Evapotranspiration estimation using feed-forward neural networks

Özgür Kişi

Civil Engineering Department, Engineering Faculty, Erciyes University, Kayseri, Turkey.
E-mail: *kisi@erciyes.edu.tr*

Received 2 May 2005; accepted in revised form 3 March 2006

Abstract Two different feed-forward neural network algorithms, Levenberg–Marquardt (LM) and conjugate gradient (CG), are used for estimation of daily reference evapotranspiration (ET) from climatic data. The performances of the LM and CG algorithms in estimating ET are analyzed and discussed and various combinations of wind speed, solar radiation, relative humidity, air and soil temperature data as inputs to the artificial neural network (ANN) models are examined in the study so as to evaluate the degree of the effect of each of these variables on ET. The LM and CG training algorithms are compared with each other according to their convergence velocities in training and estimation performances of ET. The results of the ANN models are compared with those of multi-linear regression (MLR) and the empirical models of Penman and Hargreaves. Based on the comparisons, it was found that the neural computing technique could be employed successfully in modelling evapotranspiration process from the avaliable climatic data.

Keywords Conjugate gradient; evapotranspiration; Hargreaves; Levenberg–Marquardt; neural networks; Penman

Nordic Hydrology **Vol 37 No 3 pp 247–260** © IWA Publishing 2006

Introduction

The calculation of ET loss becomes imperative in the planning and management of irrigation practices in many areas where water resources are scarce. Temperate areas of rain seasons are mainly concerned with the shortage of water in the dry seasons for the different uses. Despite this significance, evapotranspiration is one of the least measured components of the hydrologic cycle (Brutsaert 1982; Jackson 1985).

Evaporation and transpiration occur simultaneously and there is no easy way of distinguishing between the two processes (Allen *et al.* 1998). Transpiration consists of the vaporization of liquid water contained in plant tissues and the vapor removal to the atmosphere. Evaporation occurs at the topsoil if the water is available. When the crop is small, water is predominantly lost by soil evaporation; but once the crop is well developed and completely covers the soil, transpiration becomes the main process. Smith *et al.* (1997) defined the reference evapotranspiration as "the rate of evapotranspiration from a hypothetical reference crop with an assumed crop height (12 cm), a fixed crop surface resistance (70 s m^{-1}) and albedo (0.23), closely resembling the evapotranspiration from an extensive surface of green grass cover of uniform height, actively growing, completely shading the ground with adequate water". Reference evapotranspiration (ET) represents the effect of the climate on the evapotranspiration process. It represents the first step in the estimation of crop water requirements (Allen *et al.* 1998; Naoum and Tsanis 2003).

Numerous methods have been proposed for modelling evapotranspiration as described by Brutsaert (1982) and ASCE (1989). In general, the combination of energy balance/aerodynamic equations "provide the most accurate results as a result of their foundation in

doi: 10.2166/nh.2006.010

physics and basis on rational relationships" (ASCE 1989). The Food and Agricultural Organization of the United Nations (FAO) assumed the ET definition from Smith *et al.* (1997) and adopted the FAO Penman–Monteith as the standard equation for estimation of ET (Allen *et al.* 1998; Naoum and Tsanis 2003). However, the evapotranspiration models, such as those based on the equations of Shuttleworth and Wallace (1985) or simply the Penman–Monteith equation (Monteith 1973), require extensive input, such as the leaf area index and crop aerodynamic characteristics at each developmental stage, the soil moisture-dependent stomata and canopy resistances and a number of meteorological parameters like temperature, humidity, wind speed (at least at two levels if actual instead of neutral atmospheric stability conditions were to be considered) and of course net radiation and soil heat flux (Poulovassilis *et al.* 2001).

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; Tokar and Johnson 1999); river stage forecasting (Thirumalaian and Deo 1998; Campolo *et al.* 1999); streamflow prediction (Kang *et al.* 1993; Zealand *et al.* 1999; Cigizoğlu 2003*a*, *b*; Kisi 2004; Cigizoglu and Kisi 2005) reservoir operation (Jain *et al.* 1999); land drainage design (Shukla *et al.* 1996; Yang *et al.* 1998) and aquifer parameter estimation (Srinivasa 1998). There are few published works in the field of evapotranspiration modelling using ANN (Kumar *et al.* 2002; Sudheer *et al.* 2003; Trajkovic *et al.* 2003; Kisi and Yildirim 2005*a*, *b*). This provided an impetus to investigate the potential of the ANN for ET estimation. Kumar *et al.* (2002) developed ANN models using gradient descent training algorithms in estimation of ET and they compared the ANN estimates with those of the Penman method. They found that the ANN could predict ET better than the conventional method. Sudheer *et al.* (2003) and Trajkovic *et al.* (2003) indicated the performance of the radial basis neural networks in ET estimation. To the knowledge of this writer, no study has been carried out to utilize the input–output mapping capability of ANN with LM and CG training algorithms in the estimation of ET.

The performances of the LM and CG neural network algorithms in estimating ET are analyzed and discussed in the present study. Various combinations of climatic data are tried so as to evaluate the degree of effect of each of these variables on ET. The performance of the ANN is also compared with those of the Penman and Hargreaves methods.

This paper is organized as follows. The next two sections of this paper provide an overview description of the ANN algorithms and MLR. The fourth and fifth sections give brief descriptions of the data and ANN design used in the study. The applications of the models for ET estimation and results are examined in the sixth section. Finally, a last section provides concluding remarks.

Artificial neural networks

Artificial neural networks (ANNs) are based on the present understanding of biological nervous systems, though much of the biological detail is neglected. ANNs are massively parallel systems composed of many processing elements connected by links of variable weights. Of the many ANN paradigms, the back-propagation network is by far the most popular (Lippman 1987). The network consists of layers of parallel processing elements, called neurons, with each layer being fully connected to the preceding layer by interconnection strengths, or weights, W. Figure 1 illustrates a three-layer neural networks consisting of layers i, j, and k, with the interconnection weights W_{ij} and W_{jk} between layers of neurons. Initial estimated weight values are progressively corrected during a training process that compares predicted outputs to known outputs, and back-propagates any errors (from

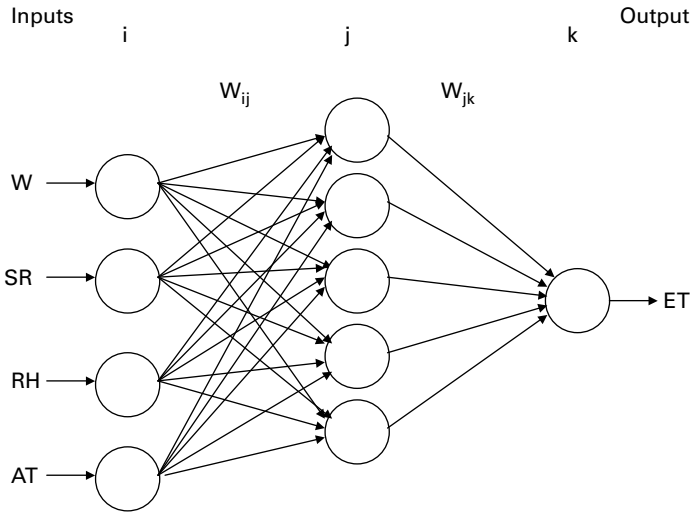

Figure 1 An ANN architecture used for ET estimation

right to left in Figure 1) to determine the appropriate weight adjustments necessary to minimize the errors. The LM and CG algorithms are used here for adjusting the weights.

Levenberg–Marquardt algorithm

The Levenberg–Marquardt algorithm was designed to approach second-order training speed without having to compute the Hessian matrix (More 1977). When the performance function has the form of a sum of squares (as is typical in training feedforward networks), then the Hessian matrix can be approximated as

$$H = J^T J \tag{1}$$

and the gradient can be computed as

$$g = J^T e \tag{2}$$

where J is the Jacobian matrix, which contains first derivatives of the network errors with respect to the weights and biases, and e is a vector of network errors.

The Levenberg–Marquardt algorithm uses this approximation to the Hessian matrix in the following Newton-like update:

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e. \tag{3}$$

When the scalar μI is zero, this is just Newton's method, using the approximate Hessian matrix. When μI is large, this becomes gradient descent with a small step size. Newton's method is faster and more accurate near an error minimum, so the aim is to shift towards Newton's method as quickly as possible. Thus, μI is decreased after each successful step (reduction in performance function) and is increased only when a tentative step would increase the performance function. In this way, the performance function will always be reduced at each iteration of the algorithm. The Levenberg–Marquardt optimization technique is more powerful than the conventional gradient descent techniques (Hagan and Menhaj 1994; El-Bakyr 2003; Cigizoglu and Kisi 2005). The application of Levenberg–Marquardt to neural network training is described in Hagan and Menhaj (1994).

Conjugate gradient algorithm

The basic back-propagation algorithm adjusts the weights in the steepest descent direction (negative of the gradient). In the conjugate gradient algorithms a search is performed along

conjugate directions, which produces generally faster convergence than steepest descent directions (Adeli and Hung 1995).

All of the conjugate gradient algorithms start out by searching in the steepest descent direction (negative of the gradient) on the first iteration

$$p_0 = -g_0. \tag{4}$$

A line search is then performed to determine the optimal distance to move along the current search direction:

$$x_{k+1} = x_k + \alpha_k g_k. \tag{5}$$

Then the next search direction is determined so that it is conjugate to the previous search directions. The general procedure for determining the new search direction is to combine the new steepest descent direction with the previous search direction:

$$p_k = -g_k + \beta_k p_{k-1}. \tag{6}$$

The various versions of conjugate gradient are distinguished by the manner in which the constant β_k is computed. For the Fletcher–Reeves update the procedure is

$$\beta_k = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}. \tag{7}$$

This is the ratio of the norm squared of the current gradient to the norm squared of the previous gradient (Fletcher and Reeves 1964). This update procedure is used in the study and denoted as CGF (conjugate gradient with Fletcher–Reeves).

Multi-linear regression (MLR)

If it is assumed that the dependent variable Y is affected by m independent variables X_1, X_2,..., X_m and a linear equation is selected for the relation among them, the regression equation of Y can be written as

$$y = a + b_1 x_1 + b_2 x_2 + \ldots + b_m x_m. \tag{8}$$

y in this equation shows the expected value of the variable Y when the independent variables take the values $X_1 = x_1$, $X_2 = x_2$,..., $X_m = x_m$.

The regression coefficients a, b_1, b_2,..., b_m are evaluated, similar to simple regression, by minimizing the sum of the e_{yi} distances of observation points from the plane expressed by the regression equation (Bayazıt and Oguz 1998):

$$\sum_{i=1}^{N} e_{yi}^2 = \sum_{i=1}^{N} (y_i - a - b_1 x_{1i} - b_2 x_{2i} - b_m x_{mi})^2. \tag{9}$$

In this study, the coefficients a, b_1, b_2,..., b_m were determined using the least squares method.

Description of data

The daily wind speed (W), solar radiation (SR), air temperature (AT), soil temperature (ST), relative humidity (RH) and evapotranspiration (ET) data belonging to the Blythe Northeast Weather Station (Latitude 33° 33′ N, Longitude 114° 40′ W) in California were used in the study. The elevation of the station is 84 m. The location of the station is shown in Figure 2. The data of December 01, 2001 to November 30, 2002 were chosen for calibration, and data for December 01, 2002 to November 30, 2003 were used for validation.

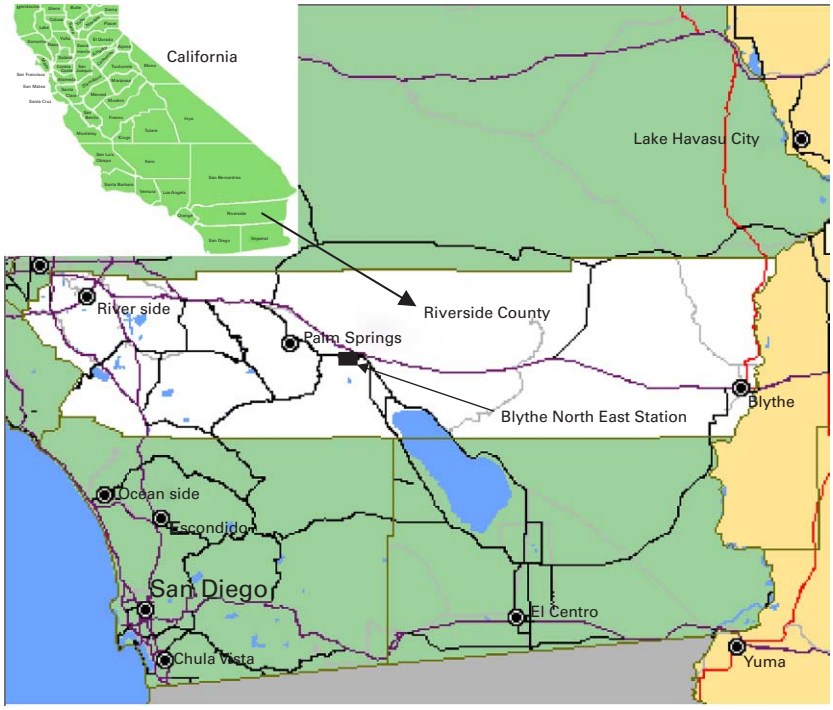

Figure 2 The location of the Blythe North East Station in Riverside County, California

The daily statistical parameters of each data are given in Table 1. In this table, x_{mean}, Sx, Cv, Csx, x_{min}, and x_{max} denote the mean, standard deviation, variation, skewness coefficient, minimum and maximum of the data, respectively. All the data sets have high correlation between ET and low skewness except the wind speed. The highest correlation with the ET (0.946) belongs to the SR. There is an inverse proportion between RH and ET with a negative correlation of 0.646.

Design of ANN

A difficult task with ANNs involves choosing parameters such as the number of hidden nodes, the learning rate, and the initial weights. Determining an appropriate architecture of a neural network for a particular problem is an important issue, since the network topology directly affects its computational complexity and its generalization capability. Here, the hidden layer node numbers of each model were determined after trying various network structures since there is no theory yet to tell how many hidden units are needed to

Table 1 The daily statistical parameters of each data set

Data set	x_{mean}	Sx	Cv (Sx/x_{mean})	Csx	x_{min}	x_{max}	Correlation with ET
W	2.15	0.95	0.44	1.07	1	7	0.327
SR	237.6	85.4	0.36	−0.08	29	391	0.946
RH	54.9	11.0	0.20	0.51	20.8	100	−0.646
AT	21.3	8.16	0.38	−0.01	4.7	36.2	0.779
ST	21.1	9.06	0.43	0.32	7.75	40	0.790
ET	4.90	2.33	0.48	−0.001	0	10.4	1.000

approximate any given function. In the training stage, the initial weights that gave the minimum error were found for each ANN network after trying fifty different initial weights. The tangent sigmoid and pure-linear functions are found appropriate for the hidden and output node activation functions, respectively.

Before applying the ANN to the data, the training input and output values were normalized using the equation

$$a\frac{x_i - x_{\min}}{x_{\max} - x_{\min}} + b \tag{10}$$

where $x_{\min}$ and $x_{\max}$ denote the minimum and maximum of the training and test data. Different values can be assigned for the scaling factors a and b. There are no fixed rules as to which standardization approach should be used in particular circumstances (Dawson and Wilby 1998). In this study the a and b were taken as 0.6 and 0.2, respectively. The learning and momentum rates were taken as 0.01 and 0.9, respectively. It is seen that choosing high values like 0.5 for the learning rate, as done by Raman and Sunilkumar (1995), throws the network into oscillations or saturates the neurons (Cigizoğlu 2003*b*). The ANN networks training were stopped after 150 and 50 epochs for the CGF and LM algorithms, respectively, since the variation of error was too small after these epochs. The error graphs for the ANN-CGF and ANN-LM models during training are shown in Figure 3.

Application and results

The weather parameters considered in this study are wind speed, solar radiation, air temperature, soil temperature and relative humidity. The study examined various combinations of these parameters as input to the model so as to evaluate the degree of effect of each of these variables on evapotranspiration. The input combinations evaluated in the present study are: (i) W; (ii) SR; (iii) AT; (iv) RH and AT; (v) W, RH and AT; (vi) SR, RH and AT; (vii) SR, RH, AT and ST; (viii) W, SR, RH and AT; (ix) W, SR, RH, AT and ST. The output layer had one neuron for the daily value of ET. The optimum ANN structures used for each input combination are given in Table 2. The node number in the hidden layer was found to vary between 2 and 5 after trial and error. Accordingly an ANN structure like ANN(3,4,1) consists of 3 inputs, 4 hidden and one output nodes. In this case the input layer covers the daily wind speed, relative humidity and air temperature and the output layer consists of the evapotranspiration.

The mean root square errors (MRSE) and mean absolute errors (MAE) are used as comparison criteria. These criteria are defined as

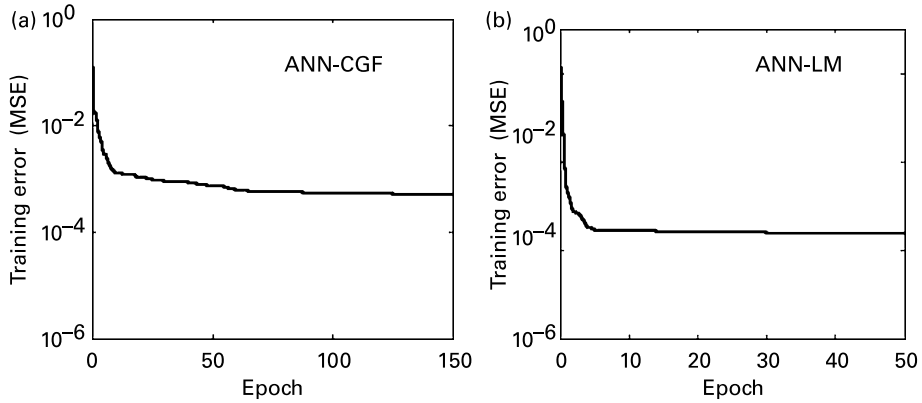

Figure 3 The training error graphs for the ANN-CGF and ANN-LM models (for the inputs W, SR, RH and AT)

Table 2 The optimum ANN structures used for each input combination

Input combination	ANN structure
(i) W	ANN(1,2,1)
(ii) SR	ANN(1,2,1)
(iii) AT	ANN(1,2,1)
(iv) RH and AT	ANN(2,4,1)
(v) W, RH and AT	ANN(3,4,1)
(vi) SR, RH and AT	ANN(3,3,1)
(vii) SR, RH, AT and ST	ANN(4,5,1)
(viii) W, SR, RH and AT	ANN(4,5,1)
(ix) W, SR, RH, AT and ST	ANN(5,5,1)

$$MRSE = \frac{\sqrt{\sum_{i=1}^{N} (Yi_{observed} - Yi_{predicted})^2}}{N} \tag{11}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} \left| Yi_{observed} - Yi_{predicted} \right| 100 \tag{12}$$

in which the N and Y denote the number of data sets and evapotranspiration data, respectively.

For each input combination, ANN was trained using two different training algorithms, that is, LM and CGF. After training was over, the weights were saved and used to test each network performance on test data. The ANN results were transformed back to the original domain, and the MRSE and MAE were computed.

The performance criteria for the test results of the ANN-LM, ANN-CGF and MLR models are given in Table 3. As can be seen from Table 3, the models have the highest MRSE and MAE and the lowest R^2 values for the input combination (i). The LM estimates for combination (i) are demonstrated in Figure 4. The model estimates are far from the corresponding observed values. Based on these results, it can be said that using only the wind speed input (combination (i)) is insufficient for modeling ET. However, adding the wind speed to the input combinations (combinations (v) and (viii)) increases the models' performances. This may be due to its (W) advection effect on ET. The second worst performance criteria were obtained for combination (iii). The LM estimates for this combination are shown in Figure 5. Using only the air temperature (combination (iii)) as in

Table 3 The performances of the models in the test period

	ANN-LM			ANN-CGF			MLR		
Input combination	**MRSE**	**MAE**	R^2	**MRSE**	**MAE**	R^2	**MRSE**	**MAE**	R^2
(i) W	0.111	1.799	0.189	0.111	1.799	0.189	0.129	2.04	0.144
(ii) SR	0.045	0.650	0.892	0.046	0.650	0.887	0.048	0.760	0.898
(iii) AT	0.080	1.273	0.605	0.080	1.268	0.607	0.084	1.348	0.580
(iv) RH and AT	0.064	0.927	0.748	0.069	1.040	0.713	0.082	1.278	0.613
(v) W, RH and AT	0.056	0.809	0.808	0.059	0.860	0.787	0.066	1.013	0.747
(vi) SR, RH and AT	0.031	0.426	0.942	0.034	0.463	0.933	0.032	0.431	0.941
(vii) SR, RH, AT and ST	0.034	0.450	0.934	0.035	0.501	0.935	0.035	0.468	0.942
(viii) W, SR, RH and AT	**0.019**	**0.269**	**0.983**	**0.023**	**0.337**	**0.972**	**0.023**	**0.347**	**0.973**
(ix) W, SR, RH, AT and ST	0.019	0.273	0.980	0.029	0.455	0.965	0.027	0.418	0.974

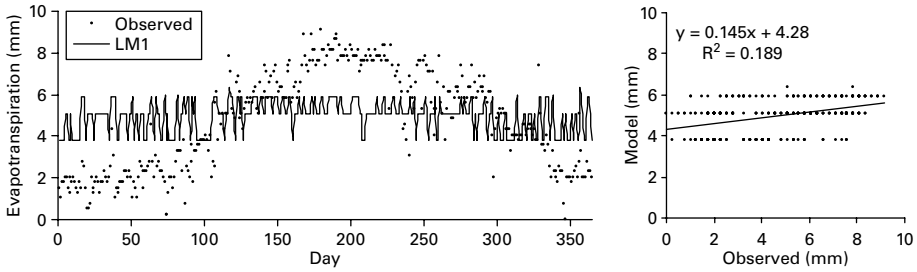

Figure 4 The observed and estimated ET values in the test period (for one input, W)

the temperature based methods (e.g. Thornthwaite 1948; Sudheer *et al.* 2003) that relate the ET to air temperature is insufficient for the estimation of ET. Out of the first five combinations, the models have the best performance criteria for combination (ii). The estimates of the LM whose input is SR are denoted in Figure 6. The estimated ET values are close to the observed values. There is a strong relationship between the solar radiation and ET. This is also confirmed by the high correlation (0.946) between these two variables (see Table 1). The LM estimates for combination (iv) are shown in Figure 7. Based on the figures (Figure 6 and Figure 7) and the criteria in Table 3, it may be said that using only solar radiation (combination (ii)) gives much better ET estimates than those obtained using the relative humidity and air temperature (combination (iv)).

According to the error statistics in Table 3, the ET estimates obtained using the input combinations (vi) and (viii) seem to be closer to the corresponding observed values than those provided using combinations (vii) and (ix), respectively. This implies that excluding the soil temperature from the input combinations provides a better estimation of ET. The models whose inputs are the wind speed, solar radiation, relative humidity and the air temperature (combination (viii)) have the best performance criteria among the input

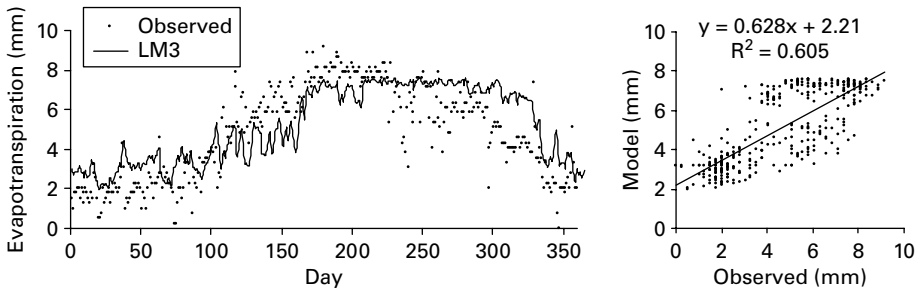

Figure 5 The observed and estimated ET values in the test period (for one input, AT)

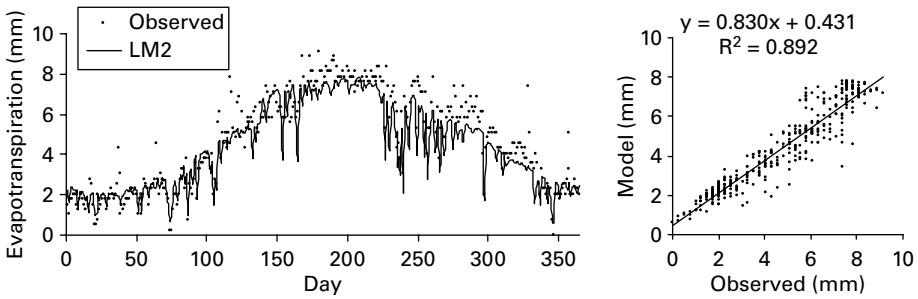

Figure 6 The observed and estimated ET values in the test period (for one input, SR)

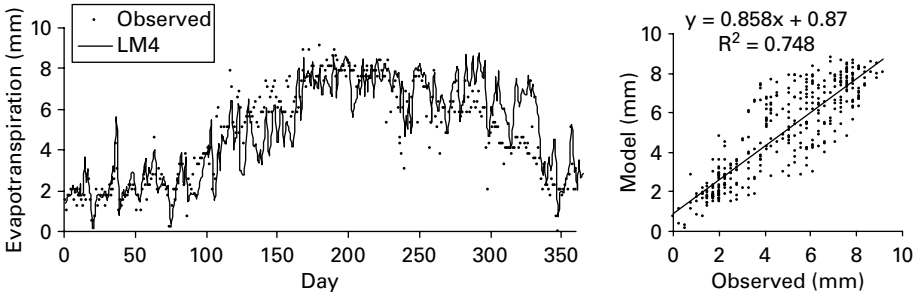

Figure 7 The observed and estimated ET values in the test period (for two inputs, RH and AT)

combinations tried in the study. It can be obviously seen from Table 3 that the ANN-LM generally has the lowest MRSE and MAE and the highest R^2 statistics. The ANN-CGF also seems to be better than MLR.

The Penman (Penman 1948) and Hargreaves (Hargreaves and Samani 1985) methods are considered for comparison in the present study. The Penman equation as described in the ASCE Manual 70 (Jensen *et al.* 1990) is

$$ET = \left(\frac{\Delta}{\Delta+\gamma}\right)R_n + \frac{\gamma}{\Delta+\gamma}\left[15.36(1+0.0062u_2)(e_w - e_a)\right]\frac{1}{\lambda} \tag{13}$$

where Δ is the slope of the vapour pressure curve, γ is the psychometric constant, R_n is the net radiation, u_2 is the wind speed at 2 m height, e_w is the saturation vapour pressure, e_a is the actual vapour pressure and λ is the latent heat of vaporization.

The original Hargreaves empirical formula (Hargreaves and Samani 1985), developed in California, calculates ET from solar radiation and air temperature:

$$ET = 0.0135\frac{R_s}{\lambda}(T+17.8) \tag{14}$$

where λ is the latent heat of vaporization, R_s is the solar radiation and T is the air temperature.

The locally calibrated Penman and Hargreaves methods are also considered for comparison. The data used for the training ANN models were used in the calibration of empirical models using the equation

$$ET = a + bET_{eq} \tag{15}$$

where ET = reference ET; ET_{eq} = ET estimated by the empirical methods; and a and b = calibration factors, respectively.

The ANN-LM and ANN-CGF models (combinations (viii)) are compared with the Penman and Hargreaves methods and their calibrated versions are denoted as C_Penman and C_Hargreaves in Table 4. The input variables used for each model are also given in this table. The ANN, MLR and Penman models use the same input variables. It can be seen from Table 4 that the ANN-LM model outperforms all other models in terms of various performance criteria. The ANN-CGF is ranked as the second best. It can be obviously seen that the calibration of the empirical models significantly improves the estimation accuracy (see C_Penman and C_Hargreaves in Table 4). Among the empirical methods, the C_Penman model performs the best.

The ET estimates of each model are represented in Figure 8 in the form of hydrographs. It is seen from the hydrographs that the ANN estimates more closely follow the observed ET values than those of the other models. The underestimations for the Penman and Hargreaves

Table 4 Values of error statistics for each model in the test period

Models	Model inputs	MRSE	MAE	R^2
ANN-LM	W, SR, RH and AT	0.019	0.269	0.983
ANN-CGF	W, SR, RH and AT	0.023	0.337	0.972
MLR	W, SR, RH and AT	0.023	0.347	0.973
Penman	W, SR, RH and AT	0.092	1.491	0.946
Hargreaves	SR and AT	0.155	2.636	0.904
C_Penman	W, SR, RH and AT	0.031	0.473	0.946
C_Hargreaves	SR and AT	0.039	0.547	0.904

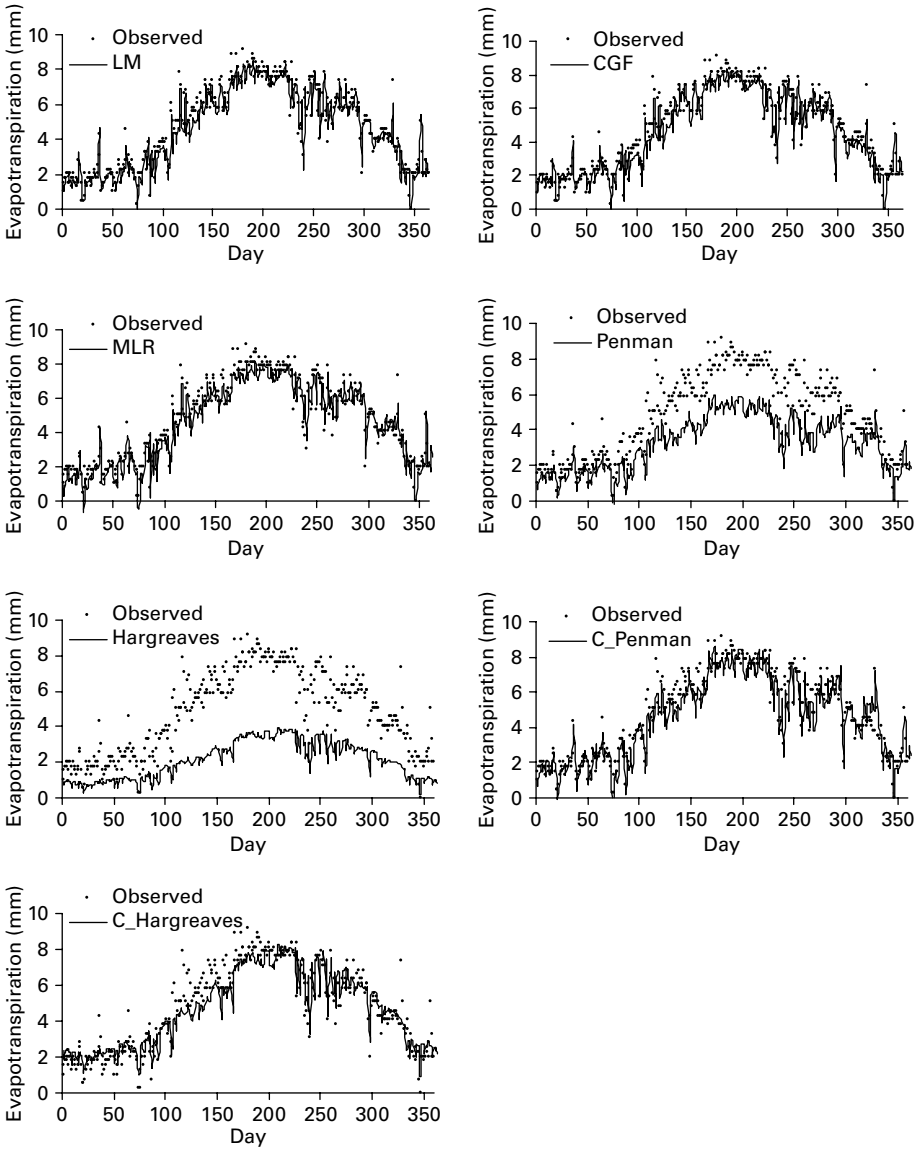

Figure 8 Plotting of ET estimates in the form of hydrographs for the test period

are obviously seen. The C_Penman and C_Hargreaves models seem to be much better than the Penman and Hargreaves models in ET estimation. The ET estimates are plotted in Figure 9 in the form of a scatterplot. As can be seen from the R^2 coefficients and fitted line equations (assume that the equations are $y = a_0x + a_1$) in the scatterplots the ANN-LM estimates are closer to the corresponding observed ET values than those of the other models. The a_0 and a_1 coefficients for the ANN-LM model are respectively closer to the 1 and 0 with a higher R^2 value of 0.983 than those of the other models. This confirms the MRSE and MAE statistics evaluated in Table 4. The MLR has better estimates than those of the Penman and Hargreaves models and their calibrated versions. However, its giving some negative ET values is a drawback for the MLR (Figure 9). The Penman and Hargreaves models seem to have good estimates from the R^2 viewpoint (as high as 0.946 and 0.904, respectively). There is a strong linear relationship between the Penman and Hargreaves estimates and observed ET values since the correlation coefficient expresses the degree of linear relationship between two variables. However, it can be seen from the fitted line equations that the Penman and Hargreaves methods considerably underestimate the ET.

The training results of the ANN-LM and -CGF that provide the best performances in the test period are represented in Table 5. As can be seen from the table, the LM has a smaller training time and iteration than the CGF. The LM is over three times faster than the CGF. The training (learning) performances of ANN-LM are also better than those of the ANN-CGF. It can be said that the LM is better than the CGF algorithm in function approximation.

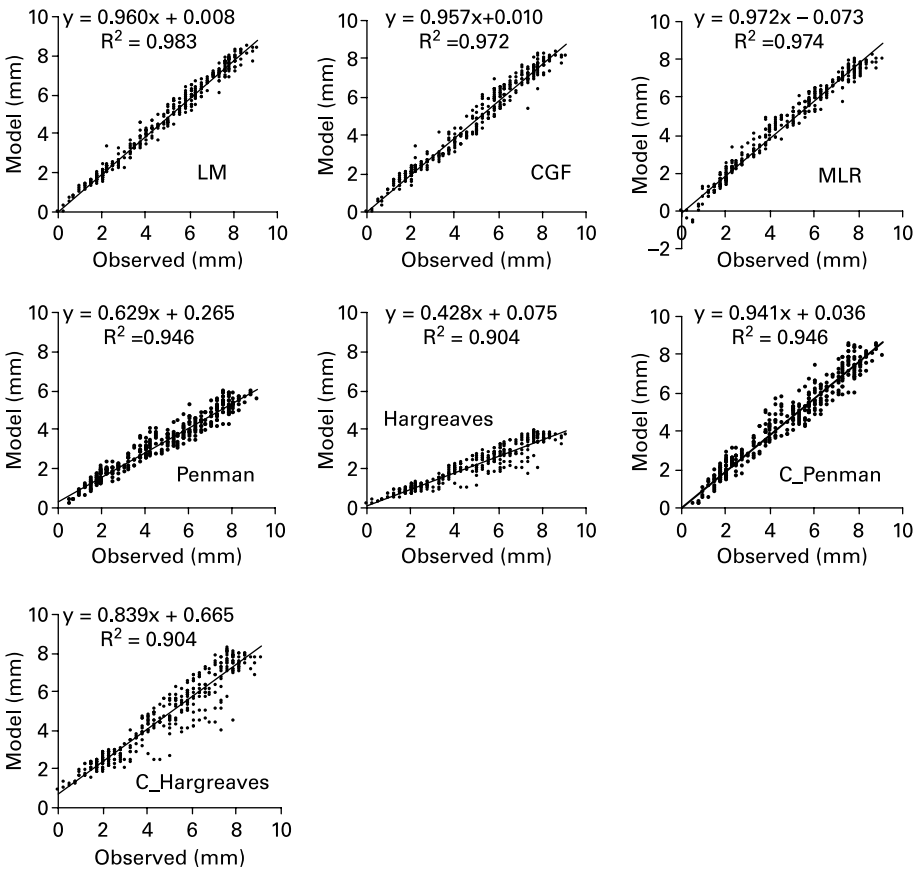

Figure 9 Plotting of ET estimates in the form of scatterplots for the test period

Table 5 Training results of the ANN training algorithms

Algorithm	Iteration	Time (s)	MRSE	MAE	R^2
Levenberg–Marquardt (LM)	50	1.26	0.014	0.203	0.987
Conjugate gradient (CGF)	150	4.34	0.019	0.279	0.977

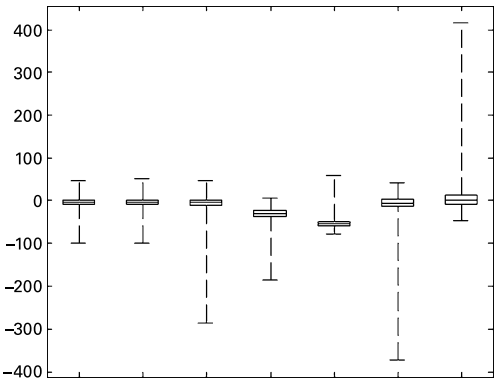

Figure 10 Box plots of accuracy in percentage for ET estimation

The ANN-LM, ANN-CGF and MLR has the 292, 248 and 249 estimates lower than the 10% error in the test period while the Penman, Hargreaves, C_Penman and C_Hargreaves models have the 19, 2, 198 and 182 estimates lower than the 10% error, respectively. Furthermore, the ANN-LM, ANN-CGF and MLR have the 169, 143 and 145 estimates lower than the 5% error while the Penman, Hargreaves, C_Penman and C_Hargreaves models have 7, 2, 96 and 104 estimates lower than the 5% error, respectively. The ANN-LM seems to perform better than the other models from the relative error viewpoint. The ANN-CGF is ranked as the second best. The results reveal that the calibration significantly increases the performance of the Penman and Hargreaves models.

The results of the models are also tested using the box plots (Figure 10) for the evaluation of the models' accuracy in percentage for ET estimation. As can be seen from the box plots in Figure 10, the ANN models have good accuracy relative to the other models. The Penman, Hargreaves and calibrated empirical models have the worst accuracy.

Concluding remarks

The main conclusions of the study can be itemized as below:

- This study indicated that modeling of the evapotranspiration is possible through the use of the ANN-LM and ANN-CGF.
- Using only the wind speed input was found to be insufficient for modeling ET. However, adding the wind speed to the input combinations increases the models' performances.
- Using only the air temperature in the temperature based methods that relate the ET to air temperature gives poor estimates.
- Excluding the soil temperature from the input combinations provides a better estimation of ET.
- The models whose inputs are the wind speed, solar radiation, relative humidity and air temperature have the best performance criteria among the input combinations tried in the study. This indicates that all these variables are needed for better evapotranspiration

modeling.

- The LM algorithm takes a small fraction of the time and iterations taken by the CGF for the training of the network.
- The calibrated empirical models C_Penman and C_Hargreaves performed much better than the Penman and Hargreaves models. It was found that the calibration significantly increased the performance of the empirical models in the estimation of ET.
- Based on the comparisons, the ANN-LM was found to be superior to the ANN-CGF, MLR, Penman, Hargreaves, C_Penman and C_Hargreaves models.

The ANN technique could be of use in the design of reservoirs and various other hydrological analyses where other models may be inappropriate. The study only used data from one area and further studies using more data from various areas are required to strengthen these conclusions.

References

Adeli, H. and Hung, S.L. (1995). *Machine Learning Neural Networks, Genetic Algorithms, and Fuzzy Systems*, John Wiley & Sons, New York.

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M. (1998). Crop evapotranspiration guidelines for computing crop water requirements. *FAO Irrigation and Drainage Paper No. 56*, Food and Agriculture Organization of the United Nations, Rome.

ASCE (1989). Evapotranspiration and irrigation water requirements. In *ASCE Manuals and Reports on Engineering Practice* M.E. Jensen, R.D. Burman and R.G. Allen (eds), vol 70, American Society of Civil Engineers, New York.

Bayazıt, M. and Oguz, B. (1998). *Probability and Statistics for Engineers*, Birsen Publishing House, Istanbul, Turkey, p. 159.

Brutsaert, W.H. (1982). *Evaporation into the Atmosphere*, D. Reidel, Dordrecht.

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

Cigizoğlu, H.K. (2003a). Estimation, forecasting and extrapolation of flow data by artificial neural networks. *Hydrol. Sci. J.*, **48**(3), 349–361.

Cigizoğlu, H.K. (2003b). Incorporation of ARMA models into flow forecasting by artificial neural networks. *Environmetrics*, **14**(4), 417–427.

Cigizoglu, H.K. and Kisi, O. (2005). Flow prediction by three back propagation techniques using k-fold partitioning of neural network training data. *Nordic Hydrol.*, **36**(1), 49–64.

Dawson, W.C. and Wilby, R. (1998). An artificial neural network approach to rainfall-runoff modeling. *Hydrol. Sci. J.*, **43**(1), 47–66.

El-Bakyr, M.Y. (2003). Feed forward neural networks modeling for K-P interactions. *Chaos, Solitons Fractals*, **18**(5), 995–1000.

Fletcher, R. and Reeves, C.M. (1964). Function minimization by conjugate gradients. *Computer J.*, **7**, 149–153.

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.

Hagan, M.T. and Menhaj, M. (1994). Training feedforward networks with the Marquardt algorithm. *IEEE Trans. Neural Networks*, **5**(6), 989–993.

Hargreaves, G.H. and Samani, Z. (1985). Reference crop evapotranspiration from temperature. *Appl. Engnh. Agric.*, **1**(2), 96–99.

Jackson, R.D. (1985). Evaluating evapotranspiration at local and regional scales. *Proc. IEEE*, **73**(6), 1086–1096.

Jain, S.K., Das, A. and Srivastava, D.K. (1999). Application of ANN for reservoir inflow prediction and operation. *J. Wat. Res. Planning Mngmnt.*, **125**(5), 263–271.

Jensen, M.E., Burman, R.D. and Allen, R.G. (1990). Evapotranspiration and irrigation water requirements. *ASCE Manuals and Reports on Engineering Practices* **70**, ASCE, New York.

Kang, K.W., Park, C.Y. and Kim, J.H. (1993). Neural network and its application to rainfall-runoff forecasting. *Korean J. Hydrosci.*, **4**, 1–9.

Kisi, O. (2004). River flow modeling using artificial neural networks. *ASCE J. Hydrol. Engng.*, **9**(1), 60–63.

Kisi, O. and Yildirim, G. (2005a). Discussion of 'Estimating actual evapotranspiration from limited climatic data using neural computing technique' by K.P. Sudheer, A.K. Gosain and K.S. Ramasastri. *ASCE J. Irrig. Drain. Engng.*, **131**(2), 219–220.

Kisi, O. and Yildirim, G. (2005b). Discussion of 'Forecasting of reference evapotranspiration by artificial neural networks' by S. Trajkovic; B. Todorovic; and M. Stankovic. *ASCE J. Irrig. Drain. Engng.*, **131**(4), 390–391.

Kumar, M., Raghuwanshi, N.S., Singh, R., Wallender, W.W. and Pruitt, W.O. (2002). Estimating evapotranspiration using artificial neural network. *ASCE J. Irrig. Drain. Engng.*, **128**(4), 224–233.

Lippman, R. (1987). An introduction to computing with neural nets. *IEEE ASSP Mag.*, **4**, 4–22.

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. (1973). *Principles of Enviromental Physics*, Edward Arnold, London.

More, J.J. (1977). In *The Levenberg-Marquardt Algorithm: Implementation and Theory. Numerical Analysis. Lecture Notes in Mathematics* G.A. Watson (ed.), vol 630, Springer Verlag, Berlin, pp. 105–116.

Naoum and Tsanis (2003). Hydroinformatics in evapotranspiration estimation. *Environ. Model. Software*, **18**, 261–271.

Penman, H.L. (1948). Natural evaporation from open water, bare soil and grass. *Proc. Royal Soc., Series A*, **193**, 120–145.

Poulovassilis, A., Anadranistakis, M., Liakatas, A., Alexandris, S. and Kerkides, P. (2001). Semi-emprical approach for estimating actual evapotranspiration in Greece. *Agric. Wat. Mngmnt*, **51**, 143–152.

Raman, H. and Sunilkumar, N. (1995). Multivariate modelling of water resources time series using artificial neural networks. *J. Hydrol. Sci.*, **40**(2), 145–163.

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.

Shuttleworth, W.J. and Wallace, J.S. (1985). Evaporation from sparse crops—an energy combination theory. *Quart. J.R. Meteorol. Soc.*, **111**, 839–855.

Smith, M., Allen, R. and Pereira, L. (1997). *Revised FAO Methodology for Crop Water Requirements*. Land and Water Development Division. FAO, Rome.

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

Sudheer, K.P., Gosain, A.K. and Ramasastri, K.S. (2003). Estimating actual evapotranspiration from limited climatic data using neural computing technique. *ASCE J. Irrig. Drain. Engng.*, **129**(3), 214–218.

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

Thornthwaite, C.W. (1948). An approach toward a rational classification of climate. *Geophys. Rev.*, **38**(1), 55–94.

Tokar, A.S. and Johnson, P.A. (1999). Rainfall-runoff modeling using artificial neural networks. *ASCE J. Hydrol. Engng.*, **4**(3), 232–239.

Trajkovic, S., Todorovic, B. and Stankovic, M. (2003). Forecasting reference evapotranspiration by artificial neural networks. *ASCE J. Irrig. Drain. Engng.*, **129**(6), 454–457.

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.

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