Estimating Evapotranspiration using Artificial Neural Network

M. Kumar[1]; N. S. Raghuwanshi[2]; R. Singh[3]; W. W. Wallender[4]; and W. O. Pruitt[5]

Abstract: This study investigates the utility of artificial neural networks (ANNs) for estimation of daily grass reference crop evapotranspiration (ETo) and compares the performance of ANNs with the conventional method (Penman–Monteith) used to estimate ETo. Several issues associated with the use of ANNs are examined, including different learning methods, number of processing elements in the hidden layer(s), and the number of hidden layers. Three learning methods, namely, the standard back-propagation with learning rates of 0.2 and 0.8, and backpropagation with momentum were considered. The best ANN architecture for estimation of daily ETo was obtained for two different data sets (Sets 1 and 2) for Davis, Calif. Using data of Set 1, the networks were trained with daily climatic data (solar radiation, maximum and minimum temperature, maximum and minimum relative humidity, and wind speed) as input and the Penman–Monteith (PM) estimated ETo as output. The best ANN architecture was selected on the basis of weighted standard error of estimate (WSEE) and minimal ANN architecture. The ANN architecture of 6-7-1, (six, seven, and one neuron(s) in the input, hidden, and output layers, respectively) gave the minimum WSEE (less than 0.3 mm/day) for all learning methods. This value was lower than the WSEE (0.74 mm/day) between the PM method and lysimeter measured ETo as reported by Jensen et al. in 1990. Similarly, ANNs were trained, validated, and tested using the lysimeter measured ETo and corresponding climatic data (Set 2). Again, all learning methods gave less WSEE (less than 0.60 mm/day) as compared to the PM method (0.97 mm/day). Based on these results, it can be concluded that the ANN can predict ETo better than the conventional method (PM) for Davis.

DOI: 10.1061/(ASCE)0733-9437(2002)128:4(224)

CE Database keywords: Neural networks; Evapotranspiration; Wave propagation.

Introduction

Evapotranspiration (ET) is one of the major components of the hydrologic cycle and its accurate estimation is of paramount importance for many studies such as hydrologic water balance, irrigation system design and management, crop yield simulation, and water resources planning and management. A common practice for estimating ET from a well-watered agricultural crop is to first estimate reference crop ET, i.e., grass reference ET (ETo) or alfalfa reference ET (ETr), from a standard surface and to then apply an appropriate empirical crop coefficient, which accounts for the difference between the standard surface and crop ET. Doorenbos and Pruitt (1977) defined reference crop evapotranspiration rate as ‘‘the rate of evapotranspiration from an extensive surface of 0.08–0.15 m tall, green grass cover of uniform height, actively growing, completely shading the ground, and not short of water.’’ In the past 5 decades, several studies have focused on the development of accurate ET estimation methods and improving the performance of the existing methods due to wide application of ET data. The research efforts are still continuing in this direction.

Evapotranspiration can be either measured with a lysimeter or water balance approach, or estimated from climatological data. However, it is not always possible to measure ET with a lysimeter because it is a time-consuming method and needs precisely and carefully planned experiments. Thus, indirect methods based on climatological data are used for ETo estimation. These methods vary from empirical relationships to complex methods based on physical processes such as the Penman (1948) combination method. The combination approach links evaporation dynamics with the flux of net radiation and aerodynamic transport characteristics of a natural surface. Based on the observations that latent heat transfer in plant stems is influenced not only by these abiotic factors, Monteith (1965) introduced a surface conductance term that accounted for the response of leaf stomata to its hydrologic environment. This modified form of the Penman equation is widely known as the Penman–Monteith evapotranspiration model.

Many scientists have studied the reliability of the Penman-Monteith (PM) method for estimating ETo (McNaughton and Jarvis 1984; Allen 1986; Allen et al. 1989; De Souza and Yoder 1994; Chiew et al. 1995). Jensen et al. (1990) analyzed the performance of 20 different methods against lysimeter measured ET for 11 stations located in different climatic zones around the world. The PM method ranked as the best method for all climatic

[1]Research Scholar, Dept. of Agricultural and Food Engineering, Indian Institute of Technology, Kharagpur WB 721 302, India.

[2]Associate Professor, Dept. of Agricultural and Food Engineering, Indian Institute of Technology, Kharagpur WB 721 302, India.

[3]Associate Professor, Dept. of Agricultural and Food Engineering, Indian Institute of Technology, Kharagpur WB 721 302, India.

[4]Professor, Dept. of Biological and Agricultural Engineering and Dept. of Hydrologic Science, Univ. of California, Davis, CA 95616.

[5]Emeritus Irrigation Engineer, Dept. of Hydrologic Science, Univ. of California, Davis, CA 95616.

Note. Discussion open until January 1, 2003. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on March 19, 2001; approved on December 13, 2001. This paper is part of the ***Journal of Irrigation and Drainage Engineering***, Vol. 128, No. 4, August 1, 2002. ©ASCE, ISSN 0733-9437/2002/4-224–233/$8.00+$.50 per page.

conditions. However, the ranking of other methods varied, depending on their local calibration and conditions.

Evapotranspiration is a complex and nonlinear phenomenon because it depends on several interacting climatological factors, such as temperature, humidity, wind speed, radiation, type, and growth stage of the crop, etc. Artificial neural networks (ANNs) are effective tools to model nonlinear systems. A neural network model is a mathematical construct whose architecture is essentially analogous to the human brain. Basically, the highly interconnected processing elements (PEs) arranged in layers are similar to the arrangement of neurons in the brain.

In the past decade, considerable attention has been focused on the application of ANNs in diverse fields including system modeling, fault diagnosis and control, pattern recognition, financial forecasting, and hydrology. Studies on ANN application in the area of hydrology include rainfall-runoff modeling (French et al. 1992; Minnes and Hall 1996); river stage forecasting (Thirumalaian and Deo 1998, Campolo et al. 1999); reservoir operation (Jain et al. 1999); land drainage design (Shukla et al. 1996; Yang et al. 1996, 1998); pesticides concentration in soil (Yang et al. 1997; Goh 1999; Tansel et al. 1999); aquifer parameter estimation (Srinivasa 1998); and optimization problems (Rogers and Dowla 1994; Wen and Lee 1998). Some of the studies (Zealand et al. 1999; Yang et al. 1996) have also shown that ANN is more accurate than conventional methods. It is evident from the literature that no study has been carried out to utilize the input–output mapping capability of ANN in the prediction of ET.

Based on the capabilities that ANNs have to simulate nonlinearity among the interacting factors in the system, the present study describes the utilization of the input–output mapping capabilities of the ANN in ETo prediction. Below, the ANN-estimated ETo values are compared with both the PM estimated and lysimeter measured ETo values. Furthermore, the performance of three ANN training methods is evaluated and the minimal ANN architecture for predicting ETo is determined.

Artificial Neural Networks

An ANN consists of input, hidden, and output layers and each layer includes an array of processing elements. A typical neural network is fully connected, which means that there is a connection between each of the neurons in any given layer with each of the neurons in the next layer. A processing element is a model whose components are analogous to the components of actual neuron. The array of input parameters is stored in the input layer and each input variable is represented by a neuron. Each of these inputs is modified by a weight whose function is analogous to that of the synaptic junction in a biological neuron. The processing element consists of two parts. The first part simply aggregates the weighted inputs; the second part is essentially a nonlinear filter, usually called the transfer function or activation function. The activation function squashes or limits the values of the output of an artificial neuron to values between two asymptotes. The sigmoidal function is the most commonly used activation function. It is a continuous function that varies gradually between two asymptotic values, typically 0 and 1 or -1 and $+1$.

Artificial Neural Network Learning

Learning is normally accomplished through an adaptive procedure or algorithm that incrementally adjusts weights of the connections so as to improve a predefined performance measure. The neural network is presented with the data patterns consisting of the input values as well as expected (or output) values. The objective is to minimize the difference between the predicted output values and expected output values using an algorithm (e.g., back-propagation algorithm). Initially, because of the random weights assigned randomly to the connections, the difference between the predicted and desired output values can be large. Learning therefore involves iteratively adjusting the connection weights to minimize these differences.

Backpropagation Training Method

Training of an artificial neural network involves two phases. In the first phase or forward pass the input signals propagate from the network input to the output. In the second phase or reverse pass, the calculated error signals propagate backward through the network, where they are used to adjust the weights. The calculation of the output is carried out, layer by layer, in the forward direction. The output of one layer is the input to the next layer. In the reverse pass, the weights of the output neuron layer are adjusted first since the target value of each output neuron is available to guide the adjustment of the associated weights. The weights in the output and hidden layer neurons can be calculated using Eqs. (1) and (2), respectively (Tsoukalas and Uhrig 1996),

$$w(N+1)=w(N)-\eta\delta\phi \tag{1}$$

$$w(N+1)=w(N)+\eta x\sum_{q=1}^{r}\delta_q \tag{2}$$

where w = weight; N = number of iteration; x = input value; η = learning rate; ϕ = output; and δ is defined as $2\varepsilon_q\partial\phi/\partial I$, I being the sum of the weighted inputs, q = neuron index of the output layer, and ε_q = error signal.

The above training method is known as the standard back-propagation training method. Since back-propagation employs a form of gradient descent, it is assumed that the error surface slope is always negative and hence, constantly adjusting weights toward minimum. However, error surfaces often involve complex, high dimensional space that is highly convoluted with hills, valleys, and folds. It is very easy for the training process to get trapped in a local minimum.

The problem of the local minima can be avoided by adding a momentum term to the weight change, to permit larger learning rates. The change of weight is then computed as follows:

$$\Delta w(N+1)=-\eta\delta\phi+\mu\Delta w(N) \tag{3}$$

where μ = momentum coefficient and $\Delta w(N+1)$ = change of weight during N to $N+1$ learning cycles. Thus, the new value of weight becomes equal to the previous value of the weight plus the weight change, which includes the momentum term. This training method is known as back-propagation with momentum.

Materials and Method

Description of Data and Grass Reference Crop Evapotranspiration Estimation

For the purpose of this study, daily climatic data of minimum and maximum temperature, minimum and maximum relative humidity, wind speed, and solar radiation for Davis California Irrigation Management Information System station were collected for the

Table 1. Artificial Neural Networks (ANNs) Training and Validation Errors using Data Sets 1a and 1b, Respectively, for Single Hidden Layer Architectures

	Std. back-propagation LR=0.2		Std. back-propagation LR=0.8		Back-propagation momentum	
Network architecture	Training (Set 1a) MSE (10^{-4})	Validation (Set 1b) MSE (10^{-4})	Training (Set 1a) MSE (10^{-4})	Validation (Set 1b) MSE (10^{-4})	Training (Set 1a) MSE (10^{-4})	Validation (Set 1b) MSE (10^{-4})
6-7-105k	2.6	1.5	1.4	8.4	1.5	1.1
6-7-11k	2.3	1.5	1.1	3.8	1.1	1.1
6-7-12k	2.0	1.4	0.8	3.4	1.1	1.2
6-7-13k	1.4	0.9	0.8	4.7	1.0	1.9
6-7-14k	1.8	1.4	1.1	1.8	1.0	1.4
6-7-15k	1.2	0.9	0.9	0.8	1.0	0.9
6-8-105k	2.5	1.6	2.1	1.4	1.6	1.8
6-8-11k	2.2	1.4	1.5	1.1	1.2	1.0
6-8-12k	1.7	1.2	1.3	0.9	1.1	1.2
6-8-13k	1.5	1.1	1.0	0.9	1.0	1.2
6-8-14k	1.2	1.0	0.9	0.9	1.0	0.8
6-8-15k	1.2	1.6	0.9	0.8	0.8	0.8
6-9-105k	2.4	1.6	4.3	2.4	1.2	1.4
6-9-11k	2.1	1.5	2.0	1.3	1.1	0.8
6-9-12k	2.0	1.4	1.1	1.3	1.0	0.9
6-9-13k	1.5	1.2	1.1	0.9	0.9	0.8
6-9-14k	1.2	1.0	1.0	0.9	1.1	1.0
6-9-15k	1.1	1.0	1.0	0.9	0.8	1.5
6-10-105k	2.4	1.7	2.0	1.3	1.3	0.9
6-10-11k	2.2	1.5	1.3	1.7	1.1	0.9
6-10-12k	2.0	1.4	1.1	1.0	1.0	1.5
6-10-13k	1.4	1.1	1.0	0.9	0.9	0.8
6-10-14k	1.2	1.0	1.0	0.9	0.9	0.9
6-10-15k	1.1	1.1	1.0	0.8	0.8	0.8
6-11-105k	2.4	1.7	1.7	1.4	1.4	1.1
6-11-11k	2.1	1.4	1.3	1.3	1.2	1.5
6-11-12k	1.9	1.4	1.2	0.9	0.9	0.9
6-11-13k	1.4	1.1	1.1	1.5	1.0	0.8
6-11-14k	1.4	1.0	0.9	1.2	0.8	0.7
6-11-15k	1.3	1.0	0.9	1.1	0.8	0.8
6-12-105k	2.4	1.6	2.0	1.3	1.5	1.1
6-12-11k	2.2	1.6	1.3	1.1	1.1	1.1
6-12-12k	2.1	1.5	1.1	0.9	1.0	0.9
6-12-13k	1.4	1.1	1.1	1.0	1.0	0.8
6-12-14k	1.2	0.9	1.0	0.9	0.9	0.8
6-12-15k	1.1	1.0	0.9	0.9	0.8	0.9

Note: LR=learning rate.

period of January 1, 1990 to June 30, 2000 (Set 1). Daily ETo values were estimated using the PM method because the lysimeter measured ETo values were not available for the period. The PM estimated ETo values were considered as standard and used for training and testing of different architectures of ANN. The PM method is considered as standard because it ranked first for both humid and arid regions (Jensen et al. 1990).

To compare the ANN predicted ETo with the PM estimated ETo, daily lysimeter measured grass evapotranspiration along with climatic data [minimum and maximum temperature, minimum and maximum relative humidity, wind speed, and solar radiation] from January 1, 1960 to December 31, 1963 (Set 2) were collected for Davis (Pruitt, personal communication, 2000). The lysimeter measured ETo corresponds to frequently mowed, frequently irrigated, ryegrass grown in the 6 m diameter weighting lysimeter at Davis. The grass height was maintained between 0.10 and 0.15 m throughout the observation period.

Normalization of Data

Prior to exporting the data to the ANN for training, the data were normalized. This was done to restrict their range within the interval of 0–1, because the PEs of the middle layer were assigned a sigmoidal activation function. The shape of this function plays an important role in ANN learning. The weight changes corresponding to a value near 0 or 1 are minimal since PE is ‘‘dull’’ whereas closer to 0.5 they respond more (Rao and Rao 1996). Keeping these facts in view, the normalization was carried out so that the mean of the data series would be equal to 0.5. The following equation was used:

$$x_{\mathrm{norm}}=0.5\left(\frac{x_0-\bar{x}}{x_{\max}-x_{\min}}\right)+0.5 \tag{4}$$

where x_{norm}=normalized value; x_0=original value; $\bar{x}$=mean; $x_{\max}$=maximum value; and $x_{\min}$=minimum value.

Table 2. Artificial Neural Networks (ANNs) Training and Validation Errors using Data Sets 1a and 1b, Respectively, for Two Hidden Layer Architectures

	Std. back-propagation LR=0.2		Std. back-propagation LR=0.8		Back-propagation momentum	
Network architecture	Training (Set 1a) MSE (10^{-4})	Validation (Set 1b) MSE (10^{-4})	Training (Set 1a) MSE (10^{-4})	Validation (Set 1b) MSE (10^{-4})	Training (Set 1a) MSE (10^{-4})	Validation (Set 1b) MSE (10^{-4})
6-4-4-105k	2.4	2.0	2.3	2.2	2.2	1.7
6-4-4-11k	2.3	2.3	2.2	2.0	2.2	1.6
6-4-4-12k	2.1	1.8	2.2	2.1	2.0	1.4
6-4-4-13k	1.9	1.7	1.7	1.6	1.9	1.4
6-4-4-14k	1.8	1.7	2.1	1.8	1.5	1.0
6-4-4-15k	2.2	2.1	1.5	1.2	1.3	0.9
6-5-5-105k	2.3	2.1	2.1	3.6	2.3	2.0
6-5-5-11k	2.3	2.1	2.1	3.5	2.3	2.1
6-5-5-12k	2.2	2.0	2.2	2.2	2.2	2.0
6-5-5-13k	2.2	2.0	1.7	1.6	2.0	2.1
6-5-5-14k	1.2	2.3	1.9	1.7	2.2	2.1
6-5-5-15k	1.5	1.3	2.2	2.1	1.6	1.4
6-6-6-105k	2.2	2.1	2.4	2.1	2.3	2.2
6-6-6-11k	2.2	2.0	2.3	2.0	2.3	2.0
6-6-6-12k	2.2	2.2	2.2	2.2	2.2	1.8
6-6-6-13k	2.1	1.9	1.7	1.6	1.7	1.5
6-6-6-14k	2.2	2.0	1.5	1.3	1.7	1.6
6-6-6-15k	2.2	2.0	1.5	1.4	1.6	1.5

One of the major advantages of neural nets is their ability to generalize. To reach the best generalization, the data set should be split into three parts, namely, training set, validation set, and test set (SNNS 1995). The training data set is used to train a neural net by minimizing the error of this data set during training. The validation data set is used to determine the performance of a neural network on patterns that are not trained during training. The test set is used for checking the overall performance of a trained and validated network. Therefore, the normalized data (Set 1) was divided into three subsets for the purpose of training, validation and testing. Data sets 1a, 1b, and 1c were comprised of 1826 (January 1, 1990 to December 31, 1994); 547 (January 1, 1995 to June 30, 1996), and 1,461 (July 1, 1996 to June 30, 2000) patterns for training, validation, and testing, respectively. Similarly, the normalized data set 2 was also divided into training (Set 2a), validation (Set 2b), and testing (Set 2c) sets. The data sets 2a, 2b, and 2c consisted of normalized data from January 1, 1960 to December 31, 1960 and January 1, 1962 to June 30, 1962; July 1, 1962 to December 31, 1962; and January 1, 1961 to December 31, 1961, respectively. Only complete records without any missing data on a given day were considered for the analysis. Consequently, the testing data set had a total of 302 patterns, after excluding the missing data between January 1, 1961 and December 31, 1961. The training and validation data sets, however, included 489 and 133 patterns, respectively.

To obtain the best ANN architecture several possibilities were considered in this study. For each ANN architecture, the number of nodes in the input and output layers were fixed at six and one, respectively. The number of nodes in the input layer corresponded to the six basic input parameters for ETo estimation by the PM method, whereas the output layer node corresponded to the PM ETo. The numbers of nodes in the hidden layer were varied from 7 to 12 for one hidden layer architecture. However, for the two hidden layers architecture three, four, and five nodes were considered in each hidden layer. Each ANN architecture was tested for 500, 1,000, 2,000, 3,000, 4,000, and 5,000 learning cycles. Furthermore, three learning methods, namely, standard back-propagation with learning rates of 0.2 and 0.8 and back-propagation momentum with a learning rate of 0.2 and a momentum term as 0.95, were used.

The ANNs tested were given names according to their architecture, the learning method, and number of learning cycles. The network architecture is described with a set of numbers separated by "−" signs. "*s*" denotes the standard back-propagation and "*b*" denotes the back-propagation with momentum as the learning method. The learning rate is denoted with the numbers 2 and 8. The alphanumeric values of 05k, 1k, 2k, 3k, 4k, and 5k represent the number of learning cycles of 500, 1,000, 2,000, 3,000, 4,000 and 5,000, respectively. Thus, the ANN 6-7-1b25k had six PEs in its input layer, seven PEs in its hidden layer, and one PE in its output layer; the back-propagation momentum learning was used with a learning rate of 0.2; the network was trained for 5,000 cycles.

A total of 162 (108 networks with single hidden layer and 54 networks with two hidden layers) ANNs were studied. The Stuttgart Neural Network Simulator version 4.1 (SNNSv4.1) distributed by the University of Stuttgart was used to implement the neural networks. Training and validation data sets (Sets 1a and 1b) were shuffled before training and validation to ensure the randomness during ANNs training and validation. The SNNS provides information on mean sum of square error (MSE) during the training and validation. Using the test input data (Set 1c), the ETo prediction performance of each network was evaluated for both peak (July) and all months. The peak month is the month when peak ETo occurred. The minimal network architecture and weighted standard error of estimate (WSEE) were used as criterion for selecting the best network for each learning method. WSEE was derived from the standard error of estimate (SEE) and adjusted SEE (ASEE) (Jensen et al. 1990). The SEE was estimated as follows:

Table 3. Artificial Neural Network Testing Error Weighted Standard Error of Estimate (WSEE) using Data Set 1c for Single Layer Architectures

Network architecture	Std. back-propagation LR=0.2 Testing (Set 1c) WSEE (mm/day)	Std. back-propagation LR=0.8 Testing (Set 1c) WSEE (mm/day)	Back-propagation momentum Testing (Set 1c) WSEE (mm/day)
6-7-105k	0.41	0.34	0.3
6-7-11k	0.36	0.46	0.29
6-7-12k	0.36	0.29	0.28
6-7-13k	0.35	0.29	0.29
6-7-14k	0.31	0.30	0.30
6-7-15k	0.29	0.29	0.27
6-8-105k	0.39	0.37	0.34
6-8-11k	0.37	0.37	0.28
6-8-12k	0.31	0.29	0.28
6-8-13k	0.30	0.31	0.28
6-8-14k	0.33	0.30	0.27
6-8-15k	0.33	0.30	0.28
6-9-105k	0.40	0.30	0.31
6-9-11k	0.35	0.30	0.34
6-9-12k	0.33	0.29	0.30
6-9-13k	0.31	0.29	0.31
6-9-14k	0.30	0.29	0.31
6-9-15k	0.30	0.29	0.29
6-10-105k	0.39	0.33	0.31
6-10-11k	0.37	0.29	0.33
6-10-12k	0.30	0.29	0.28
6-10-13k	0.30	0.29	0.31
6-10-14k	0.32	0.29	0.28
6-10-15k	0.34	0.29	0.31
6-11-105k	0.38	0.35	0.30
6-11-11k	0.38	0.31	0.30
6-11-12k	0.32	0.28	0.28
6-11-13k	0.29	0.30	0.28
6-11-14k	0.31	0.28	0.28
6-11-15k	0.29	0.30	0.29
6-12-105k	0.38	0.42	0.35
6-12-11k	0.36	0.29	0.29
6-12-12k	0.34	0.35	0.31
6-12-13k	0.31	0.28	0.27
6-12-14k	0.28	0.28	0.28
6-12-15k	0.28	0.30	0.36

$$\mathrm{SEE}=\left[\frac{\sum_{i=1}^{n}(Y_O-Y_E)^2}{n-1}\right]^{0.5} \tag{5}$$

where SEE = standard error of estimate; Y_O = ETo estimated using the standard methods (PM and lysimeter ETo represent standard method for Sets 1c and 2c, respectively); Y_E = ETo estimated using the test method (ANN, ANN, and PM represent the test methods for Sets 1c and 2c, respectively); and n = number of observations.

Linear regression analyses were made with the standard method estimated ETo as the dependent variable and the test method estimated ETo as the independent variable. The equation of regression through the origin is given below

$$Y_O=b\times Y_E \tag{6}$$

The regression coefficients b were used to adjust the ETo estimates and SEEs were recalculated for the adjusted values (ASEE), i.e., Y_E in Eq. (5) was set equal to $[b\times Y_E]$. Using the SEE and ASEE values, the WSEE was calculated as follows (Jensen et al. 1990):

$$\mathrm{WSEE}=0.7[0.67(\mathrm{SEE}_{\mathrm{all}})+0.33(\mathrm{ASEE}_{\mathrm{all}})] + 0.3[0.67(\mathrm{SEE}_{\mathrm{peak}})+0.33(\mathrm{ASEE}_{\mathrm{peak}})] \tag{7}$$

where subscripts all and peak = all months and peak months, respectively.

The results related to the effect of learning cycle, number of PEs in the hidden layer, number of hidden layers, and learning methods on ANN performance along with choice of the best ANN architecture are presented in the following section. In addition, ETo estimated with the ANN and PM models are compared with the lysimeter measured ETo values.

Results and Discussion

Training And Validation of Network

Table 1 presents training and validation errors (MSE) for all single hidden layer ANN architectures. For all learning methods, MSE during training decreased with the increase in learning cycles and did not exhibit any trend with the increase in number of PEs in the hidden layer. The standard back-propagation with amomentum learning method resulted in slightly lower training error as compared to the standard back-propagation method for the same network architecture. The optimal condition for ANN training is defined by the number of learning cycles for given ANNs where network freezes to learn further. The number of learning cycles for such a condition was determined using the validation data (Set 1b). For example, in the case of 6-8-1-b2 networks the network performance improved up to 4,000 learning cycles and remained unchanged for 5,000 learning cycles (Table 1). Similar trends were noticed with other ANNs. However, the optimal limit for cycles varied from 3,000 to 5,000.

Table 2 presents training and validation errors (MSE) for all two hidden layer network architectures. For the same number of PEs in hidden layer(s), both training and validation errors were higher for two hidden layer architectures than for a single hidden layer architectures, when trained with the same number of learning cycles (Tables 1 and 2). Therefore, networks with two hidden layers (54 cases) were not considered for further analysis.

Selection of Best Artificial Neural Network Combination

Using data set 1c, ETo was determined for all 108 single hidden layer networks (3 training methods×6 learning cycles×6 cases of PEs in the hidden layer) and was compared with the PM estimated ETo values to obtain the most promising network for each learning method for estimation of ETo. A single network for each learning method was selected based on the minimal network architecture and WSEE.

Using Eq. (7), WSEEs were estimated for all 108 networks and are presented in Table 3. The back-propagation with learning

Table 4. Statistical Summary of Artificial Neural Network (ANN) Predicted Grass Reference Crop Envirotranspiration (ETo) for Test Period (Data Set 1c)

	ANN learning scheme					
	Std Back-Propagation, LR=0.2		Std Back-Propagation LR=0.8		Back-Propagation Momentum	
Statistical parameter	ALL MONTHS	PEAK MONTHS	ALL MONTHS	PEAK MONTHS	ALL MONTHS	PEAK MONTHS
SEE[a]	0.27	0.34	0.26	0.39	0.26	0.33
(mm/day)	(0.77)	(0.70)	(0.77)	(0.70)	(0.77)	(0.70)
b[b]	0.99	0.99	1.00	0.99	0.98	0.98
	(0.98)	(1.04)	(0.98)	(1.04)	(0.98)	(1.04)
r[c]	0.99	0.97	0.99	0.95	0.99	0.97
	(0.92)	(0.80)	(0.92)	(0.80)	(0.92)	(0.80)
ASEE[d]	0.27	0.34	0.26	0.39	0.25	0.31
(mm/day)	(0.75)	(0.66)	(0.75)	(0.66)	(0.75)	(0.66)
WSEE[e]	0.29		0.29		0.27	
(mm/day)	(0.74)		(0.74)		(0.74)	

Note: Number in parenthesis is based on comparison between the PM and lysimeter ETo (Jensen et al. 1990).

[a]Standard error of estimate for ANN estimated ETo (mm/day) not adjusted by regression.

[b]Regression coefficient (slope) for regression through the origin between the PM and ANN ETo estimates.

[c]Correlation coefficient for regression through the origin between the PM and ANN ETo estimates.

[d]Standard error of estimate for ANN estimated ETo (mm/day) adjusted by regression.

[e]Weighted standard error of estimate.

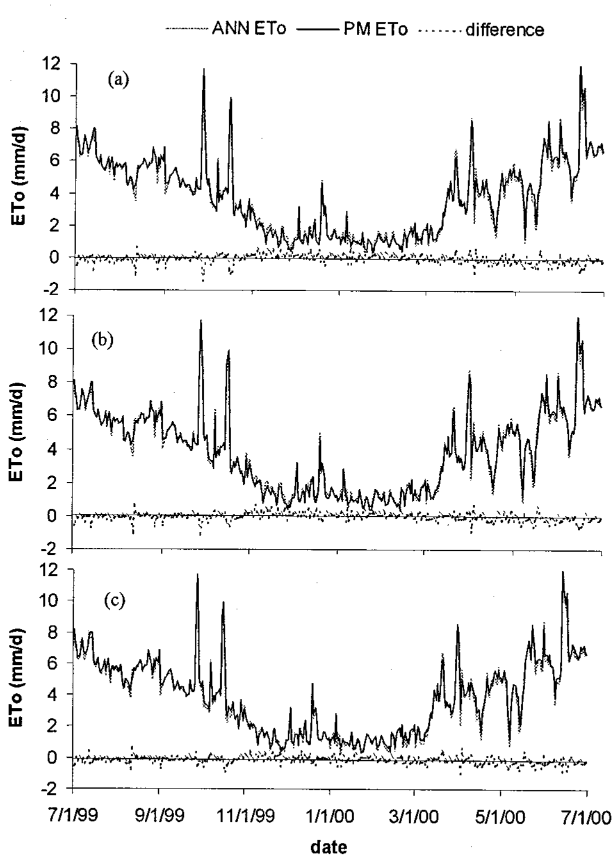

Fig. 1. Comparison between Penmann–Monteith estimated grass reference crop evapotranspiration and artificial neural network estimated grass reference crop evapotranspiration (a) 6-7-1s25k, (b) 6-7-1s85k, and (c) 6-7-1b25k for 1 year period from July 1, 1999 to June 30, 2000.

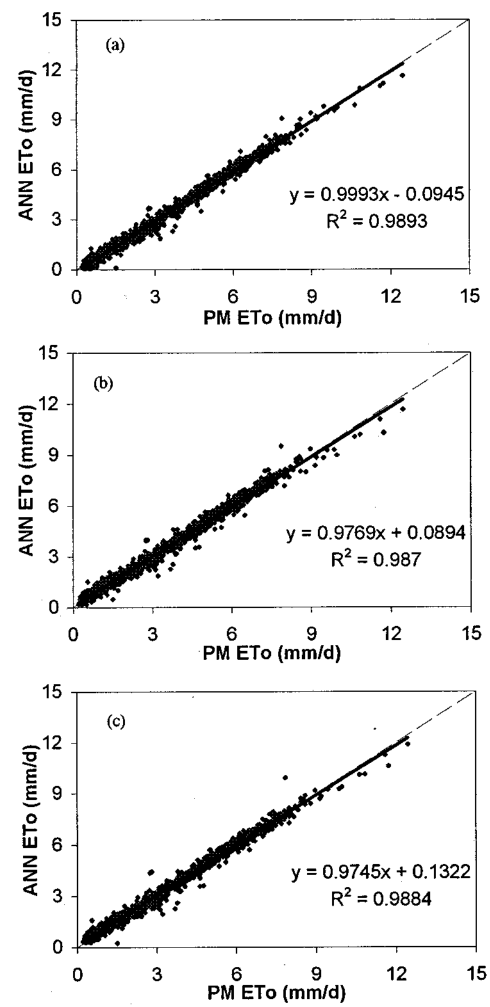

Fig. 2. Scatter plot between grass reference crop evapotranspiration estimated using Penmann–Monteith and selected artificial neural network: (a) 6-7-1b25k, (b) 6-7-1s25k, and (c) 6-7-1s85k for complete test period (July 1, 1996 to June 30, 2000).

Table 5. Artificial Neural Networks Training and Validation Errors using Data Sets 2a and 2b, Respectively, for Single Hidden Layer Architectures

	Std. back-propagation LR=0.2		Std. back-propagation LR=0.8		Back-propagation momentum	
Network architecture	Training (Set 1a) MSE (10^{-4})	Validation (Set 1b) MSE (10^{-4})	Training (Set 1a) MSE (10^{-4})	Validation (Set 1b) MSE (10^{-4})	Training (Set 1a) MSE (10^{-4})	Validation (Set 1b) MSE (10^{-4})
6-7-105k	9.1	4.9	9.6	11.6	8.2	4.1
6-7-11k	7.8	4.2	8.2	3.1	8.0	3.2
6-7-12k	7.8	3.9	8.3	3.4	7.5	2.9
6-7-13k	7.7	4.1	8.2	3.0	7.2	4.7
6-7-14k	7.2	2.9	7.1	8.0	7.2	3.4
6-7-15k	7.1	3.2	8.2	3.1	6.7	2.9

rate of 0.2 and back-propagation with momentum resulted in the maximum and minimum variation in WSEE values, respectively, for different ANN architectures. The WSEE values between ANN estimated ETo and PM ETo were lower than that reported by Jensen et al. (1990) between PM ETo and lysimeter measured ETo (0.74 mm/day). The ANN architecture 6-7-1 with 5,000 cycles gave the minimum WSEE of 0.29, 0.29, and 0.27 mm/day for the standard back-propagation with learning rates 0.2 and 0.8, and back-propagation with momentum learning methods, respectively. This network also resulted in the minimal ANN architecture. Thus, these three ANN architectures were selected as the best ones. However, the ANN architecture with the back-propagation momentum was found to be the best among the selected networks because of the minimum WSEE and higher consistency.

Statistical summary of ETo estimation performance of each learning method is presented in Table 4 for data set 1c. It may be noted that the statistics for all month and peak month were determined considering complete data set 1c and peak month data (a subset of 1c for the month of July). Furthermore, statistics given in parenthesis are based on comparison between PM ETo and lysimeter ETo (Jensen et al. 1990). All the chosen networks gave SEE, ASEE, and WSEE values less than the reported values for the PM method by Jensen et al. (1990) (Table 4). Furthermore, the regression coefficient b and correlation coefficient r were close to unity. These results indicate that if the networks were trained against the lysimeter measured ETo then their performance would have been probably better than the PM method forDavis. Most likely the nonlinearity in the interacting factors is not fully captured by the PM method.

Comparisons between the PM ETo and ANN estimated ETo using the Standard back-propagation with momentum (6-7-1b25k), standard back-propagation with learning rates of 0.2 (6-7-1-s25k) and 0.8 (6-7-1-s85k), respectively, for a 1 year period (July 1, 1999 to June 30, 2000) for Davis are presented in Fig. 1. The ANN estimated ETo values agreed with the PM ETo values and followed the same trend. In all cases, the deviation in ETo values was less than 1 mm/day (Fig. 1). Fig. 2 is a scatterplot between ETo estimated using the PM method and selected ANN architectures for the complete test period (July 1, 1996 to June 2000). The ANN estimated ETo values lie on both sides of the 1:1 line almost symmetrically. This was true for all three ANNs. All ANNs resulted in a high R^2 value, slope close to unity, and intercept close to zero. In all cases, the intercept and slope were not significantly different than zero and unity, respectively, at 5% significance level using the t-test. Therefore, any one of the networks can be used for predicting ETo.

Comparison with Lysimeter Measured Grass Reference Crop Evapotranspiration

All three ANNs were again trained and validated with the lysimeter measured ETo (Sets 2a and 2b) for six different training cycles (500, 1,000, 2,000, 3,000, 4,000, and 5,000). Training and validation errors (MSE) are presented in Table 5. The training error showed a decreasing trend with increasing learning cycles. Using test data set 2c, the ETo estimation performance of the PM and ANN was evaluated against the lysimeter ETo. WSEE values for the test period are presented in Table 6 for all three ANN architectures. The WSEE varied from 0.56 mm/day for 1,000 cycles to 0.76 mm/day for 4,000 cycles in the case of the back-propagation momentum learning method. The WSEE for the standard back-propagation with 0.8 as the learning rate was nearly stable for all cycles, whereas it fell rapidly up to 2,000 cycles and became almost stable thereafter in the case of standard back-propagation with learning rate of 0.2. The ANN architecture of 6-7-1 trained for 5,000, 5,000, and 1,000 cycles was selected for the back-propagation with learning rates of 0.2 and 0.8, and back-propagation with momentum learning method, respectively.

Grass reference crop evapotranspiration estimated using the selected ANNs was compared with both the PM estimated and lysimeter measured ETo for the test period (data set 2c) and summary statistics is presented in Table 7. The SEE and ASEE values for both all month and peak month were lower for the ANN models than for the PM method. For all ANNs, the

Table 6. Artificial Neural Networks Testing Error Weighted Standard Error of Estimate (WSEE) in Predicting Lysimeter Grass Reference Crop Evapotranspiration (Data Set 2c)

	Std. back-propagation LR=0.2	Std. back-propagation LR=0.8	Back-propagation momentum
Network architecture	Testing (Set 2c) WSEE (mm/day)	Testing (Set 2c) WSEE (mm/day)	Testing (Set 2c) WSEE (mm/day)
6-7-105k	0.72	0.59	0.60
6-7-11k	0.67	0.58	0.56
6-7-12k	0.59	0.58	0.64
6-7-13k	0.59	0.55	0.71
6-7-14k	0.56	0.56	0.76
6-7-15k	0.56	0.56	0.60

Table 7. Summary Statistics of Artificial Neural Network (ANN) Predicted Grass Reference Crop Evapotranspiration (ETo) and Penmann–Monteith (PM) Estimated ETo with Respect to Lysimeter ETo for Test Period (Data Set 2c)

	ANN learning scheme							
	Std Back-Propagation, LR=0.2		Std Back-Propagation LR=0.8		Back-Propagation Momentum		Penman–Monteith	
Statistical parameter	ALL MONTHS	PEAK MONTHS	ALL MONTHS	PEAK MONTHS	ALL MONTHS	PEAK MONTHS	ALL MONTHS	PEAK MONTHS
SEE[a] (mm/day)	0.57	0.54	0.56	0.57	0.57	0.55	1.03	0.90
b[b]	1.00	1.03	0.99	1.02	0.99	1.00	1.04	1.09
r[c]	0.97	0.95	0.97	0.95	0.97	0.95	0.91	0.83
ASEE[d] (mm/day)	0.57	0.51	0.56	0.56	0.56	0.55	1.00	0.73
WSEE (mm/day)	0.56		0.56		0.56		0.97	

[a]Standard error of estimate for ANN/PM estimated ETo (mm/day) not adjusted by regression.

[b]Regression coefficient (slope) for regression through the origin between the lysimeter ETo and ANN/PM ETo estimates.

[c]Correlation coefficient for regression through the origin between the lysimeter ETo and ANN/PM ETo estimates.

[d]Standard error of estimate for ANN/PM estimated ETo (mm/day) adjusted by regression through the origin.

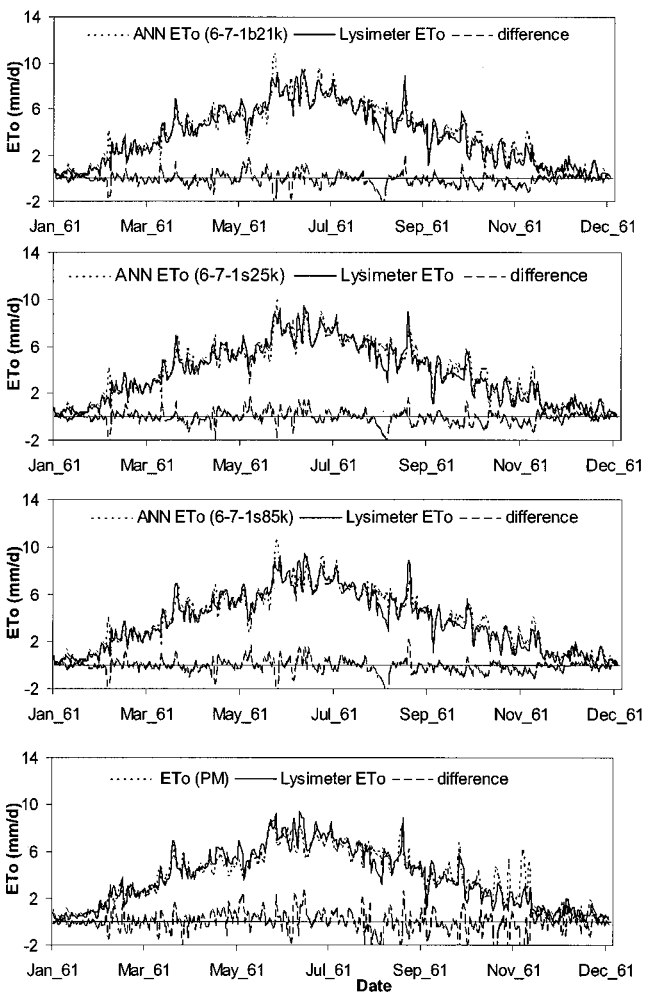

Fig. 3. Comparison between lysimeter measured grass reference crop evapotranspiration and estimated grass reference crop evapotranspiration by artificial neural network and Penmann–Monteith for test period (January 1, 1961 to December 31, 1961, data set 2c)

slope was close to unity, and correlation values were higher than for the PM method (Table 7). Furthermore, all ANNs yielded WSEE values lower than the PM method (0.97 mm/day) for the test period.

The relations between lysimeter measured and estimated ETo using the 6-7-1s25k, 6-7-1s85k, and 6-7-1b21k ANN, and PM method are shown in Fig. 3. Both the PM and ANN estimated ETo followed the trend of lysimeter ETo. However, the ANN estimated ETo is closer to the measured ETo than the PM method. All three ANNs showed a similar ETo trend and agreement with the measured ETo. The ANN and PM estimated ETo values lie on both sides of the 1:1 line (Fig. 4) almost symmetrically. This was true for all the three selected ANN architectures [Figs. 4(a–c)].However, the spread was greater in the case of the PM [Fig. 4(d)] than the ANN. Further for all ANNs, the slope was close to

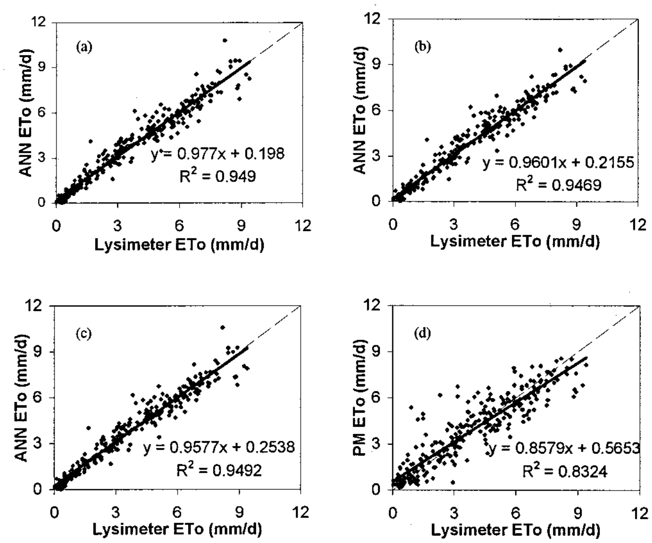

Fig. 4. Scatter plot between lysimeter grass reference crop evapotranspiration and estimated grass reference crop evapotranspiration using Penmann–Monteith (PM) and artificial neural network: (a) 6-7-1b21k, (b) 6-7-1s25k, (c) 6-7-1s85k, and (d) PM for test period (January 1, 1961 to December 31, 1961, data set 2c)

unity, the intercept was close to zero, and R^2 values were higher-than for the PM method (Fig. 4). These results clearly indicatethat the ANN models estimate ETo values better than the PM method for the Davis data.

Conclusions

The results revealed that the single hidden layer ANNs are sufficient to account for the nonlinear relationship between climatic variables and corresponding ETo. Improvement in performance ceases for higher learning cycles and PEs in the hidden layer. Several network architectures performed similarly within the same learning method. Based on the performance criteria, a network architecture of 6-7-1 trained with 5,000 learning cycles was the best for all three learning methods. All three methods gave WSEE values less than 0.30 mm/day and were lower than those obtained between the PM and measured lysimeter ETo (0.74 mm/day, Jensen et al. 1990). Thus, the results suggest that given the lysimeter measured ETo as a target, ANN predicted ETo could be better than the PM method.

The selected ANNs were trained, validated, and tested against the lysimeter measured ETo. All ANNs yielded lower WSEE than the PM method. The results of the present study show that a local ANN model can be trained to predict lysimeter ETo values better than the standard PM method. However, the PM method is a global model and thus can be applied to predict ETo even for areas for which it is not trained. Although the ANN models exhibit a tendency to obtain a generalized architecture, their application to other areas needs to be studied. The results are of significant practical use because the ANNs can be used to interpolate missing ETo data, particularly within the training data range for stations where sufficient lysimeter data for training/validation and testing exist.

Notation

The following symbols are used in this paper:

I = sum of weighted input;
N = no. of iteration;
q = neuron index in output layer;
R^2 = coefficient of determination;
r = correlation coefficient;
w = weight of link between neurons of two layer;
x = input value stored in neuron of input layer;
$\overline{x}$ = mean of input buffer x;
$x_{\max}$ = maximum value of input buffer x;
$x_{\min}$ = minimum value of input buffer x;
x_{norm} = normalized value of input buffer x;
x_0 = original value;
Y_E = grass reference crop evapotranspiration estimated using test method;
Y_O = grass reference crop evapotranspiration estimated using standard methods;
Δw = change in weight;
ε = error back-propagating;
η = learning rate;
μ = momentum coefficient; and
ϕ = normalized output.

References

Allen, R. G. (1986). ‘‘A Penman for all seasons.’’ *J. Irrig. Drain. Eng.*, 112(4), 348–368.

Allen, R. G., Jensen, M. E., Wright, J. L., and Burman, R. D. (1989). ‘‘Operational estimates of reference evapotranspiration.’’ *J. Agron.*, 81, 650–662.

Campolo, M., Andreussi, P., and Sodalt, A. (1999). ‘‘River stage forecasting with a neural network model.’’ *Water Resour. Res.*, 35(4), 1191–1197.

Chiew, F. H. S., Kamaladassa, N. N., Malano, H. M., and McMahon, T. A. (1995). ‘‘Penman-Monteith, FAO-24 reference crop evapotranspiration and class-A pan data in Australia.’’ *Agric. Water Manage.*, 28, 9–21.

De Souza, F., and Yoder, R. E. (1994). ‘‘ET estimation in the north east of Brazil: Hargreaves or Penman-Monteith equation?’’ *Proc., Tech. Paper ASAE Int. Winter Meeting*, American Society of Agricultural Engineers, St. Joseph, Mich.

Doorenbos, J., and Pruitt, W. O. (1977). ‘‘Guidelines for predicting crop water requirements.’’ *Irrigation Drain. Paper No. 24*, FAO, Rome, Italy.

French, M. N., Krajewski, W. F., and Cuykendall, R. R. (1992). ‘‘Rainfall forecasting in space and time using a neural network.’’ *J. Hydrol.*, 137, 1–37.

Goh, A. T. C. (1999). ‘‘Soil laboratory data interpretation using generalized regression neural network.’’ *Civ. Environ. Eng. Syst.*, 16, 175–195.

Jain, S. K., Das, A., and Srivastava, D. K. (1999). ‘‘Application of ANN for reservoir inflow prediction and operation.’’ *J. Water Resour. Plan. Manage.*, 125(5), 263–271.

Jensen, M. E., Burman, R. D., and Allen R. G. (1990). ‘‘Evapotranspiration and irrigation water requirements.’’ *ASCE Manual and Report on Engineering Practice No. 70*. ASCE, New York.

McNaughton, K. G., and Jarvis, P. G. (1984). ‘‘Using Penman-Monteith equation predictively.’’ *Agric. Water Manage.*, 8, 263–278.

Minnes, A. W., and Hall, M. J. (1996). ‘‘Artificial neural networks as rainfall-runoff models.’’ *Hydrol. Sci. J.*, 41(3), 399–416.

Monteith, J. L. (1965). ‘‘The state and movement of water in living organisms.’’ *Proc., Evaporation and Environment, XIXth Symp.*, Soc. For Exp. Biol., Swansea, Cambridge Univ. Press, New York, 205–234.

Penman, H. L. (1948). ‘‘Natural evaporation from open water, bare soil and grass.’’ *Proc. R. Soc. London*, 193, 120–146.

Rao, V., and Rao, H. (1996). *C++ Neural networks and fuzzy logic*, BPB, New Delhi, India, 380–381.

Rogers, L. L., and Dowla, F. U. (1994). ‘‘Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling.’’ *Water Resour. Res.*, 30(2), 457–481.

Shukla, M. B., Kok, R., Prasher, S. O., Clark, G., and Lacroix, R. (1996). ‘‘Use of artificial neural network in transient drainage design.’’ *Trans. ASAE*, 39(1), 119–124.

SNNS (1995). *User manual, version 4.1, Rep. No. 6/95*, Institute for Parallel and Distributed High Performance Systems, Univ. of Stuttgart, Stuttgart, Germany.

Srinivasa, L. (1998). ‘‘Aquifer parameter estimation using genetic algorithm and neural networks.’’ *Civ. Environ. Eng. Syst.*, 15, 125–144.

Tansel, B., Jordahl, C., and Tansel, I. (1999). ‘‘Mapping of subsurface contamination profile by neural networks.’’ *Civ. Environ. Eng. Syst.*, 16, 37–50.

Thirumalaian, K., and Deo, M. C. (1998). ‘‘River stage forecasting using artificial neural networks.’’ *J. Hydrol. Eng.*, 3(1), 26–32.

Tsoukalas, L. H., and Uhrig, R. E. (1996). *Fuzzy and neural approach in engineering*, Wiley, New York.

Wen, C. G., and Lee, C. S. (1998). ‘‘A neural network approach to multiobjective optimization for water quality management in a riverbasin.’’ *Water Resour. Res.*, 34(3), 427–436.

Yang, C. C., Lacroix, R., and Prasher, S. O. (1998). ‘‘The use of back-propagation neural networks for the simulation and analysis of time-series data in subsurface drainage system.'' *Trans. ASAE,* 41(4), 1181–1187.

Yang, C. C., Prasher, S. O., and Lacroix, R. (1996). ''Application of artificial neural network to land drainage engineering.'' *Trans. ASAE,* 39(2), 525–533.

Yang, C. C., Prasher, S. O., Sreekanth, S., Patni, N. K., and Masse, L. (1997). ''An artificial neural network for simulating pesticide concentration in soil.'' *Trans. ASAE,* 40(5), 1285–1294.

Zealand, C. M., Burn, D. H., and Simonovic, S. P. (1999). ''Short term streamflow forecasting using artificial neural networks.'' *J. Hydrol.,* 214, 32–48.