ESL-PA-96/07-03-4

Measuring Retrofit Energy Savings Using Autoassociative Neural Networks

Kyung-Jin Jang
Associate Member ASHRAE

Eric B. Bartlett, Ph.D., P.E.

Ron M. Nelson, Ph.D., P.E.
Member ASHRAE

ABSTRACT

This paper presents methods and results for the prediction of energy conservation retrofits for heating, ventilating, and air-conditioning (HVAC) systems as part of the Great Building Energy Predictor Shootout II. The predictions are based on hourly whole-building electricity, motor control center electricity, lights and equipment electricity, cooling and heating energy use, and the accompanying weather data. An autoassociative neural network was used as a preprocessor to replace the missing data and a standard feed-forward artificial neural network was utilized to predict building energy consumption. Although the prediction results were acceptable for the verification of the sample networks, actual building energy prediction could have benefited by including the holiday and weekend information during the network training as well as information from previous time steps.

INTRODUCTION

This work utilizes a combination of a feed-forward neural network and an autoassociative network to predict building energy consumption for an engineering center in central Texas and a business building located in northeast Texas. Two data sets were provided from two different buildings for the analysis (Haberl et al. 1996). Both sets contained independent variables, weather data and time stamps, and the dependent variables, including whole-building energy use. The first data set was obtained from an engineering center located at a university in central Texas, and the second data set was from the business building at a university in northwest Texas. Each data set contained training data and test data. The training data sets were used to build the base model, and test data sets were used to predict the energy savings. The method discussed in this paper uses artificial neural networks (ANNs) to predict building energy consumption.

Artificial neural networks have been found to be useful in many applications where conventional approaches have had difficulty in coming to a solution. Inspired from neuroscience, artificial neural networks have received extensive attention ranging from theory to applications. Due to their ability to model complex nonlinear processes without prior process understanding, neural networks are a useful tool. Neural networks are able to learn from examples and generalize on novel data.

The ability of ANNs to extract information from the data set is especially crucial for finding solutions to missing data. This work draws on the capability of two neural networks in predicting building energy consumption. A standard feed-forward neural network is utilized to model the functional relationship between the independent variables and dependent variables, and an autoassociative network is used for missing sensor data replacement and will be discussed here.

BACKGROUND

Due to numerous reports that describe the standard back-propagation networks (Caudill 1989; Fahlman 1988; Hecht-Nielsen 1989; Hertz et al. 1991; Kung 1993; Lawrence 1993; Lippmann 1987; Rumelhart et al. 1986; Wassermann 1993; Werbos 1974; Widrow and Lehr 1990), further discussion on standard backpropagation networks is omitted. However, the autoassociative neural network (Kramer 1991, 1992) that is used for missing data replacement for solving building energy problems will be discussed.

Autoassociative Neural Networks

Autoassociative neural networks are feed-forward networks that represent an identity mapping between the network inputs and outputs, i.e., the network outputs are the same as the network inputs. The learning procedure is similar to that of the standard back-propagation network. The salient feature of autoassociative networks is the bottleneck layer between the inputs and

Kyung-Jin Jang is a research assistant, **Eric B. Bartlett** is an associate professor, and **Ron M. Nelson** is a professor in the Mechanical Engineering Department at Iowa State University, Ames.

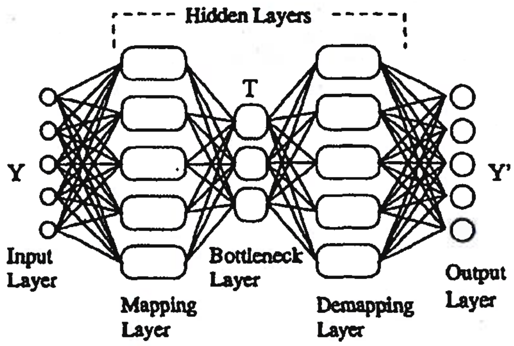

Figure 1 Architecture of the autoassociative network.

outputs. This bottleneck feature provides the compression of the information contained in the inputs and results in a correlation model of the input variables (Kramer 1991, 1992). The bottleneck layer prevents the network from learning the exact identity mapping, allowing for some transformation of the data. In turn, this allows the network to reproduce the input associations at its output using only the compressed information from the bottleneck layer. Least-squares training induces the network to model correlations and redundancies in the input data to reproduce the input data at the output with minimal distortion under the dimensional restriction of the bottleneck. Trained networks can then be utilized to detect sensor failures and estimate missing sensor data.

The architecture of the autoassociative network is shown in Figure 1. The hidden and output nodes perform the summing and transfer functions and the input nodes transfer the input values without any calculations. The independent variables that influence building energy consumption are input to the network. Hence, the input layer corresponds to the dimension of the independent variable vector, Y. The output layer results in a filtered version of the input vector, Y'. Theoretically, the autoassociative network with three hidden layers is sufficient to perform the identity mapping (Kramer 1991). The first hidden layer is called the *mapping layer*, the second hidden layer is the *bottleneck layer*, and the third hidden layer is the *demapping layer*. The dimension of the bottleneck layer must be the smallest in the entire network. The nodal transfer functions in the bottleneck layer can be linear, but the mapping and demapping layers must be nonlinear.

The autoassociative network can be viewed as a serial combination of two single hidden layer networks. The first combination network represents a nonlinear function, G, projecting the inputs to lower dimension space, feature, or factor space. The mapping is described in Equation 1.

$$T = G(Y). \tag{1}$$

T is the vector output of the bottleneck nodes and the input vector to the second combination network represented by the nonlinear function H, which reproduces the approximation of the inputs as described by Equation 2:

$$Y' = H(T). \tag{2}$$

The two equations represent a nonlinear generalization of the principal component analysis (Kramer 1991). Similar to linear principal component analysis, the loss of information is measured by the sum-of-squares difference between inputs and outputs summed over the training samples:

$$E = \sum_p \sum_i (Y_i^p - Y_i'^p)^2. \tag{3}$$

The indices p and i range over the number of training patterns and input dimensions. Using Equation 3 as the objective function to be minimized is analogous to training the network to produce the identity mapping. Minimization in E is equivalent to the minimal information loss in the principal component analysis (Kramer 1991). Since there is no set rule for determining the number of mapping and demapping nodes, Kramer suggests the following rule:

$$M_1 + M_2 \ll \frac{m(n-f)}{m+f+1}. \tag{4}$$

Here M_1, M_2, and n are the number of nodes in the mapping layer, the demapping layer, and number of training patterns, respectively. The parameters m and f represent the number of input variables and the number of bottleneck nodes. If there are too few mapping nodes, then the compressed input information is not very accurate. On the other hand, if there are too many mapping nodes, then the network has a tendency to learn the stochastic variations in the data instead of underlying functions. This tendency is referred to as *overfitting* (Kramer 1991). The autoassociative network extracts nonlinear principal components of the input variables based on the training data and outputs these principal components at the bottleneck layer. In addition, if the network is trained with sufficient training patterns and does not overfit the data, the correlation of the inputs is stored and the noise effects are rejected. Henceforth, once the network is trained, it is possible to use the network to perform missing data sensor replacement.

The autoassociative network allows the replacement of data from missing sensors with the values estimated from the remaining sensors. Because the autoassociative network performs an orthogonal mapping of the input space into a bottleneck layer of smaller dimensions in the least-squares sense, the problem of finding the value of the missing sensor data is given by

$$\min_{Y_i} |Y - Y'|. \tag{5}$$

In other words, the value of the missing value Y_i is the value that minimizes the difference between the input and output vectors. The range of the Y_i is bounded within [−1, 1]. The search is fast since the calculation is carried out only for the forward sweep of a trained network. During the search, other sensor values are kept constant and only the missing variables are adjusted. Kramer (1992) points out that as long as nonmissing variables outnumber the number of bottleneck nodes, sensor replacement is applicable to multiple sensors. This results in a multidimensional search rather than a single-dimensional search. The application of the autoassociative network is described in the "Methods" section.

PROBLEM DESCRIPTION

The purpose of this work is to evaluate the energy consumption of the engineering center and the business building before and after the installation of dual-duct variable-air-volume (VAV) air distribution systems with variable-frequency drives, new energy management, and new control systems (Haberl et al. 1996). Originally the buildings were equipped with constant-air-volume (CAV) air distribution systems.

In CAV systems, cooling or heating is accomplished by varying the zone supply air temperature and keeping the total volume of the supply air constant. The VAV systems accomplish cooling and heating by keeping the air temperature constant and varying the volume of the air supply. The VAV system provides cooling year-round, taking care of variations in all zone internal heat gains as well as skin solar gains (Lorsch 1993). Although the VAV air systems are harder to control than CAV systems, they are highly efficient, allow good room control, and are easily adaptable to economizer cycles. A potential downfall is the possibility of poor ventilation, particularly in zones with small loads. VAV systems are suitable for offices, classrooms, and many other commercial and institutional applications (McQuiston and Parker 1988).

The data affiliated with the engineering center (EC) and the business building (Buss) were provided in two files each—c.trn, c.tst and d.trn, d.tst, respectively. The data format consisted of the following:

- site number,
- time stamp,
- whole-building electricity (WBE),
- motor control center electricity (MCC),
- lights and equipment electricity (LTEQ),
- whole-building chilled water (CWE),
- whole-building hot water (HWE),
- ambient temperature (T),
- ambient relative humidity (RH),
- global horizontal solar radiation (SOL), and
- wind speed (Wspeed).

The units pertaining to electricity are in kilowatt-hours per hour (kWh/h), the water usage is in million British thermal units per hour (MBtu/h), temperature is in degrees Fahrenheit (°F), relative humidity is in percent (%), solar radiation is in watts per square meter (W/m^2), and wind speed is in miles per hour (mph). The time durations of the recorded data are listed in Table 1.

The two files, c.trn and d.trn, contain the dependent and independent variables recorded with existing equipment. The installation of the variable-frequency drives and controls were completed in November 1990 for the engineering center and in July 1991 for the business building. The dependent variables are the building energy consumptions and the independent variables are the time stamps and weather data. Building energy data for pre-retrofitted buildings are not available after the installation of the new variable-frequency drives and controls for the buildings. The contest sponsors have also deliberately removed portions of the dependent and independent variables from the training set to assess the accuracy of the contestants' predictions.

TABLE 1 Time Duration for the Data Sets

Building	Data File	Time Duration
ZEC	c.trn	1/1/90 0:00 to 11/27/90 23:00
	c.tst	11/28/90 0:00 to 12/31/92 23:00
Buss	d.trn	12/22/90 0:00 to 7/12/91 23:00
	d.tst	7/13/91 0:00 to 12/31/92 23:00

The first objective is to build the baseline models for both buildings by replacing the removed data and predicting the energy consumed by the buildings for the entire pre-retrofit training periods, estimated to be between January 1, 1990, and November 27, 1990, for the engineering center and December 22, 1990, through July 12, 1991, for the business building. The second objective is to estimate the energy savings for both buildings using the post-retrofit data, files c.tst and d.tst. The major missing periods are shown in Tables 2 and 3. There are more sparsely assigned missing data for both data files that are not specified in the tables. The final objective is to predict the building energy consumptions for the entire testing periods, including the post-retrofit time frame, using the base model.

TABLE 2 Major Missing Data for the Engineering Center File C.TRN

Variable	Time Duration
LTEQ and MCC	January 1990 to February 1990
CWE	April 1990 to November 1991
HWE	June 1991 to November 1991

TABLE 3 Major Missing Data for Business Building File D.TRN

Variable	Time Duration
CWE	December 1990 to March 1991 and few weeks in December 1992
HWE	December 1990 to March 1991 and few weeks in December 1992

In predicting the energy use, this work utilizes autoassociative neural networks to address the missing sensor data problem and a standard feed-forward neural network to estimate the buildings' energy consumption. This autoassociative neural network method had not been used in the previous predictor shootout competition.

METHODS

The neural network developed in this work utilizes the hyperbolic tangent as the transfer function with the scaled conjugate gradient methods (Moller 1993) for the training algorithm. For the hyperbolic tangent transfer function, input data need to be normalized between the ranges of [−1, 1]. The base model is developed with two neural networks—the autoassociative network and the standard network. Hence, the overall setup for the training involves feeding inputs of the data into the autoassociative network and its outputs into the standard feed-forward neural network (see Figure 2). Although MCC and LTEQ are dependent variables, they were included as inputs to the autoassociative neural network since many of these values were missing and could then be predicted by the autoassociative neural network.

Initially, from the "*.trn" sets, the missing and the nonmissing data sets were segregated. The complete sets, or the nonmissing data, consist of the nine input variables—month, day, hour, motor control center electricity (MCC), lights and equipment electricity (LTEQ), ambient temperature (Tamb), relative humidity (RH), solar radiation (sol), and wind speed (wspeed)—and three output variables—whole-building electricity (WBE), whole-building chilled water (CWE), and whole-building hot water (HWE). The training period for the c.trn data set of the ZEC ranged from April 1990 to November 1990. The 100 randomly selected test patterns selected from c.trn also follow the same period. The d.trn training and test period for the business building ranged from March 1991 to July 1991.

A program was written to normalize the data between [−0.9, 0.9] with 0 mean and unit variance to spread the values in their equal importance (for example, when squaring or taking differences of two values, a small value gets ignored compared to a

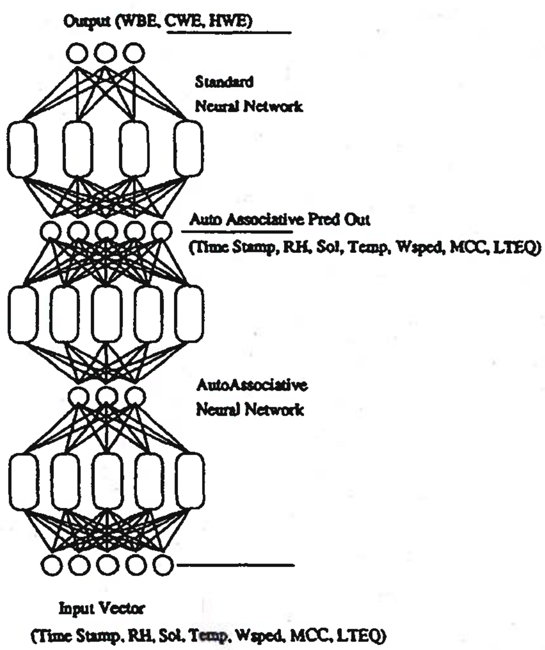

Figure 2 Two feed-forward neural networks used for analysis.

larger value). This is done using Equation 6. Here, the n is the number of patterns.

$$x_{in} = \frac{x_i - \bar{x}}{\sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}}. \tag{6}$$

The normalization range of [−0.9, 0.9] was applied to reduce the saturation limit of the node output. In contrast to the normalization range [0, 1] or [0.1, 0.9], the hyperbolic node transfer function takes advantage of twice the nodal output range of [−1, 1] or [−0.9, 0.9], eliminating the sharp threshold behavior that results in continuous transition of the activation energy. The hyperbolic tangent activation function is chosen for the convenient derivative of the function itself, as shown by Equation 7. The values are then normalized in the range of −0.9, 0.9 according to Equation 8:

$$g(h) = \tanh(h)$$
$$g'(h) = 1 - g(h)^2 \tag{7}$$

$$x_i = -0.9 + \frac{1.8(x_i - x_{min})}{x_{max} - x_{min}}. \tag{8}$$

The x_{min} and x_{max} are the low and high values pertaining to a particular variable of interest. To normalize the test pattern, the mean, variance, and the minimum and maximum values of the training set are utilized so that the normalized test patterns are of the same training data range. This will enforce the proper comparison of the generalization capability of the novel test patterns. The program is written to normalize the data directly from the raw data into the training program. When the training program is terminated, the program returns the unnormalized version of the output.

Both networks, the autoassociative network and the standard multilayered perceptron network, are trained with the nonmissing data patterns. Table 4 lists the training and recall

TABLE 4 Training Results

File	Network	Architecture	Train RMS	Recall RMS
c.trn	Autoassociate	9x12x4x12x9	0.0651	0.0635
	Std Back Prop	9x7x3	0.0647	0.0628
d.trn	Autoassociate	9x11x4x12x9	0.0767	0.1129
	Std Back Prop	9x9x3	0.0759	0.0815

results for the networks. Recall is performed using the novel test patterns to see how well the trained network is able to predict and forecast the output variables. Here, the novel test patterns are 100 randomly selected patterns from the training sets. For the c.trn, the recall RMS error was lower than the training error. This is due to the high correlation of the test patterns with those of the training examples. The RMS errors are calculated using the following equation:

$$\text{RMS error} = \sqrt{\frac{1}{nN}\sum_{\mu}^{N}\sum_{i}^{n}(\varsigma_i^{\mu} - O_i^{\mu})} \tag{9}$$

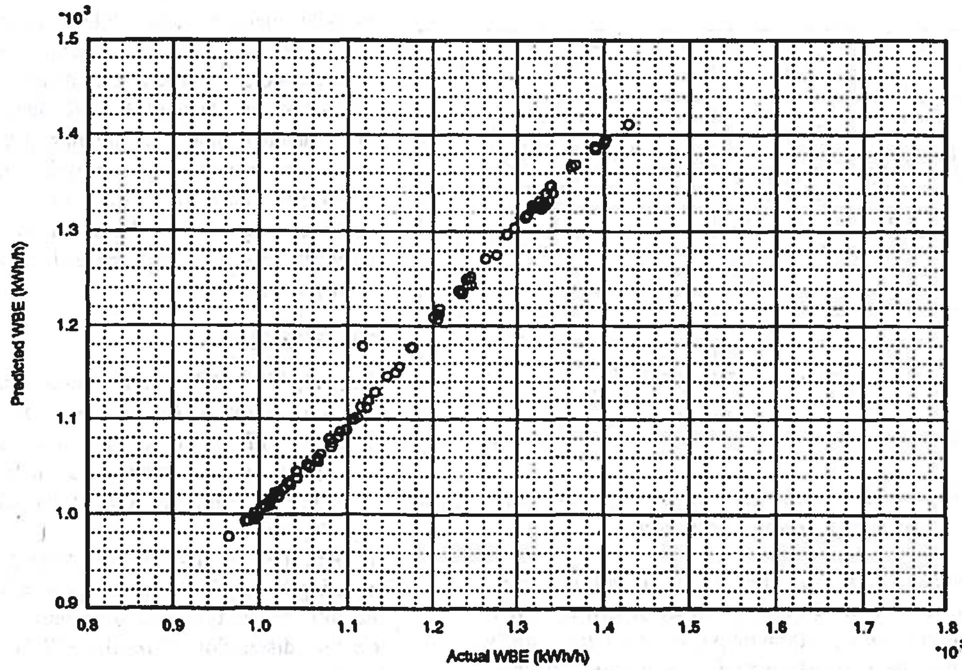

Figure 3 Test prediction of the WBE for c.trn.

where N is the number of patterns in the training set, n is the number of nodes in the output layer, ζ is the correct output of the ith output node for the μ training pattern, and O is the network output for the same node and pattern. This RMS error is a function of the existing set of weights. Once both networks have been trained using the nonmissing data, the missing independent variables can be predicted using the autoassociative network and the missing dependent variables can be predicted using the standard network.

DISCUSSION OF RESULTS

The test predictions using the 100 novel test patterns for the WBE (in kWh/h) are illustrated in Figures 3 and 4 for c.trn and d.trn. The resulting statistical correlation coefficients for the outputs are shown in Table 5. The correlation coefficients show good energy predictions for both buildings for the 100 novel test patterns..

TABLE 5 Correlation Results for the 100 Random Test Sets

File	WBE	CWE	HWE	MCC	LTEQ
c.trn	0.9961	0.9114	0.9614	0.5138	0.9428
d.trn	0.9706	0.8776	0.9371	0.9888	0.9038

The trained networks were then used in conjunction with the replacement of the missing data. Here the unknown input variables are presented to the trained autoassociative network along with the known input variables. Since the autoassociative network implements an orthogonal mapping of the input space into a smaller dimension as described in the "Autoassociative Neural Networks" section, the unknown inputs are obtained by iterating within the trained autoassociative network until the identity mapping is satisfied within a specified tolerance. The iteration in the autoassociative network is fast. Figures 5 and 6 show the autoassociative neural network predictions for MCC and LTEQ data for all 70 patterns in which they were missing. This demonstrates the ability of the autoassociative neural network to predict missing data. The autoassociative network then provides a complete set of inputs to the standard back-propagation network for the final prediction of the building energy outputs. The results of the final predictions are given in the overview paper (Haberl et al. 1996).

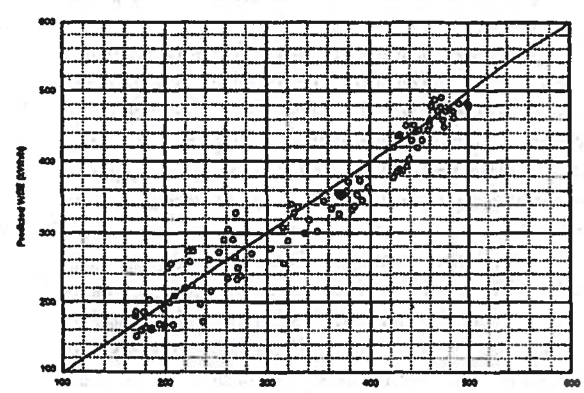

Figure 4 Test prediction of the WBE for d.trn.

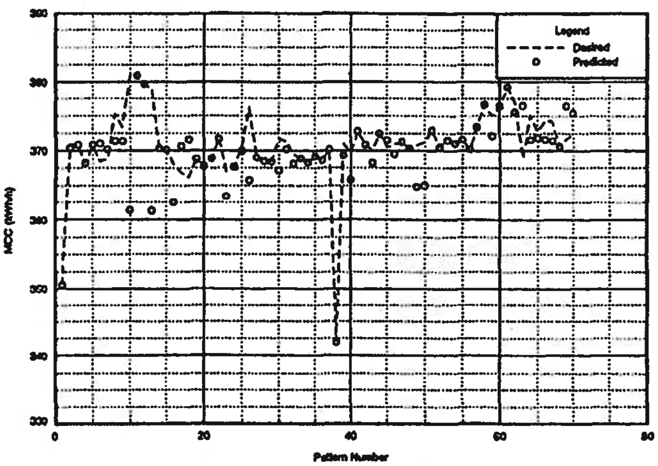

Figure 5 Test results of the associative network using missing sensors, motor control electricity.

The networks used in this study are trained with a synchronous hourly pattern, where the input of the *i*th pattern input variables correspond to the *i*th pattern output variables. It is usually advisable to train the networks including additional asynchronous information where input patterns from previous time steps are used. This would help to account for unknown occupancy, building mass, and other time-dependent effects. The addition of the asynchronous hourly schedule training would enable the network to make better predictions.

CONCLUSIONS

The main emphasis of this work was the application of an autoassociative neural network (Kramer 1992) to replace the missing sensor values for a given set of data. A standard back-propagation neural network was trained to predict the future building energy data to determine the energy savings. This approach provided good predictions for energy use. (See the overview paper for the comparison of predicted and actual results [Haberl et al. 1996].) The prediction accuracy could be further improved by using weekday/weekend/holiday information and by using information from previous time steps.

The autoassociative network effectively predicted missing sensor information. The combination of the standard feed-forward neural network using complete independent variable information supplied by the autoassociative network worked well for the prediction of building energy consumption. One of the distinctive features of an autoassociative network is that it resolves missing sensor information using nonlinear principal component analysis (Kramer 1991). This is done by training a neural network containing mapping, bottleneck, and demapping layers to achieve a general nonlinear fitting property and eliminating nonlinear redundant correlations in the data. Analogous to linear principal component analysis, the nonlinear principal component analysis performs mapping of multidimensional data into lower dimensions with minimal loss of information. This information is then retrieved from the trained autoassociative neural network to replace the missing sensor data for predicting building energy consumption. Although similar but differing in the building of the predictive model, Bayesian nonlinear modeling by MacKay (1994) has shown the relevance of the elimination of the correlated data. Overall, this neural network shows great potential for system modeling and prediction of building energy consumption.

REFERENCES

Caudill, M. 1989. Using neural nets: Representing knowledge. *AI Expert*, December, pp. 34-41.

Fahlman, S.F. 1988. An empirical study of learning speed in back propagation networks. *Carnegie Mellon University Report No. CMU-CS-88-162*, September.

Haberl, J.S., S. Thamilseran, and J. Kreider. 1996. Predicting hourly building energy use: The great energy predictor shootout II: Measuring retrofit savings—Overview and discussion of results. *ASHRAE Transactions* 102(2).

Hecht-Nielsen, R. 1989. Theory of the backpropagation neural network. *Proceedings of the International Joint Conference on Neural Networks* 1: 593-605.

Hertz, J., A. Krogh, and R.G. Palmer. 1991. *Introduction to the theory of neural computation*. Reading, Mass.: Addison-Wesley Publishing Co.

Kramer, M.A. 1991. Nonlinear principal component analysis using autoassociative neural networks. *AIChE Journal* 37(2): 233-243.

Kramer, M.A. 1992. Autoassociative neural networks. *Computers Chemical Engineering* 16(4): 313-328.

Kung, S.Y. 1993. *Digital neural networks*. Englewood Cliffs, N.J.: PTR Prentice Hall.

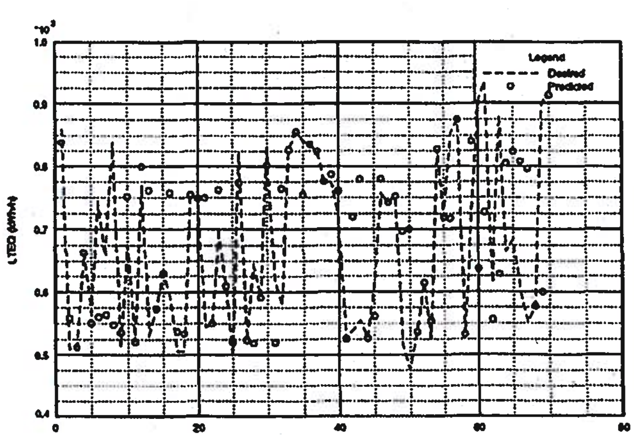

Figure 6 Test results of the associative network using missing sensors, light, and equipment electricity.

Lawrence, J. 1993. *Introduction to neural networks design, theory, and applications*. Nevada City, Calif.: California Scientific Software Press.

Lippmann, R.P. 1987. An Introduction to computing with neural nets. *IEEE Acoustics Speech and Signal Processing Magazine* 4: 4-22.

Lorsch, H.G. 1993. *Air-conditioning systems design manual*. Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

MacKay, D.J.C. 1994. Bayesian nonlinear modeling for the prediction competition. *ASHRAE Transactions* 100(2).

McQuiston, F.C., and J.D. Parker. 1988. *Heating, ventilating and air conditioning analysis and design*, 3d ed. New York: John Wiley & Sons Inc.

Moller, M.F. 1993. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks* 6: 525-533.

Rumelhart, E.D., G.E. Hinton, and R.J. Williams. 1986. Learning internal representations by error propagation. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1, pp. 318-362. Cambridge, Mass.: The MIT Press.

Wassermann, P.D. 1993. *Advanced methods in neural computing*, chapter 6, "Backpropagation and Beyond," pp. 97-117. New York: Van Nostrand Reinhold.

Werbos, P.J. 1974. Beyond regression: New tools for prediction and analysis in the behavioral sciences. Ph.D. Thesis. Cambridge, Mass.: Harvard University.

Windrow, B., and M.A. Lehr. 1990. 30 years of adaptive neural networks: Perception, madaline, and backpropagation. *Proceedings of the IEEE* 78: 1415-1441.