Neuro-short-term load forecast of the power system in Kuwait

Abdullah S. Al-Fuhaid, Mohamed A. El-Sayed, and Magdi S. Mahmoud

Department of Electrical and Computer Engineering, Kuwait University, Safat, Kuwait

This paper is concerned with short-term load forecast of the electrical power system in Kuwait. It applies artificial neural networks (ANN's) to predict the half hour total system load. The input pattern incorporates the temperature and humidity effects. Simulation results have indicated that the developed forecasting approach is flexible and efficient. © *1997 by Elsevier Science Inc.*

Keywords: artificial neural networks, load forecast, power systems

1. Introduction

Load forecasting is a very useful tool for the operation and operation planning of power systems. It is required for unit commitment, energy transfer scheduling, and load dispatch. The different types of load forecasting can be classified according to the forecast period as follows[1]:

- Short-term forecast
- Medium-term forecast
- Long-term forecast

Table 1 lists the major functions and requirements for operation planning in each priod. This paper deals with short-term load forecasting and predicting the 1/2-hr total system load with particular emphasis on the power system of Kuwait.

Short-term power system load is mainly dependent on nonlinear combinations of variables that can be classified according to their dependence on weather, social, and seasonal factors. Temperature and humidity are the most independent weather factors. Social variables that affect the system load include the human duty activities such as work, school, and entertainment. Seasonal variations result from load growth and seasonal weather changes.

There is a comprehensive bibliography[2-4] on existing methodologies for short-term load forecasting, which can be categorized into:

- Time series approaches
- Regression models
- Knowledge-based (KB) approaches
- State space and Kalman filtering

The time series approach is a non-weather-sensitive approach that uses historical load data for extrapolation to future load conditions. This approach is time consuming and numerically unstable. Regression models consider that the variables to be predicted have a cause-effect relationship with one or more independent variables. The aim of these models is to discover the dependent variables by using the future evaluations of the independent ones. This model is computationally intensive and in some cases the linearization of weather terms is unjustified.

The knowledge-based (KB) approach develops load forecasting by emulating the experience and through processes of experienced system operators. Therefore these rule-based approaches attempt to convert the logic of power system operators into mathematical equations for forecasting. The problem with KB techniques is the derivation of the rules from on-the-job training. Sometimes conversion of the operators logic to equations may be impractical.

State-space and Kalman filtering models introduce the periodic component of load as a random process. These models require historical data bases for 3 to 10 yr to model the periodic load variation and to estimate the required initial as well as system dependent variables.

Artificial neural networks (ANN's) proved to be capable of finding internal representations of interdependencies within raw data not explicitly given or even known by human experts. This typical characteristic together with the simplicity of building and training ANN's and their very short response time encouraged their application to the task of load forecasting. Because of their inherent nonlinearity, ANN's are able to deal with the complex interactions between variables that affect load levels. There is no need for complex functional models to de-

Address reprint requests to Dr. Mahmoud at the Department of Electrical and Computer Engineering, Kuwait University, P.O. Box 5969, 13060 Safat, Kuwait.

Received 26 February 1996; revised 15 November 1996; accepted 5 December 1996.

Appl. Math. Modelling 1997, 21:215–219, April

655 Avenue of the Americas, New York, NY 10010

0307-904X/97/$17.00
PII S0307-904X(96)00165-5

Table 1

Forecast problem	Short-term	Medium-term	Long-term
Time horizon	1/4 hr–24 hr	1 day–few weeks	few months–years
Forecasted value	Load curves	Load curves	Energy required
Accuracy	Exact load curves	Error ≪ Capacity*	Exact energy†
Time step	1/4 hr–1 hr	1 hr	1 hr
Operation	Economic dispatch	Unit commitment	Reserve planning
Planning	Unit commitment	Reserve planning	Capacity expansion‡

* Capacity refers to the unit.
† Exact energy refers to the unit plus approximate load curves.
‡ Capacity expansion refers to the unit, plus approximate load curves, plus cost estimation.

scribe the relationships between the input variables and the load.

Recently the ANN's technology has been proposed as a substitute for statistical approaches for classification and forecasting problems.[5-7] The above-stated advantages of ANN's in statistical applications included robustness to probability distribution assumptions, the ability to classify in the presence of nonlinear relationships, and their ability to perform reasonably well with incomplete data bases. Comparison results between ANN's and statistical techniques indicated that neural nets offer an accurate alternative to other classical methods.[7,8]

In this paper ANN's are used for short-term load forecasting. The used ANN's are trained with the back-propagation algorithm. This algorithm is a gradient descent technique that is easily applied to networks whose neurons have smooth, monotonic, differentiable transfer functions such as sigmoid and hyperbolic tangents.

2. Artificial Neural Networks

ANN's represent an attempt to simulate the biological neural systems. They are based on methods of information processing understood to exist in the brain. They consist of a large number of simple processing units, massively interconnected and operating independently from each other. Such processing architectures have the capability to create their own subsymbolic representation, to learn, to memorize, and to recall associatively. The connection weights that determine the way in which neurons interact can be modified during the network operation, providing a high degree of adaptability. This adaptation process is the so-called learning phase.

The neural structure can be configured in different ways from the signal flow point of view. When the input and intermediate signals are always propagated forward, the system is called a static feed forward network. If either states or outputs are fed back, then this is a dynamic or recurrent network. A typical feedforward NN is illustrated in *Figure 1*. A three-layered network is shown in this figure, but in principle there could be more than one layer of internal representation units. Such internal layers are known as hidden layers, which lie between the input and output layers. The nodes of the network are nonlinear elements, and their transfer function is usually chosen to be the following sigmoidal function:

$$F_j = 1/\left[1 + \exp(-ay_j)\right] \tag{1}$$

where a is the slope of the sigmoid function.

The NN maps the input vector $\mathbf{X}$ to the output vector $\mathbf{y}_j$. The output is a sum of the weighted nonlinear terms given by

$$y_j = F_j T_j \qquad j = 1,2,\ldots,M \tag{2}$$

Where $\mathbf{y}_j$ is a single component of the output vector and represents the forecasted load and M is the number of output nodes. The T_j value is given by

$$T_j = \sum_{i=1}^{H} w_{ji} F_i(t_i) + v_j \tag{3}$$

where w_{ji} $(i = 1,2,\ldots,H;\ j = 1,2,\ldots,M)$ is the weighted factor between the ith input node and the jth output of hidden node; v_j is the threshold value at the output node and

$$t_i = \sum_{k=1}^{N} (w_{ik} + x_k) + v_i \tag{4}$$

where the output of the hidden nodes is equal to $F_i(t_i)$. The learning of the network is a recursive process, which adapts the connecting weights among the nodes; the sigmoid slope and the threshold values at each iteration. The following expression describes the supervised learning al-

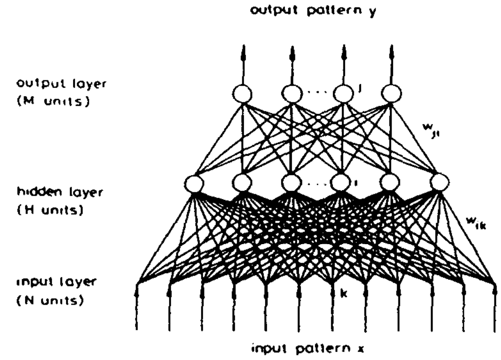

Figure 1.

gorithm by minimizing the square errors (E) defined as:

$$E = 1/2 \sum_{p=1}^{P_t} \sum_{j=1}^{M} [O_j^p - y_j^p]^2 \tag{5}$$

where p is the index of the pattern, which is equal to the number of cases used to train the network, y_j^p is the forecasted output of the network, and O_j^p is the actual (target) output of the training case.

The process of minimizing E by updating the network parameters is performed by the back propagation algorithm.[9,10] This algorithm is based on a steepest-descent technique, using the error function E. When $E < E_o$, where E_o is a given small convergence value, learning stops; when $E > E_o$ the parameter adjustment proceeds as follows:

$$\Delta w_{ji}^p = \eta(1 - y_j^p) y_j^p (O_j^p - y_j^p) F_p(T_i) \tag{6}$$

$$\Delta \upsilon_i^p = \eta(1 - y_j^p) y_j^p (O_j^p - y_j^p) \tag{7}$$

$$\Delta w_{ik}^p = \eta \, \delta_i^p x_k \tag{8}$$

$$\Delta \upsilon_i^p = \eta \delta_i^p \tag{9}$$

where

$$\delta_i^p = [1 - F^p(T_i)] \sum_{i=1}^{M} w_{ji}(1 - y_j^p) y_j^p (O_j^p - y_j^p) \tag{10}$$

The learning rate η governs the rate at which the weights are allowed to change at each iteration step. Higher η speed the convergence process, but it can result in overshooting or nonconvergence. Smaller η produces more reliable results at the expense of increased training time.

3. Case study

The electric utility in Kuwait mainly employs thermal steam turbines for the generation of power needed to cover the total demand. However, power plants also include some gas turbines that represent around 4% of total installed capacity and are usually used in emergencies and during the time of peak load. Power generation plants use different types of fossil fuels available in Kuwait such as natural gas, heavy oil, and crude oil, with a priority given to natural gas within the limits of its available quantity. After the erection of the first 30-Mw (4×7.5) steam power plant in 1954–1955, power plants' capacities have increased until they reach today 2,400 Mw (8×300) with the commissioning of Doha and Az-Zour power station.[11]

The rate of increase of electric peak load ranged about 32% in the 1950s, 26% in the 1960s, and 15% in the 1970s. Nevertheless this rate is increased during the last 3 yr in the range of 8–10%, whereas in most of the industrial countries the annual increase in electric load does not exceed 2–3%. The total load comprises mainly the seven regional loads, which represent 8.5–15% of the total load.

To examine the effectiveness of the proposed ANN's approach, load forecasting is performed using the data base for a complete previous year in order to consider the seasonal load variation. The maximum and minimum loads of that data base are found to be 4,350 Mw and 980 Mw, respectively. The average load factor of the system is approximately 0.8. The basic goal is to forecast 48 half hourly loads, which represent the 1-day ahead load curve.

4. ANN's configuration

There are few guidelines for designing the ANN's configuration. The only certain parameter is the number of neurons in the output layer. Since the load is measured over 48 half hourly intervals over a 24-hr period the outlayer constitutes 48 neurons corresponding to the daily load profile to be forecasted. On the other hand the number of neurons used for the input layer is determined by the amount and type of the available data. Generally the utilities save the historical load data along with the corresponding weather data. In addition to temperature, other weather parameters such as humidity also can be included in the set of inputs for ANN's, and its effect, especially on air-conditioning loads, can be taken into account.

For short-term load forecasting it is essential to use all of the available predicted weather parameters to improve the accuracy of the forecasting model. The three weather parameters whose forecasts are available in Kuwait are: the maximum and minimum temperatures and the maximum humidity. No forecasts are available for the minimum humidity, wind speed, or other weather parameters.

Therefore the 1-yr required data base is divided into five different sets, corresponding to the month and four day-types to be modelled: working day, thursday and friday (weekend), special event, and holiday.

4.1 Input pattern

It consists of

48	half hourly loads for the previous day	(i)
24	hourly temperature for the previous day	(i)
24	hourly humidity readings for the previous day	(i)
5	digits for the month and day-type	
2	maximum and minimum forecasted temperature readings for day	$(i + 1)$
1	maximum humidity forecasted for day	$(i + 1)$

corresponding to 104 data items.

4.2 Output pattern

It consists of 48 half hourly forecasted loads for day $(i + 1)$.

For a given forecast day the values for the input neurons are identified. These values are then compared with the historical output values for the same day-type, and the network is trained until the observed error reaches

Table 2

Test	Network configuration	Average training error
1	No. of inputs = 104 No. of outputs = 48 No. of hidden layers = 2 No. of hidden nodes (L1) = 35 No. of hidden nodes (L1) = 35 Learning rate = 0.2 No. of epochs = 4,000	0.00055

a minimum level. Finally the input values for the forecasted day are applied and a new forecast for half hourly daily loads is produced. Various tests were performed to identify the optimum number of hidden neurons and to determine a suitable value of the learning rate. Initially the ANN's were trained using 365 historical input/output training patterns from the previous year. The input and output values for each training pattern were 104 and 48, respectively. The final ANN structure, which minimized the training error, was 104-35-35-48, as shown in *Table 2.* For the selected configuration the learning rate was varied and the training errors were compared. Higher rates speed the convergence process but can result in overshooting or a large training error. Smaller rates produce more accurate results at the expense of increasing training time. For the test case the learning rate is increased from 0.2 to 0.6 and the training errors as well as the training times are compared, see *Figure 2.* A training rate of 0.2 was selected to reduce the training error with an acceptable computing time of 3 h on an IBM 386 machine.

Once the network was trained the estimated parameters such as weights and bias terms were fixed for the forecast of the following day. These parameters were updated by considering the actual data of the last forecasted day in order to keep the length of the total data base at 1 yr. This implies that daily updating was tested, which significantly improves the forecasting accuracy.

The difference between the actual and forecasted values defines the forecast error. The following forecast errors were used in testing ANN's:

$$e_i = (P_f - P_a)/P_a; \qquad |\bar{e}| = \frac{1}{n}\sum_{i=1}^{n} |P_f - P_a|/P_a$$

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(e_i - \bar{e})^2}; \tag{11}$$

$$e_M = \mathrm{Max}(e_i); \qquad i = 1, \ldots, n$$

5. Results

Simulation results of the proposed algorithm were based on a forecasting period of one complete year (1994) in order to consider the seasonal load variation. The developed approach using ANN's gave satisfactory forecasts for regular working days, however, it resulted in high forecast errors for holidays. The main reason for this is that Kuwait has 13 national and religious holidays per year. The forecast model for a holiday requires historical data extending further in the past. In *Table 3* a summary of the main results is presented. It contains a breakdown of the load forecast by month. The relevant information is: average absolute load forecast error (first column), the standard deviation (second column), and the corresponding maximum forecasting error (third column). For the year 1994 the average absolute error has been found to be 3.367%. The training of the ANN's improved the forecasting accuracy by 15–20% compared to the time series and regression techniques. In *Figure 3* a sample result of a working day (Sunday, 2 January 1994) was presented to illuminate the effectiveness of the developed approach.

From the simulation results it is evident that the maximum errors occur during extreme or rapidly changing weather conditions. This can be attributed to the strong correlation between air-conditioning load and temperature as well as humidity.

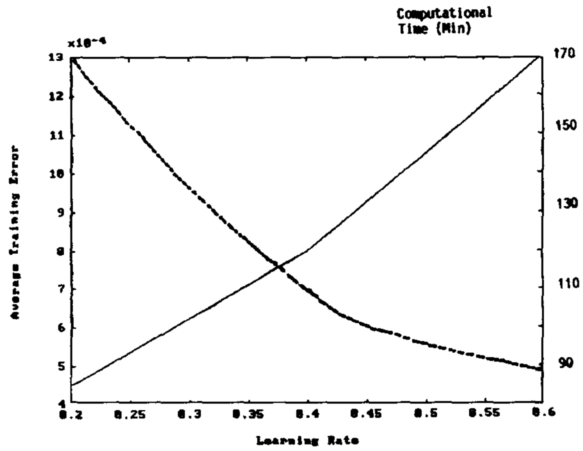

Figure 2.

Table 3. Forecast error for year 1994 (by month)

Month	ABS AVG Error (%)	Standard deviation	ABS Max Error (%)
January	3.492	1.015	5.019
February	4.788	1.156	6.287
March	2.741	0.368	3.500
April	3.870	1.029	6.606
May	4.068	2.016	8.502
June	2.438	0.973	4.240
July	2.625	0.776	3.799
August	2.440	0.690	3.863
September	1.841	0.398	2.667
October	3.545	1.547	6.267
November	4.844	2.482	9.257
December	3.715	1.600	6.090

Figure 3.

6. Conclusions

Artificial neural network-based approach has been applied to forecast the half hourly electric load of the power system of Kuwait. The ANN multilayer structure has been trained using the back-propagation techniques. This approach has been tested extensively under a wide variety of power system operating conditions, including rapidly changing weather. Furthermore the simulation results have indicated the effectiveness of the developed approach.

References

1. IEEE Committee Report. Load forecasting bibliography. *IEEE Trans. Power Apparat. Syst.* 1980, **PAS-99**, 53–58
2. Christianse, W. R. Short-term load forecasting using exponential smoothing. *IEEE Trans. Power Apparat. Syst.* 1971, **PAS-90**, 900–909
3. Srinivasan, K. and Pronovost R. Short-term load forecasting using multiple correlation models smoothing. *IEEE Trans. Power Apparat. Syst.* 1981, **PAS-100**, 3246–3253
4. Park, Y. and Park, J. K. An expert system for short-term load forecasting by fuzzy decision. *Proceedings of ESAPS*, 1989, pp. 244–250
5. Dillin, T. S. Artificial neural network applications to power systems and their relationship to symbolic methods. *Electric Power Energy Syst.* 1991, **13**, 66–72
6. Dillon, T. S., Sestito, S., and Leung. Short-term load forecasting using an adaptive neural network. *Electric Power Energy Syst.* 1991, **13**, 186–192
7. Marquez, L. T., Hill, Remus, W., and Worthley, R. Neural network models as an alternative to regression. *Proceedings of the 24th Hawaii International Conference on Systems Sciences*, Vol. 4, 1991, pp. 129–135
8. Brace, M. C., Schmidt, J., and Haldin, M. Comparison of the forecasting accuracy of neural networks with other established techniques. *Proceedings of IEEE's First International Forum on Application of Neural Networks to Power Systems*, Seattle, July, 23–26, 1991, pp. 31–35
9. Ho, K., Hsu, Y., and Yang, C. Short-term load forecasting using a multilayer neural network with an adaptive learning algorithm. *IEEE Trans. Power Syst.* 1992, **7**.
10. De Villiers, J. and Barnard, E. Back propagation neural nets with one and two hidden layers. *IEEE Trans. Neural Networks* 1993, **4**, 136–144
11. Ministry of Electricity and Water. *Statistical Year Book (Electrical Energy)*. State of Kuwait, 1994