REGRESSION BASED PEAK LOAD FORECASTING USING A TRANSFORMATION TECHNIQUE

Takeshi Haida Shoichi Muto
Computer and Communication Research Center
Tokyo Electric Power Company
4-10, Irifune, 1-Chome, Chuo-ku, Tokyo 104, JAPAN

Abstract- This paper presents a regression based daily peak load forecasting method with a transformation technique. In order to forecast the load precisely through a year, we should consider seasonal load change, annual load growth and the latest daily load change. To deal with these characteristics in the load forecasting, a transformation technique is presented. This technique consists of a transformation function with translation and reflection methods. The transformation function is estimated with the previous year's data points, in order that the function converts the data points into a set of new data points with preserving the shape of temperature-load relationships in the previous year. Then, the function is slightly translated so that the transformed data points will fit the shape of temperature-load relationships in the year. Finally, multivariate regression analysis with the latest daily loads and weather observations estimates the forecasting model. Large forecasting errors caused by the weather-load nonlinear characteristic in the transitional seasons such as spring and fall are reduced. Performance of the technique which is verified with simulations on actual load data of Tokyo Electric Power Company is also described.

Keywords- Load Forecasting, Multivariate Regression

1 INTRODUCTION

Load forecasting is one of the important tasks in a utility. Accuracy of the load forecasting has influence on security and efficiency in electric power supply. The short term load forecasting in Tokyo Electric Power Company (TEPCO) consists of the following two steps. First, the load forecasting operator at the central dispatching office predicts peak load. Next, hourly load curve in a day is calculated on the basis of the forecasted peak load. According to the result, the operating plan for generating units is scheduled taking into account total fuel costs and power system reliability. This paper concentrates on the peak load forecasting.

94 WM 206-3 PWRS A paper recommended and approved by the IEEE Power System Engineering Committee of the IEEE Power Engineering Society for presentation at the IEEE/PES 1994 Winter Meeting, New York, New York, January 30 - February 3, 1994. Manuscript submitted July 30, 1993; made available for printing January 19, 1994.

Various kinds of short-term load forecasting techniques have been proposed for last two decades [1]~[11].

Time series approach is the most widely discussed methodology[3][4][12]. In a simple application, hourly load values up to a few days ahead are modeled as an autoregressive process of past hourly load values. Box-Jenkins method takes account of both time series load data and other time independent variables such as hourly temperatures of the day. From the view point of the peak load forecasting, we need to consider time series factors such as a weekly base load pattern and the latest peak load trend.

Another popular approach is multivariate regression [5]~[8]. The load is represented as a linear combination of explanatory variables. The explanatory variables include weather factors [5]. Coefficients of explanatory variables are estimated by least squares fitting or modern regression techniques.

Neural network is a hot topic for both time series approach and the multivariate regression [9]. One significant characteristic of neural networks is to perform nonlinear modeling between input and output data. Since the modeling is not explicit and there are many parameters in the network, we must consider whether the neural network model is proper as a load model.

We use the multivariate regression as the basic forecasting method. The peak load heavily depends on temperature in daytime and is influenced by the other weather factors such as humidity. These characteristics are easily formulated by the multivariate regression. Periodical load variations can be modeled by shifting load values according to a day of the week. Weather factors for a few days before are also considered. However, a simple regression model is not adequate for precise load forecasting through a year. We should consider seasonal load change, annual load growth and the latest daily load change.

A regression based daily peak load forecasting method with a transformation technique is proposed. This technique utilizes a transformation function with translation and reflection methods. Performance of the technique which is verified with simulations on actual load data of Tokyo Electric Power Company is also described.

The paper is organized as follows. The simple regression model is reviewed in the next section. Then, the transformation technique for the regression approach is proposed in section 3. Performance evaluation using actual weather-load data and the experimental results are discussed in section 4. Finally, the work is concluded in section 5.

2 REGRESSION APPROACH

In the regression approach, relationships among loads, weather and nonweather parameters are formulated as linear or piecewise linear equations [2]. A formation of the regression model is shown as:

$$P = \beta_0 + \sum_{j=1}^{m} \beta_j X_j \qquad (1)$$

where, P is a peak load. X_j is an explanatory variable correlated with P. β_0 and β_j are regression coefficients.

The explanatory variables of the regression model are selected based on correlation analyses and experiences in utility operators. The coefficients in the forecasting model are estimated with the latest actual observation data before the forecast day. The forecast load is calculated with weather forecasts for the forecasting day and coefficients.

Since characteristics of the load is varying, later observations are more significant for the forecasting. The coefficients are estimated by the least squares method with historical exponential weight. The coefficients are updated every day.

The forecasting method is assumed to have a linear relationship between explanatory variables and peak loads in a short period. According to our research and the load forecasting experiences in TEPCO, the assumption is appropriate in ordinary summer and winter seasons. However, it is not adequate for spring and fall seasons because there is an obvious nonlinear relationship between temperatures and the peak loads as shown in Fig.1. A simple linear regression model often causes large forecasting errors in the transitional seasons as shown in Fig.2.

A polynominal regression model may represent the nonlinear relationship for the observations. However, there are some difficulties to use it.

Consider we use this year's data only. In transitional seasons, we do not have much data for a coming season. It is very difficult to estimate a proper model even if we use

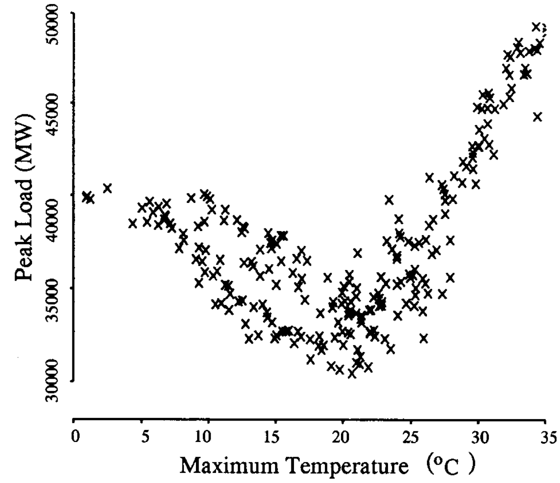

Figure 1: Maximum Temperature and Peak Load Relationship in 1990

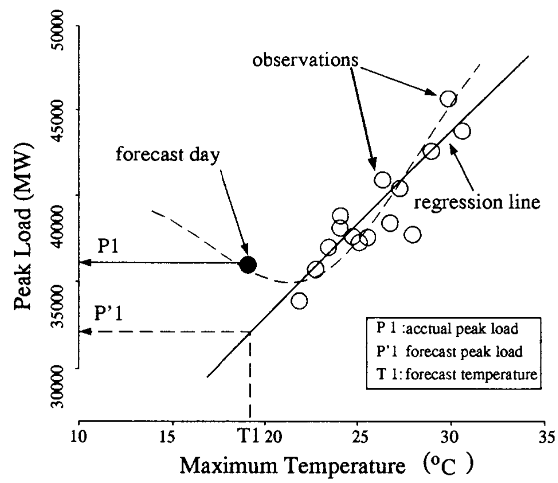

Figure 2: Large Forecasting Error caused by Seasonal Change

the polynominal model. Next, consider we use seasonal or whole year data in past years. In this case, we can make a nonlinear weather-load model. However, direct use of the model[8] is not sufficient for the daily peak load forecasting since the weather-load model will change year by year.

Thus, it is necessary to develop a more appropriate technique for the peak load forecasting.

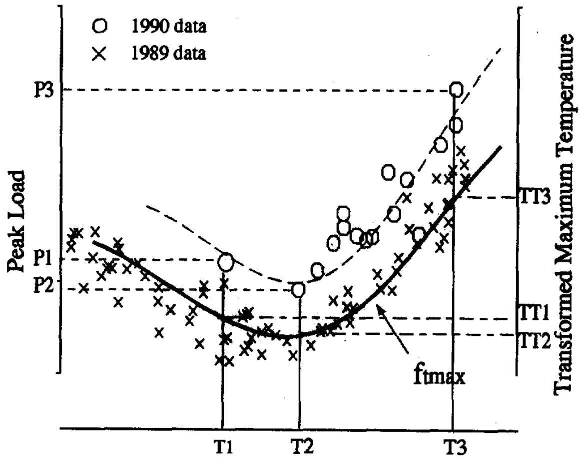

Figure 3: Maximum Temperature and Peak Load Relationship in 1989 and 1990

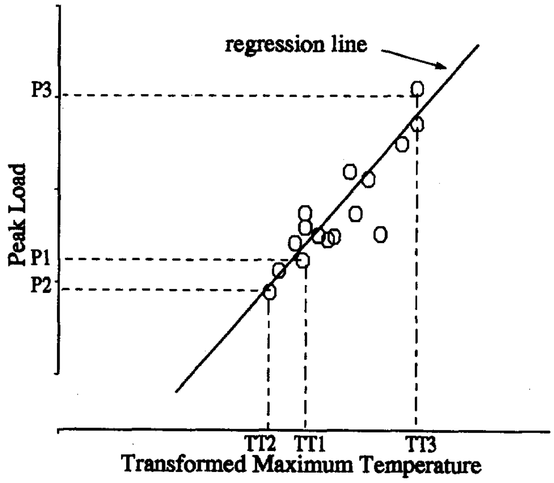

Figure 4: Transformed Maximum Temperatures and Peak Load Relationship

3 A TRANSFORMATION TECHNIQUE

Basic Idea

In order to reduce errors of the load forecasting in transitional seasons, a transformation technique is proposed. The technique makes the forecasting model which includes both the annual weather-load relationship and the latest weather-load characteristic. The technique consists of the following parts.

1. At first, the whole year or seasonal weather-peak load model in the previous year is defined as:

$$P_{prev} = C + \sum_{j=1}^{m} f_j(X_j) \tag{2}$$

   where P_{prev} is a daily peak load for the previous year. X_j is a weather parameter such as maximum temperature. f_j represents the functional relationship of the weather parameter to the peak load in the previous year. C is constant. All of the functions are expressed by polynominal equations. The equation(2) is estimated by least square fitting on previous year data. The transformation function f_j is expressed in the equation.

2. Next, using the function f_j, the model for the forecasting day in the current year is represented as:

$$P = \alpha_0 + \sum_{j=1}^{m} \alpha_j f_j(X_j) \tag{3}$$

   where P is the peak load for the day. α_0 and α_j are coefficients of the forecasting model. The observation data is converted by using function f_j. The coefficients are estimated by using the weighted least squares method with the latest several days' data before forecasting day.

3. The weather forecasts data are also converted by the functions f_j. The forecasting load is calculated with the converted data. The coefficients of the forecasting model are updated every day.

The equation(3) is meaningful for the peak load forecasting. The equation reflects both the latest load characteristic and the annual weather-load shape. Moreover, the coefficients represent two kinds of annual load growth, namely, base load growth and weather sensitive load growth. An example of the transformation technique is shown in Fig.3 and Fig.4. In Fig.3, the polynominal regression curve of the function f_{tmax} is the relationship between maximum temperatures and peak loads in 1989. Then, the observed data points and weather forecast data of a certain forecast day in 1990 are transformed to be linear through the function f_{tmax} as shown in Fig.4. For example, data points, (T1, P1), (T2, P2), (T3, P3), in Fig.3, are transformed to new data points, (TT1, P1), (TT2, P2), (TT3, P3), in Fig.4 respectively. In ideal, different season data are on one transformed linear line. As a result, in transitional seasons, the load forecasting can be almost done even if we do not have much data in a coming season. However, some problems remain in this technique. Problems and methods to meet with them are described in the next section.

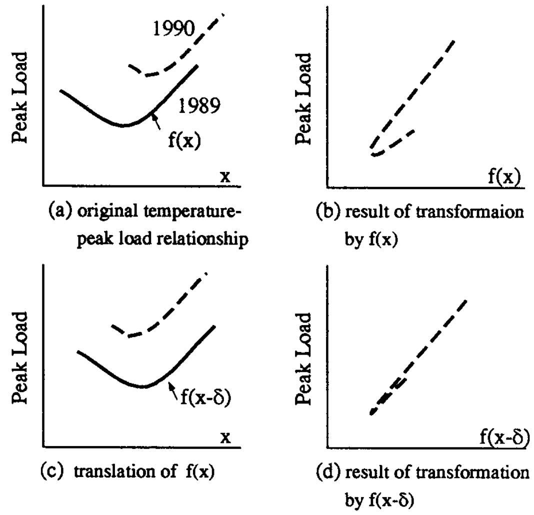

Figure 5: Effect of Translation

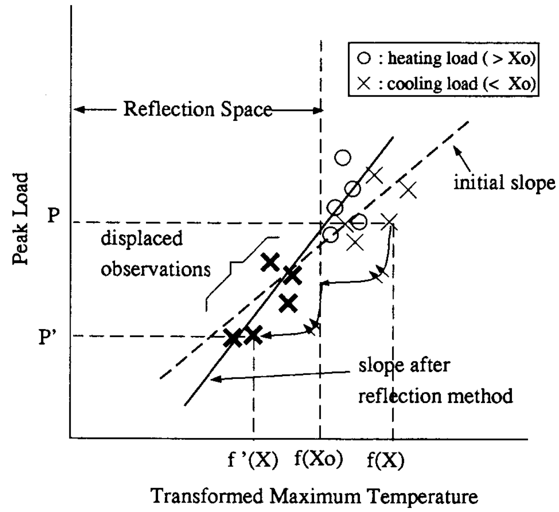

Figure 6: Effect of Reflection

Translation and Reflection Methods

The proposed technique assumes the functional similarity between a weather-load relationship of a previous year and that of a current year. However, according to the statistic analysis of the weather-load relationships using actual data, the assumption is inadequate at times. Threshold temperature of heating loads and cooling loads occasionally varies year by year. Fig.5(a) shows the case that the threshold temperature in the forecasting year is different from that in the previous year. In this case, the observed data for the forecasting day is not transformed to be on one line as shown in Fig.5(b).

The revised model is expressed by introducing the translation parameters:

$$P = \alpha_0 + \sum_{j=1}^{m} \alpha_j f_j(X_j - \delta_j) \tag{4}$$

δ_j is a translation parameter. The parameters are determined so that equation(4) gives best fitting to the data points in the current year. The translation parameters are determined as follows.

1. Several canditates for δ_j are set in advance.
2. Then, residual sum of squares of equation(4) is calculated. The model with the smallest residual sum of squares is selected as the forecasting equation.

This method adjusts the transformation function to keep to be linear in the regression plane. Though initial value of this parameter δ_j is assumed to be 0, the value is updated day by day as the number of the observation data in this year increase. This method is illustrated in Fig.5(c) and (d).

As it is possible to treat heating loads and cooling loads on a same regression plane by using the transformation function, the distribution of the observations may concentrate after transformation. Much concentration of the data may estimate an unexpected regression model. The reflection method is used in order to rearrange the concentrated observations. The procedure of the reflection method are as follows.

1. First, define the threshold temperature X'_{0j} which gives minimum point of the transformation function $f_j(X'_j)$ around middle temperatures between summer and winter. X'_{0j} is calculated by using the derived function of the $f_j(X'_j)$. The X'_j may be not the same value as the original observed data X_j if a translation parameter δ_j is determined to a certain value.
2. Second, if X'_j is lower than the X'_{0j}, the peak load P and the transformed temperature $f_j(X'_j)$ are converted into P' and $f'_j(X'_j)$ respectively according to the equation(5)(6).

$$P' = P - 2\alpha_j \left(f_j(X'_j) - f_j(X'_{0j})\right) \tag{5}$$

$$f'_j(X'_j) = 2f_j(X'_{0j}) - f_j(X'_j) \tag{6}$$

In the equations, coefficient α_j shows a slope of the transformed regression plane given by the equation(4).

Fig.6 shows an example of the reflection method. The observation is displaced into the reflection space according to the slope of the initial regression plane. The concentrated observations are displaced into the reflection space by the method. Then, the forecasting model is estimated again based on the new observations. This new regression model is more preferable for load forecasting than the initial regression model from our experiments. This method can contribute to make a stable regression model.

4 EVALUATION

Experiment Conditions

Weather and load data

The performance of the proposed technique has been evaluated using the actual weather and peak loads of weekdays, except holidays, of TEPCO in 1989 through 1992. The weather values are daily maximum temperature, minimum temperature and humidity which are available from weather forecast service companies. Each weather value is a weighted mean of weather values at ten major cities in TEPCO supply area.

Formulation of basic forecasting model

The basic forecasting equation is as follows:

$$P = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 \tag{7}$$

where P is a daily peak load. α_0, α_1, α_2 and α_3 are coefficients of the regression model. X_1, X_2 and X_3 are weather parameters.

- dependent variable:
  P = peak load
- explanatory variables:
  $X_1 =$ maximum temperature of the day
  $X_2 =$ average temperature for the latest two days
  $X_3 =$ relative humidity of the day

The explanatory variable X_2 calculated by maximum and minimum temperatures for the latest two days expresses a historical weather effect on the current peak load. This variable is derived from the data analyses and the experiences of the load forecast operators. In addition to the past weather effect, load offset which depends upon a day of the week is considered in load forecasting model. However, we omit its description in this paper because it is not concerned with the presented technique directly. The number of observations to estimate the coefficients is 20. We use 0.8 as historical weight.

The basic forecasting model shown in the equation(7) is used through a year. However, in the transitional seasons such as spring and fall, the transformation technique is used. The seasons are defined by Table 1.

Table 1: Definition of Seasons

	period
summer	June 16 ~ August 14
winter	November 16 ~ March 14
spring	March 15 ~ June 15
fall	August 15 ~ November 15

Transformation functions

The transformation functions for explanatory variables X_1, X_2, X_3 in the equation(7) are expressed as the following equations:

$$f_1(X_1) = \sum_{j=1}^{4} \gamma_j X_1^j \tag{8}$$

$$f_2(X_2) = \sum_{j=1}^{4} \epsilon_j X_2^j \tag{9}$$

$$f_3(X_3) = X_3 \tag{10}$$

where γ_j and ϵ_j are coefficients of the equations. j means order of degree for each variable. Since the relationships between the temperatures and the peak loads in winter through summer and summer through winter is slightly different, the coefficients are calculated for these two periods in the previous year. A number of order of the transformation function, f_1 and f_2, is determined to 4 based on experimental analysis. On the other hand, since a function f_3 representing a load-humidity relationship does not show an obvious nonlinear characteristic, a number of order of the transformation function, f_3, is determined to 1, and a coefficient of the previous year is not estimated.

The peak load model using the translation method is as follows:

$$\begin{aligned} P = \alpha_0 & + \alpha_1 f_1(X_1 - \delta_1) + \alpha_2 f_2(X_2 - \delta_2) \\ & + \alpha_3 f_3(X_3) \end{aligned} \tag{11}$$

As the translation method is applied to functions representing nonlinear curves, the method is applied to only f_1 and f_2 functions. The step-size and range of both δ_1 and δ_2 are 0.5 degrees and ± 2.0 degrees respectively. The reflection method is applied to some observation data which is lower than the threshold temperatures concerning with maximum temperature X_1 and average temperature X_2. Of course, since f_3 is a linear function, the reflection method is not applied to humidity data.

After using the reflection method, the load forecasting model is estimated according to the formation of equation(11) with the latest 20 observation data.

Table 2: Forecasting Results in Spring 1989~1992

	model	average absolute errors(%)	standard deviation of errors(%)
1989	model A	1.828	2.256
	model B	1.784	2.579
	model C	1.732	2.452
1990	model A	2.095	2.578
	model B	1.838	2.204
	model C	1.876	2.349
1991	model A	1.980	2.581
	model B	1.612	2.135
	model C	1.536	2.025
1992	model A	1.813	2.289
	model B	1.579	2.155
	model C	1.537	1.994
Through the four years	model A	1.931	2.426
	model B	1.704	2.272
	model C	1.677	2.210

Table 3: Forecasting Results in Fall 1989~1992

	model	average absolute errors(%)	standard deviation of errors(%)
1989	model A	1.993	3.719
	model B	1.532	2.114
	model C	1.297	1.847
1990	model A	2.249	2.988
	model B	1.510	2.130
	model C	1.457	1.941
1991	model A	2.004	2.445
	model B	1.515	2.139
	model C	1.465	1.889
1992	model A	2.393	3.546
	model B	1.654	2.116
	model C	1.505	1.755
Through the four years	model A	2.159	3.204
	model B	1.552	2.125
	model C	1.430	1.861

Results

Table 2 and 3 present forecasting results in the spring and the fall seasons from 1989 to 1992.

The tables show the average absolute errors and standard deviations of errors on actual weather data in the seasons. In the table, in order to evaluate the proposed method, the results from the three different kinds of forecasting models are described. The formation of the models are as follows:

- model A using the linear regression model represented.
- model B using the transformation function without translation and reflection methods.
- model C using the transformation function with translation and reflection methods.

It is obvious that the transformation technique is effective in the transitional seasons as shown in the tables. In the spring season, the total forecasting average absolute errors of model A, B, C, are 1.931 %, 1.704 % and 1.677% respectively. The errors in the fall season are 2.159 %, 1.552% and 1.430% respectively. Standard deviations of the forecasting errors are also reduced by using the transformation technique.

Although the effect of the translation and reflection methods is rather small in comparison with the effect only by the transformation function, these methods also improve the forecasting performance of the regression based model.

The whole year forecast accuracy using the model C from 1989 to 1992 are 1.50 % in average absolute errors, 1.96 % in the standard deviation of the errors, 140 days out of 925 days in which absolute errors exceed 1000 MW, maximum peak load record in TEPCO is 54100MW. As mentioned before, the model C using the transformation technique is used for load forecasting only in spring and fall seasons. In summer and winter seasons, the transformation technique is not used. On the same conditions, accuracy of the model A is 1.76 % in average absolute errors, 2.43% in the standard deviation, and 184 days in which errors exceed 1000 MW. Both average performance and large forecasting errors are improved by the proposed technique.

5 CONCLUSION

In this paper, we have presented the regression based daily peak load forecasting method with the transformation technique. It is necessary to use the latest data to make forecasting model. On the other hand, there is an obvious seasonal load change characterized as a nonlinear relationship between temperatures and loads. The transformation technique is presented to deal with these characteristics in the load forecasting. This technique employs a transformation function with translation and reflection methods. The performance evaluation using actual weather-load data in TEPCO shows good results.

The basic transformation technique presented in the paper is implemented and checked in the Peak Load Forecasting Assistance System at the central dispatching office in TEPCO. In the next stage, we are planning to implement the translation and the reflection methods into the assistance system. It is expected that the both methods will improve the peak load forecasting accuracy further.

The operators uses the system every morning and evening. They accept the system because the method uses regression which matches their intuitions.

6 ACKNOWLEDGMENT

The authors would like to express gratitude to Mr. Nakamura and Mr. Kawata at the central load dispatching office in TEPCO for their support and useful comments of this research. The authors also acknowledge the valuable discussion and comments of Mr. Ohashi of our Computer and Communication Research Center, for the completion of the paper.

References

[1] IEEE Committee Report, "Load Forecasting Bibliography Phase1", IEEE Trans. Power App. Sys., vol.PAS-99, 1980, pp.53-58.

[2] I. Moghram, S. Rahman, "Analysis and Evaluation of Five Short-Term Load Forecasting Techniques", IEEE Trans. Power Sys., vol.4, no.4, Oct. 1989, pp.1484-1491.

[3] M. T. Hagan, S. M. Behr, "The Time Series Approach to Short Term Load Forecasting", IEEE Trans. Power Sys., vol.PWRS-2, no.3, Aug. 1987, pp.785-791.

[4] W. M. Grady, L. A. Groce, T. M. Huebner, Q. C. Lu, M. M. Crawford, "Enhancement, Implementation, and Performance of an Adaptive Short-Term Load Forecasting Algorithms", IEEE Trans. Power Sys., vol.6, no.4, Nov. 1991, pp.1404-1410.

[5] G.T.Heinemann, D.A.Nordman and E.C.Plant, "The Relationship Between Summer Weather and Summer Loads - A Regression Analysis", IEEE Trans. Power App. Sys., vol.PAS-85, no.12, Nov. 1966, pp.1144-1154.

[6] P. C. Gupta, K. Yamada, "Adaptive Short-Term Forecasting of Hourly Loads Using Weather Information", IEEE Trans. Power App. Sys., vol.PAS-91, no.12, Nov. 1972, pp.2085-2094.

[7] B. Krogh, E. S. de Linas, ED. Lesser, "Design and Implementation of an On-Line Load Forecasting Algorithm", IEEE Trans. Power App. Sys., vol.PAS-101, no.9, Sep. 1982, pp.3285-3289.

[8] A. D. Papalexopoulos, T. C. Hesterberg, "A Regression-Based Approach to Short-Term System Load Forecasting", IEEE Trans. Power Sys., vol.5, no.4, Nov. 1990, pp.1535-1547.

[9] D. C. Park, M. A. El-Sharkawi, R. J. Marks, L. E. Atlas, M. J. Damborg, "Electric Load Forecasting Using An Artificial Neural Network", IEEE Trans. Power Sys., vol.6, no.2, May 1991, pp.442-449.

[10] S. Rahman, "Formulation and Analysis of a Rule-Based Short-Term Load Forecasting Algorithm", Proc. IEEE, vol.78, no.5, May 1990, pp.805-815.

[11] K. Jabbour, J. F. V. Riveros, D. Landsbergen, Walter Meyer, "ALFA: Automated Load Forecasting Assistant", IEEE Trans. Power Sys., vol.3, no.3, Aug. 1988, pp.908-914.

[12] Box, G. E. P. and G. M. Jenkins, "Time Series Analysis: Forecasting and Control", Holden-Day, 1976.

Takeshi haida was born in Ishikawa, Japan, on December 8, 1960. He received B.S. and M.S. degrees in electrical engineering from University of Electro-Communications, Tokyo, Japan in 1984 and 1986, respectively. Since 1989, he has been with AI Technology Department of Computer and Communication Research Center, Tokyo Electric Power Company. He is a member of the Institute of Electrical Engineers of Japan and the Society of Instrument and Control Engineers of Japan.

Shoichi Muto was born in Tokyo, Japan, on Jun 20, 1956. He received B.S. and M.S. degrees in electrical engineering from Tokyo Institute of Technology, Tokyo, Japan in 1979 and 1981, respectively. In 1992, He received Ph.D. degree in electrical engineering from the University of Tokyo, Tokyo, Japan. He joined the Tokyo Electric Power Company in 1981, serving substation department for two years. Since 1983, he is engaged in research on system technologies for power systems. From 1990 to 1991, he visited Stanford University, California, as a visiting researcher. He is a member of the Institute of Electrical Engineers of Japan and the Society of Instrument and Control Engineers of Japan.