Wavelet and Neural Network Approach to Demand Forecasting based on Whole and Electric Sub-Control Center Area

Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul

Abstract— **Whole and electric sub-control area load demand forecasting based on a wavelet transform and a neural network method that are very significant technique for a load prediction. The research used wavelet transform method in preprocessing stage; furthermore, a neural network is used to predict in forecasting stage for whole and sub-control areas prediction. The comparison results show that sub-control area forecasting has a good prediction than that the whole area forecasting based on two levels of wavelet transform. An accuracy of forecast is an essential activity for fuel reserve planning in a power system.**

Index Terms—**Whole area, electric sub-control area, wavelet transform, neural network, forecasting.**

I. INTRODUCTION

Load forecasting, can be classified into four differential types in the power system planning: very short term, a minute to an hour; short term, an hour to three months; mid term, three months to three years; long term, three years to twenty five years. Several methods have been developed for forecast that is statistical and artificial intelligent methods. In conventional, many papers have used statistical methods such as regression, multiple linear regressions (MLR), and autoregressive integrated moving average (ARIMA). These methods focus on estimating the coefficients of variables that are linear. There are prospected functional forms describing quantitative relationship between load demands and influencing factors by using the historical data. Recently, with the emergence of artificial intelligent (AI) technologies, the artificial neural network (ANN), the fuzzy logic (FZ), and the genetic algorithm (GA) have been widely implemented to improve the accuracy of load forecasting, historical data in time series are both non-linear and linear. In the research, proposes the mid term load forecasting, plays an important role in the reliability of a power system. It is vital for planning of an adequate fuel reserves to generate the electricity to the consumption demand for the future. This forecast is complicated an effect on load demand by depending two factors: the complexity of an economy and a weather factor of the whole area and each region of the country.

Several researchers preferred differential applications. The paper ref. [1] shows the hybrid model based on wavelet support vector machine and modified genetic algorithm penalizing Gaussian noises for power load forecasts, furthermore a short-term load forecasting by using similar day-based wavelet neural network [2]. In ref. [3], the intelligent hybrid wavelet models for prediction. In 2009, the combination of wavelet transform and neuro-evolutionary algorithm approach to demand forecasting [4]. In 2008, the research shows an adaptive wavelet neural network-based energy price forecasting in electricity markets [5] and in the year 2006 the wavelet based nonlinear multi-scale decomposition model [6]. In ref. [7], in the same year, the researcher proposes techniques of applying wavelet transform into combined model for demand forecasting in electricity. In research ref. [8], presents an adaptive neural wavelet model and in ref. [9] proposes a hybrid wavelet-Kalman filter method for load forecasting. Lastly, the wavelets transform and neural networks for short-term electric load forecasting are proposed [10]. All of the researches above were proposed load and price forecasting by using wavelet transform and neural network algorithm but did not present in sub-control center area forecasting.

In this research, proposes the whole area and electric sub-control center area forecasting, are implemented by using wavelet transform in preprocessing stage of all areas and neural network for forecasting in the last process. In preprocessing stage, wavelet transform is used to decompose the original signal of demand into one to four levels and after that will take it to find the relationship between factors and demand before choosing the suitable factors for feature input for neural network to prediction. Finally, this paper presents the comparison between the whole and sub-control center area forecasting based on mean absolute percentage error (MAPE).

This article proposes the five major sections. The second section presents an electricity demand in a whole and sub-control area of Thailand. The third section offers an implementation of the research stages. The fourth section shows the results and comparison of the research. Finally, conclusion is drawn in the fifth section.

Manuscript received November 16, 2011.
Pituk Bunnoon, Electrical engineering department, Engineering faculty, Prince of Songkla University, Hadyai, Songkhla, Thailand, Mobile No.+66850537350, (e-mail: add2002k@hotmail.com).
Kusumal Chalermyanont, Electrical engineering department, Engineering faculty, Prince of Songkla University, Hadyai, Songkhla, Thailand, (e-mail: kusumal.c@psu.ac.th).
Chusak Limsakul, Electrical engineering department, Engineering faculty, Prince of Songkla University, Hadyai, Songkhla, Thailand, (e-mail: chusak.l@psu.ac.th).

II. ELECTRICITY DEMAND IN WHOLE AND SUB-CONTROL CENTER AREA

A. Electricity demand in whole area

The energy consumption load demand of the Electricity Generating Authority of Thailand (EGAT) in a whole area illustrates in Fig.2 (e). It can be seen that load characters from January to December and from 1997 to 2006 are alike; load shapes for each months of a year are quite different.

The graph shows the load demand (kWh), as illustrates in Fig.2 (e) that is a behavior of an electric trend component increasing the demand every year. The demand interval year 1997 to 2006, case study of Thailand are used in the research. The trend in 1997 to 1999 was quite stable because the world economic problems had occurred and affected to an industrial sectors of the country. In 2000 to 2006, the trend of electricity demand grows at rate about 5 percent per year. The energy load demand on May is 8,749,926,336 kWh in 2000, 11,171,134,811 kWh in 2004, and 12,444,389,521 kWh in 2006 being the maximum demand in each year.

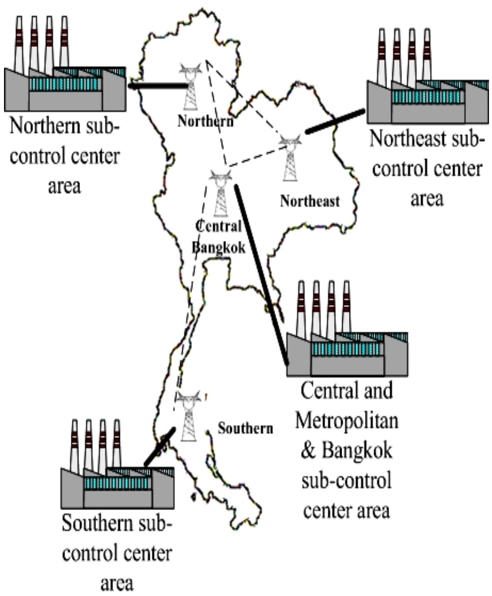


Fig.1 Map of each electric sub-control center area in Thailand.

B. Electricity demand in sub-control center area

Sub-control center area location, as illustrates in Fig.1, which shows a map of each sub-control center area through the country. Fig.2 (a)-(d) shows energy consumption demand in each sub-control center area of the country. Usually, the large energy consumption will be occurred in an industrial area such as central area, Bangkok and metropolitan can be seen in Fig.2 (a). Normally, the peak of energy consumption demand occurred between March to May in each year. In Fig. 1 illustrates the map of electric sub-control center located area of the country. The Northern, Northeast, Southern, and Central sub-control center area located at Pitsanulok, Khonkan, Krabi, and Nontaburi provinces respectively. The energy consumption load demand in these areas is illustrated in Fig.2 (a)-(d).

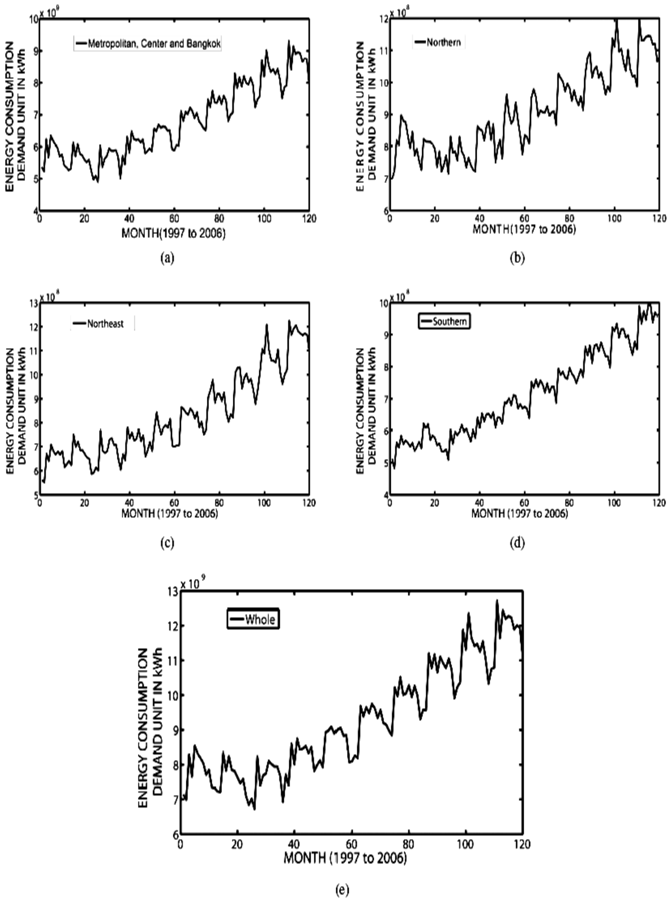


Fig.2 Electric energy consumption of a) Central b) Northern c) Northeast d) Southern e) whole area of the country.

C. Basic theory

C.1 Wavelet decomposition and reconstruction

Electricity load demand is generally complex and consists of multiple frequencies. Furthermore, features of the demand cannot be fully captured by a single neural network. Consequently, wavelet decomposition method in this research is developed and combined with the neural network for increasing the reliable forecast. The research is done by decomposing an original demand into one level, two levels, three levels, and four levels of wavelet. The feature of neural network is selected based on the correlation between affecting factors and demand.

Wavelet theory is an applicable to several subjects. It is a powerful implement which can be used for a wide range of applications, specifically; signal processing, data compression, image de-noising, speech recognition, computer graphics, and many areas of physics and engineering. All of the wavelet transform may be considered forms of time-frequency representation for continuous time (analog) signals and so are related to harmonic investigation. Almost all practically useful discrete wavelet transforms use discrete-time filter banks. These filter banks are called the wavelet and scaling coefficients in wavelet nomenclature.

This section provides a brief summary of wavelet transform method which can be divided into two categories: continuous wavelets transform (CWT) and discrete wavelet transforms (DWT). In this research, discrete wavelet transform is used. The DWT algorithm is capable of producing coefficients of fine scales for capturing high frequency information, and coefficients of coarse scales for capturing low frequency information. For a mother wavelet function ψ and for a given signal $f(t)$, a DWT can be expressed as follows [3]:

$$f(t) = \sum_k c_{j0,k}\phi_{j0,k}(t) + \sum_{j>j0}\sum_k \omega_{j,k} 2^{\frac{j}{2}}\psi(2^j t - k) \tag{1}$$

where j is the dilation or level index, k is a translation or scaling index, $\phi_{j0,k}$ is a scaling function of coarse scale coefficients, $c_{j0,k}, \omega_{j,k}$ are the scaling function of detail coefficients, and all function of $\psi(2^j t - k)$ are orthonormal.

Wavelet processing has two stages: decomposition and reconstruction. The decomposition computes the convolution between the load demand and high pass/ low pass filter, while the reconstruction calculates the convolution between the load and inverse filter. A mother wavelet based on Daubechie2 (Db2) is used for the filter's coefficients. It used to decompose an input load demand into low frequency and high frequency components. The decomposition is implemented by using multichannel filter bank: one, two, three, and four channels. The reconstructed details and approximations are true parameters of the original signal as follow [3]:

S = A1 + D1 (Level^1)
= A2 + D2 + D1 (Level^2)
= A3 + D3 + D2 + D1 (Level^3)
= A4 + D4 + D3 + D2 + D1 (Level^4)

For example, in Level 1, the coefficient vectors A1 and D1 are produced by down sampling and only half a length of the original signal. Thus, they cannot directly combine to reproduce the signal. It is necessary to reconstruct the approximations and details before combining with each other.

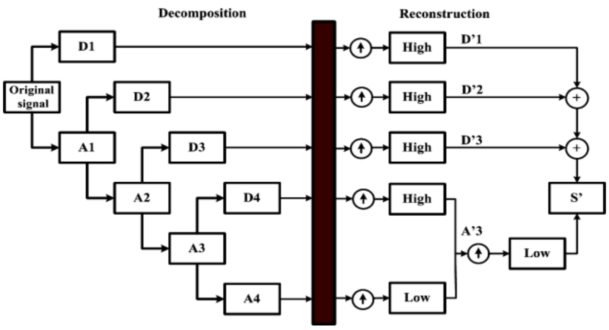


Fig.3 Wavelet decomposition and reconstruction.

C.2 Artificial intelligence-Neural network algorithm

In this article, the neural network models are proposed and used for separating the low frequency and high frequency components of demand in each level. The low frequency components, feature inputs of neural network are selected based on the correlation testing results; include high frequency components of the load demand, and weather factors such as maximum temperature, minimum temperature, mean temperature, humidity, and rainfall. Subsequently, the high frequency component, see in sub-section *C.3*, feature inputs of neural network are also selected based on correlation testing results; include the low frequency components of the load demand and economic factors like consumer price index (CPI) and industrial index (IDI).

Hence, the artificial neural network is one of a good choice to apply for the load demand forecasting problem because this technique is not requiring explicit models to represent the complex relationship between the load demand and factors. The neural network algorithm presented in this paper composes of three layers: the input layer, the hidden layer, and the output layer based on feed-forward back propagation algorithm (FFBP). The input variables come from historical and present data of factors affecting the load demand. The fundamental structure of this algorithm can be presented in Fig.4.

C.3 Correlation method

Equation (2) is the Pearson's coefficient correlation equation. It is used to evaluate correlation between variables and is one of the most familiar measures of dependence between two quantities to show how good a linear relationship among variables or factors is. In the case that the Pearson correlation is +1, it will signify a perfect linearly positive increasing correlation trend. If the Pearson correlation is -1, it will indicate a perfect linearly negative declining correlation trend. If the Pearson correlation is between 1 and -1, it will indicate a degree of linear dependence between the two variables. Lastly, if the Pearson correlation is zero, it shows that there is no relationship between the variables. Before a load forecasting, we must correlate all variables with the energy consumption load demand to choose the appropriate variables for the best feature inputs for the research model.

$$r_{XY} = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum (X_i - \overline{X})^2 . \sum (Y_i - \overline{Y})^2}} \tag{2}$$

Where X and Y are variables and r is the Pearson product moment of correlation.

III. IMPLEMENTATION

A. Case study

To demonstrate the low and high frequency demand features employs the historical electric demand data, which were recorded as monthly from the Electricity Generating Authority of Thailand (EGAT) from January 1997 to December 2007 are decomposed. Weather factor and an economic factor are used in this research.

The historical load and approximate (A) demand, the weather factor, and the economic factor for training and testing in the forecasting model were normalized into the interval 0.00 to 1.00 by using equation (3), as follows:

$$z = \frac{(y - \min(y))}{(\max(y) - \min(y))} \tag{3}$$

Where z is normalized value and y is data information.

In developing model the cyclical component (detail) was normalized into interval -1.00 to 1.00 by using equation (4), as follows:

$$z = \frac{2(y - \min(y))}{(\max(y) - \min(y))} - 1 \tag{4}$$

B. Preprocessing and forecasting stages

In Fig.4 describes the overall structure of the electric load demand forecasting for this research. The wavelet transform and neural network algorithms are used in the research. The main steps proposed for the load demand forecasting model are as follows:

1) The Northern, Northeast, Southern, and Central sub-control center area forecasting will also use six stages below:

1.1) The first, an original signal of load demand is decomposed to high and low frequency by using dB2 mother wavelet (dB2) for calculating the coefficient of the details (D) and approximate (A) components. The level of wavelet transform varies from one level, two levels, three levels, and four levels based on wavelet discrete transform (DWT). From each level decomposition by wavelet method will obtain a detail and a approximate as follows: there are a detail and a approximate components for a level, these are D1 and A1; there are the two details and a approximate components for the two levels, these are D1, D2, and A2 respectively; there are the three details and a approximate for the three levels, these are D1, D2, D3, and A3 respectively ; lastly, there are the four details and a approximate for the four levels of wavelet transform, these are D1, D2, D3, D4, and A1 respectively. All data above after decomposing are recorded and taken them to step 1.2.

1.2) Coefficient components from step 1.1 are reconstructed to the actual components using similar mother wavelet (dB2); these are an actual detail and approximate components of each level.

1.3) Actual detail and approximate components are taken them to find the relationship between each component in each level and factors: temperature, humidity, rainfall, consumer price index, and industrial index. The correlated method by using equation (2) is used.

1.4) The factor that is related with the component more than that 40 percent (up) will be chosen it being as quality inputs for a neural network model. Note that, the factor is chosen in each component of wavelet differing following the related value.

1.5) The feature inputs of a model *NND* forecast consists of the detail component and factors are selected from step 1.4 and the feature inputs of a model *NNA* forecast consists of an approximate (A) and factors are also selected from this step.

1.6) Subsequent to partial forecasts in each sub-model: for example, the 4 levels, sub-model consists of A4, D4, D3, D2, and D1 models; the output for all sub-model are integrated to the final forecasted value.

2) The output of each sub-control area will integrate together for the whole area forecasting of the country.

3) Analyses and summary of all forecast results are carried out.

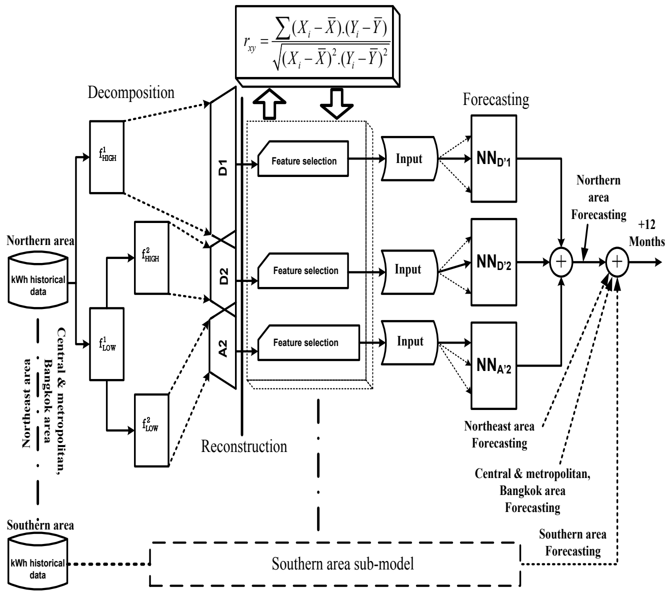


Fig.4 Architecture model of forecasting.

IV. RESULTS AND COMPARISON

The Mean Absolute Percentage Error (MAPE) was calculated by the use of equation (5). MAPE indicates the forecasting accuracy and the identifiable variable in the model, MAPE is given by the following:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|kWh_A^i - kWh_F^i\right|}{kWh_A^i} \times 100 \tag{5}$$

Where *kWhA* is the actual energy consumption load demand, *kWhF* is the forecast load demand, *N* is the number of the month forecasted.

Table I shows the results of forecasting in this research. The column number one, two, three, four, five, and six present the months, actual load demand, one level, two levels, three levels and four levels respectively. The demand is forecasted interval January to December in 2007 unit in kilo-watt hour (*kWh*). Other columns like L-no.1, 2, 3, and 4 show the percentage error in each month in year 2007 after forecasting. Hence, in this table shows the forecasted results obtained from the four models by varying the level of wavelet transform in preprocessing stage. The mean absolute percentage error (MAPE) of one level, two levels, three levels and four levels were 3.58, 2.25, 2.83, and 3.45 percent respectively, were presented in table I, in whole area forecasting. The suitable of level of wavelet has been investigated using the MAPE computation for finding the best level that is, *two levels*. Whereas table II presents sub-control center area forecasting based on same method that is used wavelet and neural network approaches to forecasts. In this table demonstrates the forecasted results obtained from the four models by varying the level of wavelet transform in pre-processing stage. The mean absolute percentage error (MAPE) of one level, two levels, three levels and four levels were 2.36, 2.07, 2.93, and 3.48 percent respectively, were presented in Table II. Furthermore, the suitable of level of wavelet in this investigation has been demonstrated using the MAPE calculation for finding the best level that also is, "*two levels"* similarly.

Therefore, the comparison between the whole area forecasting and sub-control center area forecasting can demonstrate in table I and II. The best level of wavelet transform for two methods is "two levels" whereas when we forecast by using "sub-control center area", it can show better result than "whole area forecasting".

Table I: Mean absolute percent error (MAPE) in each level of whole area forecasting.

Month	**Actual**	**L-no.**			
	kWh	**1**	**2**	**3**	**4**
January	11255610455	14.00	0.65	-0.10	-1.55
February	10903740111	-4.21	-3.41	-4.03	-5.67
March	13256202895	1.35	1.33	0.19	0.66
April	12222850690	3.12	1.90	4.47	3.32
May	12668470924	-2.63	-3.40	2.07	-1.97
June	12768025563	3.61	2.45	5.96	5.76
July	12577211640	0.92	0.53	3.37	3.88
August	12722562991	1.03	1.07	1.70	2.82
September	12504929463	3.07	3.13	3.99	5.46
October	12500303814	1.03	1.19	1.91	4.11
November	11678682183	-5.23	-4.46	-3.93	-0.90
December	11866899916	2.79	3.30	2.21	5.32
MAPE		3.58	**2.25**	2.83	3.45

Table II: Mean absolute percent error (MAPE) in each level of electricity sub-control center area forecasting.

Month	**Actual**	**L-no.**			
	kWh	**1**	**2**	**3**	**4**
January	11255610455	-1.45	-0.49	0.36	-2.02
February	10903740111	-5.22	-4.62	-3.21	-5.81
March	13256202895	0.28	0.07	0.33	0.09
April	12222850690	2.91	1.85	4.12	2.43
May	12668470924	-3.53	-3.68	-2.27	-2.71
June	12768025563	3.15	2.86	5.67	5.13
July	12577211640	0.52	0.55	6.95	7.09
August	12722562991	0.43	0.37	1.59	2.46
September	12504929463	2.34	2.34	3.44	4.85
October	12500303814	0.18	0.55	1.91	3.98
November	11678682183	-6.94	-5.66	-3.39	-0.91
December	11866899916	1.35	1.76	1.85	4.28
MAPE		2.36	**2.07**	2.93	3.48

V. CONCLUSION

The load demand forecasting based on the suitable level of wavelet transform, and a neural network method, approaches to the whole and sub-control center area forecasting, is proposed. The load demand data which employs the historical data from Electricity Generating Authority of Thailand (EGAT) is decomposed into one, two, three, and four levels of different frequencies. The correlation technique is used to choose the suitably affecting factors for each frequency component of load demand. The factors which affected in each level are chosen for neural network inputs. As a result, *"two levels"* is given the Mean Absolute Percentage Error (MAPE) better than that other level and *"sub-control center area"* can show better result than "*whole area forecasting*". The objective of load forecasting has significantly forecast for fuel reserve planning in the power system.

ACKNOWLEDGMENT

This work was funded by the Office of the Higher Education Commission. Pituk Bunnoon was supported by CHE510382 Ph.D. Scholarship. The authors would like also to thank the Thai Meteorological Department, Ministry of Transport and Communications; the Ministry of Commerce; the Office of the National Economic and Social Development Board; and lastly the Electricity Generating Authority of Thailand (EGAT) in providing her valuable data and information.

REFERENCES

[1] Qi Wu, "Hybrid model based on wavelet support vector machine and modified genetic algorithm penalizing Gaussian noises for power load forecasts," *International journal of expert systems with applications* pp.379-385, 2011.

[2] Ying Chen, Peter B. Luh, Che Guan, Yige Zhao, Laurent D. Michel et.al., "Short-term load forecasting: Similar day-based wavelet neural networks," *IEEE Trans.on power syst.*, vol.25, pp.322-330, 2010.

[3] Ajay Shekhar Pandey, Devender Singh, and Sunil Kumar Sinha, "Intelligent hybrid wavelet models for short-term load forecasting," *IEEE Trans.on power syst.*, vol. 25, pp.1266-1273, 2010.

[4] N. Amjady, and F. Keynia, "Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm," *International journal of energy,* vol.34, pp.46-57, 2009.

[5] N. M. Pindoriya, S. N. Singh, and S. K. Singh, "An adaptive wavelet neural network based energy price forecasting in electricity markets," *IEEE Trans.on power syst.,* vol.23, pp.1423-1432, 2008.

[6] D. Benaouda, F. Murtagh, J.-L. Starck, and O. Renaud, "Wavelet-based nonlinear multi-scale decomposition model for electricity load forecasting," *International journal of neuro computing,* vol.70, pp.139-154, 2006.

[7] Tai Nengling, Jurgen Stenzel, and Wu Hongxiao, "Techniques of applying wavelet transform into combined model for short-term load forecasting," *International journal of electric power systems research,* vol.76, pp.525-533, 2006.

[8] Bai-Ling Zhang, and Zhao-Yang Dong, An adaptive neural-wavelet model for short term load forecasting," *International journal of electric power systems research,* vol.59, pp.121-129, 2001.

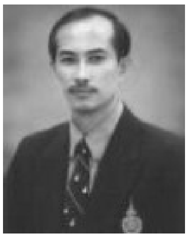
[9] Tongxin Zheng, Adly A. Girgis, and Elham B. Makram, "A hybrid wavelet-Kalman filter method for load forecasting," *International journal of electric power systems research*, vol.54, pp.11-17, 2000.

[10] S. J. Yao, Y. H. Song, L. Z. Zhang, and X. Y. Cheng, Wavelet transform and neural network for short-term electrical load forecasting, *International journal of energy conversion and management,* vol.41, pp.1975-1988, 2000.

Pituk Bunnoon *(Member IEEE),* received the B.S. degree from King Mongkut's Institute of Technology Ladkrabang, Thailand, in 1998, and the M.S. degree in electrical engineering from Prince of Songkla University, Thailand, in 2004.

His research interest is an application of artificial intelligence to power system planning and operation.

Kusumal Chalermyanont received the B.S. degree from Prince of Songkla University, Thailand, in 1993, the M.S. degree in electrical engineering from University of Colorado at Boulder in 1999, and the Ph.D. in electrical engineering from the University of Colorado at Boulder in 2003. Her research interests are power electronics, magnetic designs for power electronics, renewable energy system/management.



Chusak Limsakul received the B.S. degree from King Mongkut's Institute of Technology Ladkrabang, Thailand, in 1978, and the D.E.A. degree from INSAT France, in 1982, and Docteur Ingenieur form INSAT France, in 1985. His research interests are digital signal processing, sensors and instrumentations, and automation.