

Electric Load Forecasting Using An Artificial Neural Network

D.C. Park, M.A. El-Sharkawi, R.J. Marks II,
L.E. Atlas and M.J. Damborg

Department of Electrical Engineering, FT-10
University of Washington
Seattle, WA 98195

Abstract

This paper presents an artificial neural network(ANN) approach to electric load forecasting. The ANN is used to learn the relationship among past, current and future temperatures and loads. In order to provide the forecasted load, the ANN interpolates among the load and temperature data in a training data set. The average absolute errors of the one-hour and 24-hour ahead forecasts in our test on actual utility data are shown to be 1.40% and 2.06%, respectively. This compares with an average error of 4.22% for 24-hour ahead forecasts with a currently used forecasting technique applied to the same data.

Keywords - Load Forecasting, Artificial Neural Network

1 Introduction

Various techniques for power system load forecasting have been proposed in the last few decades. Load forecasting with lead-times, from a few minutes to several days, helps the system operator to efficiently schedule spinning reserve allocation. In addition, load forecasting can provide information which is able to be used for possible energy interchange with other utilities. In addition to these economical reasons, load forecasting is also useful for system security. If applied to the system security assessment problem, it can provide valuable information to detect many vulnerable situations in advance.

Traditional computationally economic approaches, such as regression and interpolation, may not give sufficiently accurate results. Conversely, complex algorithmic methods with heavy computational burden can converge slowly and may diverge in certain cases.

A number of algorithms have been suggested for the

load forecasting problem. Previous approaches can be generally classified into two categories in accordance with techniques they employ. One approach treats the load pattern as a time series signal and predicts the future load by using various time series analysis techniques [1-7]. The second approach recognizes that the load pattern is heavily dependent on weather variables, and finds a functional relationship between the weather variables and the system load. The future load is then predicted by inserting the predicted weather information into the predetermined functional relationship [8-11].

General problems with the time series approach include the inaccuracy of prediction and numerical instability. One of the reasons this method often gives inaccurate results is that it does not utilize weather information. There is a strong correlation between the behavior of power consumption and weather variables such as temperature, humidity, wind speed, and cloud cover. This is especially true in residential areas. The time series approach mostly utilizes computationally cumbersome matrix-oriented adaptive algorithms which, in certain cases, may be unstable.

Most regression approaches try to find functional relationships between weather variables and current load demands. The conventional regression approaches use linear or piecewise-linear representations for the forecasting functions. By a linear combination of these representations, the regression approach finds the functional relationships between selected weather variables and load demand. Conventional techniques assume, without justification, a linear relationship. The functional relationship between load and weather variables, however, is not stationary, but depends on spatio-temporal elements. Conventional regression approach does not have the versatility to address this temporal variation. It, rather, will produce an averaged result. Therefore, an adaptable technique is needed.

In this paper, we present an algorithm which combines both time series and regression approaches. Our algorithm utilizes a layered perceptron *artificial neural network* (ANN). As is the case with time series approach, the ANN traces previous load patterns and predicts (*i.e.* extrapolates) a load pattern using recent load data. Our algorithm uses weather information for modeling. The ANN is able to perform non-linear modeling and adaptation. It does not require assumption of any functional relationship between load and weather variables in advance. We can adapt the ANN by exposing it to new data. The ANN is also currently being investigated as a tool in other power system problems such as security assessment, harmonic load identification, alarm processing, fault diagnosis, and topological observability [12-18].

90 SM 377-2 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 1990 Summer Meeting, Minneapolis, Minnesota, July 15-19, 1990. Manuscript submitted August 31, 1989; made available for printing April 24, 1990.

In the next section, we briefly review various load forecasting algorithms. These include both the time series and regression approach. The generalized Delta rule used to train the ANN is shown in Section 3. In Section 4, we define the load forecasting problems, show the topologies of the ANN used in our simulations, and analyze the performance in terms of errors (the differences between actual and forecasted loads). A discussion of our results and conclusions are presented in Section 5.

2 Previous Approaches

2.1 Time Series

The idea of the time series approach is based on the understanding that a load pattern is nothing more than a time series signal with known seasonal, weekly, and daily periodicities. These periodicities give a rough prediction of the load at the given season, day of the week, and time of the day. The difference between the prediction and the actual load can be considered as a stochastic process. By the analysis of this random signal, we may get more accurate prediction. The techniques used for the analysis of this random signal include the Kalman filtering [1], the Box-Jenkins method [3,4], the auto-regressive moving average (ARMA) model [2], and spectral expansion technique [5].

The Kalman filter approach requires estimation of a covariance matrix. The possible high nonstationarity of the load pattern, however, typically may not allow an accurate estimate to be made [6,7].

The Box-Jenkins method requires the autocorrelation function for identifying proper ARMA models. This can be accomplished by using pattern recognition techniques. A major obstacle here is its slow performance [2].

The ARMA model is used to describe the stochastic behavior of hourly load pattern on a power system. The ARMA model assumes the load at the hour can be estimated by a linear combination of the previous few hours. Generally, the larger the data set, the better is the result in terms of accuracy. A longer computational time for the parameter identification, however, is required.

The spectral expansion technique utilizes the Fourier Series. Since load pattern can be approximately considered as a periodic signal, load pattern can be decomposed into a number of sinusoids with different frequencies. Each sinusoid with a specific frequency represents an orthogonal base [19]. A linear combination of these orthogonal basis with proper coefficients can represent a perfectly periodic load pattern if the orthogonal basis span the whole signal space. However, load patterns are not perfectly periodic. This technique usually employs only a small fraction of possible orthogonal basis set, and therefore is limited to slowly varying signals. Abrupt changes of weather cause fast variations of load pattern which result in high frequency components in frequency domain. Therefore, the spectral expansion technique can not provide any accurate forecasting for the case of fast weather change unless sufficiently large number of base elements are used.

Generally, techniques in time series approaches work well unless there is an abrupt change in the environmental or sociological variables which are believed to affect load pattern. If there is any change in those variables, the time series technique is no longer useful. On the other hand, these techniques use a large number of complex

relationships, require a long computational time [20] and result in a possible numerical instabilities.

2.2 Regression

The general procedure for the regression approach is: 1) select the proper and/or available weather variables, 2) assume basic functional elements, and 3) find proper coefficients for the linear combination of the assumed basic functional elements.

Since temperature is the most important information of all weather variables, it is used most commonly in the regression approach (possibly nonlinear). However, if we use additional variables such as humidity, wind velocity, and cloud cover, better results should be obtained.

Most regression approaches have simply linear or piecewise linear functions as the basic functional elements [8-11, 21-23]. A widely used functional relationship between load, L , and temperature, T , is

$$L = \sum_{i=1}^N a_i T \{U(T - T_{i1}) - U(T - T_{i2})\} + C \quad (1)$$

where

$$U(T) = \begin{cases} 1, & \text{if } T \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

and a_i , T_{i1} , T_{i2} , and C are constant, and $T_{i1} > T_{i2}$ for all i .

The variables (L , a_i , T , T_{i1} , T_{i2} , and C) are temporally varying. The time-dependency, however, is not explicitly noted for reasons of notational compactness.

After the basic functional forms of each subclass of temperature range are decided, the proper coefficients of the functional forms are found in order to make a representative linear combination of the basic functions.

Approaches other than regression have been proposed for finding functional coefficients:

1. Jabbour *et al.*[11] used a pattern recognition technique to find the nearest neighbor for best 8 hourly matches for a given weather pattern. The corresponding linear regression coefficients were used.
2. An application of the Generalized Linear Square Algorithm (GLSA) was proposed by Irisarri *et al.*[23]. The GLSA, however, is often faced with numerical instabilities when applied to a large data base.
3. Rahman *et al.*[10] have applied an expert system approach. The expert system takes the advantages of the expert knowledge of the operator. It makes many subdivisions of temperature range and forms different functional relationships according to the hour of interest. It shows fairly accurate forecasting. As pointed out in the discussion of [10] by Tsoi, it is not easy to extract a knowledge base from an expert and can be rather difficult for the expert to articulate their experience and knowledge.
4. Lu *et al.*[24] utilize the *modified Gram-Schmidt orthogonalization process* (MGSOP) to find an orthogonal basis set which spans the output signal space formed by load information. The MGSOP requires a predetermined cardinality of the orthogonal basis set

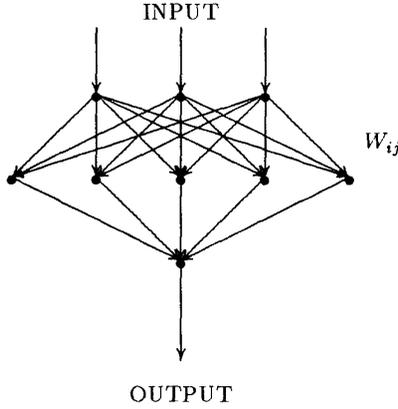


Figure 1: Structure of a Three-Layered Perceptron Type ANN

and the threshold value of error used in adaptation procedure. If the cardinality of the basis set is too small or the threshold is not small enough, the accuracy of the approach suffers severely. On the other hand, if the threshold is too small, numerical instability can result. The MGSOP also has an ambiguity problem in the sequence of input vectors. Different exposition of input vectors result in different sets of orthogonal basis and different forecasting outputs.

3 A Layered ANN

3.1 Architecture

An ANN can be defined as a highly connected array of elementary processors called *neurons*. A widely used model called the multi-layered perceptron (MLP) ANN is shown in Figure 1. The MLP type ANN consists of one input layer, one or more hidden layers and one output layer. Each layer employs several neurons and each neuron in a layer is connected to the neurons in the adjacent layer with different weights. Signals flow into the input layer, pass through the hidden layers, and arrive at the output layer. With the exception of the input layer, each neuron receives signals from the neurons of the previous layer linearly weighted by the interconnect values between neurons. The neuron then produces its output signal by passing the summed signal through a sigmoid function [12-18].

A total of Q sets of training data are assumed to be available. Inputs of $\{\vec{i}_1, \vec{i}_2, \dots, \vec{i}_Q\}$ are imposed on the top layer. The ANN is trained to respond to the corresponding target vectors, $\{\vec{t}_1, \vec{t}_2, \dots, \vec{t}_Q\}$, on the bottom layer. The training continues until a certain stop-criterion is satisfied. Typically, training is halted when the average error between the desired and actual outputs of the neural network over the Q training data sets is less than a predetermined threshold. The training time required is dictated by various elements including the complexity of the problem, the number of data, the structure of network, and the training parameters used.

3.2 ANN Training

In this paper, the *generalized Delta rule* (GDR) [25,26] is used to train a layered perceptron-type ANN. An output vector is produced by presenting an input pattern to the network. According to the difference between the produced and target outputs, the network's weights $\{W_{ij}\}$ are adjusted to reduce the output error. The error at the output layer propagates backward to the hidden layer, until it reaches the input layer. Because of backward propagation of error, the GDR is also called *error back propagation algorithm*.

The output from neuron i , O_i , is connected to the input of neuron j through the interconnection weight W_{ij} . Unless neuron k is one of the input neurons, the state of the neuron k is:

$$O_k = f(\sum_i W_{ik} O_i) \quad (3)$$

where $f(x) = 1/(1 + e^{-x})$, and the sum is over all neurons in the adjacent layer. Let the target state of the output neuron be t . Thus, the error at the output neuron can be defined as

$$E = \frac{1}{2}(t_k - O_k)^2 \quad (4)$$

where neuron k is the output neuron.

The gradient descent algorithm adapts the weights according to the gradient error, *i.e.*,

$$\Delta W_{ij} \propto -\frac{\partial E}{\partial W_{ij}} = -\frac{\partial E}{\partial O_j} \frac{\partial O_j}{\partial W_{ij}} \quad (5)$$

Specifically, we define the error signal as

$$\delta_j = -\frac{\partial E}{\partial O_j} \quad (6)$$

With some manipulation, we can get the following GDR:

$$\Delta W_{ij} = \epsilon \delta_j O_i \quad (7)$$

where ϵ is an adaptation gain. δ_j is computed based on whether or not neuron j is in the output layer. If neuron j is one of the output neurons,

$$\delta_j = (t - O_j) O_j (1 - O_j) \quad (8)$$

If neuron j is not in the output layer,

$$\delta_j = O_j (1 - O_j) \sum_k \delta_k W_{jk} \quad (9)$$

In order to improve the convergence characteristics, we can introduce a momentum term with momentum gain α to Equation 7.

$$\Delta W_{ij}(n+1) = \epsilon \delta_j O_i + \alpha \Delta W_{ij}(n) \quad (10)$$

where n represents the iteration index.

Once the neural network is trained, it produces very fast output for a given input data. It only requires a few multiplications, additions, and calculations of sigmoid function [14].

Table 1: Test Data Sets

sets	Test data from
Set 1	01/23/'89 - 01/30/'89
Set 2	11/09/'88 - 11/17/'88
Set 3	11/18/'88 - 11/29/'88
Set 4	12/08/'88 - 12/15/'88
Set 5	12/27/'88 - 01/04/'89

4 Test Cases and Results

Hourly temperature and load data for Seattle/Tacoma area in the interval of Nov. 1, 1988 - Jan. 30, 1989 were collected by the Puget Sound Power and Light Company. We used this data to train the ANN and test its performance. Our focus is on a normal weekday (*i.e.* no holiday or weekends).

Table 1 shows five sets used to test the neural network. Each set contains 6 normal days. These test data were not used in the training process of the neural network. This approach of classifier evaluation is known as a *jack-knife* method.

The ANN was trained to recognize the following cases:

- Case 1: Peak load of the day
- Case 2: Total load of the day
- Case 3: Hourly load

where

$$\text{Peak load at day } d = \max \{L(1, d), \dots, L(24, d)\} \quad (11)$$

$$\text{Total load at day } d = \sum_{h=1}^{24} L(h, d) \quad (12)$$

$L(h, d)$ is the load at hour h on day d .

The neural network structures used in this paper, including the size of the hidden layer, were chosen from among several structures. The chosen structure is the one that gave the best network performance in terms of accuracy. In most cases, we found that adding one or two hidden neurons did not significantly effect the neural network accuracy.

To evaluate the resulting ANN's performance, the following percentage error measure is used throughout this paper:

$$\text{error} = \frac{|\text{actual load} - \text{forecasted load}|}{\text{actual load}} \times 100 \quad (13)$$

4.1 Case 1

The topology of the ANN for the peak load forecasting is as follows;

Input neurons: T1(k), T2(k), and T3(k)
 Hidden neurons: 5 hidden neurons
 Output neuron: L(k)

where

k = day of predicted load,
 L(k) = *peak load* at day k,
 T1(k) = average temperature at day k,
 T2(k) = peak temperature at day k,
 T3(k) = lowest temperature at day k.

Table 2: Error(%) of Peak Load Forecasting

days	set1	set2	set3	set4	set5
day1	4.19	1.89	0.72	1.69	1.83
day2	0.24	1.85	3.03	0.31	3.25
day3	0.58	2.44	0.95	2.72	2.68
day4	2.39	3.85	3.29	2.84	1.10
day5	0.35	4.26	0.65	6.64	0.56
day6	2.81	0.13	0.63	1.40	2.04
Avg.	1.73	2.40	1.55	2.60	1.91

Table 3: Error(%) of Total Load Forecasting

days	set1	set2	set3	set4	set5
day1	0.34	0.26	2.66	1.03	0.42
day2	1.02	1.99	1.82	0.70	0.92
day3	3.47	1.03	3.25	0.66	1.42
day4	1.63	1.73	5.64	1.89	2.11
day5	1.04	0.88	4.14	0.03	0.27
day6	1.77	1.10	2.96	1.20	1.05
Avg.	1.78	1.07	3.39	1.15	1.03

Table 2 shows the error(%) of each day in the test sets. The average error for all 5 sets is 2.04 %.

4.2 Case 2

The topology of the ANN for the total load forecasting is as follows;

Input neurons: T1(k), T2(k), and T3(k)
 Hidden neurons: 5 hidden neurons
 Output neuron: L(k)

where

k = day of predicted load,
 L(k) = *total load* at day k,
 T1(k) = average temperature at day k,
 T2(k) = peak temperature at day k,
 T3(k) = lowest temperature at day k.

Table 3 shows the error(%) of each day in test sets. The average error for all 5 sets is 1.68 %.

4.3 Case 3

The topology of the ANN for the hourly load forecasting with one hour of lead time is as follows;

Input neurons: k, L(k-2), L(k-1),
 T(k-2), T(k-1), and $\tilde{T}(k)$
 Hidden neurons: 10 hidden neurons
 Output neuron: L(k)

k = hour of predicted load
 L(x) = load at hour x,
 T(x) = temperature at hour x,
 $\tilde{T}(x)$ = predicted temp. for hour x

In training stage, T(x) was used instead of $\tilde{T}(x)$. The lead times of predicted temperatures, $\tilde{T}(x)$, vary from 16 to 40 hours.

Table 4 shows the error(%) of each day in the test sets. The average error for all 5 sets is found to be 1.40 %. Note that each day's result is averaged over a 24 hour period.

Table 4: Error(%) of Hourly Load Forecasting with One Hour Lead Time

days	set1	set2	set3	set4	set5
day1	(*)	1.20	1.41	1.17	(*)
day2	1.67	1.48	(*)	1.58	2.18
day3	1.08	(*)	1.04	(*)	1.68
day4	1.40	1.34	1.42	1.20	1.73
day5	1.30	1.41	(*)	1.20	(*)
day6	(*)	1.51	1.29	1.68	0.98
avg.	1.35	1.39	1.29	1.36	1.64

(*: Predicted temperatures, \tilde{T} , are not available.)

In order to find the effect of the lead time on the ANN load forecasting, we used set 2 whose performance in Table 4 was the closest to the average. The lead time was varied from 1 to 24 hours with a 3 hour interval. The topology of ANN was as follows:

input neurons : $k, L(24,k), T(24,k), L(m,k), T(m,k), \text{ and } \tilde{T}(k)$
 hidden neurons : 1 hidden neuron
 output neuron : $L(k)$

where

k = hour of predicted load
 m = lead time,
 $L(x,k)$ = load x hours before hour k
 $T(x,k)$ = temperature x hours before hour k
 $\tilde{T}(k)$ = predicted temperature for hour k

In the training stage, $T(x)$ was used instead of $\tilde{T}(x)$. The lead times of predicted temperatures, $\tilde{T}(x)$, vary from 16 to 40 hours.

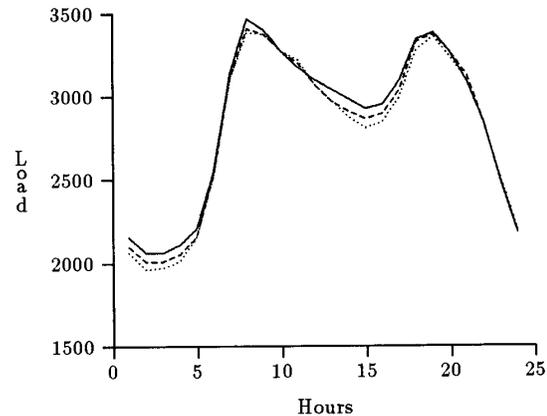
Figure 2 shows examples of the hourly actual and forecasted loads with one-hour and 24-hour lead times. Figure 3 shows the average errors (%) of the forecasted loads with different lead hours for test set 2.

From Figure 3, the error gradually increases as the lead hour grows. This is true up to 18 hours of lead time. One of the reasons for this error pattern is the periodicity of temperature and load pattern. Even though they are not quite the same as those of the previous day, the temperature and system load are very similar to those of the previous day.

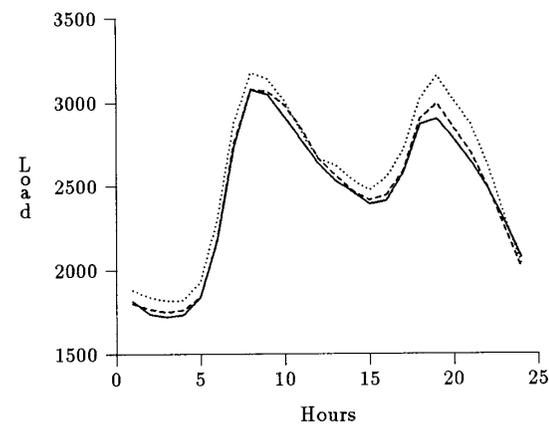
We compare our results with the prediction of Puget Sound Power and Light Co. (PSPL) in Figure 4. Since the PSPL forecasts loads with lead times of 16- to 40-hour, there are 3 overlaps(18-, 21-, and 24-hour) with our results. As shown in Figure 4, the average errors for the 18-, 21- and 24-hour lead times are 2.79, 2.65, and 2.06 %, respectively. This compares quite favorably with errors of 2.72, 6.44, and 4.22 % (18-, 21-, and 24-hour lead times) obtained by current load forecasting technique using the same data from PSPL [27]. The current load forecasting method, in addition, uses cloud cover, opaque cover, and relative humidity information.

5 Conclusions

We have presented an electric load forecasting methodology using an artificial neural network(ANN). This technique was inspired by the work of Lapedes and Farber [28]. The performance of this technique is similar to the ANN with locally tuned receptive field [29]. We find it no-



(a) Jan. 24, 1989



(b) Jan. 27, 1989

Figure 2: Hourly Load Forecasting and Actual Load (in MW) (solid: actual load, dash: 1-hour lead forecast, dot: 24-hour lead forecast)

table that Moody and Darken's technique is remarkably similar to the estimation of Gaussian mixture models.

The results shows that the ANN is suitable to interpolate among the load and temperature pattern data of training sets to provide the future load pattern. In order to forecast the future load from the trained ANN, we need to use the recent load and temperature data in addition to the predicted future temperature. Compared to the other regression methods, the ANN allows more flexible relationships between temperature and load pattern. A more intensive comparison can be found in [30].

Since the neural network simply interpolates among the training data, it will give high error with the test data that is not close enough to any one of the training data.

In general, the neural network requires training data well spread in the feature space in order to provide highly accurate results. The training times required in our experiments vary, depending on the cases studied, from 3 to 7 hours of CPU time using the SUN SPARK Station 1. However, a trained ANN requires only 3 to 10 millisec-

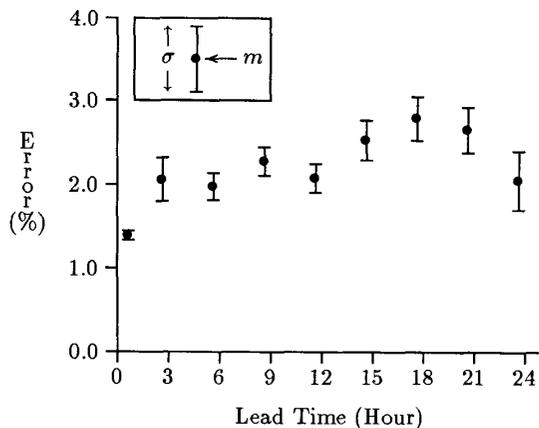


Figure 3: Mean(m) and Standard Deviation(σ) of Errors Vs. Lead Time

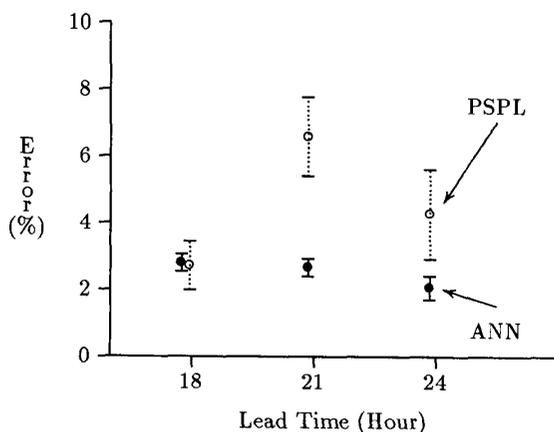


Figure 4: Mean and Standard Deviation of Errors: ANN Vs. Conventional Technique Used in PSPL

onds for testing.

The neural network typically shows higher error in the days when people have specific start-up activities such as Monday (for example, on day 1 of set 1 in Table 2), or variant activities such as during the holiday seasons (for example, on days 4 & 5 of set 3 in Table 3). In order to have more accurate results, we may need to have more sophisticated topology for the neural network which can discriminate start-up days from other days.

We utilize only temperature information among weather variables since it is the only information available to us. Use of additional weather variables such as cloud coverage and wind speed should yield even better results.

6 Acknowledgments

This work was supported by the Puget Sound Power and Light Co., the National Science Foundation, and the Washington Technology Center at the University of

Washington. The authors thank Mr. Milan L. Bruce of the Puget Sound Power and Light Co. for his contribution.

References

- [1] J. Toyoda, M. Chen, and Y. Inoue, "An Application of State Estimation to short-Term Load Forecasting, Part1: Forecasting Modeling," "— Part2: Implementation," IEEE Tr. on Power App. and Sys., vol. PAS-89, pp.1678-1688, Oct., 1970
- [2] S. Vemuri, W. Huang, and D. Nelson, "On-line Algorithms For Forecasting Hourly Loads of an Electric Utility," IEEE Tr. on Power App. and Sys., vol. PAS-100, pp.3775-3784, Aug., 1981
- [3] G.E. Box and G.M. Jenkins, *Time Series Analysis - Forecasting and Control*, Holden-day, San Francisco, 1976
- [4] S. Vemuri, D. Hill, R. Balasubramanian, "Load Forecasting Using Stochastic Models," Paper No. TPI-B, Proc. of 8th PICA conference, Minneapolis, Minn., pp.31-37, 1973
- [5] W. Christiaanse, "Short-Term Load Forecasting Using General Exponential Smoothing," IEEE Tr. on Power App. and Sys., vol. PAS-90, pp. 900 - 910, Apr., 1971
- [6] A. Sage and G. Husa, "Algorithms for Sequential Adaptive Estimation of Prior Statistics," Proc. of IEEE Symp. on Adaptive Processes, State College, Pa., Nov., 1969
- [7] R. Mehra, " On the Identification of Variance and Adaptive Kalman Filtering," Proc. of JACC (Boulder, Colo.), pp.494-505, 1969
- [8] P. Gupta and K. Yamada, "Adaptive Short-Term Forecasting of Hourly Loads Using Weather Information," IEEE Tr. on Power App. and Sys., vol. PAS-91, pp.2085-2094, 1972
- [9] C. Asbury, "Weather Load Model for Electric Demand Energy Forecasting," IEEE Tr. on Power App. and Sys., vol. PAS-94, no.4, pp.1111-1116, 1975
- [10] S. Rahman and R. Bhatnagar, " An Expert System Based Algorithm for Short Load Forecast," IEEE Tr. on Power Systems, vol.3, no.2, pp.392-399, May, 1988
- [11] K. Jabbour, J. Riveros, D. Landbergen, and W. Meyer, "ALFA: Automated Load Forecasting Assistant," IEEE Tr. on Power Systems, vol.3, no.3, pp.908-914, Aug., 1988
- [12] D. Sobajic and Y. Pao, "Artificial Neural-Net Based Dynamic Security Assessment for Electric Power Systems," IEEE Tr. on Power Systems, vol.4, no.1, pp.220-228, Feb, 1989
- [13] M. Aggoune, M. El-Sharkawi, D. Park, M. Damborg, and R. Marks II, "Preliminary Results on Using Artificial Neural Networks for Security Assessment," Proc. of PICA, pp.252-258, May, 1989
- [14] M. El-Sharkawi, R. Marks II, M. Aggoune, D. Park, M. Damborg, and L. Atlas, " Dynamic Security Assessment of Power Systems Using Back Error Propagation Artificial Neural Networks," Proc. of 2nd Sym. on Expert Systems Applications to Power Systems, pp.366-370, July, 1989
- [15] H. Mori, H. Uematsu, S. Tsuzuki, T. Sakurai, Y. Kojima, K. Suzuki, "Identification of Harmonic Loads in Power Systems Using An Artificial Neural Network," Proc. of 2nd Sym. on Expert Systems Applications to Power Systems, pp.371-377, July, 1989

- [16] E.H. Chan, "Application of Neural-Network Computing in Intelligent Alarm Processing," Proc. of PICA, pp.246-251, May, 1989
- [17] H. Tanaka, S. Matsuda, H. Ogi, Y. Izui, H. Taoka, and T. Sakaguchi, "Design and Evaluation of Neural Network for Fault Diagnosis," Proc. of 2nd Sym. on Expert Systems Application to Power Systems, pp.378-384, July, 1989
- [18] H. Mori and S. Tsuzuki, "Power System Topological Observability Analysis Using a Neural Network Model," Proc. of 2nd Sym. on Expert Systems Application to Power Systems, pp.385-391, July, 1989
- [19] N. Naylor and G. Sell, *Linear Operator Theory*, New York, Holt, Rinehart and Winston, 1971
- [20] M. Honig and D. Messerschmitt, *Adaptive Filters, Structures, Algorithms, and Applications*, Klumer Academic Publishers, Hingham, Massachusetts, 1984
- [21] J. Davey, J. Saacks, G. Cunningham, and K. Priest, "Practical Application of Weather Sensitive Load Forecasting to System Planning," IEEE Tr. on Power App. and Sys., vol.PAS-91, pp.971-977, 1972
- [22] R. Thompson, "Weather Sensitive Electric Demand and Energy Analysis on a Large Geographically Diverse Power System - Application to Short Term Hourly Electric Demand Forecasting," IEEE Tr. on Power App. and Sys., vol. PAS-95, no.1, pp.385-393, Jan., 1976
- [23] G. Irisarri, S. Widergren, and P. Yehsakul, "On-Line Load Forecasting for Energy Control Center Application," IEEE Tr. on Power App. and Sys., vol. PAS-101, no.1, pp.71-78, Jan., 1982
- [24] Q. Lu, W. Grady, M. Crawford, and G. Anderson, "An Adaptive Nonlinear Predictor with Orthogonal Escalator Structure for Short-Term Load Forecasting," IEEE Tr. on Power Systems, vol.4, No.1, pp.158-164, Feb., 1989
- [25] Y.-H. Pao, *Adaptive Pattern Recognition and Neural Network*, Addison-Wesley Pub. Co. Inc., Reading, MA., 1989
- [26] D. Rumelhart, G. Hinton, and R. Williams, "Learning Internal Representations by Error Propagation," in *Parallel Distributed Processing Explorations in the Microstructures of Cognition, vol.1: Foundations*, pp.318-362, MIT Press, 1986
- [27] S. Mitten-Lewis, *Short-Term Weather Load Forecasting Project Final Report*, Puget Sound Power and Light Co., Bellevue, Washington, 1989
- [28] A. Lapedes and R. Farber, *Nonlinear Signal Processing Using Neural Networks: Prediction and System Modeling*, Technical Report, Los Alamos National Laboratory, Los Alamos, New Mexico, 1987
- [29] J. Moody and C. Darken, "Learning with Localized Receptive Fields," Proc. of the 1988 Connectionist Models Summer School, Morgan Kaufmann, 1988
- [30] L. Atlas, J. Connor, D. Park, M. El-Sharkawi, R. Marks II, A. Lippman, and Y. Muthusamy, "A Performance Comparison of Trained Multi-Layer Perceptrons and Trained Classification Trees," Proc. of the 1989 IEEE International Conference on Systems, Man, and Cybernetics, pp.915-920, Nov. 1989

Dong C. Park received his B.S. Degree in Electronic Engineering in 1980 from Sogang University and the M.S. degree in Electrical Engineering in 1982 from the Korea Advanced Institute of Science and Technology, Seoul, Korea. From 1982 through 1985, he had been with the Goldstar Central Research Laboratory. Since September 1985,

he has been working toward the Ph.D. degree in the Department of Electrical Engineering at the University of Washington. His research interests include artificial neural network application to nonlinear system modeling, signal processing and optical computing.

M. A. El-Sharkawi (SM'76-M'80-SrM'83) was born in Cairo, Egypt, in 1948. He received his B.Sc. in Electrical Engineering in 1971 from Cairo High Institute of Technology, Egypt. His M.A.Sc and Ph.D. in Electrical Engineering were received from University of British Columbia in 1977 and 1980 respectively. In 1980 he joined University of Washington as a faculty member where he is presently an associate professor. He is the Chairman of IEEE Task Force on "Application of Artificial Neural Networks for Power Systems". His major areas of research include neural network applications to power systems, electric devices, high performance tracking control, power system dynamics and control. Most of his research in these areas are funded by the US government, and by public and private industrial organizations.

Robert J. Marks II received his Ph.D. in 1977 from Texas Tech University in Lubbock. He joined the faculty of the Department of Electrical Engineering at the University of Washington, Seattle, in December of 1977 where he currently holds the title of Professor. Prof. Marks was awarded the Outstanding Branch Councillor award in 1982 by IEEE and, in 1984, was presented with an IEEE Centennial Medal. He is President of the IEEE Council on Neural Networks and former Chair of the IEEE Neural Network Committee. He was also the co-founder and first Chair of the IEEE Circuit & Systems Society Technical Committee on Neural Systems & Applications. He is a Fellow of the Optical Society of America and a Senior Member of IEEE. He has over eighty archival journal and proceedings publications in the areas of signal analysis, detection theory, signal recovery, optical computing and artificial neural processing. Dr. Marks is a co-founder of the Christian Faculty Fellowship at the University of Washington. He is a member of Eta Kappa Nu and Sigma Xi.

Les E. Atlas (Member, IEEE) received his B.S.E.E. degree from the University of Wisconsin and his M.S. and Ph.D. degrees from Stanford University. He joined the University of Washington College of Engineering in 1984 and is currently an Associate Professor of Electrical Engineering. He is currently doing research in speech processing and recognition, neural network classifiers, and biologically-inspired signal processing algorithms and architectures. His research in these areas is funded by the National Science Foundation, the Office of Naval Research, and the Washington Technology Center. Dr. Atlas was a 1985 recipient of a National Science Foundation's Presidential Young Investigator Award.

M. J. Damborg received his B.S. Degree in Electrical Engineering in 1962 from Iowa State University, and the M.S. and Ph.D. degrees from the University of Michigan in 1963 and 1969, respectively. Since 1969, Dr. Damborg has been at the University of Washington where he is now Professor of Electrical Engineering. His research interests concern analysis and control of dynamic systems with emphasis on power systems.

Discussion

O. A. Mohammed (Florida International University Miami, FL): The authors are to be thanked on their excellent work applying this new ANN technique to load forecasting. I would like the authors to clarify or explain the followings points:

1. The authors presented a new method for load forecast which shows a promise for providing accurate forecasts. This discussor feels that the ANN method would be adequate for providing the base forecast which might be combined with an expert system approach to fine tune the load forecast for additional factors.
2. If one experiments with additional factors which may affect the load forecast such as humidity, load inertia, wind velocity, etc., how much additional training time would be required compared with the data size.
3. The authors presented results for hourly load forecast for weekdays but not weekends because of the variation in load pattern. Will this be handled by a separate neural network? and if so, how would it be combined with previous day forecasts. For example, to forecast Monday's load.
4. Have the authors experimented with different ANN architectures other than the ones explained in the paper. It seems to this discussor that the proposed architectures will not work all the time or it may yield larger errors because of the continual change in weather and load information. May be a methodology which updates the weights of the ANN based on the new short term weather and load information.

Manuscript received August 13, 1990.

M. A. El-Sharkawi and M. J. Damborg: The authors would like to thank the discussor for their interest and encouraging comments. The research work reported in this paper is preliminary. Several key issues, such as those raised by the discussor, need to be carefully addressed before a viable electric load forecasting system is deployed. The purpose of the paper, however, is to investigate the potentials of the Neural Network

(NN) in load forecasting. Future work should certainly address questions related to weather conditions, distinct load profiles, cold snaps, etc.

To respond to the specific issues raised by the reviewer, we would like to offer the following comments:

1. The role of expert systems in NN environment, and vice versa, is a topic that is being proposed for several applications. In load forecasting applications, as an example, the selection of relevant training sets from load and weather data base is currently accomplished manually and off-line. Also, the convergency of the NN is currently observed and controlled at only discrete training steps. These functions, for example, may be effectively accomplished by a supervisory layer employing a rule-based system.
2. Other weather variables such as wind speed and humidity may result in more accurate load forecasting. The problem, however, is that the forecasting errors of these variables are usually high which may lead to a biased training or erroneous network.
3. Except for Tuesday to Thursday, the load profile of the each other day of the week is distinct. For example, the profile of Monday morning include the "pickup loads". Due to these differences in load profiles, we have used one NN for the days with similar load profiles and one NN for each day with distinct load profile.

When we forecasted the electric loads of Saturday, Sunday or Monday, we used weather and load data obtained up to Friday morning (9:00 am) to conform with Puget Power practice.

4. We have tried several architectures for load forecasting. The key issue in selecting a particular NN configuration is to achieve low training error without "memorization". This can be accomplished by first selecting an over sized network then "prune" the network to eliminate any memorization problem that might exist without jeopardizing the training accuracy.

Manuscript received September 23, 1990.