

Short-term Electric Load Forecasting Based on Wavelet Transform and GMDH

Bon-Gil Koo*, Heung-Seok Lee* and Juneho Park[†]

Abstract – The group method of data handling (GMDH) algorithm has proven to be a powerful and effective way to extract rules or polynomials from an electric load pattern. However, because it is nonstationary, the load pattern needs to be decomposed using a discrete wavelet transform. In addition, if a load pattern has a complicated curve pattern, GMDH should use a higher polynomial, which requires complex computing and consumes a lot of time. This paper suggests a method for short-term electric load forecasting that uses a wavelet transform and a GMDH algorithm. Case studies with the proposed algorithm were carried out for one-day-ahead forecasting of hourly electric loads using data during the years 2008–2011. To prove the effectiveness of our proposed approach, the results were evaluated and compared with those obtained by Holt-Winters method and artificial neural network. Our suggested method resulted in better performance than either comparison group.

Keywords: Discrete wavelet transform, GMDH, Short-term electric load forecasting

1. Introduction

The importance of short-term electric load forecasting (STLF) has been growing over the years. STLF is one of the significant factors in the operation of an electricity grid. Traditionally, accurate STLF can lead to the proper selection of the amount of reserve margin, which contributes to the improved efficiency of the power supply. STLF also helps in the controlling and scheduling of power systems. By contrast, inaccurate STLF can cause a residual or short supplement and may lead to overinvestment or shortage of electric power.

STLF is necessary not only for the stable and efficient operation of grid but also for appropriately accommodating new industrial restructuring. Nowadays, as the electric power market experiences privatization and deregulation, accurate STLF has become a matter of common interest to market participants. In the near future, demand response will be available owing to technical improvements in metering and real-time communication. Tools such as advanced metering infrastructure (AMI) will help customers access time-varying electricity prices [1]. Thus, electric load forecasting will have a key role in electric market.

Various models have been established for STLF over the course of several decades. Each model can be classified into three categories by their base method, statistical method, intelligent method and hybrid method. For statistical method, the regression method are quite common. To meet the specific purpose, improved regression algorithm, such

as ARMA, ARIMA, have been presented [2-3] However, functional relationship between electric load and various weather variables or other social variables. Nowadays, the methods based on intelligence algorithm such as Artificial Neural Network, Fuzzy algorithm can be alternatives to overcome the finding function. They also flexible and robust for nonlinear problem. So, during recent decade, these methods are focused to solve nonlinear relationship between meteorological data and electric load data [4-7]. There are several hybrid method to complement the single algorithm model. [8-9]

Electric load data are time series data that have a certain period. However, time series data are the sum of information from various periods. The period for time series data consists of a weekly day-type pattern. Time series data also have a shorter-term pattern that comes from industrial load, and a longer-term pattern that comes from seasonal or weather changes. For this reason, it is helpful to analyze the origin of data after we divide the data into some frequencies. To express the data in the frequency domain, the Fourier transform is one of most popular methods. However, practical electric load data are nonstationary, so a Wavelet transform is more suitable than the Fourier transform. Many papers have been published that explore Wavelet decomposition. Most of these papers have featured an approach that uses artificial neural networks [4-7]. This approach showed relatively high accuracy compared to conventional forecasting methods [7].

2. Wavelet Transform

Using a wavelet means that certain seismic signals can be suitably modeled by combining translations and dilations

[†] Corresponding Author: Dept. of Electrical and Electronic Engineering, Pusan National University, Korea. (parkjh@pusan.ac.kr)

* Dept. of Electrical and Electronic Engineering, Pusan National University, Korea. (koobg@pusan.ac.kr, zznada@naver.com)

Received: May 1, 2014; Accepted: December 10, 2014

of a simple oscillatory function of finite duration [10]. Unlike trigonometric functions, this wave enables us to analyze a nonstationary signal using a pulse that is located in time and has finite energy that integrates to zero.

2.1 Wavelet decomposition and reconstruction

Wavelet decomposition is a method to gather information from the breaking up of a signal into shifted and scaled versions of the mother wavelet. As it travels, the original signal becomes a lower-resolution signal. There are two types of filters (low-pass and high-pass) that can decompose or reconstruct the signal. Both the low- and high-pass filters form a system. The original signal *S* is decomposed into details (abbreviated as *D1*) and approximations (abbreviated as *A1*) by going through each filter. The coefficient *D1*, obtained from the high-pass filter, contains a high-frequency component that describes the short-term period pattern. The other coefficient, *A1*, is obtained from the low-pass filter and contains a low-frequency component that describes the long-term period pattern. If the decomposition level is two or higher, only an approximate coefficient is decomposed. The coefficient *A1* is decomposed into *A2* and *D2*. The next level of decomposition is achieved by repeating this sequence. The decomposition can continue only until the individual details consist of a single sample or pixel [11]. After this process is complete, the signal can be smoothed by eliminating the high frequency or by separating it into high frequency and low frequency. This is shown in Fig. 1.

Wavelet-decomposed components can be reconstruct without loss of data. This process is called reconstruction, and the procedure is shown in Fig. 2.

The major characteristic of reconstruction is that it can reconstruct the signal from the coefficients of the details and approximations. The original signal can be recon-

structed from the coefficient vectors of the approximations and details. During this process, reconstruction only needs coefficient vectors, which are produced by downsampling the length of the signal by half. Therefore, before reconstruction, the coefficient should be modified by allocating zeros between samples.

Using multiresolution analysis theory, a fast discrete wavelet transform known as the Mallat algorithm was developed. When the signal is decomposed, it uses different resolution levels.

Fig. 1 shows this process using a signal that has been decomposed into three resolution levels. As shown in Fig. 1, the Mallat algorithm focuses on a careful decomposition of the low-frequency space of the signal, resulting in better resolution in the low-frequency series.

Compared with the original signal, the series decomposed by the Mallat algorithm shows a great reduction in the nonstationary characteristics. Because the series decomposed by wavelet decomposition is an approximate stationary series, wavelet decomposition is useful in improving the forecasting accuracy.

A three-level wavelet transform can be divided into three details and three approximations, as seen below:

$$\begin{aligned} S &= A1 + D1 \\ &= A2 + D2 \\ &= A3 + D3 + D2 + D1 \end{aligned}$$

2.2 Continuous and discrete wavelet transforms

Wavelet Transform is consist of a continuous Wavelet and a discrete wavelet transform (DWT). $W(a, b)$ of signal $f(x)$, with respect to a mother wavelet $\psi(x)$, is described as [12]

$$W(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) dt \tag{1}$$

where a and b are real, and $*$ denotes complex conjugation. The scale parameter a controls the spread of the wavelet, and translation parameter b determines its central position.

The $W(a, b)$ coefficient represents how well the original signal $f(x)$ and the scaled / translated mother wavelet match. Thus, for all a, b associated with a particular signal, the set of all wavelet coefficients $W(a, b)$ is the wavelet representation of the signal with respect to the mother wavelet. The CWT provides a redundant representation of the signal in the sense that the entire support of $W(a, b)$ need not be used to recover $f(t)$. Instead of using that approach, the mother wavelet can be scaled and translated using certain scales and positions usually based on powers of two. This is known as the DWT. This scheme is more efficient than, and is as accurate as, the CWT [10]. The DWT is described in the following equation:

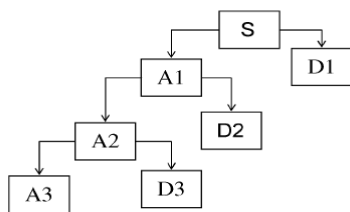


Fig. 1. Process of a signal being decomposed into three resolution levels.

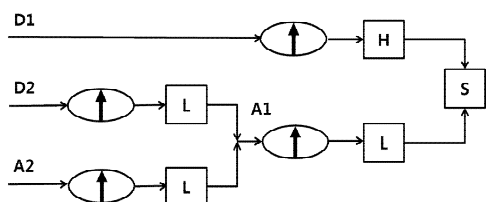


Fig. 2. Reconstruction tree of decomposed signal.

$$DWT(m, k) = \frac{1}{\sqrt{a_0^m}} \sum_n f(n) \psi\left(\frac{k - nb_0 a_0^m}{a_0^m}\right) \quad (2)$$

where a, b are same in (1) and can be expressed in m. In (2), k is an integer variable that refers to a particular point of the input signal, and n is the discrete time index.

3. GMDH(Group Method of Data Handling) Algorithm

The GMDH algorithm was first developed by A.G. Ivakhnenko. It is a heuristic self-organization method. The GMDH can formulate or identify a complex system without tracking the path of input-output. Generally, the relationship between inputs and output in a nonlinear function can be expressed by a complicated polynomial series in the form of the Volterra series:

$$y(t) = \int_0^t h_1(\tau)x(t-\tau)d\tau + \iint_0^t h_2(\tau_1\tau_2)x(t-\tau_1)x(t-\tau_2)d\tau_1d\tau_2 + \iiint_0^t h_3(\tau_1\tau_2\tau_3)x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)d\tau_1d\tau_2d\tau_3 + \dots \quad (3)$$

The discrete form of (3) can be expressed by (4) named Kolmogorov-Gabor polynomial [13]

$$y = a_0 + \sum_i^N a_i x_i + \sum_i^N \sum_j^N a_{ij} x_i x_j + \sum_i^N \sum_j^N \sum_k^N a_{ijk} x_i x_j x_k + \dots \quad (4)$$

where x represents the input of the system, N is the number of inputs, and a represents coefficients.

In the case of double variables and the quadratic form, (4) can be simplified by (5)[14]

$$y_k = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2 \quad (5)$$

To obtain coefficient a, the algorithm follows this procedure.

- Step 1:** Select m input variables affecting the output variable. If required, normalize input data. This input data calls x_1, x_2, \dots, x_m .
- Step 2:** Divide N input-output data ($x_{1i}, x_{2i}, \dots, x_{ni}, y_i, i = 1, 2, \dots, N$) into training data and testing data. $N = N_{train} + N_{test}$
- Step 3:** Select input data x_p, x_q among n input data. Establish double-variable, second-order partial descriptions (PD).

$$G_k = a_0 + a_1 x_p + a_2 x_q + a_3 x_p x_q + a_4 x_p^2 + a_5 x_q^2 \quad (6)$$

where a_0, a_1, \dots, a_5 are coefficients, and G_k is a intermediate-variables

- Step 4:** Estimate coefficients a_0, a_1, \dots, a_5 , using linear regression and its input training data. Coefficients can be obtained by minimizing E in the equation (8).

$$E_L = \sum_{i=1}^{N_{train}} (y_i - G_{Li})^2 \quad L = 1, \dots, m(m-1)/2 \quad (7)$$

- Step 5:** Using coefficients obtained from the previous step, calculate the square error E_L . Among $E_1, E_2, \dots, E_{n(n-1)/2}$, the lowest error E is assigned E_{best} . If the polynomials and intermediate variables satisfy the condition $E_{best} \leq E_c$, terminate the algorithm, or arrange organized polynomials by error E_L from lowest to highest. Select m polynomials in order of lowest error.

- Step 6:** Using the selected polynomial in the previous step, construct the next layer using new input-output data, and repeat the process beginning with Step 3.

This procedure can be shown as Fig. 3.

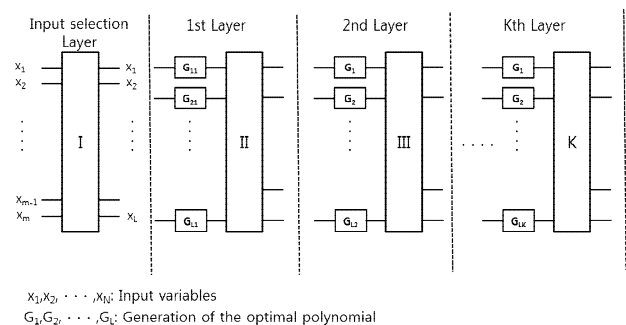


Fig. 3. GMDH network.

4. Case Study

4.1 Electric load classification

Before electric load forecast, electric load data has to be classified to their day-type. To do this, normally has been used calendar based method. However, electric load data does not always follows their day-type, it has possibility to obtain incorrect data with respect to their genuine day-type characteristic. To avoid error, this paper classified electric load data into seasonal and day-type using K-mean and k-NN respectively.

4.2 Electric load decomposition

For our proposed approach, we selected the DWT with

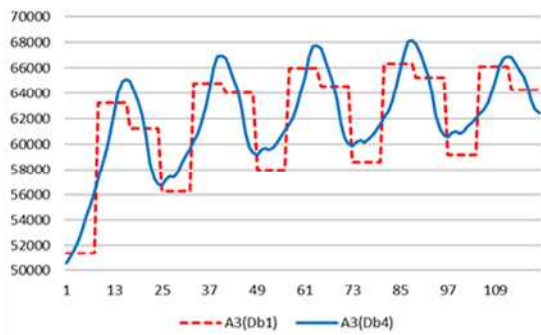


Fig. 4. Comparison of results using Db1 and Db4

Mallat method to perform the load decomposition. For basis function $f(t)$, we selected Db1 (Daubechies wavelet of order 1). Some of the existing research with the wavelet transform in electric load forecasting and price forecasting usually used Db4 (Daubechies wavelet of order 4), but this approach was not appropriate for GMDH. Generally, GMDH has inputs at the utmost 2nd order polynomial, so it is better to fit the curve of the load as a square rather than as a quadratic or higher order curve.

Fig. 4 shows different results of the three-layer wavelet decomposition using Db1 and Db4, respectively. In Fig. 4, we see that Db1 disassembles the original load into a suitable wave to be fit by lower order polynomials than Db4 does.

4.3 Electric load forecasting and development of whole forecasting model

GMDH has two input units, which describe the previous load and temperature trend by the hour. The layer is modeled to be created repeatedly, until there is no more improvement in the lowest error. Data groups utilize the previous 72 h, divided into 48 h of training data and 24 h of test data. Each model can have different numbers of layers, but the maximum number is set at 100.

The variation of data structure of the proposed forecasting system is shown in Fig. 5. Before applying the wavelet transform, the electric load is classified by its day-type and seasonal type. We divide the electric data into four seasons commonly used and five day-types such as Monday, Weekdays, Saturday, Sunday, Special holiday. Each classified electric load is used to establish forecasting models except for special holiday(summer vacation, new year’s day, etc). The load is decomposed using the wavelet transform and is forecast using GMDH, Holt-Winters method, and an

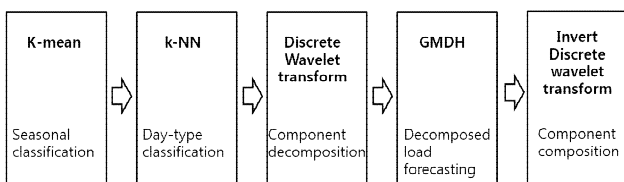


Fig. 5. Structure of overall forecasting system

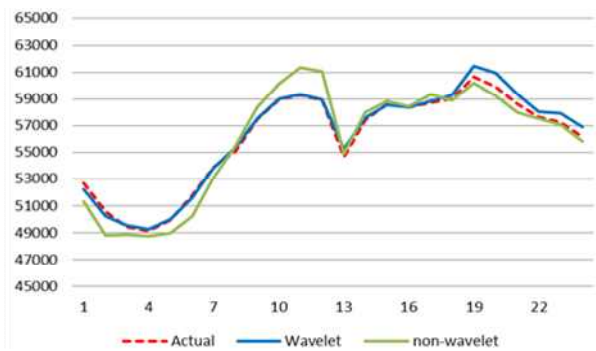


Fig. 6. Comparison of applying wavelet and non-wavelet forecasted

artificial neural network.

4.4 Test and discussion

Case studies for our proposed algorithm involved a one-day-ahead forecasting of hourly electric loads using the data from 2008–2011. To compare our data with the data from other papers, we evaluated our results using the following indices:

Mean Absolute Percentage Error(MAPE)

$$E_{MAPE}(i) = \frac{1}{N} \left[\sum_{t=1}^N \frac{|\hat{y}(i,t) - y(i,t)|}{y(i,t)} \right] \times 100 \quad (8)$$

Standard deviation

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N \{y(i,t) - \hat{y}(i,t)\}^2} \quad (9)$$

To prove the validity of our proposed method, we compared our results with the same data from the applications of Holt-Winters multiplicative method and the artificial neural network. We found the smoothing constant $\alpha=0.5$, $\beta=1$, $\gamma=0.1$ by increasing each value from 0.1 to 1. We selected the multilayer perceptron type with 26 input nodes, 2 layers(one of which is hidden) and 24 output neurons for Artificial Neural Network.[15] It was trained by back propagation algorithm, while the selected iteration tolerance after the error is less than 0.005. Table 1 shows day-ahead forecasting results, for each day types in MAPE. Table 2 lists the standard deviations and MAPE of the prediction errors, for each hour and for all types, in 2008 and 2011.

According to Table 1, for spring season forecasting, our proposed method improves the accuracy in MAPE difference over Holt-Winters, ANN by 1.592%, 0.312%, respectively, for summer season forecasting, by 1.191%, 0.294%, respectively, for fall season forecasting, by 1.9%, 0.317%, respectively, for winter season forecasting, by 0.568%, 0.277%, respectively. For all day-types in each

Table 2. Comparison of forecasting results in day-type

		Holt- winters MAPE	ANN MAPE	GMDH MAPE
Spring	Mon	2.227	1.172	0.875
	Week-days	2.200	1.184	0.877
	Sat	2.948	0.864	0.521
	Sun	2.620	1.004	0.688
	total	2.385	1.105	0.793
Summer	Mon	2.022	1.265	0.994
	Week-days	1.937	1.297	1.005
	Sat	2.653	0.990	0.664
	Sun	2.435	1.089	0.796
	total	2.118	1.221	0.927
Fall	Mon	2.146	1.207	0.913
	Week-days	2.153	1.207	0.921
	Sat	2.807	0.929	0.616
	Sun	2.598	1.015	0.698
	total	2.259	1.161	0.868
Winter	Mon	1.513	1.478	1.209
	Week-days	1.561	1.460	1.189
	Sat	2.262	1.157	0.849
	total	1.693	1.403	1.125
Total	Sun	2.025	1.262	0.972
	Mon	1.941	1.296	1.016
	Week-days	1.924	1.303	1.014
	Sat	2.633	0.999	0.676
	Sun	2.383	1.108	0.807

Table 2. Comparison of forecasting results in hour

Hour	Holt-winters		ANN		GMDH	
	MAPE	σ	MAPE	σ	MAPE	σ
1	2.321	0.700	1.132	0.299	0.823	0.325
2	2.650	0.756	0.993	0.320	0.673	0.340
3	2.838	0.796	0.920	0.365	0.600	0.379
4	2.923	0.811	0.878	0.338	0.561	0.343
5	2.873	0.814	0.899	0.340	0.581	0.349
6	2.719	0.789	0.964	0.332	0.643	0.351
7	2.475	0.715	1.067	0.302	0.752	0.326
8	2.294	0.705	1.145	0.298	0.838	0.322
9	1.996	0.805	1.272	0.339	0.982	0.364
10	1.778	0.849	1.366	0.358	1.088	0.380
11	1.662	0.824	1.416	0.346	1.143	0.369
12	1.646	0.813	1.423	0.341	1.150	0.364
13	1.932	0.760	1.300	0.319	1.013	0.341
14	1.736	0.790	1.384	0.332	1.106	0.356
15	1.676	0.811	1.409	0.341	1.134	0.366
16	1.715	0.809	1.393	0.349	1.116	0.386
17	1.717	0.811	1.391	0.343	1.114	0.369
18	1.721	0.780	1.389	0.333	1.110	0.361
19	1.669	0.697	1.411	0.298	1.132	0.329
20	1.627	0.603	1.429	0.258	1.151	0.287
21	1.681	0.552	1.406	0.236	1.125	0.263
22	1.792	0.517	1.358	0.221	1.072	0.246
23	1.776	0.507	1.365	0.217	1.079	0.242
24	1.861	0.605	1.329	0.259	1.040	0.287
Mean	2.045	0.734	1.252	0.311	0.959	0.334

season, our proposed algorithm outperforms other methods.

As shown in Table 2, for overall mean performance, our proposed method results in the lowest forecasting error but not in standard deviation. For all hourly data, it shows better performance than other two methods. The results

shows, it represents 53% and 23% higher accuracy with respect to other two algorithms. Most of all, we can find its competitive performance for peak load period (2-3 P.M. in summer, 10 A.M. in winter). Whereas, for all hourly results, our proposed method doesn't show better performance than ANN in standard deviation.

Two additional test cases — weekday data only (case 1) and weekend data included (case 2) — showed case 1 with 0.96% and case 2 with 1.01%, respectively. There was not much of a difference.

5. Conclusion

This paper proposed a new algorithm using DWT and the GMDH algorithm. The forecasting was performed after we classified the electric loads by their day-type and seasonal type, using a data mining technique. Refined electric load data were decomposed into their frequency. We then used GMDH to perform load forecasting, and the data were reconstructed by using an inverted DWT.

Regarding accuracy, our proposed algorithm showed a lower error rate over the tested period, not only in the comparison group but also in reference paper. Most of all, its accuracy for peak time was competitive. Whereas, it is necessary to improve in standard deviation. Including additional meteorological factors such as wind speed, solar irradiance, and rainfall rate can help to decrease the error rate.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF-2013R1A1A2013798).

References

- [1] U. S. Department of Energy, "Energy policy Act of 2005," *section 1252*, Feb. 2006.
- [2] Amjady, N., "Short-term hourly load forecasting using time-series modeling with peak load estimation capability," *IEEE Transactions on Power Systems*, Vol. 16, no. 3, 2001, pp. 498-505
- [3] Shyh-Jier Huang, Kuang-Rong Shih, "Short-term load forecasting via ARMA model identification including non-Gaussian process considerations," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 673-679, May 2003.
- [4] D. Singh and S. P. Singh, "A self selecting neural network for short-term load forecasting," *Electric Power Component System*, vol. 29, pp. 117-130, 2001.
- [5] D. K. Ranaweera, N. F. Hubele, and G. G. Karady, "Fuzzy logic for short-term load forecasting," *Electric Power Energy System*, vol. 18, no. 4, pp. 215-222,

- 1996.
- [6] D. Srinivasan, A. C. Liow, and C. S. Chang, "Forecasting daily load curves using a hybrid fuzzy neural approach", *IEE Proceedings Generation Transmission Distribution*, vol. 141, no. 6, pp. 561-567, Nov. 1994.
 - [7] J. H. Park, Y. M. Park, K. Y. Lee, "Composite Modeling for Adaptive Short-term Load Forecasting", *IEEE Transactions on Power Systems*, vol. 6, no. 2, pp. 450-457, May, 1991.
 - [8] A.A El Desouky, Elkateb.M.M, "Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA" *Generation, Transmission and Distribution, IEE Proceedings*, vol.147, no.4, pp.213-217, 2000.
 - [9] Dai-Zheng Huang, Ren-Xi Gong, Shu Gong, "Prediction of Wind Power by Chaos , BP Artificial Neural Networks Approach Based on Genetic Algorithm" *Journal of Electrical Engineering & Technology*, vol.10, no.1, pp.41-46, Jan., 2015
 - [10] Agnaldo J.Rocha Reis, Alexandre P.Alves da Silva, "Feature Extraction via Multiresolution Analysis for Short-term Load Forecasting", *IEEE Transactions on Power Systems*, vol. 20, no. 1, Feb. 2005
 - [11] Ajay Shekhar Pandey, Devender Singh, Sunil Kumar Sinha, "Intelligent Hybrid Wavelet Models for Short-Term Load Forecasting", *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1266-1273, Aug., 2010
 - [12] Yanqiu Bi, Jianguo Zhao, Dahai Zhang, "Power Load Forecasting Algorithm Based on Wavelet Packet Analysis", *Power System Technology – POWERCON*, vol. 1, pp. 987-990
 - [13] Aiyun Zheng, Weimin Liu, Fanggeng Zhao, "Double Trends Time Series Forecasting Using a Combined ARIMA and GMDH Model" *Control and Decision Conference (CCDC), 2010 Chinese*, pp. 1820-1824, May. 2010
 - [14] Ehab. E. Elattar, John Yannis Goulermas, "Generalized Locally Weighted GMDH for Short Term Load Forecasting", *IEEE Transactions on Systems, Man and Cybernetics-part C: Applications and Reviews*, Vol. 42, no. 3, pp. 345-356, May 2012
 - [15] Henrique Steinherz Hippert, Carlos Eduardo Pedreira, Reinaldo Castro Souza, "Neural Networks for Short-Term Load Forecasting:A Review and Evaluation" *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 44-55, Feb. 2001



Bon-Gil Koo He received his B.S. and M.S. Degrees from Pusan National University, Busan, Korea in 2008 and 2010, respectively. He is currently PH.D candidate at Pusan National University. His research interests include electric load forecasting and intelligent systems applications to power systems.



power systems.

Heung-Seok Lee He received his B.S. Degrees from Ulsan University, Ulsan, Korea in 2012. He received his M.S. Degrees from Pusan National University, 2014. He is currently PH.D candidate at Pusan National University. His research interests include electric load forecasting and also electric ship



include intelligent systems applications to power systems. Dr. Park has been a member of the IEEE Power Engineering Society.

June Ho Park He received the B.S., M.S. and Ph.D. degrees from Seoul National University, Seoul, Korea in 1978, 1980, and 1987, respectively, all in electrical engineering. He is currently a Professor at the School of Electrical Engineering, Pusan National University, Busan, Korea. His research interests