International Journal of Engineering & Technology, 7 (4) (2018) 3515-3520

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i4.17372
Research paper

Forecasting electricity consumption by multiple linear regression

K. G. Tay [1] *, Y. Y. Choy [2], Audrey Huong [3]

[1] Department of Communication Engineering, Faculty of Electrical and Electronic Engineering, UTHM
[2] Department of Mathematics, Faculty of Applied Science and Technology, UTHM
[3] Department of Electronic Engineering Faculty of Electrical and Electronic Engineering, UTHM
**Corresponding author E-mail: tay@uthm.edu.my*

Abstract

Electricity consumption forecasting is crucial for effective operation, planning and facility expansion of the power system. An accurate forecasts can save operating and maintenance costs. As a result, increased the reliability of power supply and delivery system. Universiti Tun Hussein Onn Malaysia (UTHM) is a developing Malaysian Technical University. There is a great development of UTHM infrastructure since its formation in 1993. The development will be accompanied with the increasing demand for electricity. Hence, there is a need to forecast UTHM electricity consumption for future decisions on generating electric power, load switching, and infrastructure development. The UTHM load demand was forecasted by using multiple linear regression (MLR). The monthly load demand from January 2011 to August 2018 was used to forecast January to August 2019 monthly load demand. MLR can forecast the UTHM load demand quite well with mean absolute percentage error (MAPE) of 10.62%. MLR was then compared with curve fitting methods from an Excel spreadsheet.

Keywords: *Forecasting; Electricity Consumption; UTHM; Multiple Linear Regression; MAPE.*

1. Introduction

Forecasting is predicting future values based on past and current time series data. Forecasting for future load demand is essential for future power system planning and control.

Load forecasting can be divided into short-term load forecasting (STLF), medium-term load forecasting (MTLF) and long-term load forecasting (LTLF). STLF up to one day or one week at most, MTLF ranges from one day to several months while LTLF forecasts more than a year ahead [1]. STLF is used for scheduling the generation and transmission of electricity, MTLF is used to plan the fuel purchases, whereas LTLS is aimed to develop the power supply and delivery system (generation units, transmission system, and distribution system) [2].

Universiti Tun Hussein Onn Malaysia (UTHM) is one of the Malaysian technical university located in Parit Raja, Batu Pahat, Johor Malaysia. It was formerly known as Pusat Latihan Staf Politeknik (PLSP) which was established in1993. There is a great development of UTHM infrastructure after UTHM has been upgraded from Institut Teknologi Tun Hussein Onn (ITTHO) in 1996, Kolej Universiti Teknologi Tun Hussein Onn (KUiTTHO) in July 2001 to UTHM in January 2007.

There are four new buildings in UTHM there are Faculty of Technical and Vocational Education (FPTV), Faculty of Technology Management & Business (FPTP), Faculty of Civil and Environmental Engineering (FKAAS), Faculty of Computer Science and Information Technology (FSKTM) and several new buildings (Complex of Faculty of Electronic and Engineering (FKEE) and Faculty of Mechanical and Manufacturing (FKMP), as well as new library, were built. However, in August 2017, three faculties of UTHM which is Faculty of Applied Science and Technology (FAST), Faculty of Technology Engineering and Diploma Studies Centre were moved to Pagoh, Johor. On top of that, there is a new building of FKEE, which is completed in 2018.

The development of new buildings and movement of the above three faculties to Pagoh will affect the electricity demand. Hence, there is a need to forecast UTHM electricity consumption for future decisions on generating electric power, load switching, and infrastructure development.

The multiple linear regression (MLR) method is a popular forecasting method due to its simplicity. It is often used in load forecast affected by a number of factors ranging from meteorological effects, per capita growth, electricity prices, economic growth etc [3].

Perry [4] used MLR to forecast STL of an electric utility where current temperature, the temperature two hours ago, the time of the day, a dummy variable for whether it is a weekend or weekday, and the electric demand (kW) two hours ago were used as input variables.

A MLR model was used by Mohamed and Bodger [5] to forecast the electricity consumption of New Zealand where the independent variables were a gross domestic product (GDP), electricity price and population.

Kandananond [6] forecasted electricity demand in Thailand using the autoregressive integrated moving average (ARIMA), artificial neural network (ANN) and MLR. The independent variables of Kandananond [6] MLR model were population, stock index, GDP and export.

Amral, Özveren and King [7] conducted STLF of Sulewesi Island– Indonesia's power system by using MLR where independent variable was temperature.

STLF was carried on by Kumar, Mishra and Gupta [8] by using MLR where the independent variables were time, mean tempera-

ture, dew point, humidity, wind speed, the day of the week, date and holiday.
Kaytez et.al [9] forecasted electricity energy consumption of Turkey by using ANN, MLR and least squares support vector machines. in their MLR model. Installed capacity, gross electricity generation, population, and total subscribership were chosen as independent variables.
Hahn et al [10] reviewed methods used on electric load forecasting ranging from regression-based models, time series approach, neural networks, support vector machines, hybrid and other approaches.
Kyriakides and Polycarpou [11] presented various approaches from conventional to computational intelligence methods to the STLF.
In this paper, we have only UTHM monthly historical electricity consumption from the January 2011 to August 2018 and wish to utilize MLR method to forecast UTHM 2019 monthly electricity consumption. The MLR will be compared with curve fitting methods from Excel.

2. Multiple linear regression (MLR)

The general form of multiple linear regression model follows [12] is shown as below:

$$y_i = \beta_o + \beta_1 t_i + \beta_2 M_{1i} + \cdots + \beta_{13} M_{12i} + \varepsilon_i \tag{1}$$

where y represents the real electricity consumption, t is time period, $\beta_0, \beta_1, \beta_2, ...\beta_{13}$ are the coefficients of the actual electricity consumption,

$$M_{ji} = \begin{cases} 1, & \text{if month j} \\ 0, & \text{if other month} \end{cases}$$

ε_i is the residual, i is the i-th data.
The predicted response is given as below:

$$\hat{y}_i = b_o + b_1 t_i + b_2 M_{1i} + \cdots + b_{13} M_{12i} \tag{2}$$

Where $\hat{y}$ represents the predicted electricity consumption, t is time period, $b_0, b_1, b_2, ...b_{13}$ are the coefficients of the forecasted electricity consumption of MLR
The error between the observed and the predicted is

$$\varepsilon_i = y_i - \hat{y}_i = y_i - b_o - b_1 t_i - b_2 M_{1i} - \cdots - b_{13} M_{12i} \tag{3}$$

The sum of square of the errors (SSE) in Equation (3) is

$$SSE = \sum_{i=1}^{n} \left[y_i - \left(b_0 + b_1 t_i + b_2 M_{1i} + \cdots + b_{13} M_{12i} \right) \right]^2,$$

Where n is number of data. A good MLR model of thirteen independent variables is obtained by minimizing the SSE by differentiating SSE with respect to parameters b_0, b_1, … and b_{13}. This yields to

$$\begin{aligned}
& nb_0 + b_1 \sum_{i=1}^{n} t_i + b_2 \sum_{i=1}^{n} M_{1i} + \cdots + b_{13} \sum_{i=1}^{n} M_{12i} = \sum_{i=1}^{n} y_i \\
& b_0 \sum_{i=1}^{n} t_i + b_1 \sum_{i=1}^{n} t_i^2 + b_2 \sum_{i=1}^{n} t_i M_{1i} + \cdots + b_{13} \sum_{i=1}^{n} t_i M_{12i} = \sum_{i=1}^{n} t_i y_i \\
& b_0 \sum_{i=1}^{n} M_{1i} + b_1 \sum_{i=1}^{n} M_{1i} t_i + b_2 \sum_{i=1}^{n} M_{1i}^2 + \cdots + b_{13} \sum_{i=1}^{n} M_{1i} M_{12i} = \sum_{i=1}^{n} M_{1i} y_i \\
& \vdots \\
& b_0 \sum_{i=1}^{n} M_{12i} + b_1 \sum_{i=1}^{n} M_{12i} t_i + b_2 \sum_{i=1}^{n} M_{12i} M_{1i} + \cdots + b_{13} \sum_{i=1}^{n} M_{12i}^2 = \sum_{i=1}^{n} M_{12i} y_i
\end{aligned}$$

By solving the above system of linear equation, the coefficients of $b_0, b_1, b_2, ...b_{13}$ can be obtained. Here, the coefficient of the MLR can be found easily using the Data Analysis tool in Excel spreadsheet as the following steps:
Step 1: Install the Data Analysis ToolPak in Microsoft Excel by clicking 'File' at the menu bar >> select 'Options' >> click 'Add-Ins' >> click 'Analysis ToolPak' >> click 'Go…' >> tick the 'Analysis ToolPak' at the 'Add-Ins' table >> click 'Ok'.
Step 2: Click 'Data' menu >> click 'Data Analysis' >> choose 'Regression' >> click 'Ok'.
Step 3: Selecting the 'Actual load' label and its data as 'Input Y Range' while 'Period' until 'M1 to M12' labels and their data as 'Input X Range'. Then check on the 'Labels' and Confidence Levels checkbox. Next, click the 'Output Range' and type in any desired cell in EXCEL. Finally, click 'Ok'.

3. Evaluation Performance

The performance of the MLR can be measured by mean absolute error (MAPE) as below:

$$MAPE = \frac{\sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i}}{n} \times 100\%,$$

where $y_i, \hat{y}_i$ are real and forecasted load respectively, n is the number of real data.

4. Results and discussion

UTHM electricity consumption patterns versus month for years 2011-2018 is shown in Figure 1. It is noticed that electricity consumption has increased since the year 2011. The electricity consumption fluctuates for each month. The lowest value is in February 2011 (1683.617 MWh), while the highest value is in March 2015 (3228.53MWh). On the other hand, the electricity consumption in December 2017 is the minimum if compared to the same December month as three faculties of UTHM moved to Pagoh since August 2017. The year 2015 has the highest electricity consumption, while the year 2011 has the minimum electricity consumption if compared to other years in most of the months. The electricity consumption for certain months are low than usual may because that month is mostly semester break of UTHM. There are fewer students in the campus and therefore, the electricity consumption will be lesser if compared to the months that are not semester break.

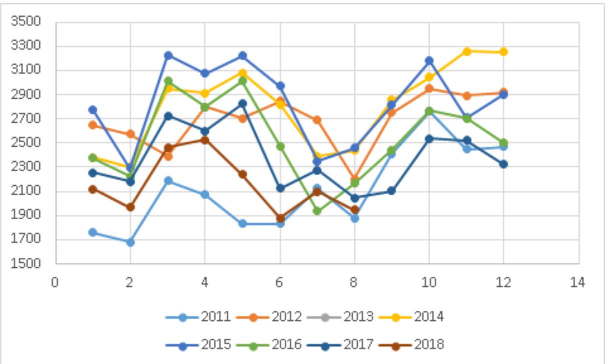

Fig. 1: Actual Electricity Consumption by Years.

Figure 2 shows the time series of UTHM electricity consumption continuously from January 2011- August 2018. The electricity consumption is range from 1700 MWh to 3300 MWh. The time series seems is not stationary and is increasing till 2015 then decreasing. The electricity consumption from January 2011 to August 2019 will be forecasted by using MLR.

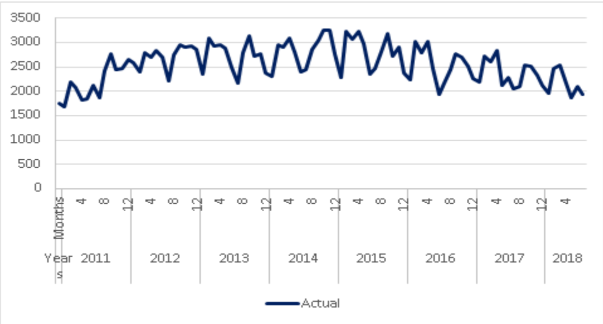

Fig. 2: UTHM Electricity Consumption for All Years.

By performing MLR using Equation (1) following the steps in Section 2, MLR model was obtained as

$$\begin{aligned}\hat{y} &= 2625.65 - 0.712t - 199.642M_1 - 398.016M_2 + 162.069M_3 \\ &+ 121.148M_4 + 140.498M_5 - 114.615M_6 - 296.458M_7 \\ &- 427.688M_8 + 0M_9 + 317.427M_{10} + 157.681M_{11} + 143.506M_{12}\end{aligned}$$

The actual UTHM electricity consumption (blue colour) from January 2011 to August 2018 and forecasted electricity consumption (red colour) from January 2011 to August 2019 were depicted in Figure 3. It shows MLR can forecast the UTHM electricity consumption according to the actual electricity consumption pattern closely if compared to linear, quadratic, cubic, exponential, logarithmic and power trends as displayed in Figures 4-9 respectively with MAPE 10.16%.

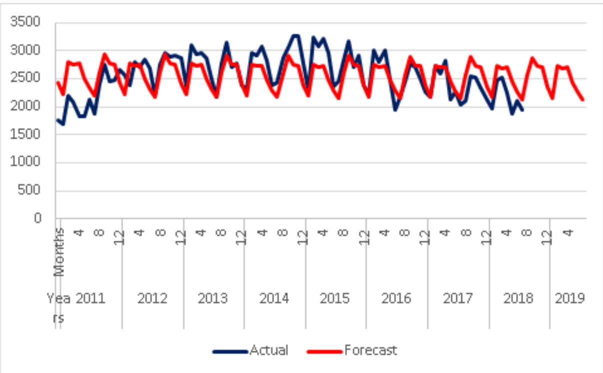

Fig. 3: Actual and Forecasted Electricity Consumption.

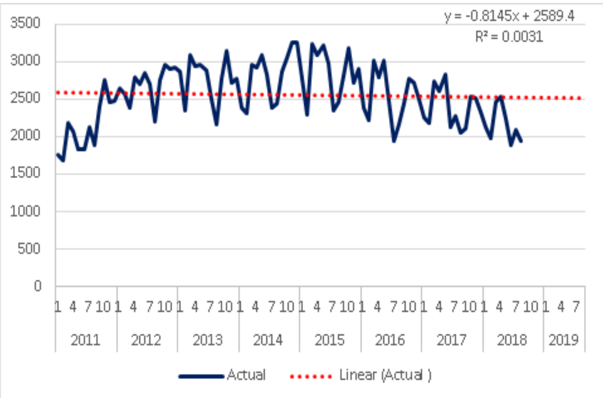

Fig. 4: Actual Electricity Consumption and Linear Trend.

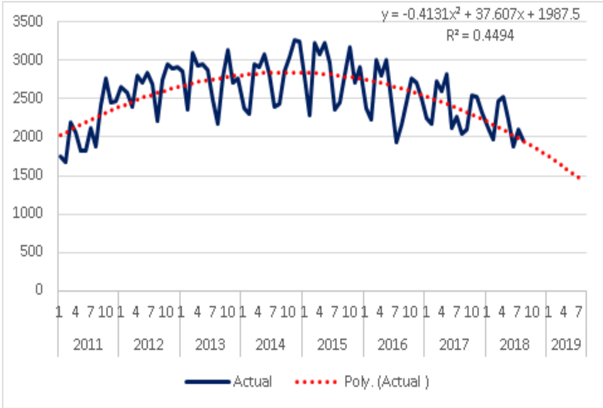

Fig. 5: Actual Electricity Consumption and Quadratic Trend.

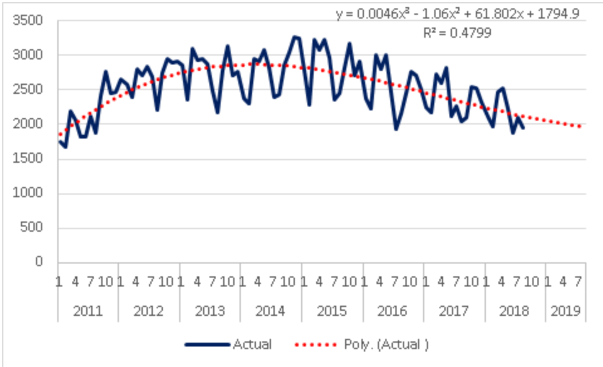

Fig. 6: Actual Electricity Consumption and Cubic Trend.

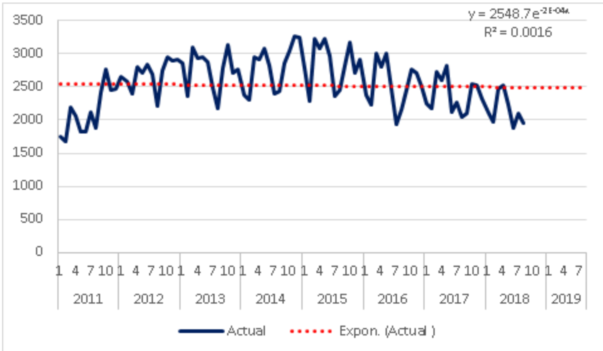

Fig. 7: Actual Electricity Consumption and Exponential Trend.

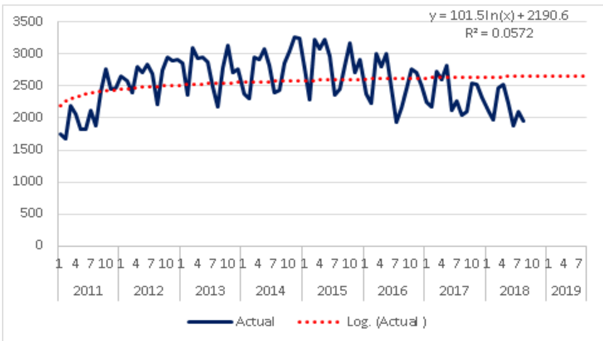

Fig. 8: Actual Electricity Consumption and Logarithmic Trend.

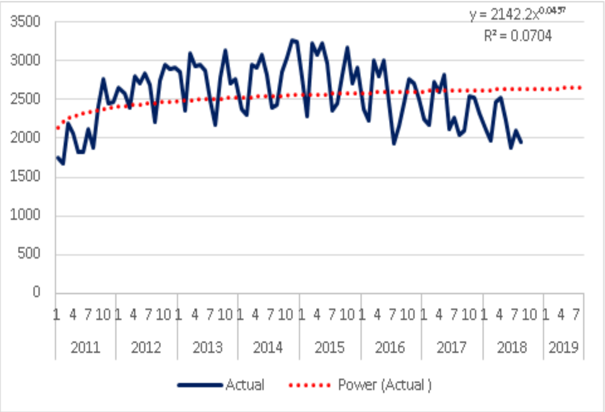

Fig. 9: Actual Electricity Consumption and Power Trend.

Table 1 gives the correlation coefficients, R^2 and MAPE values for MLR, linear, quadratic, cubic, exponential, logarithmic and power trends. The cubic trend has the highest R^2 and smallest MAPE followed by quadratic and MLR. A perfect model should has the maximum $R^2 = 1$ and 0% MAPE value, from the value of R^2 and MAPE seems like the cubic trend is the best model, but MLR model produces similar pattern and trend if compared to cubic and quadratic trends. Hence, MLR should be a comparable good model in terms of R^2, MAPE and pattern.

Table 1: R^2 and MAPE

	R^2	MAPE (%)
MLR	0.3755	10.16
Linear	0.0031	13.61
Quadratic	0.4494	10.08
Cubic	0.4799	9.56
Exponential	0.0016	13.69
Logarithmic	0.0572	13.47
Power	0.0704	13.41

Table 2 displays the monthly data for actual and forecasted load demands for UTHM from January 2011 to December 2018. The load is in MWh.

Table 2: Actual and Forecasted Electricity Consumption by MLR

Years	Months	Actual	Forecasted
2011	1	1757.133	2425.2966
	2	1680.617	2226.2103
	3	2187.953	2785.5835
	4	2071.238	2743.9508
	5	1829.331	2762.5881
	6	1831.845	2506.7636
	7	2123.014	2324.2086
	8	1876.535	2192.2668
	9	2410.113	2619.2426
	10	2761.56	2935.9579
	11	2448.26	2775.4992
	12	2470.5	2760.6124
2012	1	2646.807	2416.7528
	2	2572.654	2217.6665
	3	2389.16	2777.0397
	4	2797.864	2735.4070
	5	2704.19	2754.0443
	6	2843.698	2498.2198
	7	2689.711	2315.6648
	8	2202.056	2183.7230
	9	2752.318	2610.6989
	10	2948.825	2927.4141
	11	2893.017	2766.9554
	12	2919.263	2752.0686
2013	1	2855.407	2408.2091
	2	2350.236	2209.1227
	3	3090.031	2768.4959
	4	2932.166	2726.8632
	5	2956.513	2745.5006
	6	2875.193	2489.6761
	7	2494.238	2307.1211
	8	2170.468	2175.1792
	9	2777.709	2602.1551
	10	3135.956	2918.8704
	11	2711.885	2758.4116
	12	2772.699	2743.5248
2014	1	2379.815	2399.6653
	2	2302.285	2200.5789
	3	2947.008	2759.9521
	4	2914.361	2718.3194
	5	3077.515	2736.9568
	6	2820.286	2481.1323
	7	2389.876	2298.5773
	8	2438.02	2166.6354
	9	2854.192	2593.6113
	10	3046.04	2910.3266
	11	3258.29	2749.8679
	12	3250.95	2734.9810
2015	1	2774.32	2391.1215
	2	2291.53	2192.0351
	3	3228.53	2751.4084

	4	3077.08	2709.7756
	5	3219.13	2728.4130
	6	2968.51	2472.5885
	7	2349.65	2290.0335
	8	2458.26	2158.0916
	9	2811.92	2585.0675
	10	3177.03	2901.7828
	11	2,710.42	2741.3241
	12	2904.37	2726.4372
2016	1	2376.343	2382.5777
	2	2224.489	2183.4913
	3	3012.959	2742.8646
	4	2797.192	2701.2318
	5	3011.201	2719.8692
	6	2476.527	2464.0447
	7	1936.78	2281.4897
	8	2163.858	2149.5478
	9	2444.475	2576.5237
	10	2768.369	2893.2390
	11	2706.516	2732.7803
	12	2502.642	2717.8934
2017	1	2256.096	2374.0339
	2	2179.734	2174.9475
	3	2725.83	2734.3208
	4	2597.001	2692.6880
	5	2826.637	2711.3254
	6	2122.3	2455.5009
	7	2275.021	2272.9459
	8	2044.741	2141.0040
	9	2104.552	2567.9799
	10	2534.506	2884.6952
	11	2520.687	2724.2365
	12	2324.443	2709.3496
2018	1	2117.226	2365.4901
	2	1968.911	2166.4037
	3	2463.971	2725.7770
	4	2525.478	2684.1442
	5	2236.962	2702.7816
	6	1876.524	2446.9571
	7	2096.153	2264.4021
	8	1944.97	2132.4602
	9		2559.4361
	10		2876.1514
	11		2715.6927
	12		2700.8058
2019	1		2356.9463
	2		2157.8600
	3		2717.2332
	4		2675.6005
	5		2694.2378
	6		2438.4133
	7		2255.8583
	8		2123.9165

5. Conclusion

MLR was applied to forecast UTHM electricity consumption from January 2011 to August 2018. It can forecast UTHM electricity consumption from January 2011 to August 2019 quite well with MAPE of 10.62% if compared to other curves fitting methods in Excel. The predicted results may be served as a catalyst for possible actions by the management team. For instance, if the management team find the electricity increases drastically then they may educate users to save the usage of electricity consumption.

Acknowledgement

We would like to thank Mr. Shukur Saleh and Mr. Abd Rashid Puteh from Development and Maintenance Office, UTHM for providing us UTHM electricity consumption data, UTHM ORICC research fund, and UTHM Tier 1 2018 research grant vote number H258 for financial support of this project.

References

[1] Pedregal DJ & Trapero JR (2010), Mid-term hourly electricity forecasting based on a multi-rate approach. *Energy Conversion and Management* 51, 105–111. https://doi.org/10.1016/j.enconman.2009.08.028.

[2] Almeshaiei E & Soltan H (2011), a methodology for Electric Power Load Forecasting. *Alexandria Engineering Journal* 50, 137–144. https://doi.org/10.1016/j.aej.2011.01.015.

[3] Singh AK, Ibraheem, Khatoon S & Muazzam Md (2013), An Overview of Electricity Demand Forecasting Techniques. *National Conference on Emerging Trends in Electrical, Instrumentation & Communication Engineering* 3, 38-48.

[4] Perry C, (1999), Short-Term Load Forecasting; Using Multiple Regression Analysis. *Rural Electric Power Conference,* 1-8. https://doi.org/10.1109/REPCON.1999.768683.

[5] Mohamed Z & Bodger P (2005), Forecasting electricity consumption in New Zealand using economic and demographic variables. *Energy* 30, 1833–1843. https://doi.org/10.1016/j.energy.2004.08.012.

[6] Kandananond K (2011), Forecasting Electricity Demand in Thailand with an Artificial Neural Network Approach. *Energies* 4, 1246-1257. https://doi.org/10.3390/en4081246.

[7] Amral N, Özveren CS & King D (2007), Short Term Load Forecasting Using Multiple Linear Regression. *Universities Power Engineering Conference*, 1192-1198. https://doi.org/10.1109/UPEC.2007.4469121.

[8] Kumar S, Mishra S & Gupta S (2016), Short Term Load Forecasting Using ANN and Multiple Linear Regression. *2016 Second International Conference on Computational Intelligence & Communication Technology*, 184-186.

[9] Kaytez F, Taplamacioglu MC, Cam E & Hardala F (2015), Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. *Electrical Power and Energy Systems* 67, 431–438. https://doi.org/10.1016/j.ijepes.2014.12.036.

[10] Hahn H, Meyer-Nieberg S & Pickl S (2009), Electric load forecasting methods: Tools for decision-making. *European Journal of Operational Research*. 199, 902-907. https://doi.org/10.1016/j.ejor.2009.01.062.

[11] Kyriakides E & Polycarpou M (2007), Short-term electric load forecasting: A tutorial. In: Chen, K., Wang, L. (Eds.), Trends in Neural Computation, Studies in Computational Intelligence, Springer, 35, 391–418 (Chapter 16).

[12] Lepojević V & Anđelković-Pešić M (2011), Forecasting Electricity Consumption by Using Holt-Winters and Seasonal Regression Models. *Facta Universitatis Series: Economics and Organization* 8, 421 – 431.