Universal Journal of Electrical and Electronic Engineering 6(5B): 26-36, 2019 http://www.hrpub.org
DOI: 10.13189/ujeee.2019.061605

Electricity Consumption Forecasting Using Nonlinear Autoregressive with External (Exogeneous) Input Neural Network

K. G. Tay[1,*], Hassan Muwafaq[1], Shuhaida Binti Ismail[2], Pauline Ong[3]

[1]Faculty of Electrical and Electronics Engineering, Universiti Tun Hussein Onn Malaysia, Malaysia
[2]Faculty of Applied Science and Technology, Universiti Tun Hussein Onn Malaysia, Malaysia
[3]Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, Malaysia

Received August 1, 2019; Revised October 5, 2019; Accepted December 12, 2019

Abstract Forecasting is prediction of future values based on historical data. Electricity consumption forecasting is crucial for utility company to plan for future power system generation. Even though there are previous works of electricity consumption forecasting using Artificial Neural Network (ANN), but most of their data is multivariate data. In this study, we have only univariate data of electricity consumption from January 2009 to December 2018 and wish to do a prediction for a year ahead. On top of that, our data consist of autoregressive component, hence Nonlinear Autoregressive with External (Exogeneous) Input (NARX) Neural Network Time Series from Matlab R2018b was used. It gives the mean absolute percentage error (MAPE) between actual and predicted electricity consumption of 1.38%.

Keywords ANN, NARX, Electricity Consumption, Forecasting

1. Introduction

Forecasting is predicting future values based on historical data. Electricity consumption forecasting is important for Utilities Company to plan for future power generation and distribution. Overestimation of electricity demand will cause the wasting of resources, while underestimation will lead to higher operation cost [1]. 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, MTF ranges from one day to several months while LTF forecasts more than a year ahead [2]. 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) [3].

Various techniques have been applied in electricity consumption forecasting including holt-winters and seasonal regression [4], time series models [5], first-order fuzzy time series [6], multiple linear regression [7-8], autoregressive integrated moving average (ARIMA) [9], seasonal ARIMA (SARIMA) [10], artificial neural network (ANN) [8, 10-14], support vector machine (SVM) [14], Least square SVM (LSSVM) [1], support vector regression [11], etc.

The gross domestic product, number of population, maximum ambient temperature and electrical power demand were used as inputs in a neural network to predict electricity consumption in Thailand. [8]

Research in [10] employed monthly electricity consumption data from 1970 to 2009 as training data and 2010 consumption data as test data. Besides, years and months of data were selected as input while the consumption was considered as output.

Bayesian Regularization algorithm in the autoregressive neural network was utilized in [11] for electricity consumption forecasting and they obtained 95% accuracy.

Kandananond [13] applied multilayer perceptrons (MLP) neural network on historical data of population, gross domestic product (GDP), stock index (SET index), total revenue from exporting industrial products (export) and electricity consumption from 1986 to 2010 to predict future electricity consumption. He obtained a mean absolute percentage error (MAPE) of 0.996%

ANN seems to be the recently popular machine learning methods used in electricity consumption forecasting. However, in ANN literature, researchers [8,13] adopted multivariate data, where electricity is the dependent variable and other independent variables which influenced the electricity consumption such as temperature, population, GDP, etc as input.

Even though consumption data in [10-11] is univariate, research in [10] added year and month index to transform univariate data to multivariate data. Whereas, lag 12 and lag 24 of consumption data, seasonal, time and month indexes were added in [11] to make it multivariate. A study in [11] obtained 13.8% and 3.9% of MAPE for training and testing data respectively.

Our data only consists of monthly electricity consumption of University Tun Hussein Onn Malaysia (UTHM) which is univariate similar to [10-12], but we will make it multivariate as shown in Table 1. On top of that, the structure of our input data has an autoregressive component, hence a nonlinear autoregressive neural network with External (Exogeneous) Input (NARX) will be built to forecast UTHM future electricity consumption by using Matlab R2018b.

Table 1. Data structure

Input					Output
M_1	M_2	M_3	$\cdots$	M_{12}	M_{13}
M_2	M_3	M_4	$\cdots$	M_{13}	M_{14}
M_3	M_4	M_5	$\cdots$	M_{14}	M_{15}
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
M_{107}	M_{108}	M_{109}	$\cdots$	M_{118}	M_{119}
M_{108}	M_{109}	M_{110}	$\cdots$	M_{119}	M_{120}

2. Artificial Neural Network (ANN)

ANN is one type of machine learning inspired by the biological neural network. ANN consists of one input layer, at least one hidden layer and one output layer as seen in Figure 1.

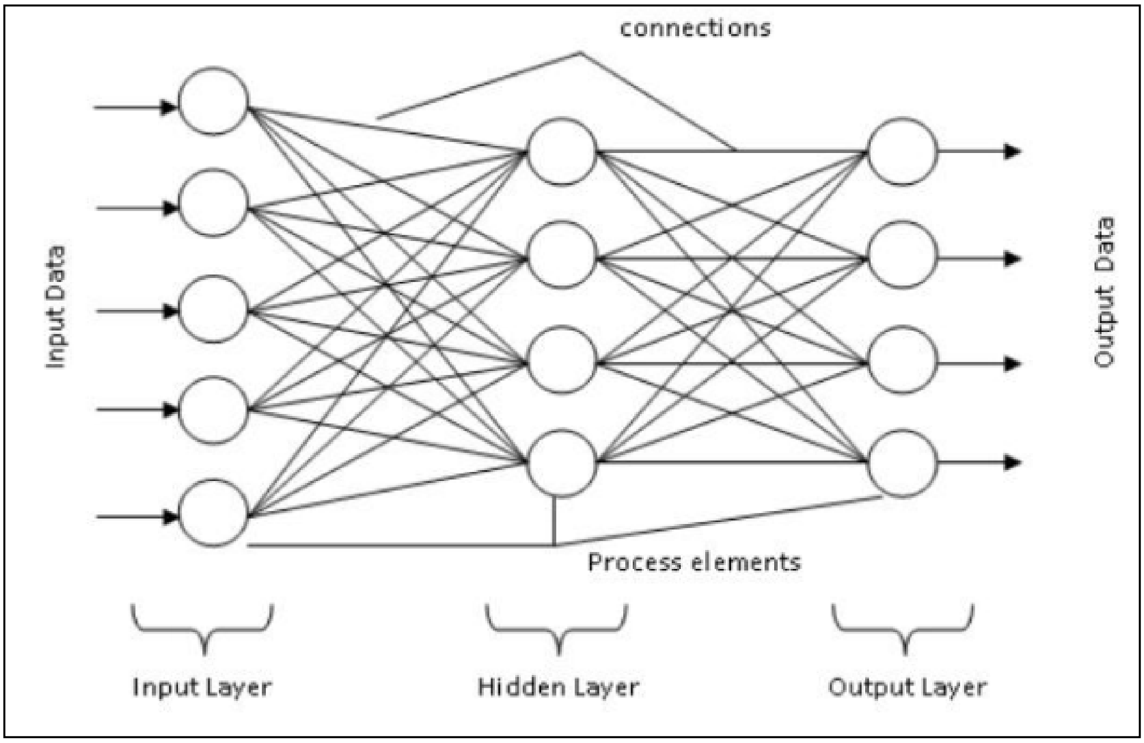

Figure 1. Architectural of ANN [10]

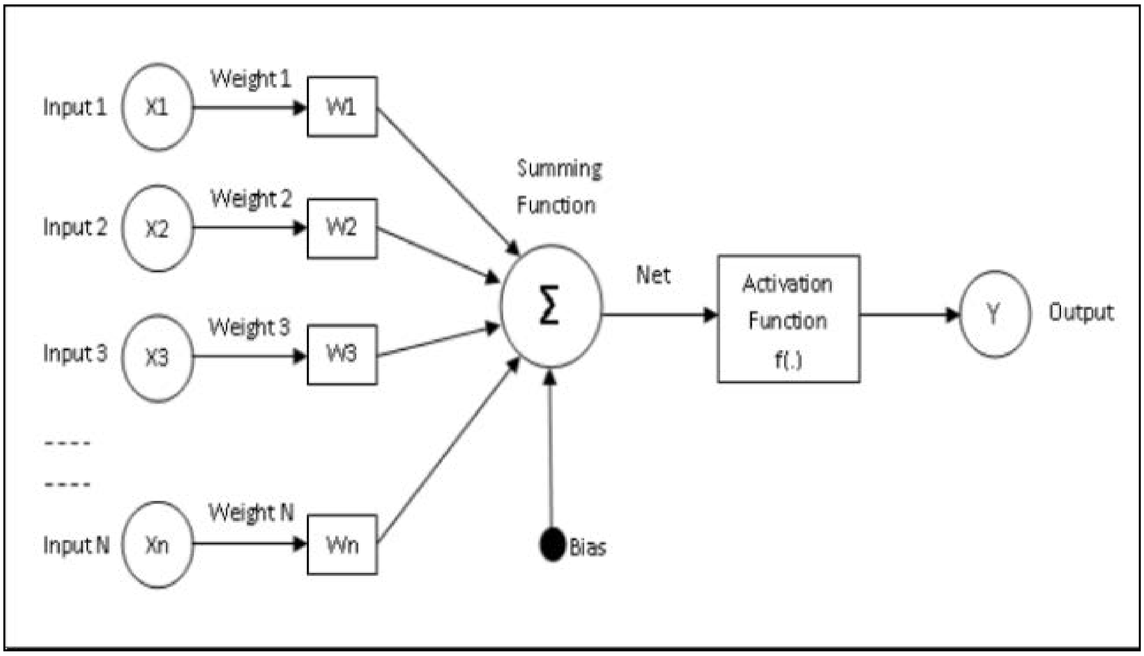

Figure 2. NN working principle [10]

ANN is a supervised machine learning where input and output must be supplied to the neural network (NN). The NN will learn the relationship between the input and output and then give a prediction to the output. It will learn in such a way to minimize the error of prediction.

Figure 2 shows each input of ANN is processed in nodes/neurons. Each neuron X_i has it own weight W_i . Summation of the product of each neuron by its weight $(\sum W_i X_i)$ is called a net. The net is inputted in an activation function and gives the prediction output:

$$Y_i = f(\sum W_i X_i)\ Y_i = f(\sum W_i X_i) \tag{1}$$

where Y_i is the output of node i, $f(.)$ is the activation function. One of the common activation functions is logistic sigmoid function [15]

$$f(x) = \frac{1}{1+e^{-x}} \tag{2}$$

3. Material and Methods

The monthly UTHM electricity consumption data from January 2009 to December 2018 was used to forecast UTHM monthly electricity consumption for the year 2019.

3.1. Data Structure

The 120 data was arranged in Table 1 such that the input is 108 rows by 12 columns, while the output is the last column of dimension 108 rows by 1 column starting from month 13 to month 120 as given in Table 1. Here M_i is the month i of UTHM electricity consumption data from January 2009 to December 2018.

3.2. Methodology

In this study, Artificial Neural Network (ANN) Apps (*Neural Net Time Series*) from Matlab R2018b was used to do forecasting. The Nonlinear Autoregressive with External (Exogeneous) Input (NARX) NN (the first option) as seen in Figure 3 was used as it gives the best prediction. NARX is a recurrent dynamic network with feedback connections enclosing several layers of the network [16]. In this case, $x(t)$ is input consisting of 12 columns with 108 rows, while output is $y(t)$ that is 1 column with 108 rows.

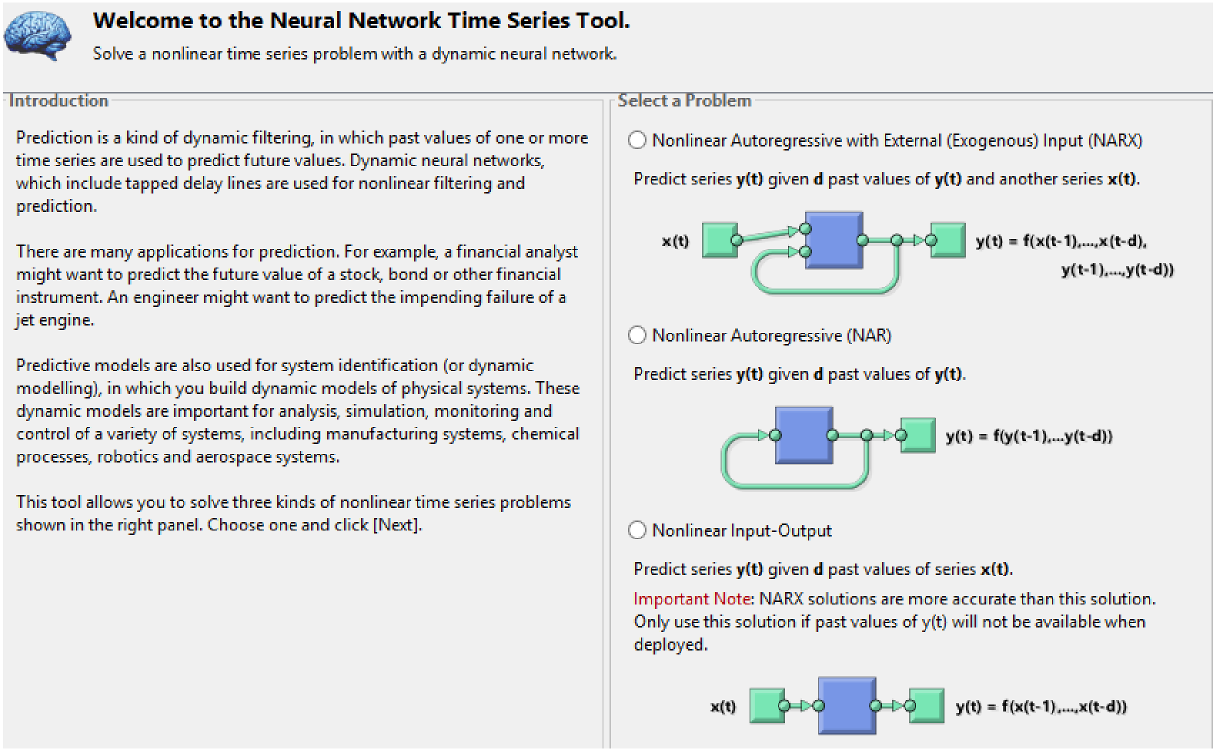

Figure 3. Neural Net Time Series

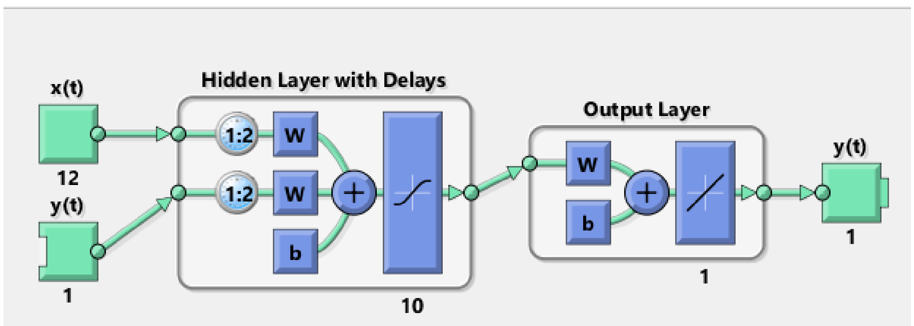

Figure 4. NARX structure

Figure 4 shows the NARX NN structure which consists of 2 inputs ($x(t)$ and $y(t)$), 1 hidden layer with 2 delays, 1 output layer and 1 output, $y(t)$. A bipolar sigmoid activation function and a linear activation function were used in the hidden layer and output layer respectively.

The data was divided to 70%, 15% and 15% for training, validation and testing respectively. 10 hidden neurons and 2 delays were used in this NN. *Bayesian Regularization* algorithm was used to train the NN as it gives the best performance.

3.3. Performance

The performance of the NARX was measured in terms of mean absolute percentage error as given below:

$$MAPE = \frac{\sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i}}{n} \times 100\%, \tag{3}$$

where n is the number of data, y_i and $\hat{y}_i$ are real and predicted values correspondingly.

4. Results and Discussions

Figure 5 reveals UTHM monthly electricity consumption versus month by year from 2009 to 2018. The lowest consumption is the year 2009 followed by the year 2010. The highest consumption year is 2015. The maximum consumption is 3228.53 MWh in March 2015, whereas the minimum is 496.379 MWh in January 2009. December 2014 is the highest among the month December.

The UTHM monthly electricity consumption plotted in time sequence is shown in Figure 6. It shows the consumption time series is fluctuating and not stationary.

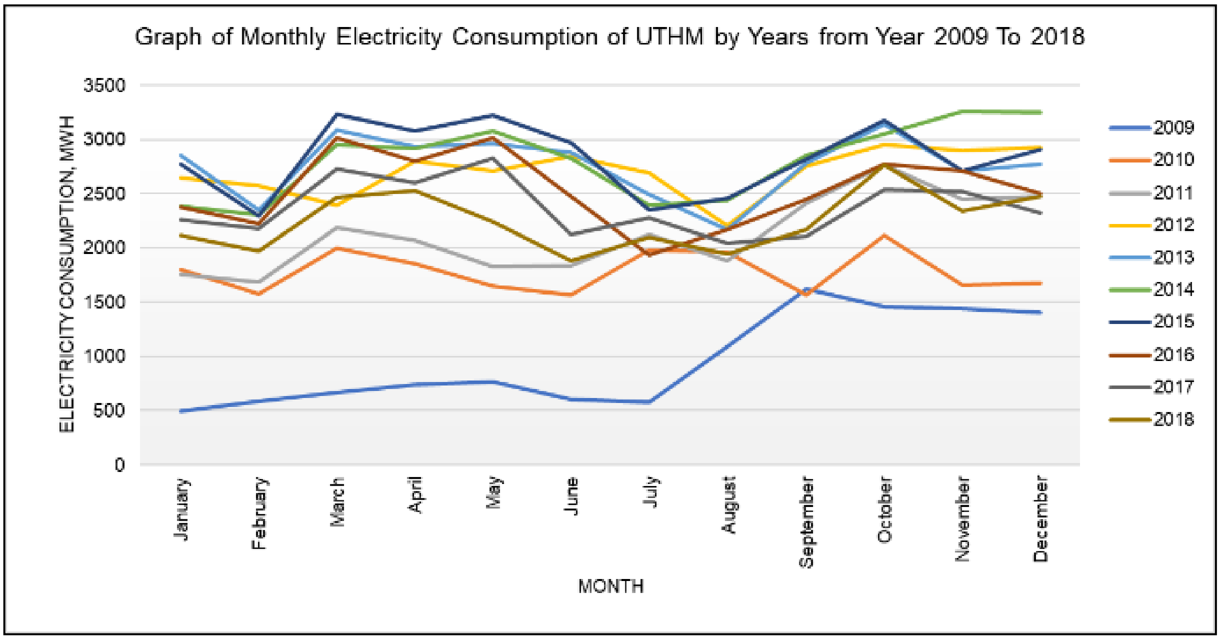

Figure 5. Monthly electricity by year

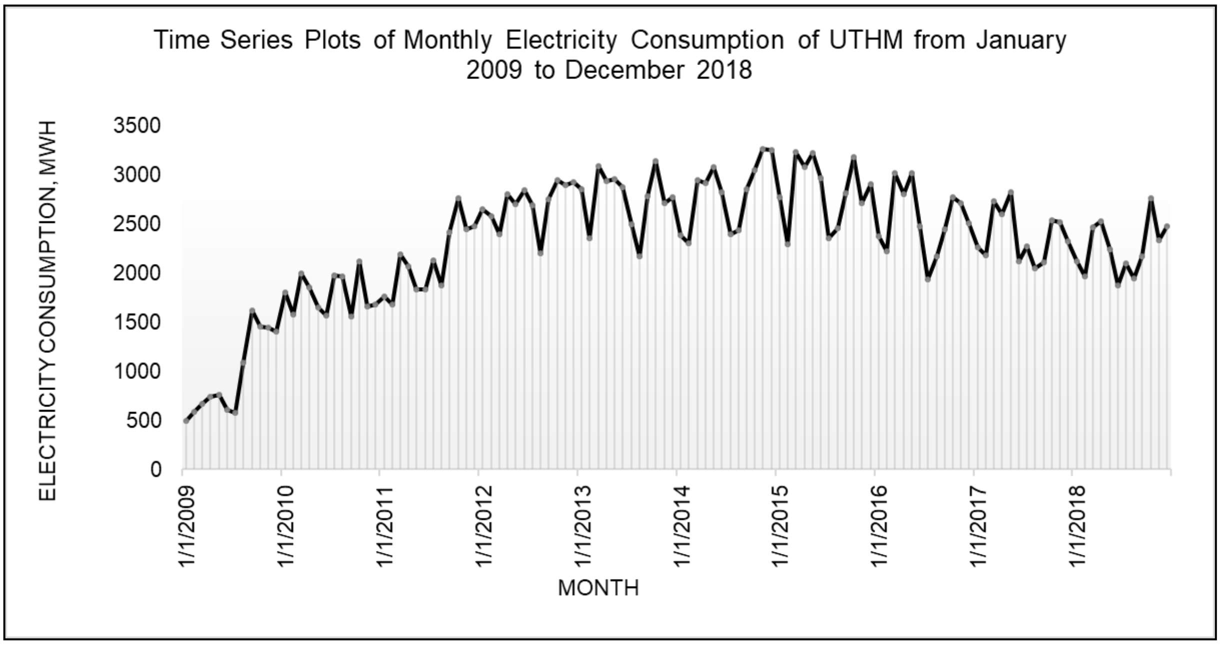

Figure 6. Monthly electricity

The performance of NARX is displayed in Figure 7. The best training performance is at epoch 686 with mean square error (MSE) of 2.58×10^{-8}.

The regression graphs *R* between the prediction (target) and the actual value (output) for training, test and all (training and test) were depicted in Figure 8. The Pearson correlation coefficient *R* for training, test and all are 1, 0.86 and 0.98 respectively. The *R* value ranges from -1 to 1. -1 is a perfect inverse relation, while 1 is perfect relation. The Pearson correlation coefficient *R* for training is perfect 1 shows that NN learn the nonlinear pattern of the historical UTHM electricity consumption with zero error while training, but when testing with some errors. In overall (training and testing data), it shows a good correlation between actual and predicted values for our NN implementation.

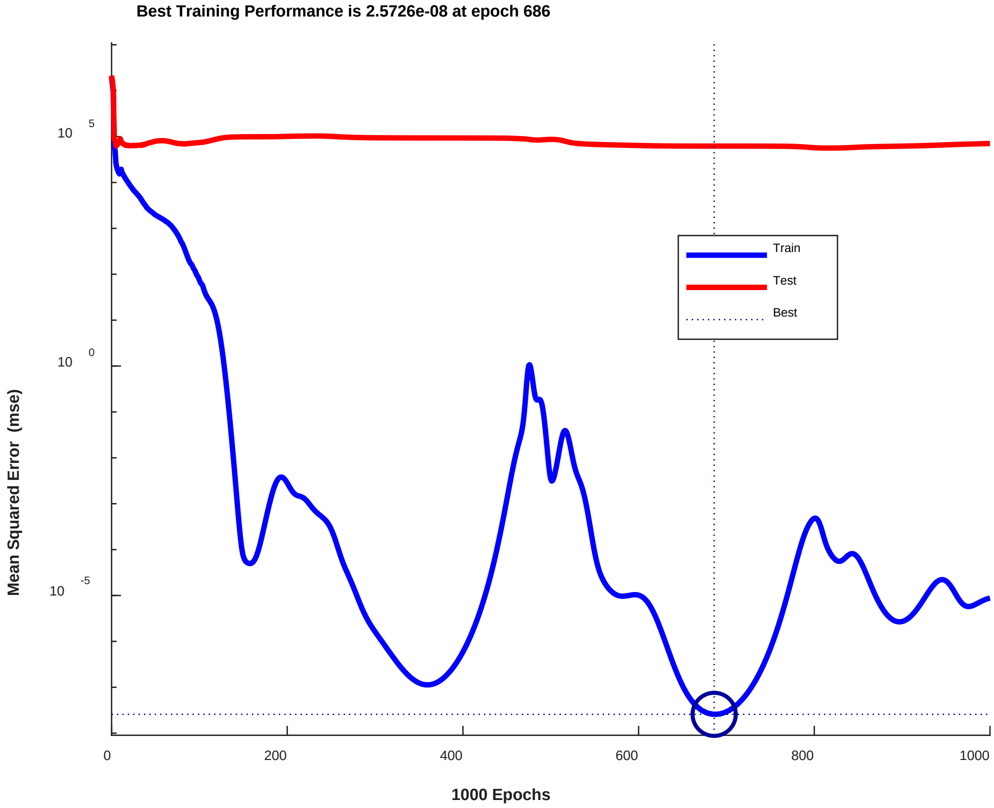

Figure 7. Performance of NARX

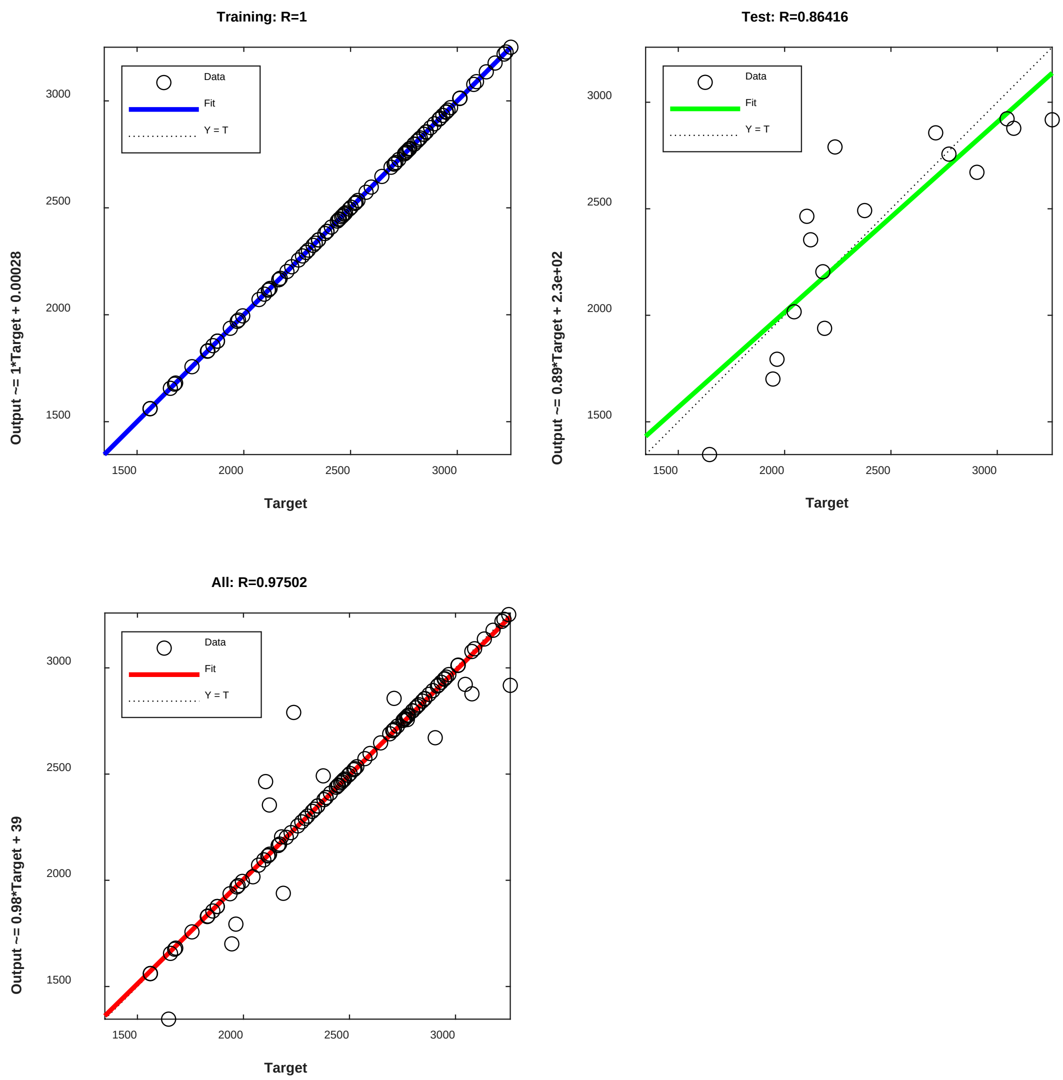

Figure 8. Target versus Output

The time series response for the training target, training output, test target and test output with their respective errors is shown in Figure 9. Target is actual y_t, whereas output is predicted $\hat{y}_t$. It shows that the errors between target and output are quite small.

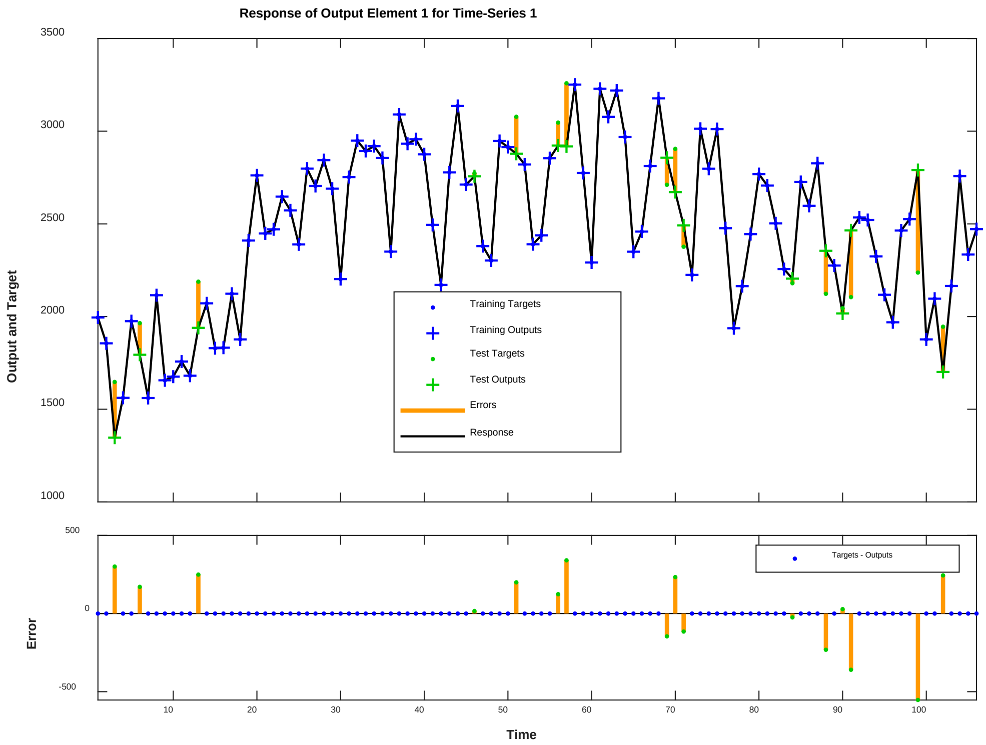

Figure 9. Time series response

Error histogram in Figure 10 shows the errors distributed from -530.8 to 318.2. Autocorrelation error in Figure 11 is bounded from -2000 to 2000.

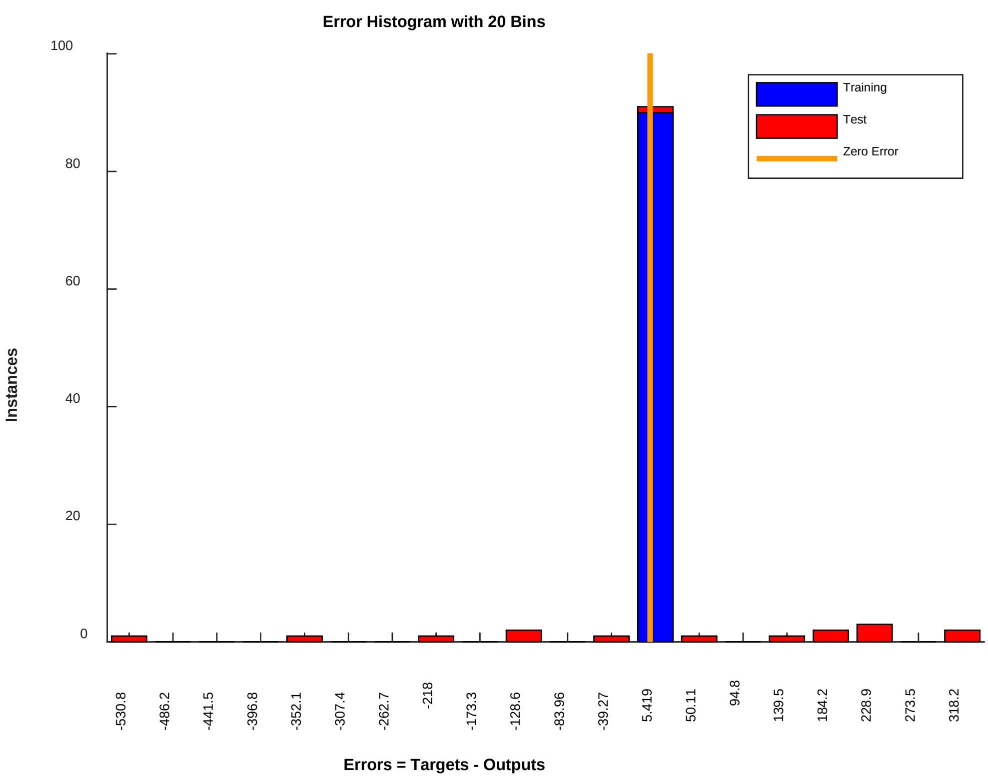

Figure 10. Errors histogram

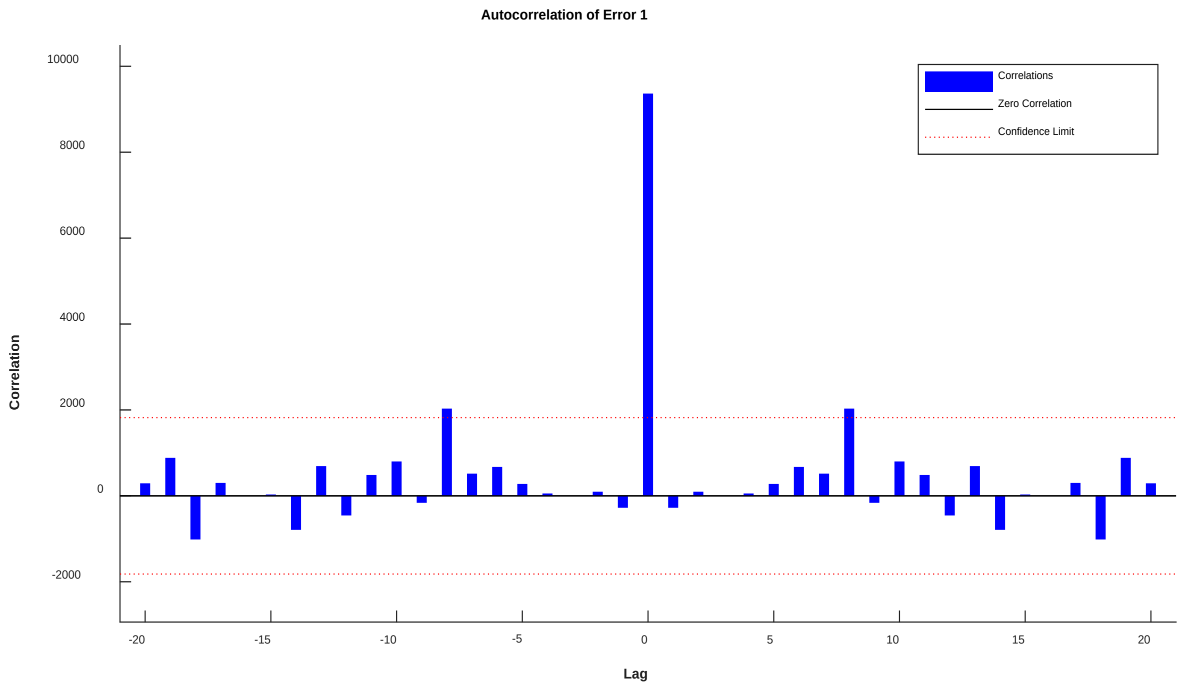

Figure 11. Error autocorrelation

The actual electricity consumption from January 2009 to December 2018 is shown in red colours in Figure 12. It is noticed that the forecasted values used previous 14 months values to predict for month 15 ahead. Hence, the forecasted electricity consumption is displayed in blue colour in Figure12 from Mar 2010 to December 2019. The future 12 months electricity consumption was forecasted as seen in Figure 13. The MAPE between actual data and forecasted results from Mar 2010 to December 2018 was calculated based on Eq (3) and was obtained as 1.38%.

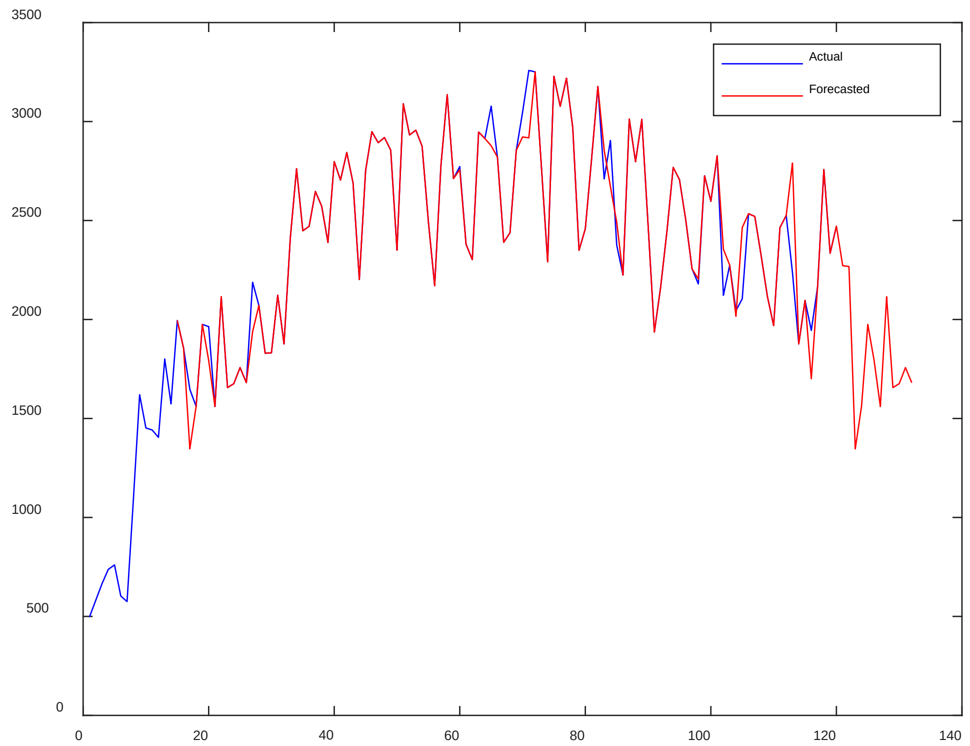

Figure 12. Actual and forecasted results

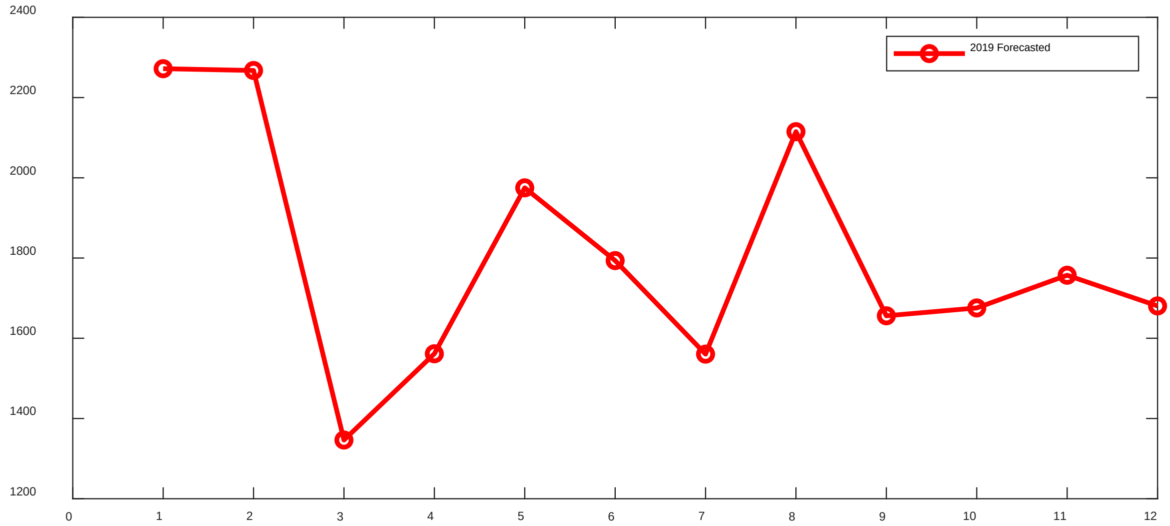

Figure 13. Forecasted results for year

5. Conclusions

Univariate monthly UTHM electricity consumption from January 2009 to December 2018 was transformed to multivariate data. This multivariate data was later inputted to NARX NN in Matlab R2018b to forecast electricity consumption for 2019. The Bayesian Regularization training algorithm with 10 hidden neurons and 2 delays was employed. The data was divided to 70%, 15% and 15% for training, validating and testing. The MAPE between actual and predicted values is 1.38%. This MAPE value is quite low if compared to the prediction of UTHM electricity consumption study done using time series model [5], fuzzy time series [6] and multiple linear regression[7]. The MAPE for [5-7] are 11.14%, 5.74%, and 10. 62% respectively.

Acknowledgements

We wish to thank Mr. Shukur Saleh and Mr. Abd Rashid Puteh from Development and Maintenance Office, UTHM for providing us with UTHM electricity consumption data, UTHM Tier 1 2018 research grant vote number H258 and Fundamental Research Grant Scheme (FRGS) vote K095 granted by Ministry of Education (MOE) Malaysia for financial support of this project.

REFERENCES

[1] F. Kaytez, M. C. Taplamacioglu, E. Cam & F. Hardalac, Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines, Electrical Power and Energy Systems, Vol. 67, 431-438, 2015.

[2] D. J. Pedregal, & J. R. Trapero, Mid-term hourly electricity forecasting based on a multi-rate approach, Energy Conversion and Management, Vol. 51, 105–111, 2010.

[3] E. Almeshaiei, & H. Soltan, A methodology for Electric Power Load Forecasting. Alexandria Engineering Journal, Vol. 50, 137–144, 2011.

[4] V. Lepojević & M.Anđelković-Pešić, Forecasting Electricity Consumption by Using Holt- Winters And Seasonal Regression Models. Series: Economics and Organization, Vol. 8, No.4, 424-431, 2011.

[5] Y. W. Lee, K. G. Tay & Y. Y. Choy, Forecasting Electricity Consumption Using Time Series Model. International Journal of Engineering and Technology, Vol. 7, No. 4.30, 218-223, 2018.

[6] K. G. Tay, Y. Y. Choy, & C. Y. Chew, Forecasting Electricity Consumption Using Fuzzy Time Series. International Journal of Engineering and Technology, Vol. 7, No.4.30, 342-346, 2018.

[7] K. G. Tay, Y. Y. Choy, & A. Huong, Forecasting Electricity Consumption Using Multiple Linear Regression. International Journal of Engineering and Technology, Vol. 7, No.4, 3515-3520, 2018.

[8] K. Panklib, C. Prakasvudhisam & D. Khummonqkol, Electricity Consumption Forecasting in Thailand Using an Artificial Neural Network and Multiple Linear Regression. Vol. 10, No.4, 427-434. 2015.

[9] P. K. Jain, W. Quamer, & R. Rajendra Pamula, Electricity Consumption Forecasting Using Time Series Analysis. International Conference on Advances in Computing and Data Sciences, 327-335. 2018.

[10] A.H. Erol, N. Özçelikkan, A. Tokgöz, S. Özel, S. Zaim, & O. F. Demirey, Time Series Methods and Neural Network, 118-127, 2012.

[11] G. Ogcu, O.F Demirey & S. Zaim, Forecasting Electricity

Consumption with Neural Networks and Support Vector Regression. Procedia - Social and Behavioral Sciences Vol. 58, 1576 – 1585, 2012.

[12] A. Bâra, and S. V. Oprea, Electricity Consumption and Generation Forecasting with Artificial Neural Networks, Advanced Applications Artificial Neural Networks, 119-139, 2017.

[13] K. Kandananond, Forecasting Electricity Demand in Thailand with an Artificial Neural Network Approach. Energies. Vol. 4, 1246-1257, 2011.

[14] K. P. Amber, R. Ahmad, M. W. Aslam, A. Kousar, M. Usman & M. S. Khan, Intelligent techniques for forecasting electricity consumption of buildings, Vol. 157, No.15, 886-893, 2018.

[15] L. Fausett, Fundamentals of Neural Networks: Architectures, Algorithms and Applications, Prentice Hall, 1994.

[16] https://www.mathworks.com/help/deeplearning/ug/design-time-series-narx-feedback-neural-networks.html;jsessionid=5d2ebb31f4e3e98d0ea1a2214b41