Forecasting of Electricity Consumption and Supply for Campus University using Time Series Models

Rosnalini Mansor#1, Bahtiar Jamili Zaini#2, Chong Shi Yee*3

#School of Quantitative Sciences, Universiti Utara Malaysia, 06010 UUM Sintok Kedah Malaysia

*Gamuda Land (T12) Sdn. Bhd., 47820 Petaling Jaya, Selangor, Malaysia

1rosnalini@uum.edu.my

2bahtiar@uum.edu.my

3sychong92@hotmail.com

***Abstract*— Electricity is an important energy source in university as lecture classes need electricity supply to function. It is also important for the development of the university. Since electricity consumption is a necessity of a university's operation, the forecast of electricity consumption on the university campus should be made. This is essential for the development of the university as the treasury department can manage the funding from the government according to the value forecasted to make full use of the funding in the university's development. There are several forecasting methods used in this study, including time series regression, seasonal exponential smoothing, Box-Jenkins (SARIMA), decomposition and the naïve method. Error measurements used to evaluate the performance of forecasting model were mean square error (MSE), root mean square error (RMSE), mean absolute percentage error (MAPE) and geometric root mean square error (GRMSE). The results of this study showed that the seasonal exponential smoothing model was the best in the 1-step ahead and 2-step ahead forecasting while SARIMA $(0,2,2)(0,2,1)_{12}$ was the best in the 3-step ahead forecast. The overall performance of seasonal exponential smoothing was the best in this study. Throughout this study, suggestions were made for the next study regarding electricity consumption in university to consider factors such as semester breaks and students' activities in order to examine its effect in electricity consumption.**

Keywords*— *Forecasting, Electricity Consumption, Univariate Time Series, Box-Jenkins

1. Introduction

Electrical consumption is the total amount of energy used represented by kilowatt hours (kWh). It is different from load demand which means the immediate rate of that consumption (kW). For example, for a light bulb using 100 watts of electricity that is switched on for 10 consecutive hours, the consumption is 1kWh. Alternatively, ten 100 watts light bulbs switched on at the same time for an hour has the same consumption (1kWh), but its load demand is 1kW of electricity to operate. The forecast electricity consumption in terms of kWh is due to the policy in Malaysia. Electricity tariffs will change, therefore, forecasting based on electricity charges in Ringgit Malaysia (RM) is not meaningful. Electricity consumption (kWh) is more representative as it shows the actual usage every month without the influence of electricity tariff.

Electricity is an important energy source in each country. It is also important for the development of a university. Other than basic facilities such as lecture halls, student residential halls, and the library, a university also provides facilities such as the sports center, health center, food court, smart reading room, gym room, swimming pool, and many more. Without the supply of electricity, electronic components such as lights, fans, air-conditioners, projectors and computers cannot function. This will affect the development and learning process of students in that campus. The monthly expense for the electricity bill is different as the activities in the university would affect it. Electricity tariff also changes from time to time, thus, studies regarding electricity consumption are more representative. Therefore, our objective for this study is to evaluate the performance of several forecasting models by using the four error measures to then forecast the monthly electricity consumption.

2. Literature Review

Articles related to electricity consumption in Malaysia and foreign countries were studied. Researchers used different forecasting methods in their study according to data type and regional factors. Chujai, Kerdprasop and Kerdprasop [1] forecasted the electricity consumption in an individual household by performing the Autoregressive Integrated Moving Average (ARIMA) and Autoregressive Moving Average (ARMA) model. The data was from December 2006 to November 2010. The result shows that ARIMA is suitable for monthly and quarterly forecasting, while ARMA is suitable for daily and weekly forecasting. Besides that, Box-Jenkins Seasonal Autoregressive Integrated Moving Average (SARIMA) was carried out to forecast electricity consumption in Malaysia [2].

International Journal of Supply Chain Management
IJSCM, ISSN: 2050-7399 (Online), 2051-3771 (Print)

Researchers also used ARIMA in an empirical quantitative evaluation to conduct very short-term forecasting to compare two customer types [3].

The Holt-Winters method has two approaches that are additive and multiplicative. It can be used to deal with time series with both trend and seasonal variations. The forecasting of Electric Consumption in a semiconductor plant using Winter's method was conducted by Boonkham and Surapatpichai [4] in Thailand. Lepojevic and Andelkovic-Pesic [5] also predicted the electricity consumption covering an area in Serbia using the Holt-Winters method. Besides that, there was also an evaluation of some classical methods in forecasting electricity usage for the Washington Water Power by Javedani et al. [6]. Data were analyzed using naïve, Winter's, decomposition, regression and ARIMA. The result showed that Winter's method was the best forecast.

The forecast of electrical consumption of Elektroistok Ltd. Nis by using Holt-Winters and seasonal regression models was carried out and it found that seasonal regression adapted better to empirical data, therefore it was considered more reliable [5]. Imtiaz et al. [7] also conducted a research to forecast the long-term electricity consumption in Malaysia by using the multivariate time series regression method. Gross domestic product (GDP), real electricity prices and population were taken as the factors. Although the electricity load demand data was different to electricity consumption data, however, the effect factors can be used for both types of data. For example, factor temperature was considered in both consumption and demand forecasting [8], [9] and holiday effect [10], [11]. However, most of the research used electricity data only in their forecasting method due to efficiency of time and cost.

A comparison between neural network and Box-Jenkins models was carried out to model and forecast the electricity in Malaysia as the neural network was able to analyze the unseen part correctly even when it contains noisy information [2]. Results showed that neural network with data pre-processing performed better than the Box-Jenkins method. Zhang et al. [12] conducted a review of modelling issues of neural network forecasting, and the overall performance was satisfied. Neural network was also used in a comparative analysis to forecast electricity consumption at University Malaysia Sarawak [13].

3. Methodology

3.1 Data Collection and Partition

Data was collected from the Development and Maintenance Department of a university and was divided into the estimation part and evaluation part. Data from January 2009 to July 2014 was classified under the estimation part while data from August 2014 to December 2014 was under the evaluation part. Sixty-seven data were used in the estimation part to carry out its process while five data were used in the evaluation part to determine the error measure using four different error measures. For the estimation part, this study will discuss the ways to identify the four components in a time series which are trend, seasonal, cyclical and irregular fluctuation. The most suitable component to our data will then be identified.

3.2 Method of Data Analysis

Firstly, the forecasting method chosen are time series regression, exponential smoothing, Box-Jenkins (ARIMA), decomposition and naïve method. If the collected data satisfies all the assumptions, the forecast of electricity consumption will be conducted. Then, error measures such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE) and Geometric Root Mean Squared Error (GRMSE) will be carried out. It will be used to evaluate the forecasting performance. Then, the best model to forecast electricity consumption will be chosen. Figure 1 shows the study framework.

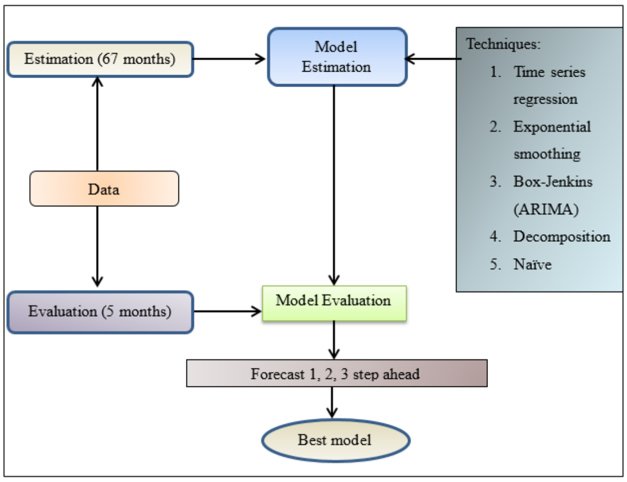

Figure 1. Framework of Study

After collecting the historical data, the identification of time series component needed to be done before proceeding to the next step of this study. An observed time series can be decomposed into four components which are trend, seasonal, cyclical and irregular.

Lastly was to evaluate the error for each model. We used four error measurements such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute

Percentage Error (MAPE) and Geometric Root Mean Squared Error (GRMSE). According to Lazim [14], these four error measurements are considered as the most popular among researchers and practitioners to measure forecasting model accuracy. From the error measure, the model with the smallest error measure was selected as the best forecasting model.

4 Results and Discussion

After collecting the historical data, this study proceeded to analyze the data components. The determination of data component is relatively important in determining the performance between methods. There are four types of time series components which are trend, seasonal, cyclical and irregular. Figure 2 shows that there is a slightly downward tendency ($consumption$ = 3960657.6 - 633.88004t) in the historical data. From the analysis of variance, we can see that the significance level at 0.8677 (p = 0.87) is greater than 0.05. Since the result is shown to be insignificant, this study excludes the trend component from the model.

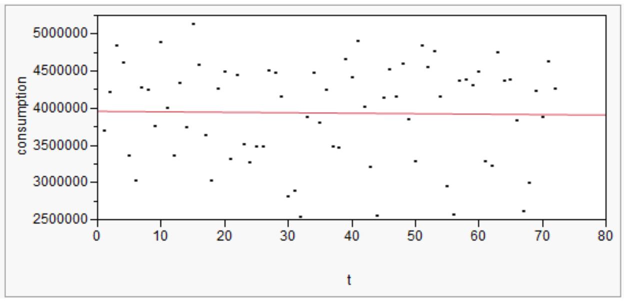

Figure 2. Trend line of electricity consumption

4.1 Seasonal Component

Seasonal component is a fluctuation that occurs within a period of time with the same regulatory pattern. It can be identified by observing the fluctuation on a yearly basis with a consistent pattern. It can also be identified by plotting the autocorrelation function (ACF) and partial auto correlation function (PACF) or by calculating the seasonal indices. From Figure 3, it can be observed that the diagram fluctuated left and right from 0 with different autocorrelation values which can be concluded that the seasonal component existed.

4.2 Cyclical Component

Cyclical component refers to the rise and fall of the series over an unspecified time. Normally, it occurs due to economic fluctuations which is an economic-business cycle phenomenon. The residual method was used to identify the cyclical component [14]. However, the residual method result showed that the relative cyclical residuals were mostly in a negative value. Positive values that occurred could not be assumed as cyclical as it only occurred once each time. Therefore, there was no cyclical component in this study.

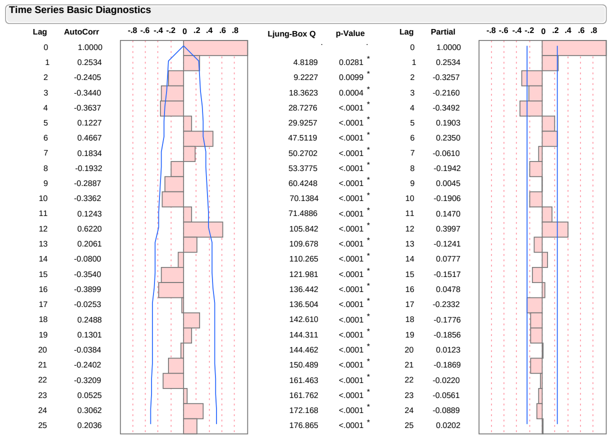

Figure 3. ACF and PACF graph

4.3 Irregular Component

Irregular component is an erratic movement in a time series that is unpredictable. From Figure 4, we can see that there was an abnormal scenario for the consumption between July-August for year 2009 and 2010. The electricity consumption between July-August for these two years were higher than the same period compared to other research years. The difference of electricity consumption for August 2010 (4483337.00 kWh) and August 2011 (2536287.25 kWh) was 43.43%. However, the difference was caused by changes in the academic calendar in the university, so it was not an irregular component.

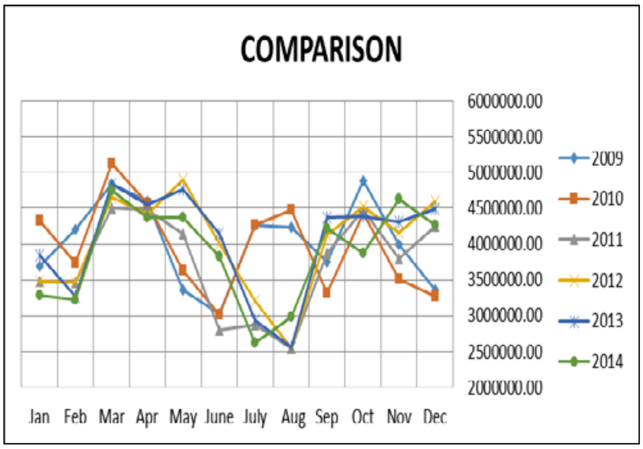

Figure 4. Comparison of electricity consumption over 6 consecutive years

4.4 Model Estimation

After all the time series components in the data were identified, forecasting models that suited the data behavior were developed and the results were analyzed. All the forecasting models will be explained in next sub-section.

4.4.1 Time Series Regression

The JMP software was used to run this model. It gave the model, *consumption* = 3960658 – 633.88*t*. From the analysis of variance, it showed that the significance level was 0.87 (p=0.87), which is greater than 0.05. This indicates that the test is not significant. Therefore, it can be concluded that the time series regression is not suitable to forecast university electricity consumption.

4.4.2 Exponential Smoothing- Seasonal

Seasonal exponential is more suitable for this study as it has no trend but has the seasonal component only. From the result of JMP in Table 1, the model is significant since the seasonal smoothing weight is less than 0.05. Thus, it can be concluded that seasonal additive exponential smoothing model is suitable to forecast university electricity consumption.

Table 1. Parameter Estimates for Winters' Additive

Term	Estimate	Std Error	t Ratio	Prob >|t|
Level Smoothing Weight	0.09228	0.051808	1.78	0.0802
Seasonal Smoothing Weight	1	0.146776	6.81	0.0001

4.4.3 Box-Jenkins (SARIMA)

The first step in Box-Jenkins is to determine the existence of a seasonal component and determine whether it is stationary. From the ACF and PACF graph shown in Figure 3, the result showed that the seasonal component existed in our data and it was stationary. There were seven spikes in the ACF and PACF graph. Therefore, the Seasonal Autoregressive Integrated Moving Average (SARIMA) was used.

This study used JMP to run the ARIMA model group, which is a function to determine the best Box-Jenkins model by setting p (autoregressive order), d (differencing order) and q (moving average order) from the range of 0 to 2 while the periods per season was 12. Table 2 shows 15 outcomes of the SARIMA model with its AIC and SBC values. The AIC and SBC values were used to determine the best models among the Box-Jenkins model.

Out of 729 models that JMP analyzed, SARIMA $(0,2,2)(0,2,1)_{12}$ was the best as its AIC and SBC value was 1373.602, and 1380.917 respectively was the lowest among all models. From Table 3, it shows that the model is significant.

Table 2. SARIMA Model Outcome

Model	DF	AIC	SBC
Seasonal ARIMA $(0, 2, 2)(0, 2, 1)_{12}$	42	1373.602	1380.917
Seasonal ARIMA $(1, 2, 2)(0, 2, 1)_{12}$	41	1374.095	1383.238
Seasonal ARIMA $(0, 2, 2)(0, 2, 2)_{12}$	41	1375.481	1384.624
Seasonal ARIMA $(0, 2, 2)(1, 2, 1)_{12}$	41	1375.486	1384.630
Seasonal ARIMA $(1, 2, 2)(0, 2, 2)_{12}$	40	1375.550	1386.522
Seasonal ARIMA $(1, 2, 2)(1, 2, 1)_{12}$	40	1375.554	1386.526
Seasonal ARIMA $(2, 2, 2)(1, 2, 1)_{12}$	39	1377.033	1389.834
Seasonal ARIMA $(0, 2, 2)(1, 2, 2)_{12}$	40	1377.249	1388.221
Seasonal ARIMA $(0, 2, 2)(2, 2, 1)_{12}$	40	1377.472	1388.444
Seasonal ARIMA $(0, 2, 2)(2, 2, 0)_{12}$	41	1377.514	1386.658
Seasonal ARIMA $(1, 2, 2)(2, 2, 1)_{12}$	39	1377.554	1390.354
Seasonal ARIMA $(1, 2, 2)(2, 2, 0)_{12}$	40	1377.736	1388.708
Seasonal ARIMA $(1, 2, 2)(1, 2, 0)_{12}$	41	1377.834	1386.978
Seasonal ARIMA $(1, 2, 2)(1, 2, 2)_{12}$	39	1377.862	1390.662
Seasonal ARIMA $(0, 2, 2)(1, 2, 0)_{12}$	42	1377.918	1385.233

Table 3. Parameter Estimates for SARIMA

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob >|t|
MA1,1	1	1	1.722	0.243	7.1	<.0001
MA1,2	1	2	-0.722	0.197	-3.67	0.0007
MA2,12	2	12	1	0.415	2.41	0.0206
Intercept	1	0	-2421.4	2188.87	-1.11	0.2749

4.4.4 Decomposition Method

Decomposition is popular among forecasters because it is easy to understand. It decomposes the time series into sub-components to study its effect. First, we needed to discover the adjusted seasonal component of the data and calculate the trend estimation, where the formula used was $\hat{y}_t = T_t + S_t$. The calculation in Microsoft Excel is shown in Appendix 5. The fitted value of decomposition is shown in Table 4.

Table 4. Fitted Value for Decomposition

		Fitted Value		
t	Actual Value	1-step	2-step	3-step
68	2988000	3889092.7		
69	4222971	3834803.3	3887803.9	
70	3867548	3854871.3	3832363.1	3886515.9
71	4624101	3853612.8	3852881.9	3829891.4
72	4263452	3895067.8	3851654.4	3850908.1

4.4.5 Naïve Model

The naïve model was performed. It is the simplest method which assumes the current value that would be the next forecasted value. The fitted value of the naïve model is shown in Table 5.

Table 5: Fitted Value of Naïve Model

		Fitted Value		
t	Actual Value	1-step	2-step	3-step
68	2988000	2613322		
69	4222971	2988000	2613322	
70	3867548	4222971	2988000	2613322
71	4624101	3867548	4222971	2988000
72	4263452	4624101	3867548	4222971

4.5 Model Evaluation

The previous section had obtained the fitted value of 1-step ahead, 2-step ahead and 3-step ahead forecast for each model. Fitted value can be evaluated by using four different error measurement methods which are MAPE, MSE, RMSE and GRMSE. The model with the lowest error measure value is the best. In this study, the performance of each model is ranked in order to determine the best model for 1-step ahead, 2-step ahead and 3-step ahead forecast. The comparison between model for 1-step, 2-step and 3-step is made below. The model with the lowest ranking is the best.

4.5.1 One-step Ahead Forecast

The 1-step ahead forecast was used to forecast the electricity consumption for the next month. Table 6 shows the comparison between models for the 1-step ahead forecast. They had been ranked according to their error measure performance followed by their ranking that was highlighted.

4.5.2 Two-step Ahead Forecast

The 2-step ahead forecast was used to forecast the electricity consumption for the second month. Table 7 shows the comparison between models for the 2-step ahead forecast. They had been ranked according to their error measure performance followed by their ranking that had been highlighted.

Table 6. Comparison between Models for 1-Step Ahead Forecast

One-step Ahead Forecast					
	MSE	RMSE	MAPE	GRMSE	Total Rank
SARIMA	5.33E+11	729900.3	16.376	456542.8	
Rank	4	4	4	3	15
Seasonal Exponential Smoothing	1.49E+11	386199.7	8.831	222354.9	
Rank	1	1	1	1	4
Decomposition	3.4E+11	583184.4	13	264586.8	
Rank	2	2	2	2	8
Naïve	4.99E+11	706300.6	15.159	537522.8	
Rank	3	3	3	4	13

Table 7. Comparison between models for 2-step ahead forecast

Two-step Ahead Forecast					
	MSE	RMSE	MAPE	GRMSE	Total Rank
SARIMA	1.429E+11	377997.59	8.153	288835.96	
Rank	2	2	2	3	9
Seasonal Exponential Smoothing	9.201E+10	303327.61	5.693	147534.54	
Rank	1	1	1	1	4
Decomposition	2.195E+11	468489.88	8.796	247382.92	
Rank	3	3	3	2	11
Naïve	9.206E+11	959455.55	19.705	688599.08	
Rank	4	4	4	4	16

Table 8. Comparison between models for 3-step ahead forecast

Three-step Ahead Forecast					
	MSE	RMSE	MAPE	GRMSE	Total Rank
SARIMA	9.99E+10	316122.5	7.21	283050.8	
Rank	1	1	2	3	7
Seasonal Exponential Smoothing	1.06E+11	325299.4	7.053	264293.5	
Rank	2	2	1	2	7
Decomposition	2.67E+11	516824	9.114	183854.8	
Rank	3	3	3	1	10
Naïve	1.42E+12	1190455	22.92	436327.3	
Rank	4	4	4	4	16

4.5.3 Three-step Ahead Forecast

The 3-step ahead forecast was used to forecast the electricity consumption for the third month. Table 8 showed the comparison between models for the 3-step ahead forecast. They had been ranked according to their error measure performance followed by their ranking that had been highlighted.

4.6 Forecast Electricity Consumption

Based on results from tables 6 to 8, the seasonal exponential smoothing model was the best forecasting model to forecast the future value of electricity consumption. This study applied the best forecasting model to forecast 1-step ahead, 2-step ahead and 3-step ahead forecast of electricity consumption. Table 9 shows the result of electricity consumption forecast.

Table 9. Electricity consumption forecast

Forecast	Fitted Value
1-step ahead	3222815.263
2-step ahead	3157280.276
3-step ahead	4688927.013

5 Conclusion

In conclusion, the seasonal component is the only component that exists in the historical data of this study. This study ran the data with five models and evaluated it with four types of error measurements. Time series regression was not suitable for this study, therefore, it was excluded. Four models were tested and the results showed that seasonal exponential smoothing was the best in 1-step ahead and 2-step ahead forecast since error measurement such as MSE, RMSE, MAPE and GRMSE had the least value and had the lowest ranking, while SARIMA $(0,2,2)(0,2,1)_{12}$ and seasonal exponential smoothing had the same ranking in the 3-step ahead forecast. Overall, the performance of seasonal exponential smoothing was the best among the four models.

Acknowledgments

The authors are appreciative for the financial support received from Universiti Utara Malaysia under the Research Generating University Grant (S/O Code: 13878). We also thank the anonymous reviewers for their useful recommendations.

References

[1] P. Chujai, N. Kerdprasop, and K. Kerdprasop, "*Time Series Analysis of Household Electric Consumption with ARIMA and ARMA Models*", in Proceedings of the International MultiConference of Engineers and Computer Scientists, vol. I, 2013

[2] M. Y. Lai and H. C. Ong, "*Neural networks forecasting on electricity consumption in Malaysia*", The 2nd IMT-GT Regional Conference on Mathematics, Statistics and Their Applications. 2005

[3] S. Aman, M. Frincu, C. Chelmis, M. Usman Noor, Y. Simmhan, and V. Prasanna, "*Empirical Comparison of Prediction Methods for Electricity Consumption Forecasting*", USC Tech. Rep. 2014.

[4] P. Boonkham and S. Surapatpichai, "*Forecasting of Electric Consumption in a Semiconductor Plant using Time Series Methods*", Journal of Energy Technologies and Policy, vol. 3, no. 11, pp. 170–177, 2013.

[5] V. Lepojevic and M. Andelkovic-Pesic, "*Forecasting electricity consumption by using holt-winters and seasonal regression models*", Economic and Organization, vol. 8, no. 4, pp. 421–431, 2011.

[6] H. Javedani, M. H. Lee, and Suhartono, "*An Evaluation of Some Classical Methods for Forecasting Electricity Usage on Specific Problem*", Sciences New York, vol. 2010, no. June, pp. 47–56, 2010.

[7] A. K. Imtiaz, N. B. Mariun, M. Amran M R, M. Saleem, N. I. a Wahab, and Mohibullah, "*Evaluation and forecasting of long term electricity consumption demand for Malaysia by statistical analysis*", in First International Power and Energy Conference, (PECon 2006) Proceedings, pp. 257–261, 2006,

[8] J. B. Jorgensen and F. Joutz, "*Modelling and Forecasting Residential Electricity Consumption in the U . S . Mountain Region*", Research Program and forecasting working paper No. 003, pp. 1–42, 2012.

[9] Z. Ismail and R. Mansor, "*Fuzzy Logic Approach for Forecasting Half-hourly Malaysia Electricity*", in International Institute of Forecasters (ISF) 2011 Proceedings, 2011

[10] S. Ismail, M. Zulkifli, R. Mansor, M. Mat Yusof, and M. I. Ismail, "*The Role of Exploratory Data Analysis (EDA) in Electricity Forecasting*", Pertanika Journal of Social Sciences & Humanities, vol. 24 (S), pp. 111–118, 2016.

[11] F. A. Razak, A. H. Hashim, I. Z. Abidin, and M. Shitan, "*Moving holidays' effects on the Malaysian peak daily load*", 2010 IEEE Internatioal Conference of Power Energy, pp. 906–910, Nov. 2010.

[12] G. Zhang, B. E. Patuwo, and M. Y. Hu, "*Forecasting with artificial neural networks: The state of the art*", International journal of Forecasting, vol. 14, no. 1, pp. 35–62, 1998.

[13] P. Ozoh, S. Abd-Rahman, J. Labadin, and M. Apperley, "*A Comparative Analysis of Techniques for Forecasting Electricity Consumption*". *International Journal of Computer Applications.*, vol. 88, no. 15, pp. 8–12, 2014.

[14] M. A. Lazim, *Introductory Business Forecasting: a practical approach*, 3rd Edition. UiTM Press, 2013.