

## Short-Term Load Forecasting Using System-Type Neural Network Architecture

Shu Du+, Kwang Y. Lee++

+Graduate Student ([Shu\\_Du@baylor.edu](mailto:Shu_Du@baylor.edu))

++ Professor and Corresponding Author ([Kwang\\_Y\\_Lee@baylor.edu](mailto:Kwang_Y_Lee@baylor.edu))

Department of Electrical and Computer Engineering  
Baylor University

### Abstract

This paper presents a methodology for short-term load forecasting using a system-type neural network based on semigroup theory. A technique referred to as algebraic decomposition is proposed for the modeling of electric power load demand in terms of the coefficient vector and the basis vector, and a new learning algorithm based on semigroup theory is put forward for extrapolation of the coefficient vector. Due to the non-stationary attribute of the load, the actual load is preprocessed by regression to become better correlated to daily time and temperatures. A rearrangement method based on the hourly temperature is developed to solve the problem of the roughness of the coefficient vector. With the algebraic decomposition of the rearranged regression load, a much smoother coefficient curve can be obtained. Based on the smoothness, interpolation and extrapolation can be achieved for each hour using the historical hourly temperatures and the hourly temperature forecast. The interpolated or extrapolated coefficient vector is recombined with the basis vector for each hour, and the recombined hourly load are grouped to form the final load forecast of the target day. A moving window slides through the whole year to perform the day-ahead load forecasting. Load data from New England Independent System Operator (ISO) is used to verify the capability of the proposed approach.

### I. Introduction

Accurate load forecasting is very important for electric utilities in a competitive environment created by the electric industry deregulation. In order to supply high quality electric energy to the customer in a secure and economic manner, an electric company faces many economical and technical problems in operation, planning and control of an electric energy system<sup>1</sup>. Load forecasting helps an electric utility to make important decisions including decisions on purchasing and generating electric power, load switching, and infrastructure development. Load forecasting is also important for energy suppliers, financial institutions, and other participants in electric energy generation, transmission, distribution, and markets<sup>2</sup>.

The load is a non-stationary process which is affected by two main factors: time of the day and weather conditions. The time dependence of the load reflects the existence of a

daily load pattern, which may vary for different weekdays and seasons. Temperature is the primary weather factor affecting load. Humidity and wind speed may also influence power consumption. In general, the load has two distinct patterns: weekday and weekend patterns. In addition, holiday patterns are different from non-holiday patterns.

Section II describes the proposed approach. In Section III, the proposed approach will be applied to the forecasting problem. Finally, some conclusions are drawn and presented in Section IV.

## II. Proposed Method

### A. System-Type Neural Network Method

In previous papers<sup>3-8</sup>, a system-type neural network which implemented extrapolation was proposed. In this method, the distributed parameter system (DPS) surface determined by a given data set was expanded along one axis. The load is in general a function,  $Load = f(Day, Hour, Weather, Customer\ classes)$ , and is often considered as  $Load = f(Day, Hour)$ , which is parameterized by weather and customer classes. In this paper, however, the load is treated as  $Load = f(Temperature, Hour)$  due to the importance of the temperature among many factors that affect the load. In the next section, it will be shown that the load can be represented in the following form:  $L(Temperature, Hour) = C(T)^T E(H)$ , where  $E(H)$  is a vector of linearly independent orthonormal basis functions, and  $C(T)$  is a preliminary coefficient vector as a function of the hourly temperatures. To obtain  $E(H)$  and  $C(T)$ , a technique referred to as algebraic decomposition, which aims to approximate and model load data set, is involved. This process at first requires load profile  $L(T, H)$  to be parameterized as  $\{L_T(H)\}, T = T_1, \dots, T_l$ , then converts  $n$  chosen members of  $\{L_T(H)\}$  using Gram-Schmidt process to obtain an orthonormal basis set  $E(H) = [e_1(H), e_2(H), \dots, e_n(H)]$ . Finally the least squares method is performed to determine the coefficient vector  $C(T)$ , where  $L(T, H) = C(T)^T E(H)$  and  $C(T) = [c_1(T), c_2(T), \dots, c_n(T)]$ .

Neural networks are being used for systems described by partial differential equations (PDEs)<sup>9</sup>. The system-type architecture is shown in Fig. 1, which implements an arbitrary function  $L(T, H)$ . The proposed architecture reflects a system-type approach using two neural network channels, a Function Channel and a Semigroup Channel. The Semigroup Channel supplies the Function Channel with a coefficient vector  $C(T)$  as a function of the index  $T$ . The Function Channel provides a vector of basis functions  $E(H)$ . These two channels realize a semigroup-based implementation of the mapping  $L(T, H)$ . The Function Channel can have a Radial Basis Function (RBF) architecture<sup>10</sup>. It consists of  $n$  RBF networks, each representing one of the  $n$  orthonormal basis functions in  $E(H)$ . The outputs of the orthonormal vectors are (internally) linearly summed so that the channel spans an  $n$ -dimensional function space. The Semigroup Channel can be adapted from the Diagonal Recurrent Neural Network (DRNN) or the Elman architecture<sup>8</sup>, in which the

input is split into a dynamic scalar component  $T$  and one static vector component, the initial vector  $C(0)$ . The output is a vector  $\tilde{C}(T)$ , which is related to the dynamic input  $T$  and to the static input  $C(0)$  by the semigroup property:  $\tilde{C}(T) = \Phi(T)C(0)$ , where  $\Phi(T_1 + T_2) = \Phi(T_1)\Phi(T_2)$ .

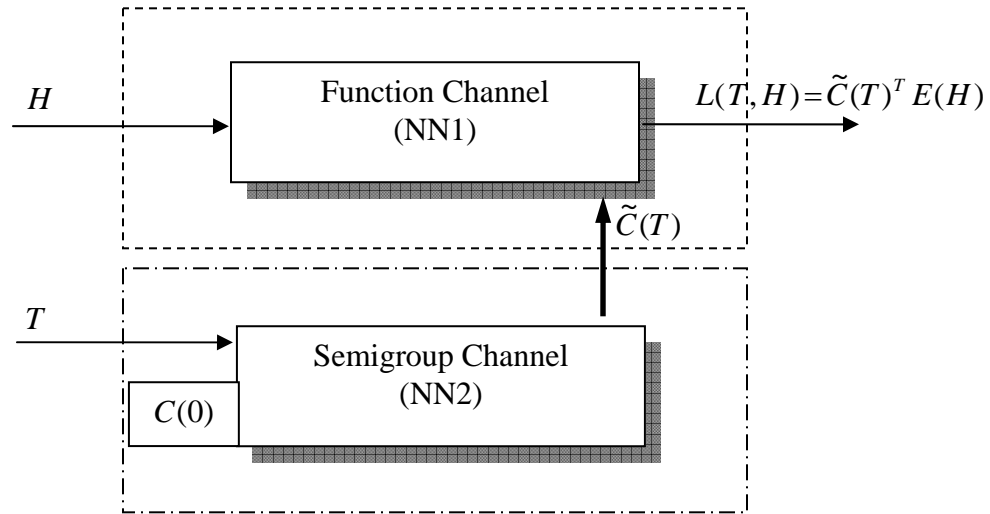


Figure 1. System-Type Architecture.

## B. Learning Algorithm of Proposed System-type Neural Network

The second component, the Semigroup Channel, can be trained in a successive way illustrated in Fig. 2. During training, the Semigroup Channel receives a preliminary coefficient vector  $C(T)$  as input and produces a smoothened coefficient vector  $\tilde{C}(T)$  as output. That is, the primary objective of training is to replicate (and, if necessary, to smoothen) the vector  $C(T)$  with a vector  $\tilde{C}(T)$ , which has the following semigroup property:  $\tilde{C}(T) = \Phi(T)C(0)$ , where  $\tilde{C}(T) = [\tilde{c}_1(T), \tilde{c}_2(T), \dots, \tilde{c}_n(T)]^T$  and  $\Phi(T)$  is an  $n \times n$  matrix that satisfies:  $\Phi(T_1 + T_2) = \Phi(T_1)\Phi(T_2)$ . However, there is a secondary objective of training: the channel must also “replicate” the semigroup property of the trajectory by gradually acquiring a semigroup property of its own, in the weight space. The existence of this acquired semigroup property in the weight space becomes the basis for extrapolation. In order to elicit this gradual acquisition of the semigroup property, it is necessary that the training in this second step (semigroup tracking) occur in a gradual manner, as shown in Fig. 2.

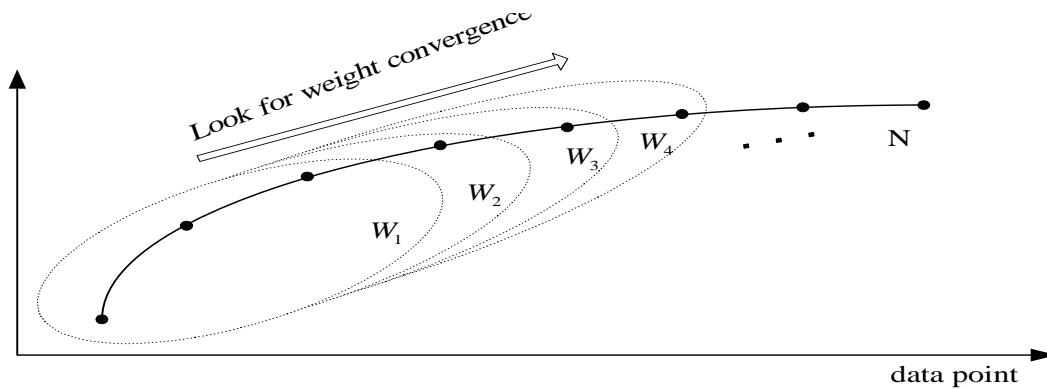


Figure 2. Overview of New Training Algorithm.

### C. Regression

Regression method is one of the most widely used approaches for load forecasting<sup>11</sup>. However, it is not playing the leading role to perform forecasting in this paper. The electric load is a complex system influenced by many factors. It can be expressed as the following form:

$$L_{total} = L_{base} + L_{weather} + L_{other\ factors}$$

where  $L_{base}$  is the base load which is caused by time factor;  $L_{weather}$  is the weather sensitive load which is due to weather variables, temperature being usually dominant among various weather variables; and  $L_{other\ factors}$  is a component of the load resulting from other factors. Therefore, regression is used here to filter the load and remove the load component caused by other factors.

### D. Rearrangement of Load

Assuming the load changes smoothly with temperature, the filtered load is rearranged according to the hourly temperature. If the data surface of the rearranged load is smooth, the extrapolation along a single coordinate (temperature) can be performed. Because the temperature is changing hourly, it is necessary to rearrange the load with respect to the temperature at each hour so that a smooth load surface can be obtained.

### E. Interpolation and Extrapolation

Interpolation and extrapolation involve only the coefficient vector modeled in the Semigroup Channel. Fig. 3 illustrates the interpolation and extrapolation procedures for a given hour. Red circles represent the historical load and green circles represent the forecasting load.

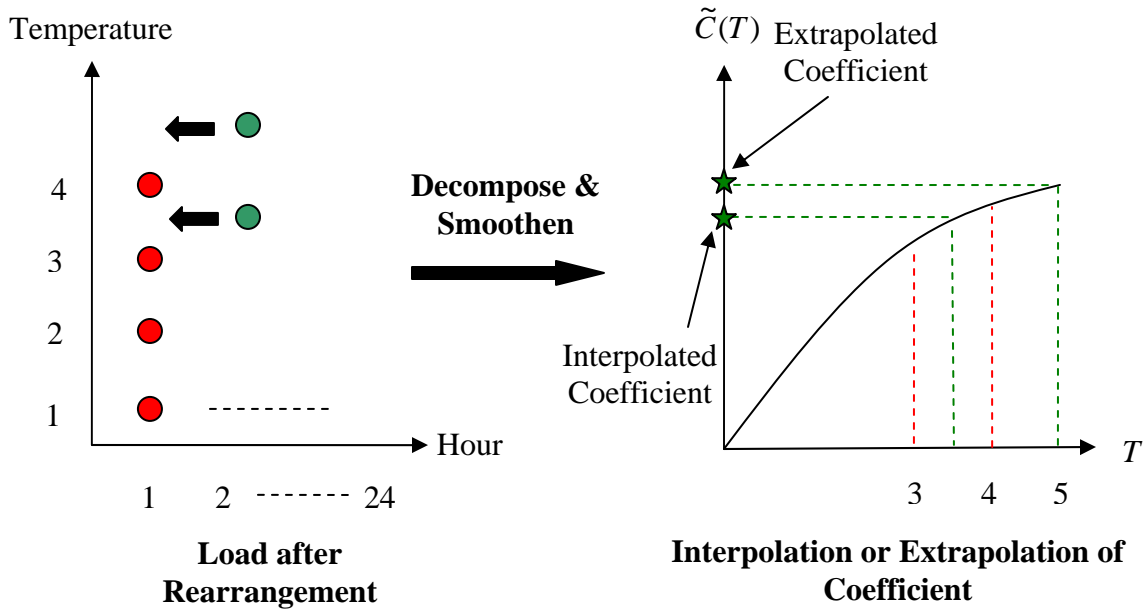


Figure 3. Interpolation and Extrapolation of the Coefficient.

Because the hourly temperatures for a forecasting day are already known, the coefficient vector can be found by interpolation if the forecasted temperature at a given hour does not exceed the temperature bounds experienced in previous (historical) days. Then the load forecasting at this hour is achieved by recombining the basis vector and the interpolated coefficient vector. If the temperature forecast exceeds the temperature bounds, then the coefficient vector can be obtained by extrapolating the coefficient curves. The load forecasting at this hour is achieved by recombining the basis vector and the extrapolated coefficient vector.

### III. Simulation Results

The proposed forecasting approach is tested by using the past load data obtained from New England Independent System Operator (ISO). The hourly temperatures of each day are weighted average values of 8 weather stations in the New England area in degrees Fahrenheit. In the simulation, load data for the year 2002 is chosen for demonstration. The load is classified into two groups: weekdays and weekends load. For each group, the simulation uses data in a window of previous four weeks. For each forecasting day, actual raw load data in the moving window passes through the regression filter, and then the filtered data is decomposed into a basis vector and a coefficient vector of dimensionality  $n$ , where  $n$  is set to two. The results are analyzed by the following formulas:

i) Standard Deviation

$$\sigma = \sqrt{\frac{1}{N} \sum_{d=1}^N [L(d, h) - \hat{L}(d, h)]^2}$$

where  $L(d,h)$  is empirical load data for a given day ( $d$ ) and hour ( $h$ ),  $\hat{L}(d,h)$  is the corresponding load forecast.

ii) Percent Error

$$Error = |L(d,h) - \hat{L}(d,h)| / L(d,h) \times 100$$

Here an arbitrary weekday is chosen as the forecasting day to show the simulation results. Fig. 4 shows the filtered load before the rearrangement and after the rearrangement. After the rearrangement, the load is decomposed into a basis vector and a preliminary coefficient vector. The Semigroup Channel is trained with the preliminary coefficient vector using the proposed successive training algorithm. Figure 5 shows the smoothed coefficient vector plotted in two components. When a forecasted temperature exceeds the temperature bounds of the moving window, extrapolation of the coefficient vector is required, which is illustrated in Fig. 6.

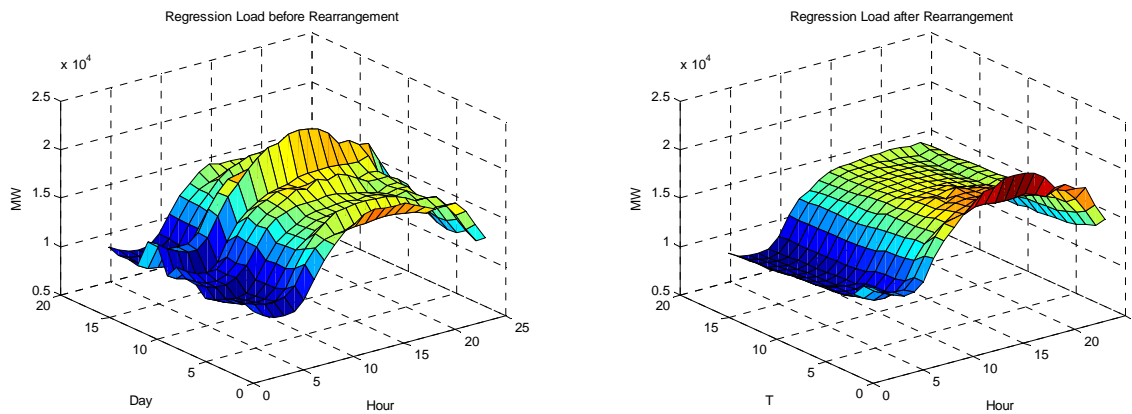


Figure 4. Regression Load before Rearrangement (left) and after Rearrangement (right).

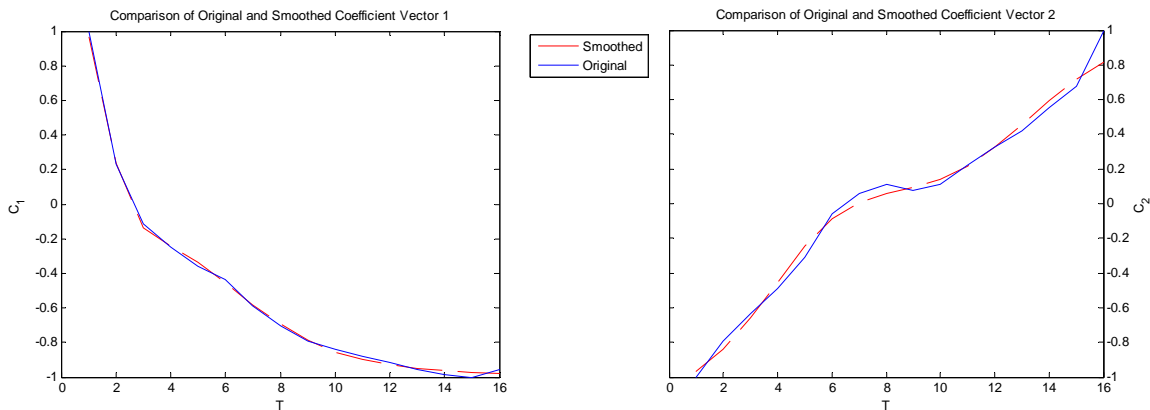


Figure 5. Comparison of Original and Smoothed Coefficient Vector  $C_1$  and  $C_2$ .

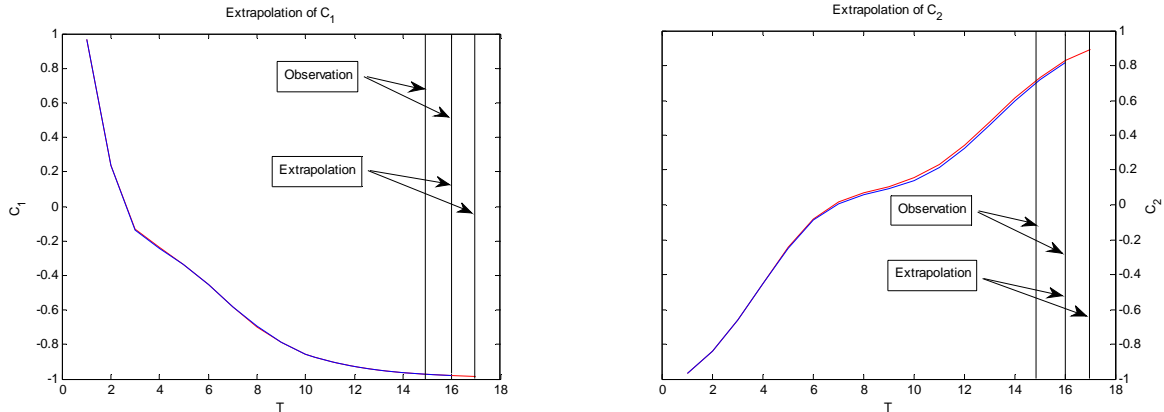


Figure 6. Extrapolation of  $C_1$  and  $C_2$ .

Here three typical weeks are selected to show the performance of the proposed method, winter, spring and summer weeks. The results include percent error and standard deviation at each hour of each day, both being calculated from the regression load and the actual load. In the first week which is a typical winter week, the highest error is 7.29% at 3:00 on Wednesday and the lowest error is 0.01% at 13:00 on Sunday for the regression load. On the other hand, for the actual load in the same week, the highest error is 29.87% at 8:00 on Tuesday and the lowest error is 0.04% at 24:00 on Friday. It should be noticed that Tuesday is the New Year's Day, but we still have treated it as a normal weekday. Therefore the errors for this day are relatively large. In the 11<sup>th</sup> week which is a typical spring week, the highest error is 2.39% at 17:00 on Thursday and the lowest error is nearly 0% at 9am on Sunday for the regression load. With respect to the actual load, the highest error is 7.95% at 7:00 on Sunday and the lowest error is 0.01% at 21:00 on Tuesday as well as at 12:00 on Wednesday. In the 28<sup>th</sup> week, which is a typical summer week, the highest error is 7.06% at 20:00 on Friday and the lowest error is 0.03% at 9:00 on Saturday for the regression load. For the actual load, the highest error is 17.27% at 20:00 on Friday and the lowest error is 0.01% at 10:00 on Tuesday. As shown above, high hourly percent errors with respect to the regression load may appear for some hours, e.g., 7.29% in the first week. This is due to relatively larger hourly temperature changes in the forecasting day compared to the historical temperatures of previous days.

Table 1 shows the results of the regression load averaged for the whole year. Friday shows the best forecasting result and Saturday shows the worst. Based on the actual load, Thursday shows the best result and Saturday shows the worst. Large errors are in late afternoons for weekdays and mornings for weekends.

## IV. Conclusions

In this paper, a new neural network approach is proposed to perform load forecasting. Instead of actual empirical load data, the filtered load which is already preprocessed by regression is used. The filtered load rearranged along the temperature axis provides the smoothness of the coefficient vector to carry out interpolation and extrapolation. The results of the whole year show that the proposed method achieves satisfactory forecasting results. However, the errors resulting from the use of actual load are generally larger than



the short-term load forecasting requirement. This is because the load may also depend on other unknown factors. If the given load and temperature data are highly correlated to each other, it is expected that much better results can be achieved and the regression procedure can be removed.

## Acknowledgement

The work is supported in part by National Science Foundation under Grant #0758385.

## References

1. K. Y. Lee, Y. T. Cha, and J. H. Park, "Short-term load forecasting using an artificial neural network," IEEE Trans. on Power Systems, vol. 7, pp. 124-132, Feb. 1992.
2. J. H. Chow, Applied Mathematics for Restructured Electric Power Systems: Optimization, Control, and Computational Intelligence. New York: Springer-Verlag, 2005, ch. 12.
3. K. Y. Lee, J. P. Velas, and B. H. Kim, "Development of an intelligent monitoring system with high temperature distributed fiberoptic sensor for fossil-fuel power plants," IEEE Power Engineering Society General Meeting, pp. 1350-1355, Jun 6-10, 2004.
4. B. H. Kim, J. P. Velas, and K. Y. Lee, "Development of intelligent monitoring system for fossil-fuel power plants using system-type neural networks and semigroup theory," IEEE Power Engineering Society General Meeting, pp. 2949-2954, 2005.
5. B. H. Kim, J. P. Velas, and K. Y. Lee, "Semigroup based neural network architecture for extrapolation of enthalpy in a power plant," Proceedings of the ISAP, pp. 291-296, 2005.
6. B. H. Kim, J. P. Velas, and K. Y. Lee, "Semigroup based neural network architecture for extrapolation of mass unbalance for rotating machines in power plants," International Federation of Automatic Control (IFAC) Symposium on Power Plants and Power Systems Control, in CD, 2006.
7. B. H. Kim, J. P. Velas, and K. Y. Lee, "Short-term load forecasting using system-type neural network architecture," International Joint Conference on Neural Networks 2006, pp. 2619-2626.
8. N. J. Hobbs, B. H. Kim, and K. Y. Lee, "Long-term load forecasting using system type neural network architecture," International Conference on Intelligent Systems Applications to Power Systems 2007, pp. 1-7, Nov, 2007.
9. R. Padhi and S. N. Balakrishnan, "Proper orthogonal decomposition based feedback optimal control synthesis of distributed parameter systems using neural networks," Proceedings of the 2002 American Control Conference 2002, pp. 4389-4394, Vol. 6, pp. 8-10, May 2002.
10. R. Jacobs and M. Jordan, "A competitive modular connectionist architecture," Advances in Neural Information Processing Systems, Vol. 3, p. 767-773, 1991.
11. I. Moghram and S. Rahman, "Analysis and evaluation of five short-term load forecasting techniques," IEEE Trans. on Power Systems, vol. 4, pp. 1484-1491, Nov. 1989.

### SHU DU

Shu Du received his B.S. degree in Automation from Beijing University of Chemical Technology in 2003. He is currently enrolled in the Master program in Electrical Engineering at Baylor University, Texas. His research interests are in neural networks and power systems.

### KWANG Y. LEE

Dr. Lee received his B.S. degree in Electrical Engineering from Seoul National University, Korea, in 1964, M.S. degree in Electrical Engineering from North Dakota State University, Fargo, in 1968, and Ph.D. degree in System Science from Michigan State University, East Lansing, in 1971. He has been with Michigan State, Oregon State, Univ. of Houston, Penn State, and Baylor University, where he is now a Professor and Chair of Electrical and Computer Engineering. His interests include power system control, operation, planning, and intelligent system applications to power systems. Dr. Lee is a Fellow of IEEE,



former Associate Editor of IEEE Transactions on Neural Networks, and Editor of IEEE Transactions on Energy Conversion. He is also a registered Professional Engineer.

Table 1. Statistics of the Whole Year Average Forecasting Results (Regression Load)

Hour	Mon.		Tue.		Wed.		Thr.		Fri.		Sat.		Sun.	
	Per. Err.	Std. Dev.	Per. Err.	Std. Dev.	Per. Err.	Std. Dev.	Per. Err.	Std. Dev.	Per. Err.	Std. Dev.	Per. Err.	Std. Dev.	Per. Err.	Std. Dev.
1	1.45	238	1.73	378	1.76	454	1.73	418	1.47	299	2.39	425	1.97	366
2	1.27	216	1.72	350	1.81	456	1.68	380	1.47	294	2.55	424	1.83	341
3	1.42	221	1.80	363	1.95	460	1.65	352	1.28	250	2.44	429	1.76	299
4	1.54	226	1.93	369	2.03	464	1.70	339	1.32	261	2.19	408	1.64	278
5	1.57	230	2.13	389	2.06	468	1.90	341	1.39	276	2.17	387	2.05	315
6	1.53	252	1.97	424	2.04	458	1.71	316	1.40	258	1.85	490	1.81	261
7	1.54	305	1.74	400	1.69	390	1.64	379	1.34	259	1.79	624	1.88	288
8	1.41	330	1.22	324	1.40	374	1.47	451	1.17	289	1.87	1137	1.64	302
9	1.21	291	1.39	333	1.27	387	1.49	403	1.10	307	1.93	1105	1.49	316
10	1.18	295	1.36	320	1.32	415	1.30	406	1.10	345	1.69	629	1.21	294
11	1.01	277	1.06	288	1.28	441	1.11	391	1.01	335	1.89	599	1.27	292
12	0.89	213	0.99	294	1.33	464	1.07	346	1.01	326	2.00	562	1.49	342
13	0.97	244	1.17	350	1.48	547	1.11	320	1.08	271	2.10	602	1.74	393
14	1.09	290	1.33	416	1.55	565	1.15	343	1.19	291	2.09	608	1.87	457
15	1.07	282	1.52	408	1.62	582	1.19	407	1.29	338	1.98	609	1.63	367
16	1.19	324	1.54	426	1.73	663	1.73	631	1.10	310	2.19	648	1.79	415
17	1.51	413	1.57	505	1.56	616	1.73	589	1.22	317	2.30	673	1.88	467
18	1.59	439	1.52	499	1.77	562	1.48	622	1.62	439	2.33	871	1.66	411
19	1.41	408	1.75	607	1.79	541	1.29	375	1.61	437	1.92	662	1.47	346
20	1.39	409	1.42	464	1.66	480	1.17	332	1.50	421	1.61	559	1.51	336
21	1.34	387	1.47	462	1.62	492	1.22	379	1.26	391	1.57	612	1.56	371
22	1.42	357	1.55	525	1.60	474	1.51	447	1.36	391	1.85	565	1.74	401
23	1.34	295	1.95	587	1.58	486	1.47	374	1.17	289	2.41	739	1.97	400
24	1.48	284	2.20	583	2.00	544	1.46	333	1.16	238	2.16	554	2.16	403
Avg	1.33	301	1.58	419	1.66	491	1.46	403	1.28	318	2.05	622	1.71	352