Fuzzy neural network and fuzzy expert system for load forecasting

P.K. Dash
A.C. Liew
S. Rahman

Indexing terms: Neural networks, Load forecasting, Expert systems, Power system planning

Abstract: A hybrid neural network fuzzy expert system is developed to forecast short-term electric load accurately. The fuzzy membership values of the load and other weather variables are the inputs to the neural network, and the output comprises the membership values of the predicted load. An adaptive fuzzy correction scheme is used to forecast the final load by using a fuzzy rule base and fuzzy inference mechanism. Extensive studies have been performed for all seasons, and a few examples are presented in the paper, including average, peak and hourly load forecasts.

1 Introduction

The short-term load forecast is very important to an electric utility. The quality of control of a power system, and economy of operation, are highly sensitive to forecasting error. A sound basis for load predictions is generalisation of past, known cases. The science of statistics provides a range of tools for this purpose. They are based on the idea of fitting a particular class of models to data and then hypothesising that future events will conform to the fitted model. Many approaches have been applied to electric load forecasting, including linear regression, exponential smoothing, stochastic process and state space methods. While each of these methods demonstrates success in forecasting, they have serious disadvantages: reliance on large historical databases with possible obsolete and irrelevant data, assumptions about static load shapes and parameters etc. A detailed comparison of various statistical approaches is found in [1].

One of the most promising application areas of the artificial neural network (ANN) is load forecasting [2–13]. The neural network is able to perform nonlinear modelling and adaptation and does not rely on the explicitly expressed relationship between input variables and forecast load. When using neural networks for load forecasting, one needs to consider only the selection of variables as the network input. The relationship between the input variables and predicted load will be formulated by a training process. Several authors have attempted to apply the backpropagation learning algorithm to train the ANNs for forecasting time series. The fuzzy expert system approach [14] has also been applied to forecasting where the advantage of an operator's expert knowledge is used. However, the fuzzy decision system for load forecasting requires detailed analysis of data and the fuzzy rule base has to be developed heuristically for each season. The rules fixed in this way may not always yield the best forecast. On the other hand, hybrid solutions [15, 16] have been proposed for short-term forecasting of electric loads, whereby the functionality of the fuzzy expert system and the learning capabilities of the neural network can be merged to yield a forecasting system more powerful than either of its components alone.

The present work is aimed at achieving the said objective of a robust load forecast with improved accuracy using a fuzzy neural network for initial forecast and a fuzzy expert system (FES) producing load corrections to yield the final forecast. For the neural network to be called a FNN, the signal and/or the weights should be fuzzified. This type of FNN is based on the multilayer perceptron, using the backpropagation algorithm. The fuzzified input vector consists of the membership values of the past load and weather parameters and the output vector is defined in terms of fuzzy class-membership values of the forecast load. A simple fuzzy-inferencing mechanism is used to yield the magnitude of the forecast load during the initial phase. In the final phase a fuzzy expert system is used to produce load corrections.

The input vector to the fuzzy expert system (FES) consists of differences in the weather parameters between the present and the forecasted instant. The output of the FES gives the load correction which, when added to the initial forecast, yields the final forecast. Thus, by using a hybrid approach the load-forecasting errors are expected to reduce considerably. However, as the lead time increases from 1 h to 24h or 48h, the forecasting error increases, necessitating the use of an adaptive fuzzy load error correction scheme. Several examples presented in this paper include 24h-ahead average, peak and hourly load forecasts using the hybrid approach. The effects of both linear and nonlinear adaptive load-correction schemes are shown.

IEE Proceedings online no. 19960314
Paper received 26th October 1994

P.K. Dash is with the Department of Electrical Engineering, Regional Engineering College, Rourkela, India
A.C. Liew is with the Department of Electrical Engineering, National University of Singapore, Singapore
S. Rahman is with the Bradley Department of Electrical Engineering, Virginia Polytechnic Institute & State University, USA

2 Fuzzy pattern representation in linguistic form

The approach used in this paper is aimed at improving the prediction by handling uncertain and ambiguous variations of load patterns and weather parameters using a fuzzy linguistic approach. Since it is easier to convert exact information into linguistic form than vice versa, we consider the major linguistic properties small, medium, and large as the attributes of the input feature. Any input feature value can be described in terms of some combination of membership values for these properties. In traditional two-state classifiers, an element x either belongs or does not belong to a given class A; thus, the characteristic function is expressed as

$$\mu_a(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

In real-life problems such as load forecasting, the classes are often ill defined, overlapping, or fuzzy and a pattern point may belong to more than one class; in such situation, fuzzy set theoretic techniques can be very useful.

We use the modified π-function [17, 18], lying in the range [0,1] to assign membership values for the input features corresponding to the linguistic properties small, medium, and large.

The membership function of the input feature

$$\mu(x_i) = \begin{cases} 2\left(1 - \dfrac{x_i - c_i}{\lambda_i}\right)^2 & \text{for } \dfrac{\lambda_i}{2} \le (x_i - c_i) \le \lambda_i \\ 1 - 2\left(1 - \dfrac{x_i - c_i}{\lambda_i}\right)^2 & \text{for } 0 \le (x_i - c_i) \le \dfrac{\lambda_i}{2} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $\lambda_i > 0$ is the radius of the π-function with c_i as the central point at which $\mu(x_i) = 1$. This is shown in Fig. 1.

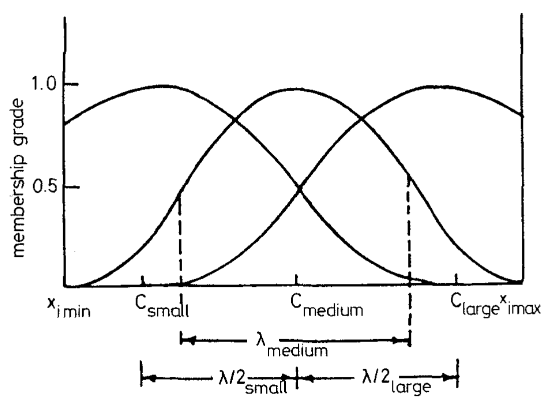

Fig. 1 *π-function representation*

2.1 Choice of parameters for the π-function

Let x_{imax} and x_{imin} denote the upper and lower bounds of the observed range of feature x_i in all L pattern points, considering numeric values only. Then, for the three linguistic property sets, the following are used:

$$\begin{aligned} \lambda_{medium}(x_i) &= 0.5(x_{i\,max} - x_{i\,min}) \\ c_{medium}(x_i) &= x_{i\,min} + \lambda_{medium}(x_i) \\ \lambda_{small}(x_i) &= (1/f_d)\{c_{medium}(x_i) - x_{i\,min}\} \\ c_{small}(x_i) &= c_{medium}(x_i) - 0.5\lambda_{small}(x_i) \\ \lambda_{large}(x_i) &= (1/f_d)\{x_{i\,max} - c_{medium}(x_i)\} \\ c_{large}(x_i) &= c_{medium}(x_i) + 0.5\lambda_{large}(x_i) \end{aligned} \tag{3}$$

where $0.5 \le f_d \le 1.0$ is a parameter controlling the extent of overlapping. The π-function representation permits a more compact and meaningful representation of each pattern point in terms of its linguistic properties, and ensures better handling during both the training and testing phases of the proposed fuzzy neural network model.

3 Fuzzy neural network for initial forecast

The present work attempts to build a fuzzy neural network model based on the multilayer perceptron using the gradient descent based backpropagation algorithm by incorporating concepts from fuzzy sets at various stages. Fig. 2 shows the fuzzy neural network model for obtaining the initial forecast. The fuzzy sets for load, temperature and humidity parameters are shown in Figs. 3–5. The input to the fuzzy neural network comprises the membership values to the overlapping partitions of linguistic properties small, medium, and large corresponding to each input feature such as past load, temperature, humidity etc. This provides scope for incorporating linguistic information in both the training and testing phases of the said model and increases robustness in tackling imprecise or uncertain input specifications. The components of the output layer consist of the membership values to the overlapping partitions of linguistic properties small, medium and large corresponding to the forecast load magnitude. During training, the network backpropagates the errors with respect to the desired membership values at the output nodes. After a number of cycles, the neural network converges to a minimum error solution by using a gradient descent algorithm as shown in the Appendix.

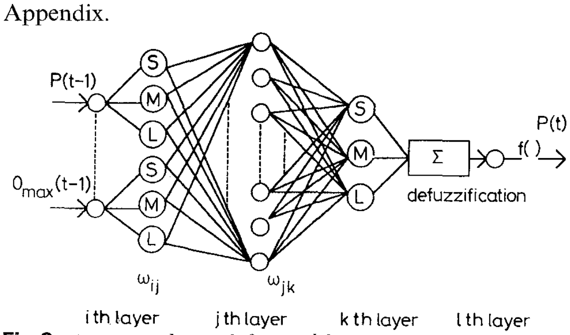

Fig. 2 *Fuzzy neural network for initial forecast*

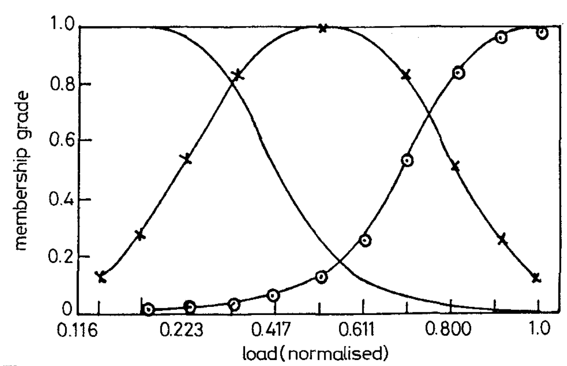

Fig. 3 *Fuzzy membership grade for load*
—— small
—⊙— medium
—×— large

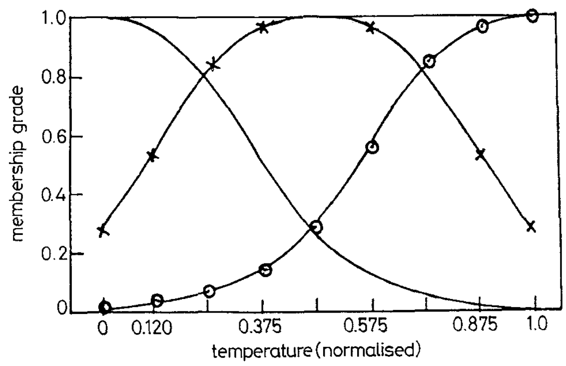

Fig. 4 *Fuzzy membership grade for temperature*
—— small
—⊙— medium
—×— large

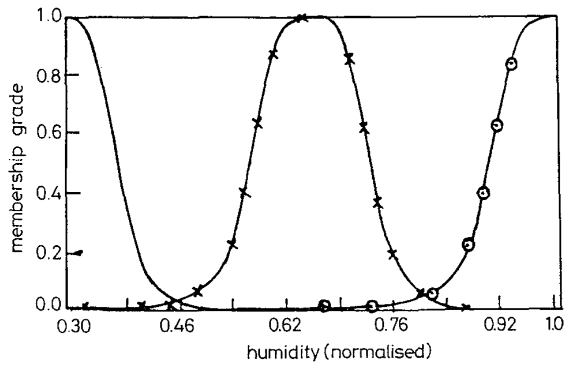

Fig. 5 *Fuzzy membership grade for humidity*
—— small
—⊙— medium
—×— large

After the learning phase is over, when separate load and weather patterns are presented at the input layer, the output nodes automatically generate the membership values corresponding to the linguistic properties small, medium and large. Thus a centroid defuzzification technique [19] is used to obtain the initial load forecast from the membership values and the corresponding loads obtained from the π-functions.

4 Selection of training patterns

The utility data studied here are susceptible to large and sudden changes in weather and load, so selection of appropriate training cases plays a vital role in training the network. Several techniques for the selection of training patterns have been suggested in [11–13]. The present paper discusses a different training scheme for the selection of training patterns for hourly load forecast. To predict hourly loads the following load model is chosen:

$$y(i,t) = f\{y(i-m), y(i,t-m-1), \ldots, y(i,t-m-n_1), \\ z(i,t-m), z(i,t), \ldots, z(i-t-n_2)\} \quad (4)$$

and

$$m = n_2$$

where y and z are the load and weather variables, respectively; i and t indicate the day and the hour, respectively, m indicates the lead time for the hourly load forecast, (i.e. $m = 1$ for 1h-ahead forecast, $m = 24$ for 24h-ahead forecast, $m = 48$ for 48h-ahead forecast, $m = 168$ for 168h-ahead forecast); n_1 indicates the data length for load; and n_2 indicates the data length for temperature.

For hourly load forecasting, eqn. 4 is used to select the training patterns. Various lengths of the past historical load and temperature values are used and their effects on the load-forecast accuracy are studied. It is found that, with $n_1 > 0$ and $n_2 > 0$, there are no marked improvements in the results for the utility data used in this paper. Also the training time increases considerably with larger values of n_1 and n_2. Therefore, $n_1 = 0$ and $n_2 = 0$ are chosen in this paper.

Using the above scheme, the network is trained for 14 days (i = 1, ..., 14) at time t and load is predicated for the next 14 days (i = 15, ..., 28) at time t. Hence to predict for all 24h of a given day, 24 different neural networks are used, each one trained separately with the same parameters. This is desirable because the training set is small for each neural network consisting of a few patterns (14 patterns only in this case) with irrelevant data for other hours being discarded. Further, depending upon the difference of the load responses on the day of the week, a day-of-the-week indicator is introduced along with the input vector to the network.

As the weather variable temperature is the most important parameter in short-term load predictions, an all-temperature model is used to obtain the hourly forecasts. The training data used for 1 h-ahead predictions are:

Input pattern:

$P(i,t)$ = power at tth instant of ith day,

$\theta(i,t)$ = temperature at $(t+1)$th instant of ith day,

$\theta(i,t+1)$ = temperature at $(t+1)$th instant of ith day

$wd(i)$ = day of the week indicator i.

Output pattern:

$P(i,t+1)$ = power at $(t+1)$th instant of ith day.

The training data for 24h-ahead, 48h-ahead and 168h-ahead predications are

Input patterns:

$P(i,t)$ = power at tth instant of ith day,

$\theta(i,t)$ = temperature at tth instant of ith day,

$\theta(i+m,t)$ = temperature at tth instant of $(i+m)$th day.

Output pattern:

$P(i+m,t)$ = power at tth instant of $(i+m)$th day

where m = 1, 2, 7 for 24h-, 48h- and 168h-ahead predictions, respectively.

However, for average daily and peak-load predictions, the following training data are used:

Input and output patterns:

$P(i-1)$ = average load on $(i-1)$th day,

$\theta_{min}(i-1)$ = minimum temperature of $(i-1)$th day,

$\theta_{max}(i-1)$ = maximum temperature of $(i-1)$th day,

$\theta_{min}(i)$ = minimum temperature of ith day,

$\theta_{max}(i)$ = maximum temperature of ith day,

$P(i)$ = average load for ith day.

The same data can be used for the peak load forecast. Although an all-temperature model will produce an accurate forecast in most seasons, the other weather parameters like humidity and wind speed affect the forecasting accuracy during summer and winter, respectively. Thus, if humidity records are available in a particular season, they may be included in a training pattern.

5 Fuzzy expert system for final forecast

During training, the neural net produces an initial forecast with a set of initial load, temperature and humidity data expressed in terms of fuzzy membership values. The error between the actual load $A(i)$ and predicted load $P(i)$ for a given hour or a given day (ith hour or day) is expressed as

$$\Delta P(i) = A(i) - P(i) \tag{5}$$

In a similar way, temperature and humidity errors are expressed as

$$\begin{aligned} \Delta\theta(i) &= \theta(i) - \theta(i-1) \\ \Delta H(i) &= H(i) - H(i-1) \end{aligned} \tag{6}$$

However, if maximum and minimum temperatures are used,

$$\begin{aligned} \Delta\theta_{max}(i) &= \theta_{max}(i) - \theta_{max}(i-1) \\ \Delta\theta_{min}(i) &= \theta_{min}(i) - \theta_{min}(i-1) \end{aligned} \tag{7}$$

The errors in the weather parameters and load-correction values are fuzzified using six fuzzy sets such as SP (small positive), MP (medium positive), LP (large positive), SN (small negative), NM (medium negative) and LN (large negative). Figs. 6 and 7 show the fuzzy sets for temperature and humidity errors. These sets are obtained using the Π-function given in eqn. 2. Both positive and negative fuzzy sets are symmetrical about the origin.

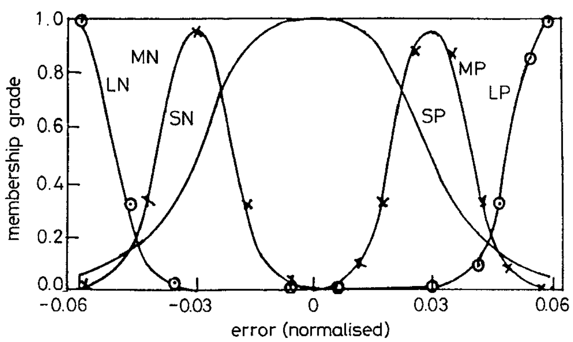

Fig. 6 *Fuzzy membership grade for temperature errors*
—— small
—⊙— medium
—×— large

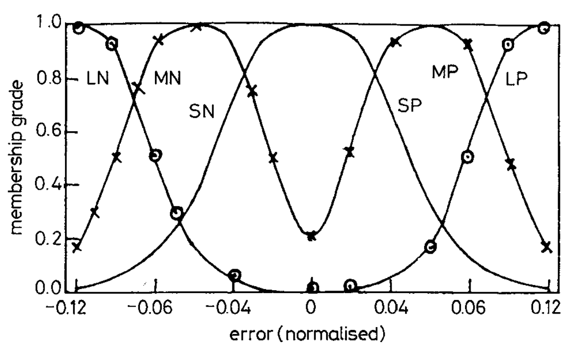

Fig. 7 *Fuzzy membership grade for humidity errors*
—— small
—⊙— medium
—×— large

The load correction output sets have six members and use a linear fuzzification principle for obtaining membership grades as shown in Figs. 8 and 9.

Nonlinear membership grades can also be used for obtaining load corrections for the final forecast. The sets for load corrections are classified as SPC (small positive correction), MPC (medium positive correction), LPC (large positive correction) etc. The corresponding negative fuzzy sets are SNC, MNC, and LNC, respectively.

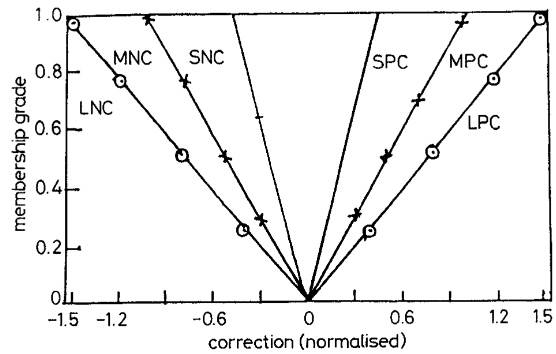

Fig. 8 *Membership grades for load correction (linear)*
—— small
—⊙— medium
—×— large

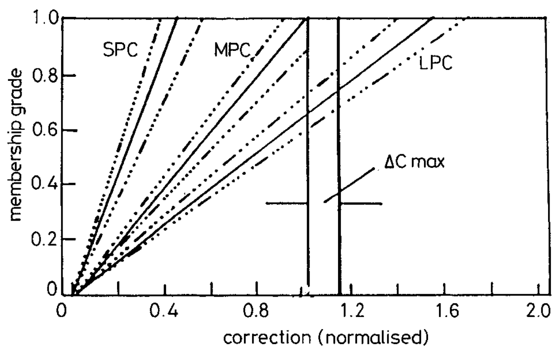

Fig. 9 *Membership grades for load correction (linear)*

The membership values of the load correction ΔP_c output is given by

$$\mu\{\Delta P_c(i)\} = \frac{\Delta P_c(i)}{C_{max}} \tag{8}$$

where C_{max} is the slope of the load error correction and e_{LC} is the maximum load error correction for the corresponding linguistic set (for which the membership value becomes unity). The value of C_{max} for different output linguistic sets SPC, MPC and LPC is C^1_{max}, C^2_{max}, C^3_{max}, respectively, and these values are obtained by observing the load prediction errors over a two-week period prior to forecasting.

The fuzzy rule base is formed by trial and error to reduce the load correction to a very small value during training. However, a fuzzy basis function approach [19] can be used to select the appropriate rules automatically out of a large number of possible combinations. Two sample rules for an all-temperature model for average load forecast will be

Rule 1: IF $\Delta\theta_{max}(i)$ is LN and $\Delta\theta_{min}(i)$ is SP, THEN $\Delta P_c(i)$ is MPC

Rule 2: IF $\Delta\theta_{max}(i)$ is MP and $\Delta\theta_{min}(i)$ is LP, THEN $\Delta P_c(i)$ is LPC

In a similar way the two-sample rule for the peak-load forecast are

Rule 1: IF $\Delta\theta_{max}(i)$ is SP THEN $\Delta P_c(i)$ is SPC

Rule 2: IF $\Delta\theta_{max}(i)$ is LN THEN $\Delta P_c(i)$ is MNC.

However, if a load forecasting model with both temperature and humidity parameters is used, the rules are of the form

Rule 1: IF $\Delta\theta(i)$ is SP and $\Delta H(i)$ is MP THEN $\Delta P_c(i)$ is MPC

Rule 2: IF $\Delta\theta(i)$ is SN and $\Delta H(i)$ is LN THEN $\Delta P_c(i)$ is LNC.

The total number of production rules in the fuzzy knowledge base using two variables and six categories of sets is 36.

Because of partial matching of the fuzzy rules and the fact that preconditions do overlap, more than one fuzzy control rule can fire at a time. For the fuzzy rules used for load forecasting, the truth values of the preconditions are (considering a temperature–humidity model)

$$\alpha_i = \wedge\left[\mu\{\Delta\theta(i)\}, \mu\{\Delta H(i)\}\right] \tag{9}$$

$i = 1, 2, ..., k$ and k = number of rules fired for a given value of $\Delta\theta(i)$, and $\Delta H(i)$ belonging to fuzzy classifiers SP, MP, LP etc. and $\wedge$ denotes a conjunction operator (usually a minimum operator). The output of rule (i) is calculated by applying the matching strength of its precondition on its conclusion as

$$\Delta P_c(i) = \alpha_i C_{max}^{A_j} \tag{10}$$

$j = 1, 2, ..., 6$ and c_{max}^{Aj} is the value of C_{max} for the output set A_j. If two rules have the same consequent output set, the Lukasiewickz OR rule is used as

$$\alpha_i = \min\left[1, \mu\{\Delta\theta(i)\} + \mu\{\Delta H(i)\}\right] \tag{11}$$

A centroid defuzzification technique is used to yield the load correction as

$$\Delta P_c(i) = \sum \alpha_i \Delta P_c \Big/ \sum \alpha_i$$
$$= \sum \alpha_1^2 C_{max}^{A_j} \Big/ \sum \alpha_i \tag{12}$$

However, by taking the load correction as C_{max}^{Aj}, for which the output membership function is unity, the value of $\Delta P_c(i)$ is obtained as

$$\Delta P_c(i) = \sum \alpha_1^2 C_{max}^{A_j} \Big/ \sum \alpha_i \tag{13}$$

The final value of forecast load is thus obtained by summing the ANN output, and the output from the fuzzy expert system is

$$P_f(i) = P(i) + \Delta P_c(i) \tag{14}$$

The block diagram for the integrated fuzzy neural network and fuzzy expert system is shown in Fig. 10.

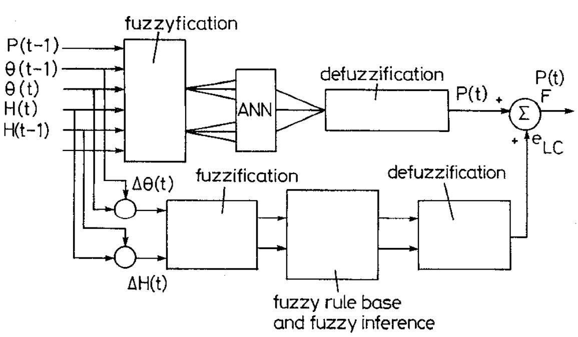

Fig. 10 *Integrated fuzzy neural network–fuzzy expert system*

5.1 Adaptive load correction

For small load correction eqns. 10 and 11 are adequate to produce accurate forecasts (usually for a lead time from 1 – 6h). However, as the lead time increases to 24, 48, 72 or 168h, the membership function for the load-error correction is adapted as

$$\mu\{\Delta P_c(i)\} = \frac{1}{C_{max} \pm \Delta C_{max}} \Delta P_c(i) \tag{15}$$

the value of ΔC_{max} is obtained as a function of load-correction errors from the training cases using different lead times for predictions. Fig. 11 shows the linear adaptive correction ΔC_{max} as a function of the normalised error. The nonlinear adaptive version of the load-error correction is shown in Fig. 12. For adaptive corrections, eqns. 12 and 13 are rewritten as

$$\Delta P_c = \sum (C_{max}^{A_j} \pm \Delta C_{max}) \alpha_i^2 \Big/ \sum \alpha_i \tag{16}$$

and

$$\Delta P_c = \sum (C_{max}^{A_j} \pm \Delta C_{max}) \alpha_i \Big/ \sum \alpha_i \tag{17}$$

Some of the results are given below for the various models considered in this paper.

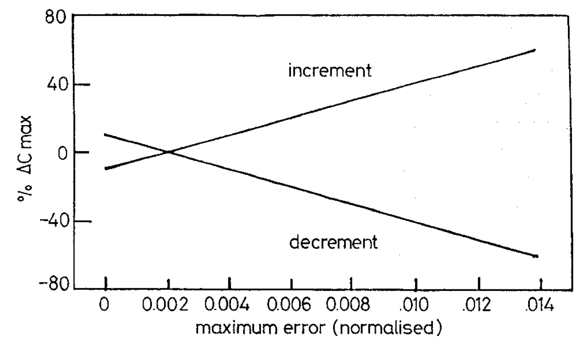

Fig. 11 *Linear adaptive load correction*

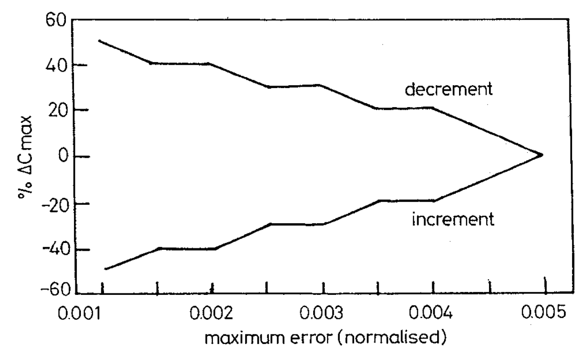

Fig. 12 *Nonlinear adaptive load correction*

6 Forecasting results

6.1 Average load forecasting

To evaluate the fuzzy ANN and fuzzy expert system approach, load forecasting is performed on the load data collected at the Virginia Polytechnic Institute and State University. The fuzzy neural network (FNN) is compared with the combined FNN and fuzzy expert system model using ordinary backpropagation algorithm for obtaining one-day-ahead average load forecasting during a 14 day period in the month of May. The input layer of the FNN comprises 15 neurons for five input features and the output layer has three neurons. The number of neurons in the hidden layer is fixed as 17 for this particular forecast, to obtain the best results. If the day-of-the-week indicator is used, one more neuron is added to the input layer. Backpropagated errors are assigned appropriate weightage for weight updating depending on the membership values at corresponding outputs. The learning rate η and momentum coefficient α are gradually decreased to prevent oscillations as the neural network converges to a minimum-error solution in a maximal number of sweeps through the training set.

Note that the parameters η and α traverse a range of values in the course of computations and one may choose $0.6 < \alpha < 1.0$, and $0.0001 < \eta < 0.1$. The values of β and γ are chosen for best performance as

$$0 < \beta < 0.02$$

$$0.2 < \gamma < 0.6$$

For the present study the initial learning rate η and momentum α are

$$\eta = 0.01, \quad \alpha = 0.8$$

and

$$\beta = -0.015, \quad \gamma = 0.4$$

The percentage error PE is evaluated as

$$\mathrm{PE} = \frac{\text{forecast load} - \text{actual load}}{\text{forecast load}} \times 100$$

and percentage absolute error PAE = |PE|

The average load profile for the 14 day period in May is shown in Fig. 13. Fig. 14 presents the PAE for 24h period ahead forecast. The maximum PAE for a 24h-ahead forecast during a 14 day period in May is 1.75 using the FNN and fuzzy expert system in comparison with 3.40 for the FNN only. Weekdays, weekends and Sundays are all included in this model.

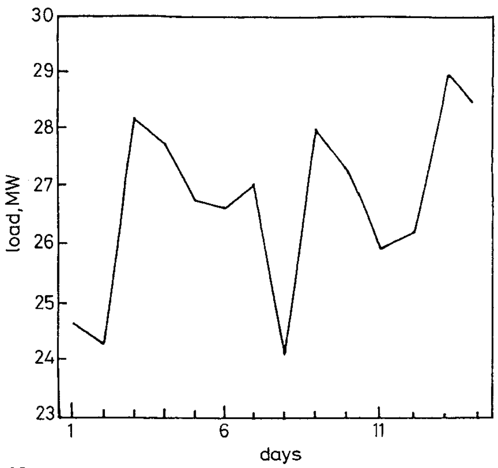

Fig. 13 *Average load profile forecast for the month of May*

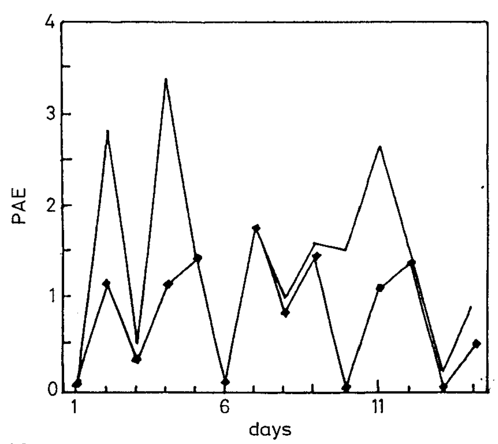

Fig. 14 *Percentage absolute error in average load profile forecast for the month of May*
——— FNN only
—◆— FNN with fuzzy corrections

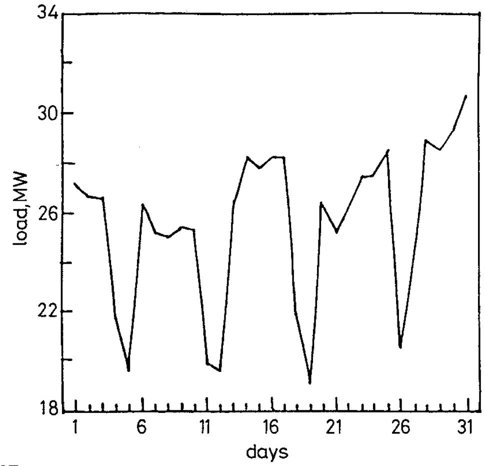

Fig. 15 *Peak load profile forecast for the month of May*

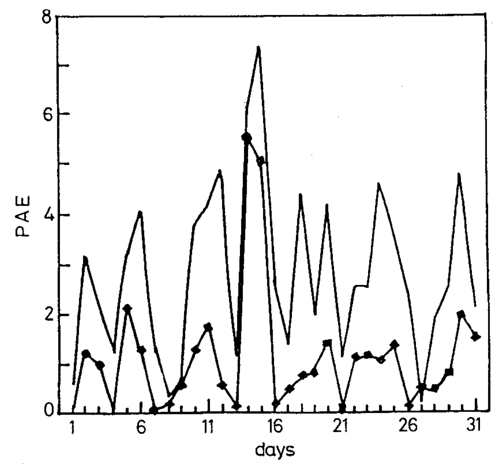

Fig. 16 *Percentage absolute error in peak load profile forecast for the month of May*
——— FNN only
—◆— FNN with fuzzy corrections

6.2 Peak-load forecast

Figs. 15–18 show the actual peak loads and PAEs for a 31 day period during May and December. The maximum temperature errors are used for fuzzy corrections. The combined model produces a maximum PAE of about 1.65 for a 24h-ahead peak-load forecast during the month of December. During May, the maximum PAE is 7.5 for a 24h-ahead peak-load forecast using the FNN model. However, using fuzzy correction, the maximum PAE is reduced to 5.5. Both the forecasts shown in these Figures include weekdays, and weekends etc. Special holidays like Christmas etc. are included in obtaining the peak-load forecast for the month of December. These errors can be reduced further by using adaptive load-correction schemes.

6.3 Hourly load forecasts

The data from a utility in Virginia is used to produce 24h- and 48h-ahead hourly forecasts. For this utility, both temperature and humidity records are available during all seasons of the year. However, the forecasting errors for 21 and 22 January over a 24h period are shown in Figs. 19–22. The maximum PAEs for 24h- and 48h-ahead forecasts are 2.8 and 3.5 without fuzzy corrections and 0.69 and 1.72 with fuzzy corrections, respectively. The corresponding load profiles are also shown in these Figures.

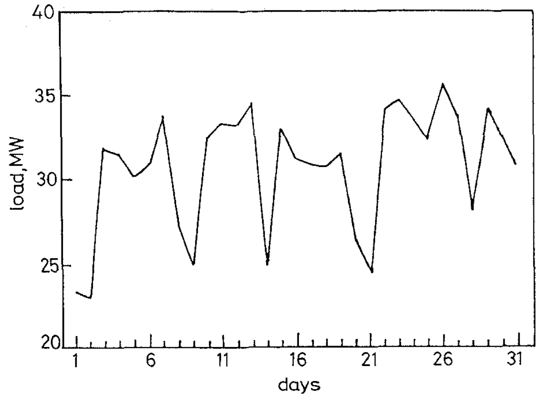

Fig. 17 *Peak load profile forecast for the month of December*

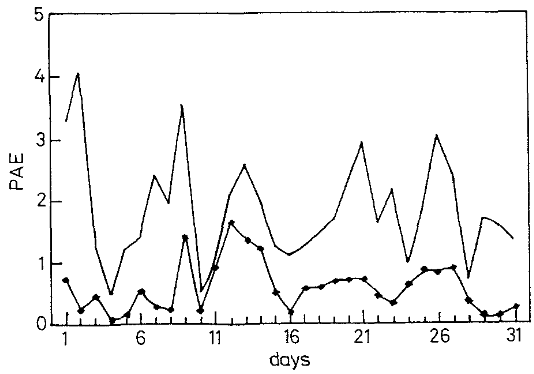

Fig. 18 *Percentage absolute error in peak load profile forecast for the month of December*
——— FNN only
—◆— FNN with fuzzy corrections

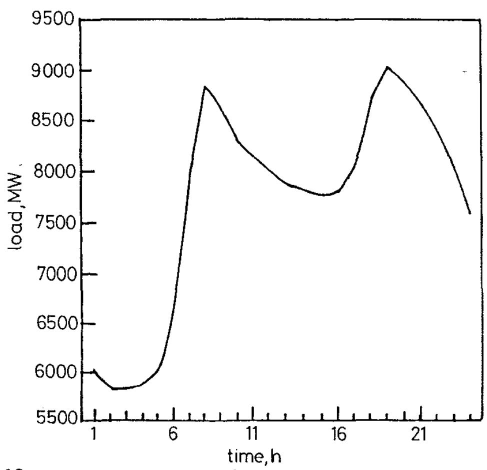

Fig. 19 *24h ahead hourly load forecasts with temperature and humidity records for utility data*

6.4 Hourly load forecasts using adaptive correction

Figs. 23 and 24 show the 24h ahead forecast errors for a summer day (27 May) using an all-temperature model and the data from the experimental setup at the Virginia Polytechnic. Both linear and nonlinear adaptive fuzzy corrections are used to provide a more accurate forecast. From the Figure it is found that the percentage absolute-load-prediction error over the entire 24h period comes down significantly using both adaptive-fuzzy-correction formulations. However, the difference between the two versions is not very significant, and thus the PAE with the nonlinear version is shown in the Figure. The nonlinear version is preferred, as it is expected to produce significant accuracy for 1 to 2h ahead load forecast. The hourly-load-forecast accuracy is significant in this case as the University load profile (shown in the Figure) does not show significant changes during the 24h period.

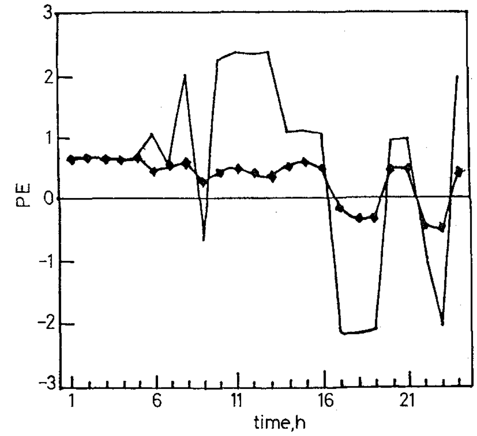

Fig. 20 *24h ahead hourly load forecasts with temperature and humidity records for utility data*
——— without fuzzy correction
—◆— with fuzzy correction

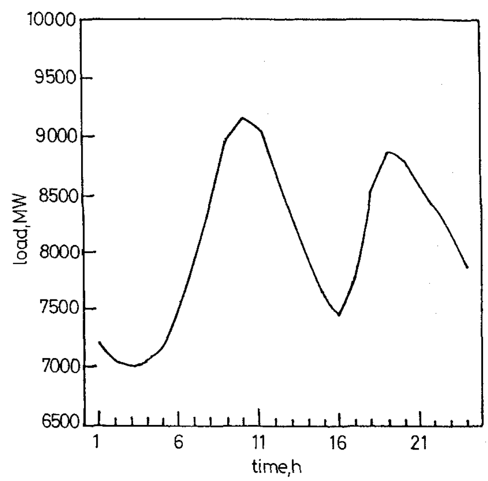

Fig. 21 *48h ahead hourly load forecast with temperature and humidity records for utility data*

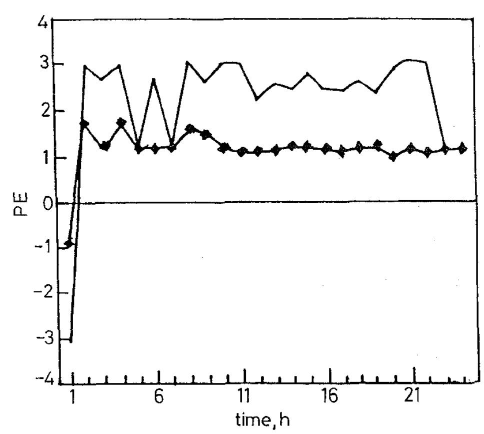

Fig. 22 *48h ahead hourly load forecasts with temperature and humidity records for utility data*
——— without fuzzy correction
—◆— with fuzzy correction

Figs. 25 and 26 show the 24h-ahead forecast errors for 15 February (winter day) using the nonlinear version of the fuzzy adaptive correction scheme along with FNN and FNN with fuzzy corrections. Significant accuracy in the hourly forecast is also obtained in this case using adaptive corrections. The above two days are nonspecial days and are chosen to illustrate the accuracy of the adaptive correction scheme.

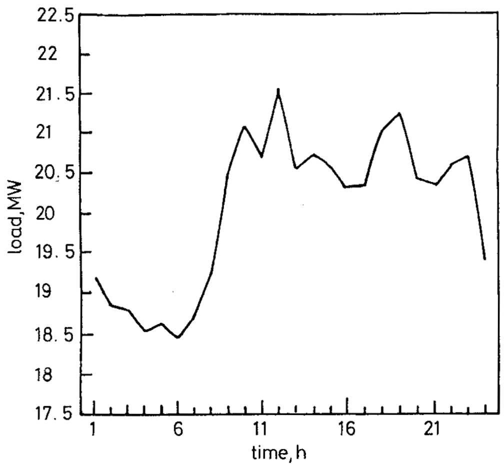

Fig. 23 *Hourly load forecast with adaptive corrections for 27 May (summer day)*

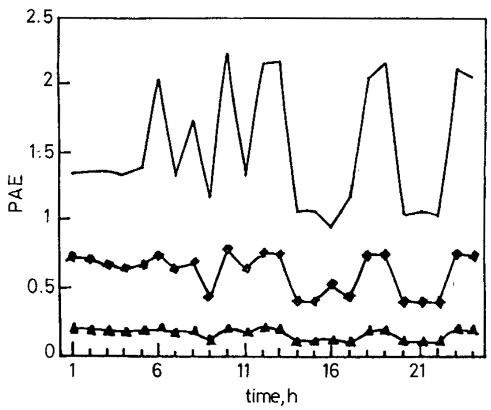

Fig. 24 *Hourly load forecast with adaptive corrections for 27 May (summer day)*
——— without fuzzy correction
—◆— with fuzzy correction
—▲— with nonlinear adaptive correction

7 Discussion

The proposed fuzzy-neural-network/fuzzy-expert-system approach is found to be very powerful and robust for short-term load predictions.

Although the results for two seasons of the year are presented in this paper for validating the effectiveness of this approach, extensive tests have been conducted for other seasons, Sundays, holidays and special days of the year. From the results presented in this paper, it is observed that significant accuracy can be achieved for 24h ahead hourly load forecasts and the PEs could be even less than 1%. Adaptive fuzzy load correction schemes enhance the accuracy of the predictions in most cases. However, if the lead time increases to 48h, the percentage error will be around 2%. The accuracy of load predictions using both adaptive corrections will be the highest with loads which do not show large hourly variations. However, significant accuracy can still be achieved with highly stochastic load variations, if a fuzzy basis function approach is used to arrive at adaptive corrections instead of the trial-and-error method presented in this paper.

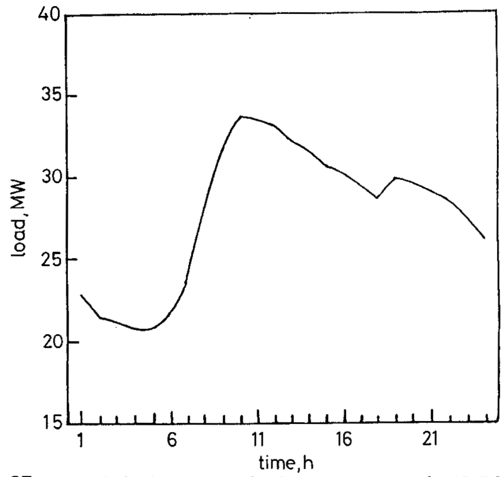

Fig. 25 *Hourly load forecast with adaptive corrections for 15 February (winter day)*

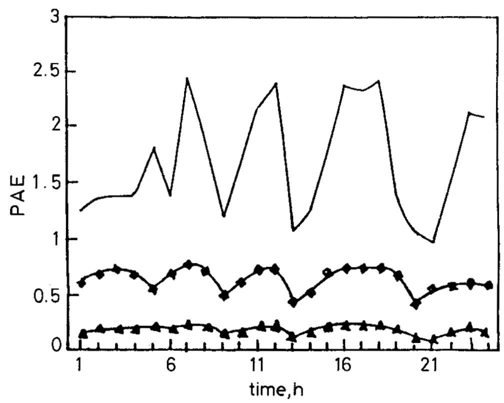

Fig. 26 *Hourly load forecast with adaptive corrections for 15 February (winter day)*
——— without fuzzy correction
—◆— with fuzzy correction
—▲— with nonlinear adaptive correction

The results for average load forecast are very encouraging and the maximum absolute percentage error is around 1.8. From the results for peak-load forecast presented in this paper, it is observed that except for 14 and 15 May, the PAE is less than 2 for most of the days. The mean absolute error is evaluated for this month and is found to be 2.824 with the FNN model only and 1.150 using fuzzy corrections.

The large errors for the above two days are probably due to other factors which have not been used in the training patterns. Although the results for 168 h-ahead load forecast using the above models have not been reported in this paper, the computations reveal that the mean absolute percentage error is around 2 and the maximum PE is 4.28 using the FNN model. With fuzzy corrections, the maximum PE is reduced to 2.16. This is quite comparable with the results for the ANN-based load forecasting technique presented by Karady *et al.* [12, 13].

Although the studies reported here have utilised a few simple examples and models, they are extremely valuable in identifying a promising hybrid forecasting methodology which can be investigated for larger number of inputs, more weather parameters, special load patterns, seasonal load changes, peak and total load forecasts one week ahead etc.

8 Conclusion

A new hybrid model integrating an artificial neural network and a fuzzy expert system is developed for 24 h ahead average and peak load forecasts and tested with historical load data. The hybrid model uses a fuzzy neural network for obtaining the initial forecast from the fuzzified input data and a fuzzy expert system generating the load correction to produce the final forecast. The simulation results of the proposed method using historical data show that the forecasting errors are less than 2% for both 24h ahead average and peak load predictions. The selection of training pattern presented in this paper does not classify the patterns to weekday and weekend day and thus the hybrid model is a promising approach for short-term load forecast for all days of the year. The paper also presents an adaptive fuzzy correction scheme to minimise the forecast errors more precisely and 24h ahead hourly forecasts using this scheme yield significant accuracy. The fuzzy neural network and fuzzy expert system approach has also been applied to one week ahead load forecast, and the results are found to be very promising.

9 Acknowledgment

The authors acknowledge funds from the US National Science Foundation (NSF grants INT-9209103 and INT-9117624) for undertaking this research.

10 References

1 RAHMAN, S., and BHATNAGAR, R.: 'An expert system based algorithm for short-term load forecast', *IEEE Trans.*, 1988, **PWRS-3**, (2), pp. 392–399
2 PARK, J., PARK, Y., and LEE, K.: 'Composite modelling for adaptive short-term load forecasting', IEEE Power Systems Engineering Summer Meeting, 1990
3 HUBELE, N.F., and CHENG, C.S.: 'Identification of seasonal short-term forecasting models using statistical decision functions', *IEEE Trans.*, 1990, **PWRS-5**, (1), pp. 40–45
4 PAPALEXOPOULOS, A.D., and HESTERBERG, T.C.: 'A regression-based approach to short-term system load forecasting', *IEEE Trans.*, 1990, **PWRS-5**, (4), pp. 1535–1547
5 CAMPO, R., and RUIZ, P.: 'Adaptive weather-sensitive short term load forecast', *IEEE Trans.*, 1987, **PWRS-2**, (3), pp. 592–600
6 VEMURI, S.V., HILL, E.F., and BALASUBRAMANIAN, R.: 'Load forecasting using stochastic models', IEEE Power Industry Computer Applications Conference, 1973, pp. 31–37
7 HAGAN, M.T., and BEHR, S.M.: 'The time series approach to short term load forecasting', *IEEE Trans.*, 1990, **PWRS-2**, (3), pp. 785–791
8 PARK, D., EL-SHARKAWI, M., MARKS, R., ATLAS, A., and DAMBORT, M.: 'Electric load forecasting using an artificial neural network', *IEEE Trans.*, 1991, **PWRS-6**, (2), pp. 442–449
9 EL-SHARKAWI, M.A., OH, S., MARKS, R.J., DAMDORG, M.J., and BRACE, C.H.: 'Short term electric load forecasting using an adaptively trained layered perceptron', Proceedings of the first international forum on Applications of Neural Networks to Power Systems, 1991, pp. 3–6
10 CONNER, J.T., ATLAS, L.E., and MARTIN, D.: 'Recurrent neural networks and load forecasting', Proceedings of the first international forum on Applications of Neural Networks to Power Systems, 1991, pp. 22–25
11 LEE, K.Y., CHA, Y.T., and KU, C.C.: 'A study of neural networks for short-term load forecasting', Proceedings of the first international forum on Applications of Neural Networks to Power Systems, 1991, pp. 26–30
12 PENG, T.M., HUBELE, N.F., and KARADY, G.G.: 'Advancement in the application of neural networks for short-term load forecasting', IEEE Power Engineering Society Summer Meeting, San Diego, 1991
13 PENG, T.M., HUBELE, N.F., and KARADY, G.G.: 'An adaptive neural network approach to one-week ahead load forecasting', *IEEE Trans.*, 1993, **PWRS-8**, (3), pp. 1195–1202
14 LAMBERT TORRES, G., VALIQUETTE, B., and MUKHADKAR, D.: 'Short-term feeder load forecasting using fuzzy logic', IFAC symposium on Power System and Power Plant Control, Korea, 1989, pp. 22–25
15 KIM, K.-H., PARK, D., and PARK, J.: 'A hybrid model of artificial neural network, and fuzzy expert system for short-term load forecast', Proceedings of ESAP-IV, Australia, 1993, pp. 164–168
16 SRINIVASAN, D., LIEW, A.C., and CHANG, C.S.: 'Fuzzy neural network for generation of daily load profiles', Proceedings of ISDEM Conference, Singapore, 1993, pp. 470–477
17 ZADEH, L.A.: 'Fuzzy sets', *Informat. Control*, 1965, **8**, pp. 338–358
18 PAL, S.K., and MAJUMDAR, D.D.: 'Fuzzy mathematical approach to pattern recognition' (Halstead Press, New York, 1986)
19 WANG, L.-X., and MENDEL, J.M.: 'Fuzzy basis functions, Universal approximation, and orthogonal least-square learning', *IEEE Trans.*, 1992, **NN-3**, (5), pp. 807–814

11 Appendix

Weight updating using error backpropagation for fuzzy neural network

The least-mean-square error in output vectors is minimised by using a gradient-descent algorithm by starting with any set of weights and repeatedly updating each weight by an amount

$$\Delta\omega_{ji}(n+1) = -\eta\frac{\partial E}{\partial\omega_{ji}} + \alpha\Delta\omega_{ji}(n) + \beta\Delta\omega_{ji}(n-1) \quad (18)$$

where

η = learning rate

α, β = momentum coefficients

n = iteration number

E = error cost function

Further, the learning rate η is adaptively varied as

$$\eta(n+1) = \eta(n) - \gamma\frac{\partial E}{\partial\eta} = \eta(n) - \gamma\left\{\frac{E(n)-E(n-1)}{\eta(n)-\eta(n-1)}\right\} \quad (19)$$

where γ determines the tuning of learning rate η.