Indian Journal of Science and Technology, Vol 9(46), DOI: 10.17485/ijst/2016/v9i46/107141, December 2016 ISSN (Print) : 0974-6846 ISSN (Online) : 0974-5645

Correlation and Wavelet-based Short-Term Load Forecasting using Anfis

M. Mustapha, M. W. Mustafa*, S. N. Khalid, I. Abubakar and Abdirahman M. Abdilahi

Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Johor Bahru, 81310, Johor, Malaysia; mamunu33@gmail.com, wazir@fke.utm.my, nizam@fke.utm.my, aisiyaku2@live.utm.my, abdirahman@fkegraduate.utm.my

Abstract

Objective: This paper addresses the issue of model inputs selection before the forecasting exercise. Appropriate data analysis is one of the basic steps in obtaining accurate load forecast. It shapes the forecasting data in to working data by reducing the variation between the individual forecasting variables, or reduces the number of the model inputs. Also, the information received from data analysis determines the method to be used, or how to use it. **Methods/Statistical Analysis:** It employs the use of correlation analysis to select the forecasting variables, and wavelet transforms to decompose the selected data in to a number of approximations. The purpose is to select the actual forecasting variables, and to limit the variation between them (model inputs). ANFIS was used to forecast the load using the processed data. **Findings:** From the result obtained, it was observed that selecting the data based on correlation analysis, and wavelet transform improve the accuracy of the forecast, and enhanced the forecasting speed. **Applications/Improvements:** Improving the forecasting accuracy will save the utility economically, and improving the speed will enhance the time taken to make crucial decisions in power system operation

Keywords: ANFIS, Correlation Analysis, Short-term Load Forecasting, Wavelet Transform

1. Introduction

The rapid changing in Electrical grid which led to introduction of smart grid and Demand Response (DR) programs, with aim to make the grid more reliable and cost effective, make it necessary to determine the demand ahead of time. Moreover, general utility operations depend on the demand forecast. Accuracy obtained from the demand forecasting algorithms plays a vital role in the energy market. As of 1997, it was reported that reduction in 1% of the forecasting accuracy saved British Power Corporation £10m[1].

In the recent researches, electrical load forecasting is divided in to three. Short-Term Load Forecasting (STLF), which is usually from one hour to one week, Medium or Mid-Term Load Forecasting (MTLF) which is from a week to a year, and Long-Term Load Forecasting (LTLF) which is more than a year[2]. The nature as well as the forecasts for different time zone are important for different operation within a utility company[2]. STLF can help to estimate load flows and to make decisions that can prevent over loading.

This can lead to the improvement of network stability and reduce frequent equipment failures and blackouts. Decisions on capital expenditure such as future load plan, plant extension and infrastructural development, is based on LTLF[2]. On the hand, MTLF provides the primary information with regards to safety and reliability of the

*Author for correspondence

power system during operation, and also serve as a tool in determining the optimal generators and station utilization of the power system[3].

Good feature selection and data analysis make a very good forecast, and improve the forecasting accuracy and speed. That is why researchers are giving more emphasis in this area. To obtain an accurate forecast, there is need to make a good data processing. Electricity utility companies give more emphasis on storing these data, because of its significance in their operational activities[4]. Depending on the forecasting method and the forecaster, many data processing algorithms ware presented[5–9], up to now there is no reason of choosing one method over the other. Most of the researchers use the methods based on their knowledge. It is important to note that information received from data analysis determines the best method to be used or how to use the method. It is also easy to determine when (the time at which) the consumption is low or high in the load profile, or relationship between the consumption and the forecasting variables[10]. Variables (model inputs) selection determines the relevant inputs to forecasting model, thus make it an area of interest in the machine learning or Artificial intelligence modelling[9].

The main goals of this article is to device a means of selecting the forecasting variable using correlation analysis, and decompose the selected data using wavelet transform. Decomposition of the data will results to a well processed data which is free from noises and outliers, and the variations between the variables is reduced. This will improve the forecasting accuracy.

2. Correlation Analysis

Correlation is the measure of interrelation between the changes in two variables. It describes the relation between two pairs of data. Correlation coefficient (R) ranges from -1 to +1. Any value within this range determines the relationship between the variables[11]. High value (close to +1 or -1) indicates strong relationship and value close to zero indicates low correlation. Zero correlation coefficient shows that there is no any relation between the two pair of the data. Correlation coefficient R, between two set of data can be calculated using the following formula;

$$R_{xy} = \frac{\operatorname{cov}(x, y)}{\delta_x \delta_y} \tag{1}$$

Where $\operatorname{cov}(x, y)$ is the population covariance, δ_x and δ_y are the population individual standard deviation.

In this analysis, relationship between the load consumption and weather variables is considered. Weather variables that comprised temperature, dew point, relative humidity, wind speed and wind direction of three weekdays (Tuesday, Wednesday and Thursday) on hourly basis is used. Table 1 shows the correlation coefficient between the load consumption and the weather variables.

Table 1. Correlation coefficients between the load and the model inputs (variables)

S/N	Variables	Correlation coefficient (R)(load and model inputs)
1	Last day load (MW)	0.8696
2	Last two days load (MW)	0.7967
3	Current day temperature (°C)	-0.6381
4	Last day temperature (°C)	-0.6377
5	Last two days temperature (°C)	-0.6376
6	Current day dew point (°C)	-0.6355
7	Last day dew point (°C)	-0.6736
8	Last two days dew point (°C)	-0.6692
9	Current day relative humidity (%)	-0.2099
10	Last day relative humidity (%)	-0.2517
11	Last two days relative humidity (%)	-0.6045
12	Current day wind direction (10s deg)	0.1895
13	Last day wind direction (10s deg)	0.0095
14	Last two days wind direction (10s deg)	0.1509
15	Current day wind speed (km/h)	0.5295
16	Last day wind speed (km/h)	0.4221
17	Last two days wind speed (km/h)	0.1940

3. Wavelet Transform (WT)

Wavelet is used to decompose time series signal in to approximate and details components. The load series is decomposed in to low and high coefficients. This is to extract high frequencies from the load series and reduce the variation between the load data[12,13]. The data is decomposed in to three levels, using db2 as presented in equation (2)

$$l(t) = A_3(t) + D_3(t) + D_2(t) + D_1(t) \tag{2}$$

where *l(t)* is the load series, $A_3(t)$ is the approximate component and *D1(t), D2(t)* and $D_3(t)$ are detail components. Generally, a wavelet transform of time series signal is given by equation (3)[14]

$$WT_{(a,b)} = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t)\psi(\frac{t-b}{a})dt \tag{3}$$

where $\psi(t)$ is the mother wave let, *a* is scale factor and *b* is the time-shift parameter. Following the decomposition of the load series as presented in[12], it can be observed that the approximate part describes the load pattern and the details part presents the most important components of the load series. To reconstruct the load data back after the forecasting an inverse of the same wavelet was used. The expression of the inverse wavelet is given in equation (4), with all the parameters maintaining their original definition.

$$f(t) = \frac{1}{c_\psi^2} \int_a \int_b \psi_\psi(a,b) \frac{1}{a^2} \psi(\frac{t-b}{a})dadb \tag{4}$$

4. ANFIS

ANFIS was developed by J. S. Roger in 1993 through combining the advantages of fuzzy systems and neural-networks[15,16]. ANFIS is a network-based structure that uses the Sugeno-type 'IF…..THEN' rules and Neural Network (NN) to match inputs with their corresponding outputs. The Least Square Algorithm (LSA) is being used to calculate the consequent parameters of the ANFIS output in the forward pass, and Gradient Decent Algorithm (GDA) computes the premise parameters associated with fuzzy membership functions in the backward pass. Figure 1 shows the general structure of ANFIS network, with only two inputs (x and y). The structure consists of five layers with several nodes (depending on the inputs).

For Sugeno-type fuzzy system with only two inputs, the rules are[17];

1. If *x* is A_1 and *y* is B_1, then $f_1 = p_1x + q_1y + r_1$
2. If *x* is A_2 and *y* is B_2, then $f_2 = p_2x + q_2y + r_2$

where p_i, q_i and r_i are the consequent parameters.

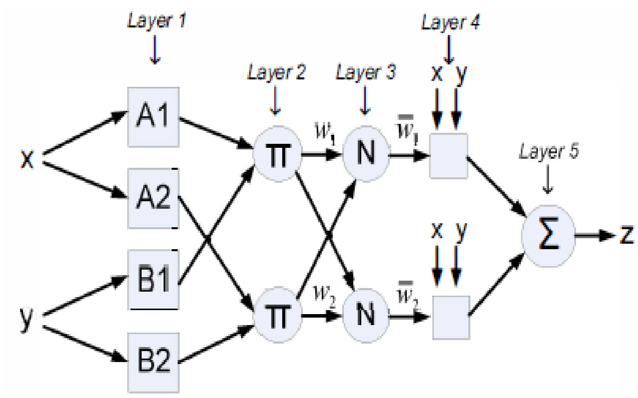

Figure 1. Typical ANFIS structure.

Now we can define the function of each node. Let o_i^j be the output of node *i* in layer *j*[18];

Layer 1: In this layer, all the nodes are adaptive nodes. Output of a node is determined by the membership function. Therefore, for a node A_1 the output is given by [18]

$$o_i^1 = \mu_{A_i}(x_i) \tag{5}$$

where μ_{A_i} is the membership function (MF) of node *A*. Depending on the complexity of the problem, there are available MFs such as linear MF, triangular MF, Gaussian MF, trapezoidal MF, pi MF and bell-shape. All these exhibit different expression with different parameters.

Layer 2: Output of this layer is the firing strength of all the signals entering the node from the previous layer. In other words, the output is the **AND** or **OR** operation of all the MFs entering the node from the previous layer[18]. Thus;

$$o_{\pi_i}^2 = w_i = \mu_{A_i} \times \mu_{B_i} \times \mu_{C_i} \ldots.. \tag{6}$$

where μ_{A_i} is the membership function of node *A* in the input layer, μ_{B_i} is the membership function of node *B* in the input layer, μ_{C_i} is the membership function of node *C* in the input layer.

Layer 3: Normalization takes place in this layer. The output of each node is given by the ratio of the node's fairing strength to the sum of all the firing strengths of the other nodes entering this node[18]. Thus;

$$\overline{o_i^3} = \overline{w_i} = \frac{w_i}{w_1 + w_2 + \ldots..} \tag{7}$$

where w_i is the firing strength of node *i*.

Layer 4: Output of each node in this layer is the product of the normalised signal from previous node and the linear combination of the input signals (variables)[18], thus

$$o_i^4 = \overline{w_i} f_i = \overline{w_i}(p_i x + q_i y + r_i) \tag{8}$$

where $\overline{w_i}$ is the normalised signal of node i (in layer 3), p_i, q_i and r_i are consequent parameters of the defuzzification node (in layer 4).

Layer 5: This layer produces the overall output through summing up all the signals from the nodes of the previous Layer 4[18], thus

$$o_i^5 = \sum_i \overline{w_i} f_i \tag{9}$$

where f_i is the linear combination of the consequent parameters of layer 4.

Three methods are used to determine the membership functions. Grid partition[19], Fuzzy C-means Clustering[20](FCM) and Subtractive Clustering[21]. Limitations of these methods vary from one method to another. Fuzzy c-means clustering and subtractive clustering are introduced to reduce the number of membership functions being generated. In this work, FCM is considered because it improves the speed of the algorithm (ANFIS)[22].

5. Load Forecasting Implementation

In this work, data sets (available online) from Nova Scotia region are used. Being the smallest province in Canada, and not more than 67km from the ocean, the weather is controlled by the ocean. It is therefore difficult to determine exact variables that will affect the load consumption. Also, data collected during spring season is considered, because the weather in this season is uniform compared to other seasons. Temperature, Dew Point, relative Humidity, Wind speed, Wind direction and hourly load are used in this experiment. Spring starts from middle of March to the Middle of June. First eight weeks (middle of March and complete April) for training and next four weeks (May) for testing. Meaning that two months data for training and the subsequent for testing. Data of three days (Tuesday, Wednesday and Thursday) in each week was used for both training and testing. Tuesday and Wednesday data were used to forecast Thursday. This is because the load pattern is similar over these days Figure 2.

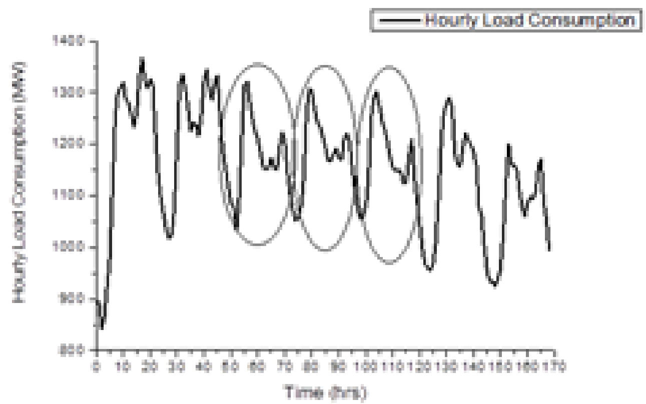

Figure 2. Daily load pattern for one week.

Before the forecasting exercise, the data is divided in to three groups. First group comprised all the available data. Second group is also all the available data decomposed using wavelet. Last group is the selected data based on correlation analysis and then decomposed using wavelet.

As presented in Table 1, eleven variables are selected and six variables (shaded) are rejected from the total of Seventeen variables, because their correlation coefficients are less than ±0.4000. A db2 of the wavelet family is used to decompose the input variables in to approximate and detail coefficients. This will reduce the volatility of the variance and reduce the effect of outliers in the forecasting data[13]. Two level decomposition is used, thus one approximation and two detailed coefficients are used as depicted in Figure 3.

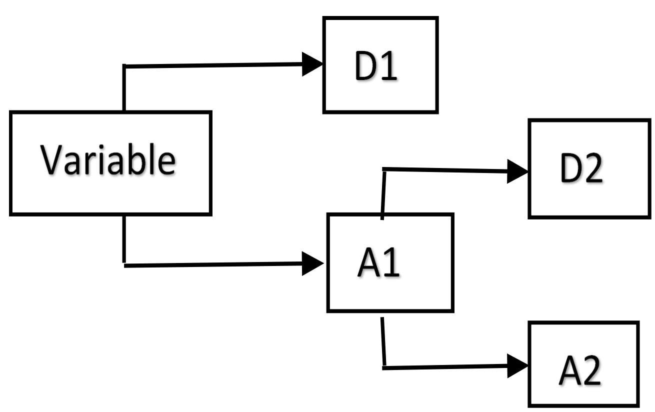

Figure 3. Schematic diagram of signal decomposition using wavelet transform.

After the data analysis all the three groups of data are used to forecast the load using ANFIS. The error measure and the time taken to forecast for each case are recorded in Table 2.

Table 2. Forecasting results for the three cases

Data selection	Error measurement			Forecasting time (sec)
	MSE	RMSE	MAPE (%)	
Case 1	920.70	30.34	2.7	46.79
Case 2	827.85	28.77	2.5	39.74
Case 3	508.59	22.55	2.1	25.26

6. Results and Discussion

In this experiment three different case studies are presented based on the type of data used. Each case is represented by each group of data described in section 5. Mean Square Error (MSE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) are used to evaluate the accuracy in each case, and the results obtained are as follows;

Case 1: Here all the data sets are used and the results are shown in Figures 4-7. An MSE of 920.70, RMSE of 30.34 and MAPE of 2.7% are obtained.

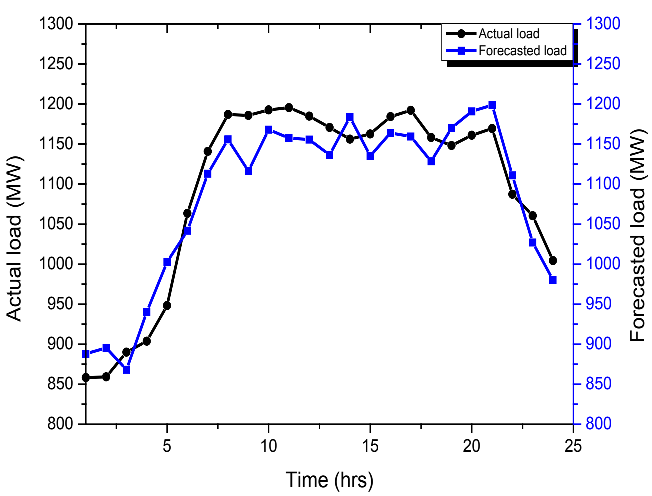

Figure 4. Forecasted and actual load of thursday, 8th May, 2014, for Case 1.

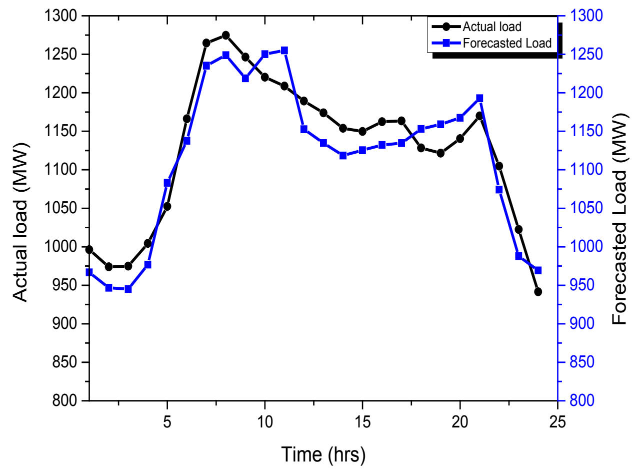

Figure 5. Forecasted and actual load of Thursday, 15th May, 2014, for Case 1.

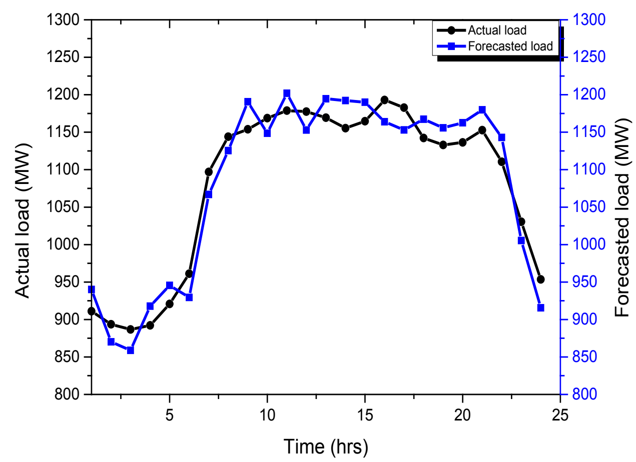

Figure 6. Forecasted and actual load of Thursday, 22th May, 2014, for Case 1.

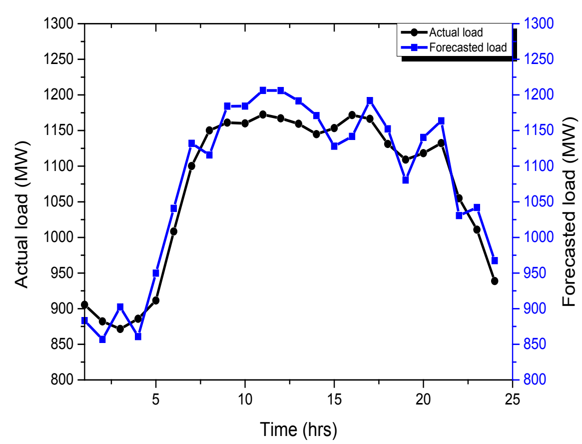

Figure 7. Forecasted and actual load of Thursday, 29th May, 2014, for Case 1.

Case 2: Here all the data sets are used, but decomposed at level 2 using db2 mother wavelet. Figures 8-11 shows the plot of actual load and predicted load. An MSE of 827.85, RMSE of 28.77 and MAPE of 2.5% are obtained.

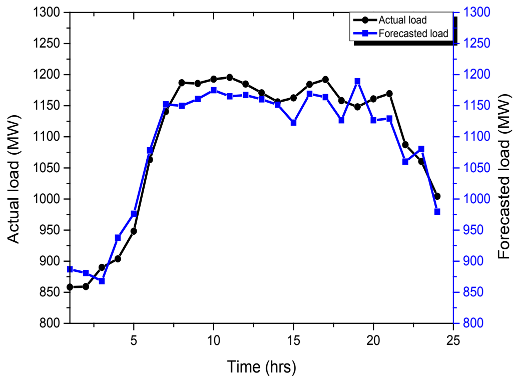

Figure 8. Forecasted and actual load of Thursday, 8th May, 2014, for Case 2.

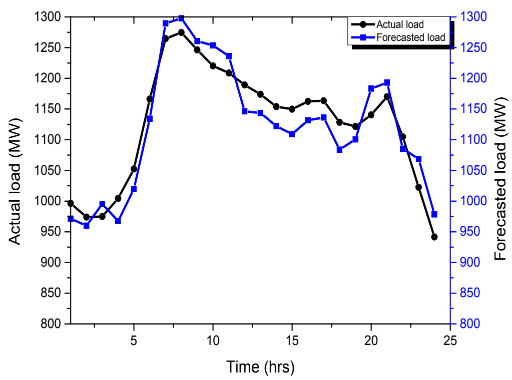

Figure 9. Forecasted and actual load of Thursday, 15th May, 2014, for Case 2.

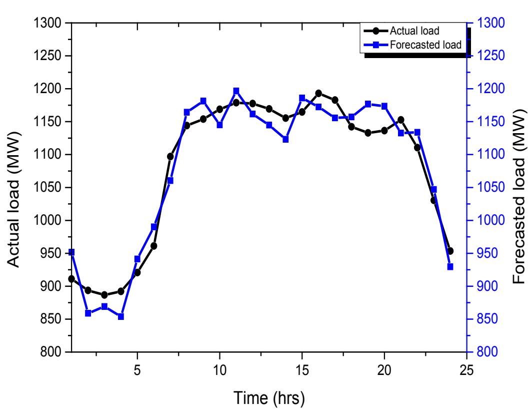

Figure 10. Forecasted and actual load of Thursday, 22nd May, 2014, for Case 2.

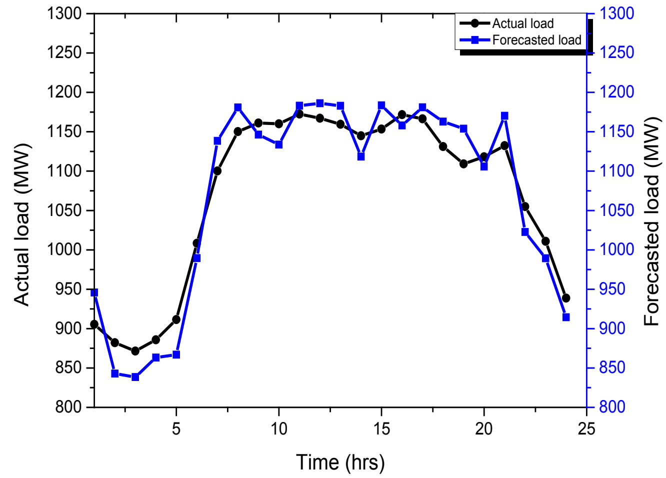

Figure 11. Forecasted and actual load of Thursday, 29th May, 2014, for Case 2.

Case 3: Here, data is selected based on the results of the correlation coefficients are used. Out of the 17 tested variables, 11 are selected based on correlation coefficient value greater than or equals to ±4.000 (refer to Table 1). The selected variables are decomposed at level 2, using db2 mother wavelet. Figures 12-15 shows the plot of actual load and predicted load. An MSE of 508.59, RMSE of 22.55 and MAPE of 2.1% are obtained.

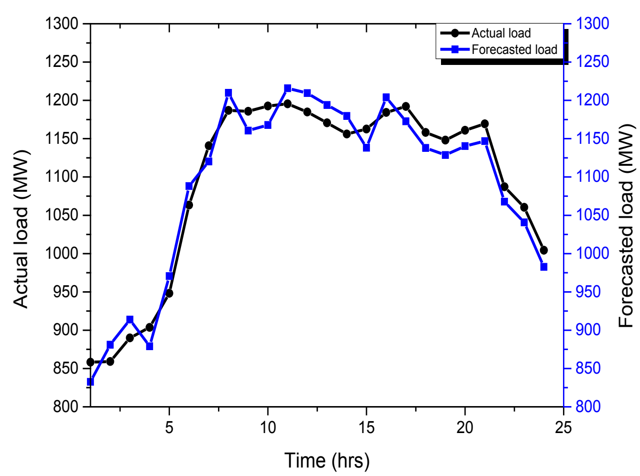

Figure 12. Forecasted and actual load of Thursday, 8th May, 2014, for Case 2.

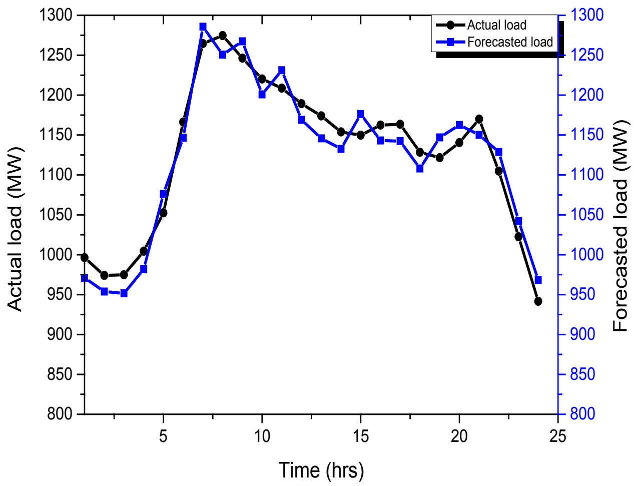

Figure 13. Forecasted and actual load of Thursday, 15th May, 2014, for Case 2.

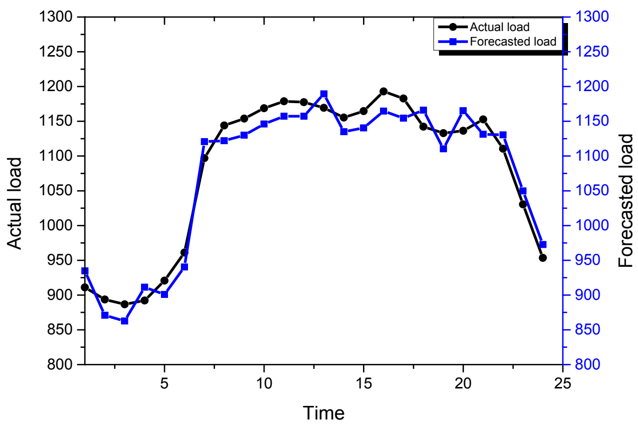

Figure 14. Forecasted and actual load of Thursday, 22nd May, 2014, for Case 2.

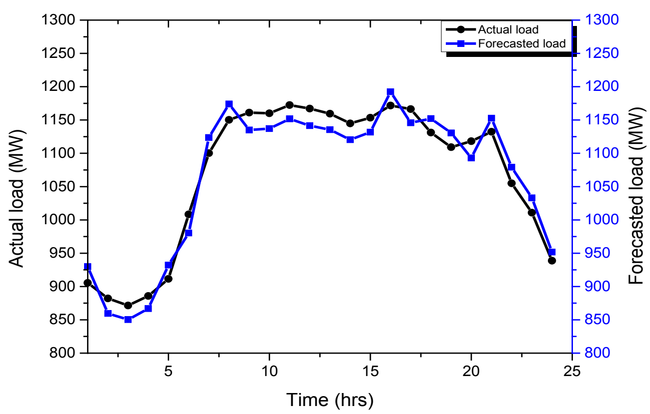

Figure 15. Forecasted and actual load of Thursday, 29th May, 2014, for Case 2.

The results obtained for the three cases is summarised in Table 2. It can be observed that case 3 presented more accurate results, followed by case 2 and then case 1.Also, it was observed that using correlation analysis and wavelet transform on the forecasting data improves the forecasting accuracy, and improved the speed of the forecasting. As presented by[1], reduction in the forecasting error improve the economy and operation of power system, therefore, case 3 is a very good approach. Also, result of the hourly forecasting for next day should be presented at a certain time of the current day[23], our approach can improve the submission time, so that energy can be purchased as early as possible.

7. Conclusion

In this work correlation analysis and wavelet transform are used to analyse data for load forecasting using ANFIS. From the data analysis results, three different case studies were formulated using historical load and weather data. All the three cases were used to determine the best approach in processing load forecasting data. The results obtained shows that selecting forecasting variables using correlation analysis, and then decompose the selected data using WT is a good practice in load forecasting. Also when the actual variables are selected the speed of the forecasting is increased, and this will help in submitting the forecasting results within the recommended period of time. In further experiments we will consider effect of seasonal weather variation using same approach.

8. Acknowledgment

The authors acknowledge Malaysian Ministry of Higher Education and Universiti Teknologi Malaysia for supporting this work.

9. References

1. Ranaweera DK, Karady GG, Richard CF. Economic impact analysis of load forecasting. IEEE Transaction on Power System 1997; 12(3). p. 1388–92.
2. Feinberg EA, Genethliou D, Chao, JH, Wu FF, Momoh JA. Load Forecasting. In: editor. Applied Mathematics for restructured Electric Power System. 2006. p. 269–85.
3. Dragomir OE, Dragomir F, Gouriveau R, Monica E. Medium term load forecasting using ANFIS predictor. In:

18th Mediterranean Conference on Control & Automation. 2010. p. 551–6.

4. Chakravorty A, Rong C, Wlodarczyk E P, Wiktor T. A Distributed Gaussian-Means clustering algorithm for forecasting domestic energy usage. In: International Conference on Smart Computing. IEEE; 2014. p. 229–36.
5. Chen Y, Luh PB, Guan C, Zhao Y, Michel LD, Coolbeth MA. Short-term load forecasting : Similar day-based wavelet neural networks. IEEE Transaction on Power System. 2010; 25(1). p. 322–30.
6. Mahmoud T, Lachowicz SW, Member TSM, Habibi D, Member S, Bass O. Load demand forecasting: Model inputs selection load demand forecasting: Model Inputs Selection. IEEE Conference and Innovation Smart Grid Technology Asia; 2011.
7. Sovann N, Nallagownden P, Baharudin Z. A method to determine the input variable for the neural network model of the electrical system. 5th International Conference on Intellectual Advanced System. Jun 2014. p. 1–6.
8. Quilumba FL, Lee W, Huang H, Wang DY, Member S, Szabados RL. Using smart meter data to improve the accuracy of intraday load forecasting considering customer behaviour similarities. IEEE Transaction on Smart Grid. 2015; 6(2):911–8.
9. Koprinska I, Rana M, Agelidis VG. Correlation and instance based feature selection for electricity load forecasting. Knowledge-Based System. Elsevier B.V. 2015; 82:29–40.
10. Tiong SK, Ahmed SK. Electrical Power Load Forecasting using Hybrid Self-Organizing Maps and Support Vector Machines. In: The 2nd International Power Engineering Optimization Conference (PEOCO). Selangor: IEEE. 2008. p. 51–6.
11. Hernandez L, Baladron C, Aguiar JM, Calaavia L, Carro B, Esguevillas AS. A study of the relationship between weather variables and electric power demand inside a smart grid/smart world framework. SENSORS. 2012; 12:11571–91.
12. Bashir ZA, El-Hawary ME. Applying wavelets to short-term load forecasting using PSO-based neural networks. IEEE Transaction on Power System. 2009; 24(1):20–7.
13. Li S, Wang P, Goel L. A Novel Wavelet-based ensemble method for short-term load forecasting with hybrid neural Networks and Feature Selection. IEE Transaction on Power System. 2016; 31(3):1788–98.
14. Pousinho HMI, Mendes VMF. Hybrid Wavelet-PSO-ANFIS approach for short-term wind power forecasting in Portugal. IEEE Transaction Sustainable Energy. 2010; 2(1):50–9.
15. Jang JR. ANFIS: Adaptive-network-based fuzzy inference system. IEEE Transaction on System Management Cybernetic. 1993; 23(3):65–85.
16. Ghomsheh VS, Shoorehdeli MA, Teshnehlab M. Training ANFIS structure with modified PSO algorithm. In: Mediterranean Conference on Control and Automation. Athens. 2007. p. 1–6.
17. Cheng C-H, Wei L. One step-ahead ANFIS time series model for forecasting electricity loads. Optimal Engineering. 2010; 11(2):303–17.
18. Nguyen T, Liao Y. Short-Term Load Forecasting Based on Adaptive Neuro-Fuzzy Inference System. Journal of Computing. 2011; 6(11):2267–71.
19. Jang JR, Sun C-T, Mizutani E. Neuro-Fuzzy and Soft Computing. NJ: Prentice Hall: New York. 1997. p. 611–4.
20. Bezdek JC, Ehrlich R, Full W. FCM : The fuzzy c-means clustering algorithm. Computer Geoscience. 1984; 10(2):191–203.
21. Chiu SL. Fuzzy model identification based on cluster estimation. Journal of Intellectual Fuzzy System. 1994; 2:267–78.
22. Mustapha M, Mustafa MW, Khalid SN. Data selection and fuzzy-rules generation for short- term load forecasting using ANFIS. Telekomnika. 2016; 14(3):791–9.
23. Hong T, Wang P. Fuzzy interaction regression for short term load forecasting. Fuzzy Optimal Decision Making. Sep 2013; 13(1):91–103.