	920.24
	6	Predictive Value	1014.92
	7	Predictive Value	1086.35
	8	Predictive Value	1094.32
	9	Predictive Value	1012.37
	10	Predictive Value	867.95
	11	Predictive Value	981.35
	12	Predictive Value	1066.64

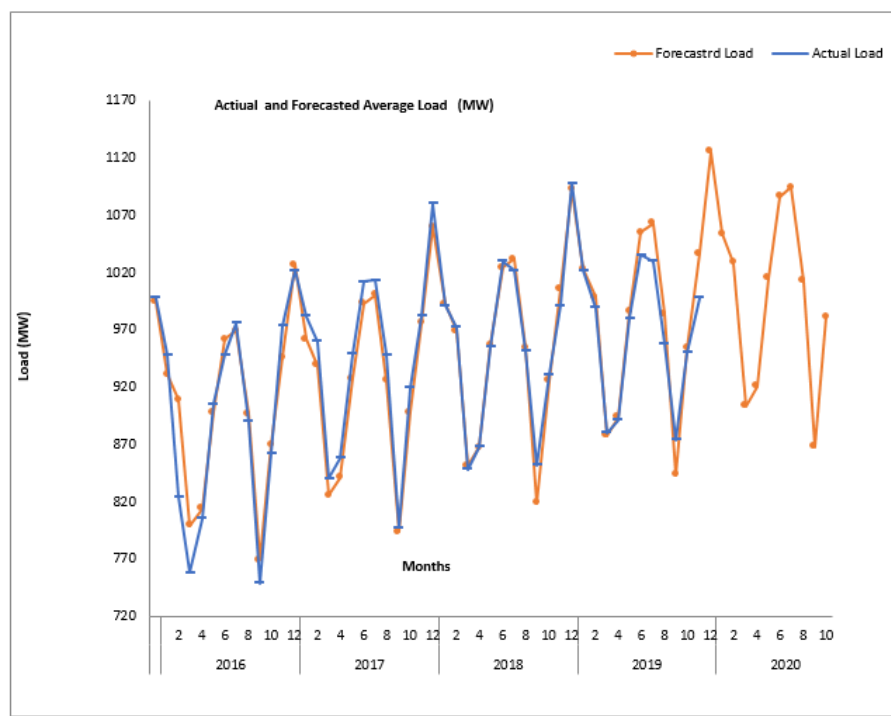


Fig.2 Comparison between the forecasted and actual load demands of Erbil from January 2016 – December 2020 using the Linear Regression Method.

Table II: The list of ARIMA potential models

2) Results Using the auto-regressive integrated moving average ARIMA Model

Table II indicates the computed ARIMA model for forecasting the load with their respective values of AIC.

From Table II, ARIMA (7, 1, 1) has the minimum AIC indicates that it is the most optimum model among the other ARIMA models.

The best satisfactory model for ARIMA forecasting can be proven by utilizing the accuracy criteria such as MAE and RMSE, which are given by the respective equations (7 and 8), and the result was MAE=17.05 and MASE =19.32. Table III represents the actual electrical load and the forecast average load values utilizing ARIMA model from 2016 to 2109 and forecasting load for 2020, while Figure 3 shows the forecast values from January 2016 to December 2020 and upper and lower limit of forecasted load for 2020.

ARIMA Models	(AIC)
(1,0,1)	22.74197
(1,1,1)	22.87393
(0,1,1)	22.64192
(2,1,1)	21.81226
(2,1,2)	21.46635
(1,1,3)	21.60749
(2,1,3)	21.57948
(1,1,4)	21.53301
(5,1,1)	20.36306
(7,1,1)	20.11954

Table III: shows actual and forecast load values by ARIMA.

Year	Months	Actual Load (MW)	Forecasting load (MW)
2016	1	998	980.2
	2	948	937.7
	3	824	812.9
	4	758	782.8
	5	805	803.4
	6	905	897.1
	7	948	938.6
	8	976	960.8
	9	890	885.4
	10	749	774.4
	11	862	831
	12	973	991.6
2017	1	1022	1013.8
	2	982	978.1
	3	960	889.2
	4	840	875.7
	5	858	839.7
	6	949	941.8
	7	1012	993.7
	8	1013	1026
	9	948	936.5
	10	797	833.7
	11	919	900.5
	12	982	1007.9
2018	1	1080	1019.6
	2	991	1044.1
	3	972	990.4
	4	848	885.1
	5	868	893.8
	6	955	971.1
	7	1030	1027.5
	8	1022	1018.6
	9	952	905.9
	10	852	798.1
	11	931	814.9
	12	991	882.7

2019	1	1098	963.3
	2	1022	937.1
	3	989	891
	4	880	822.7
	5	891	851.9
	6	980	933.1
	7	1035	999.3
	8	1030	943.5
	9	958	912.3
	10	874	839.8
	11	950	844.4
	12	998	908.7
2020	1	Predictive Value	939.7
	2	Predictive Value	1019
	3	Predictive Value	965.3
	4	Predictive Value	910
	5	Predictive Value	964.9
	6	Predictive Value	995.5
	7	Predictive Value	1042
	8	Predictive Value	1008.9
	9	Predictive Value	987.4
	10	Predictive Value	948.2
	11	Predictive Value	950.7
	12	Predictive Value	990.7

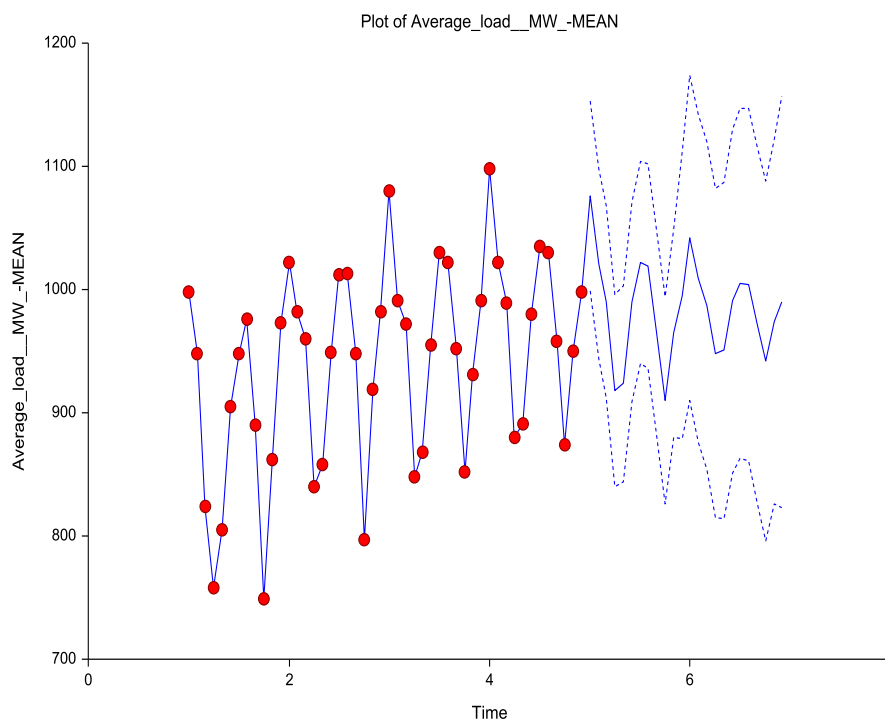


Fig.3 Forecasts data from January-2016 to Desember-2020using ARIMA (note: 1, 2, 3, 4.... Represent number of year)

5.CONCLUSION

The present research has been carried out using the peculiarity of Erbil governorate in Kurdistan, in which data was collected from the Electricity Control Center (ECC) of the Kurdistan region of Iraq to discover which type of load forecasting method between the two methods described above, has the most positively respond to the electrical load data presented. As a result, we have compared the values of MAE and the RMSE, one can conclude that ARIMA method is much better to use for the electrical load forecasting than the first method of regression analysis method because of the following reasons: First, the ARIMA was capable to forecast the electrical load data in spite of the fall (decrease) in load demands in May and April. ARIMA did not forecast electrical load only, but forecasts the future electrical load demands with a much minimized error if the results are compared to the actual electrical load demands. Second, Because of its high accuracy, and great precision, ARIMA considered being more robust to forecast electrical

load demand. This will be helpful in future research studies to perform load forecasting for a long term forecasting.

Third, ARIMA method produces results much faster than regression analysis because of the direct arithmetical calculations while regression analysis needs some mathematical computations before it can begin forecasting electrical load data.

REFERENCES

- Amjady N. "Short-term Hourly load Forecasting Using Time Series Modeling with Peak Load Estimation Capability", Institute of Electrical and Electronics Engineers Transactions on Power Systems, 2002 Aug, 16(4),798–805.
- Bowerman BL, O'Connell RT, Koehler AB, "Forecasting, Time Series, and Regression: an Applied Approach. 4th edition Thomson Brooks/Cole, California; 2003..
- Brockwell P.J. , Davis, R.A, "Introduction to Time Series and Forecasting", Springer Texts in Statistics. Second Edition. New York, 2002, Springer-Verlag.
- Bruhns A, Deurveilher G, Roy J-S, "A non-linear Regression Model for Mid-term load Forecasting and Improvements in Seasonality" 15th Power

- Systems Computation Conference (PSCC), Liege; 2005,p. 1–8.
- Charles A., Greenhall, A. D., et al, “Total Variance, an Estimator of Long-term Frequency Stability “, IEEE Transactions on Ultrasonic, Ferroelectrics and Frequency Control, (1999), 46(5), 1183–1191.
- Dara H. Amin Mohammed, Mohammed A. Husain, and Ismael K. Saeed, “Voltage Stability Analysis and its Enhancement Using Static Shunt and Series Capacitors of KRPS” ,Zanco Journal of Pure and Applied Sciences, Salahaddin University-Erbil, Vol. 25, No. 2, 2013.
- Feinberg E.A. , Genethlion, D,”Load Forecasting Applied Mathematics for Restructured Electric Power Systems: Optimization, Control, and Computational Intelligence”, 2005,pp. 269-285. Springer.
- Ismael Kareem Saeed, “Artificial Neural Network Based on Optimal Operation of Economic Load Dispatch in Power System”, Zanco Journal of Pure and Applied Sciences, Salahaddin University-Erbil, Vol. 31, No. 4, 2019.
- Li Xiaojing ,CHEN Jiabin, YONG Shangguan. “A Method to Analyse and Eliminate Stochastic Noises of Fog Based on ARMA and Kalman Filtering Method”, Intelligent Human Machine Systems and Cybernetics (IHMSC),2014.
- N. Amral C. S. Ozveren and D. King, “Short Term Load Forecasting using Multiple Linear Regression”, 42nd Universities Power Engineering International Conference (UPEC 2007), 1192-1198 ,2007.
- Nur Adilah Abd Jalil, Maizah Hura Ahmad and Norizan Mohamed, “Electricity Load Demand Forecasting Using Exponential Smoothing Methods”, World Applied Sciences Journal 22, 1540-1543 ,2013
- Okolobah VA, Ismail Z,” New Approach to Peak load Forecasting Based on EMD and ANFIS”, Indian Journal of Science and Technology, 2013 ,Dec; 6(12):5600–6.
- Samuel IA, Felly-Njoku CF, Adewale AA, Awelewa AA. “Medium term load Forecasting of Covenant University Using the Regression Analysis Methods”, International Journal of Energy Technologies and Policy (IISTE), 2014, 4(4):10–6.