Hydrological Sciences Journal

ISSN: 0262-6667 (Print) 2150-3435 (Online) Journal homepage: http://www.tandfonline.com/loi/thsj20

Generalized regression neural networks for evapotranspiration modelling

ÖZGÜR KIŞI

Generalized regression neural networks for evapotranspiration modelling

ÖZGÜR KIŞI

Erciyes University, Engineering Faculty, Civil Eng. Dept., Hydraulics Division, 38039, Kayseri, Turkey
kisi@erciyes.edu.tr

Abstract The potential is investigated of the generalized regression neural networks (GRNN) technique in modelling of reference evapotranspiration (ET_0) obtained using the FAO Penman-Monteith (PM) equation. Various combinations of daily climatic data, namely solar radiation, air temperature, relative humidity and wind speed, are used as inputs to the ANN so as to evaluate the degree of effect of each of these variables on ET_0. In the first part of the study, a comparison is made between the estimates provided by the GRNN and those obtained by the Penman, Hargreaves and Ritchie methods as implemented by the California Irrigation Management System (CIMIS). The empirical models were calibrated using the standard FAO PM ET_0 values. The GRNN estimates are also compared with those of the calibrated models. Mean square error, mean absolute error and determination coefficient statistics are used as comparison criteria for the evaluation of the model performances. The GRNN technique (GRNN 1) whose inputs are solar radiation, air temperature, relative humidity and wind speed, gave mean square errors of 0.058 and 0.032 mm^2 day^{-2}, mean absolute errors of 0.184 and 0.127 mm day^{-1}, and determination coefficients of 0.985 and 0.986 for the Pomona and Santa Monica stations (Los Angeles, USA), respectively. Based on the comparisons, it was found that the GRNN 1 model could be employed successfully in modelling the ET_0 process. The second part of the study investigates the potential of the GRNN and the empirical methods in ET_0 estimation using the nearby station data. Among the models, the calibrated Hargreaves was found to perform better than the others.

Key words evapotranspiration; generalized regression neural networks; Hargreaves method; Penman method; Ritchie method

Réseaux de neurones de régression généralisée pour la modélisation de l'évapotranspiration

Résumé Le potentiel de la technique des réseaux de neurones de régression généralisée (RNRG) est étudié pour modéliser l'évapotranspiration de référence (ET_0) obtenue par application de l'équation FAO Penman-Monteith (PM). Différentes combinaisons de données climatiques journalières, en l'occurrence de rayonnement solaire, température de l'air, humidité relative et vitesse du vent, sont utilisées en entrée du réseau de neurones artificiel afin d'évaluer l'influence de chacune de ces variables sur ET_0. Dans la première partie de cette étude, une comparaison est faite entre les estimations fournies par le RNRG et celles produites par les méthodes de Penman, Hargreaves et Ritchie, telles qu'elles sont implémentées dans le California Irrigation Management System (CIMIS). Les modèles empiriques ont été calés grâce aux valeurs standard FAO PM de l'ET_0. Les estimations du RNRG ont également été comparées avec celles des modèles calés. Les statistiques d'erreur quadratique moyenne, d'erreur absolue moyenne et de coefficient de détermination ont été utilisées comme critères de comparaison pour évaluer les performances du modèle. La technique RNRG (RNRG 1) dont les entrées sont le rayonnement solaire, la température de l'air, l'humidité relative et la vitesse du vent, ont donné des erreurs quadratiques moyennes de 0.058 et 0.032 mm^2 jour^{-2}, des erreurs absolues moyennes de 0.184 et 0.127 mm jour^{-1}, et des coefficients de détermination de 0.985 et 0.986 pour les stations de Pomona et de Santa Monica (Los Angeles, USA), respectivement. Sur la base des comparaisons, il est apparu que le modèle RNRG 1 pourrait être utilisé avec succès pour modéliser le processus ET_0. La deuxième partie de l'étude examine le potentiel du RNGR et des méthodes empiriques pour estimer l'ET_0 en utilisant les données de la station voisine. Parmi les modèles, celui de Hargreaves calé est apparu donner de meilleurs résultats que les autres.

Mots clefs évapotranspiration; réseaux de neurones de régression généralisée; méthode de Hargreaves; méthode de Penman; méthode de Ritchie

INTRODUCTION

Evapotranspiration (ET) is the sum of the volume of water used by vegetation (transpired), evaporated from the soil and the intercepted precipitation on vegetation (Singh, 1988) and it plays an important role in the environment at global, regional and local scales. Water entering the evaporation phase of the hydrological cycle becomes unavailable and cannot be recovered for further use (Brutsaert, 1982). In many areas where water resources are scarce, the calculation of this loss becomes imperative in the planning and management of irrigation practices. The main concern in temperate regions is the shortage of water, required for various uses, during the dry seasons. Despite this significance, evapotranspiration is one of the less understood components of the hydrologic cycle (Brutsaert, 1982; Naoum & Tsanis, 2003).

Numerous methods have been proposed for modelling evapotranspiration, as described by Brutsaert (1982) and Jensen *et al.* (1990). In general, the combination of energy balance/aerodynamic equations "provides the most accurate results as a result of their foundation in physics and basis on rational relationships" (Jensen *et al.*, 1990). The Food and Agricultural Organization of the United Nations (FAO) assumed the ET definition from Smith *et al*. (1997) and compiled the so-called FAO Penman-Monteith method as the standard for estimation of ET (Allen *et al*., 1998; Naoum & Tsanis, 2003); this is also known as the FAO-56 method (from the number of the related technical report).

In the past decade, considerable attention has been focused on the application of artificial neural networks (ANN) in diverse fields including system modelling, fault diagnosis and control, pattern recognition, financial forecasting and hydrology. Studies on ANN application in the area of hydrology include rainfall–runoff modelling (Minns & Hall, 1996; Dawson & Wilby, 1998; streamflow prediction (Cigizoğlu, 2003; Sahoo & Ray, 2006) and suspended sediment modelling (Tayfur, 2002; Kisi, 2005). However, the application of ANN to evapotranspiration modelling is limited in the literature. Kumar *et al.* (2002) used a multi-layer perceptron ANN (MLP) with back-propagation training algorithm for the estimation of ET_0. They used various ANN architectures and found that the ANN gives accurate ET_0 estimates. Sudheer *et al.* (2003) used radial basis ANN in modelling ET_0 using limited climatic data. Trajkovic *et al.* (2003) developed a radial basis type ANN in forecasting ET_0. Trajkovic (2005) used temperature-based radial basis ANN for modelling FAO-56 PM ET_0. He compared the ANN results with those of the Hargreaves, Thornthwaite and reduced PM methods and concluded that the radial basis ANN generally performs better than the other models in modelling ET_0. To the knowledge of the author, no work has been reported in the literature that addresses the application of generalized regression neural networks (GRNN) to ET_0 estimation. This provided an impetus to investigate the potential of the GRNN for better mapping of the process.

The potential of the GRNN in modelling of ET_0 is investigated and discussed in the present study. The ET_0 values are obtained using the standard FAO-56 Penman-Monteith equation. Various combinations of daily climatic data are tried so as to evaluate the degree of effect of each of these data on ET_0. The performance of the GRNN is compared with those of the commonly used CIMIS Penman, Hargreaves and Ritchie empirical methods.

GENERALIZED REGRESSION NEURAL NETWORKS (GRNN)

The GRNN architecture (Specht, 1991) subsumes the radial-basis function (RBF) method. It approximates any arbitrary function between input and output vectors, drawing the function estimate directly from training data. It looks much like the common feedforward topology used with back-propagation training; however, its operation is fundamentally different. The GRNN is based on nonlinear regression theory for function estimation. The training set consists of values of inputs x, each with a corresponding value of an output y. This regression method produces the estimated value of y, which minimizes the squared error.

The GRNN can be viewed as the normalized RBF network in which there is a unit centred at every training case. The network architecture is a one-pass learning algorithm with a highly parallel structure. Even with sparse data in a multi-dimensional measurement space, the algorithm provides smooth transitions from one observed value to another. The algorithmic form can be used for any regression problem in which an assumption of linearity is not justified. The GRNN is a universal approximator for smooth functions, so it is capable of solving any smooth function approximation problem (Disorntetiwat, 2001).

The feedforward back-propagation method performance is very sensitive to randomly assigned initial weight values. However, this problem was not faced in GRNN simulations (Cigizoğlu, 2005). The GRNN does not require an iterative training procedure as in the back-propagation method explained by Specht (1991). The local minima problem was not faced in GRNN simulations. For these reasons, the GRNN was preferred instead of feedforward back-propagation method.

A hidden neuron layer is created to hold the input vector. The weight between the newly created hidden neuron and the output neuron is assigned the target value.

GRNN is based on the following formula:

$$E[y|x] = \frac{\int_{-\infty}^{\infty} y f(x,y)\,\mathrm{d}y}{\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y} \tag{1}$$

where y is the output of the estimator, x is the estimator input vector, $E[y|x]$ is the expected value of output, given the input vector x, and $f(x,y)$ is the joint probability density function (pdf) of x and y. The function value is estimated optimally as follows:

$$y_j = \frac{\sum_{i=1}^{n} h_i w_{ij}}{\sum_{i=1}^{n} h_i} \tag{2}$$

where w_{ij} is the target output corresponding to input training vector x_i and output j; $h_i = \exp[-D_i^2/(2\sigma^2)]$ is the output of a hidden layer neuron; $D_i^2 = (x - u_i)^{\mathrm{T}}(x - u_i)$ is the squared distance between the input vector xi and the training vector ui; and σ is a constant controlling the size of the perceptive region. The major difference between GRNN and RBF neural networks is the method that the weights w_{ij}, are determined. Instead of training weights, the GRNN assigns the target value directly to the weights, w_{ij}, from the training set associated with input training vector and a component of its corresponding output vector.

The GRNN architecture is selected in this study because of its fast learning and convergence to the optimal regression surface. In essence, GRNN is a method for estimating $f(x,y)$, given only a training set. Due to the pdf being derived from the training data with no preconceptions about its form, the system is perfectly general. There is no problem if the functions are composed of multiple disjoint non-Gaussian regions in any number of dimensions, as well as those of simpler distributions (Wasserman, 1993). Particular to the GRNN is the use of the smoothing factor, σ, which alters the degree of generalization of the network. High smoothing factors, approaching 1, will straighten the path of the prediction line, while smoothing factors approaching 0 essentially create a dot-to-dot map. While high smoothing factors increase the network's ability to generalize, they may also degrade the error of prediction. Conversely, low smoothing factors degrade the network's ability to generalize, or make predictions at all. Therefore, a range of smoothing factors and methods for selecting smoothing factors are tested in the study in order to determine the optimum smoothing factors for model inputs.

CASE STUDY

The daily climatic data of two automated weather stations, Pomona Station (34°03′N; 117°48′W) and Santa Monica Station (34°02′N; 118°28′W) operated by CIMIS are used in the present study. The elevations are 222 and 104 m for the Pomona and Santa Monica stations, respectively. The California Irrigation Management System was developed in 1982 by the California Department of Water Resource and the University of California at Davis to assist California's irrigators manage their water resources efficiently. The CIMIS manages a network of over 120 automated weather stations in the state of California. The following weather data used in the present study are measured at the CIMIS weather stations: the total incoming solar radiation is measured using pyranometers at height of 2.0 m above the ground; air temperature is measured at a height of 1.5 m above the ground using a thermistor; relative humidity is measured by a sensor sheltered in the same enclosure with the air temperature sensor at 1.5 m above the ground; and wind speed is measured using three-cup anemometers at 2.0 m above the ground. These measured daily climatic data and the ET_0 values calculated using the CIMIS Penman were downloaded from the CIMIS web server (http://www.ipm.ucdavis.edu/WEATHER/wxretrieve.html).

The data sample consisted of four years (2001–2004) of daily records of solar radiation, air temperature, relative humidity and wind speed. For each station, the first three years (2001–2003) of data were used to train the GRNN and the remaining data were used for testing.

APPLICATION AND RESULTS

First, the ET_0 values for the Pomona and Santa Monica stations were calculated using the FAO-56 PM method as described in Allen *et al*. (1998):

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \dfrac{900}{T+273} U_2 (e_a - e_d)}{\Delta + \gamma(1 + 0.34U_2)} \tag{3}$$

where ET_0 is the reference evapotranspiration (mm day^{-1}); Δ is the slope of the saturation vapour pressure function (kPa °C^{-1}); R_n is the net radiation (MJ m^{-2}day^{-1}); G is the soil heat flux density (MJ m^{-2} day^{-1}); γ is the psychometric constant (kPa °C^{-1}); T is the mean air temperature (°C); U_2 is the average 24-h wind speed at 2 m height (m s^{-1}); e_a is the saturation vapour pressure (kPa); and e_d is the actual vapour pressure (kPa).

Then, the input and output data were normalized to fall in the range [0,1]. The ET_0 obtained by the FAO-56 PM was standardized by the following formula:

$$ET_{0s} = ET_0 / ET_{0\max} \tag{4}$$

where ET_{0s} is standardized ET_0; and $ET_{0\max}$ is the maximum of the ET_0 values. The other climatic data were also standardized in a similar way. Finally, these normalized data were used for the calibration of GRNN models. A program code, including ANN toolboxes, was written in MATLAB language for the GRNN simulations. Mean square errors (MSE), mean absolute error (MAE) and determination coefficient (R^2) statistics were used as comparison criteria. The R^2 measures the degree to which two variables are linearly related; MSE and MAE provide different types of information about the predictive capabilities of the model. The MSE measures the goodness-of-fit relevant to high ET_0 values, whereas the MAE yields a more balanced perspective of the goodness-of-fit at moderate ET_0 (Karunanithi *et al.*, 1994). The MSE and MAE are defined as:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(ETi_{\text{observed}} - ETi_{\text{predicted}}\right)^2 \tag{5}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|ETi_{\text{observed}} - ETi_{\text{predicted}}\right|100 \tag{6}$$

in which the N denotes the number of data sets.

Next, the potential of the GRNN is tested for two different applications. In the first application, the ET_0 data of two stations are estimated separately and the GRNN estimates are compared with those of the empirical models. In the second application, the ET_0 data of one station are estimated using the data from the other station and the GRNN test results are compared with those of the empirical models and multilinear regression (MLR).

Estimation of the Pomona and Santa Monica station ET_0 data

The weather parameters considered in the study are solar radiation (R_s), air temperature (T), relative humidity (RH) and wind speed (U_2). The study examined four combinations of these parameters as inputs to the GRNN models. Building the GRNN model several times with one different variable added into the input combination per time. Thus, the degree of affect of each of these variables on ET_0 was evaluated in respect of the reduction or increment in error statistics. Accordingly, the input combinations evaluated in the present study are: (i) R_s; (ii) R_s and T; (iii) R_s, T and RH; and (iv) R_s, T, RH and U_2. The output is the daily value of ET_0 obtained by the FAO-56 PM. When using GRNN models, very few user decisions are required. Among the decisions that

Table 1 The GRNN spreads used for each input combination.

Input combinations	Pomona Station	Santa Monica Station
(i) R_s	0.01	0.01
(ii) R_s and T	0.02	0.02
(iii) R_s, T and RH	0.02	0.04
(iv) R_s, T, RH and U_2	0.04	0.04

Table 2 The performance statistics of the GRNN models in the test period—Pomona Station.

Input combinations	MSE (mm^2 day^{-2})	MAE (mm day^{-1})	R^2
(i) R_s	0.164	0.297	0.960
(ii) R_s and T	0.085	0.212	0.981
(iii) R_s, T and RH	0.062	0.178	0.983
(iv) R_s, T, RH and U_2	0.058	0.184	0.985

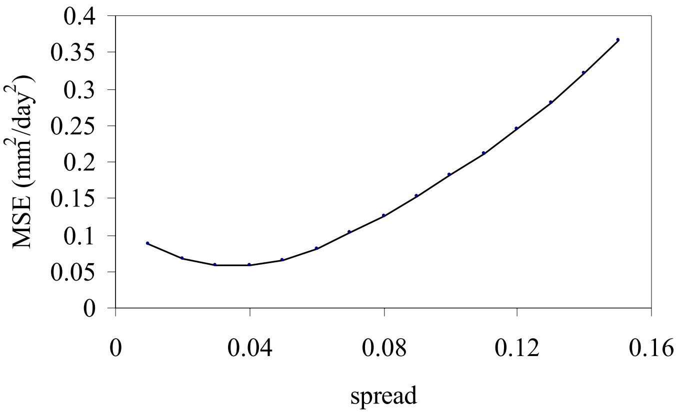

Fig. 1 The effect of spread constant on GRNN performance.

are required is the selection of the appropriate smoothing factors to be applied to each of the model inputs. For each input combination, the optimum spread for the GRNN model was determined according to the MSE criterion. The determined spread parameters for each combination are given in Table 1. The spreads were found to vary between 0.01 and 0.04. Figure 1 shows the variation of MSE statistic *vs* spread parameter for the input combination (iv) of the Pomona Station. It appears from the figure that the GRNN has the minimum MSE with the spread value of 0.04. As can be seen from Table 1, in general, the spread values are increasing in parallel to the number of input variables.

The GRNN is first used for Pomona Station. For this station, the MSE, MAE and R^2 statistics of each GRNN model in the test period are given in Table 2. The table indicates that the GRNN model whose inputs are the R_s, T, RH and U_2 (input combination (iv)) has the smallest MSE (0.058 mm^2 day^{-2}) and the highest R^2 (0.985). This emphasizes the factors influencing ET_0, since the model considered all the parameters. As may be seen from Table 2, using only the solar radiation (input combination (i)) input gives accurate estimates from the various statistics viewpoints. For the Pomona Station, T seems to be more effective than RH and U_2 in estimation of ET_0

because of the fact that adding T into the input combination (input combination (ii)) significantly increases the model performance (the reduction in the MSE (48%), MAE (29%) and the increment in R^2 (2.2%) are the largest). It is seen from the error statistics of combinations (iii) and (iv) in Table 2 that RH seems to be more effective on ET_0 than U_2.

The CIMIS Penman (Snyder & Pruitt, 1985), Hargreaves (Hargreaves & Samani, 1985) and Ritchie (Jones & Ritchie, 1990) methods are considered for comparison in the present study. The CIMIS Penman equation employs the modified Penman equation (Pruitt & Doorenbos, 1977) with a wind function that was developed at the University of California, Davis. The method uses hourly average weather data as an input to calculate hourly ET_0. The 24-h ET_0 values for the day (midnight-to-midnight) are then summed to produce estimates of daily ET_0. The hourly PM equation that CIMIS uses to estimate hourly PM ET_0 is the FAO version that is described in Irrigation and Drainage Paper no. 56 (Allen *et al.*, 1998). The CIMIS Penman equation, described in detail in Hidalgo *et al.* (2005) (see http://wwwcimis.water.ca.gov/cimis/ infoEtoCimisEquation.jsp), is:

$$ET_0 = \left(\frac{\Delta}{\Delta+\gamma}\right)R_n + \left(1-\frac{\Delta}{\Delta+\gamma}\right)(e_a - e_d)f_U \tag{7}$$

where f_U is a wind function (m s^{-1}) and all other symbols have been explained with equation (3). Daily ET_0 equals the sum of 24-h ET_0 (mm).

The Hargreaves empirical formula (Hargreaves & Samani, 1985), one of the simplest to estimate ET_0, is:

$$ET_0 = 0.0023R_a\left(\frac{T_{\max}+T_{\min}}{2}+17.8\right)\sqrt{T_{\max}-T_{\min}} \tag{8}$$

where $T_{\max}$ and $T_{\min}$ are the maximum and minimum temperature (°C) and R_a is extra-terrestrial radiation (expressed as equivalent depth of evaporated water, mm day^{-1}).

The Ritchie method, as described by Jones & Ritchie (1990), is:

$$ET_0 = \alpha_1\left[3.87\times10^{-3}R_s\left(0.6T_{\max}+0.4T_{\min}+29\right)\right] \tag{9}$$

where R_s is solar radiation (total electromagnetic radiation emitted by the sun, MJ m^{-2} day^{-1}); α_1 is a coefficient whose value depends on $T_{\max}$ as follows: $\alpha_1 = 1.1$ when $5 < T_{\max} \leq 35$°C; $\alpha_1 = 1.1 + 0.05(T_{\max} - 35)$ when $T_{\max} > 35$°C; and $\alpha_1 = \exp[0.18(T_{\max} + 20)]$ when $T_{\max} < 5$°C.

The temperature-based methods mostly underestimated or overestimated ET_0 obtained by the FAO-56 PM method. In those cases, Allen *et al.* (1994) recommended that empirical methods be calibrated using the standard PM method; ET_0 is calculated as:

$$ET_0 = a + bET_{eq} \tag{10}$$

where ET_0 is grass reference ET defined by the FAO-56 PM equation; ET_{eq} is ET estimated by the temperature-based methods; and a and b are calibration factors, respectively.

The data used for the training GRNN models were used for calibration of temperature-based methods. The CIMIS Penman method was also calibrated. Thus, the

Table 3 The performance statistics of the models in test period—Pomona Station.

Models	Model inputs	MSE (mm^2 day^{-2})	MAE (mm day^{-1})	R^2
GRNN 1	R_s, T, RH and U_2	0.058	0.184	0.985
GRNN 2	R_s and T	0.085	0.212	0.981
CIMIS Penman	R_s, T, RH and U_2	0.439	0.361	0.970
Hargreaves (1985)	R_s and T	27.43	4.62	0.981
Ritchie (1990)	R_s and T	1.381	0.756	0.985
C_CIMIS Penman	R_s, T, RH and U_2	0.109	0.253	0.970
C_Hargreaves (1985)	R_s and T	0.078	0.216	0.981
C_Ritchie (1990)	R_s and T	0.071	0.205	0.985

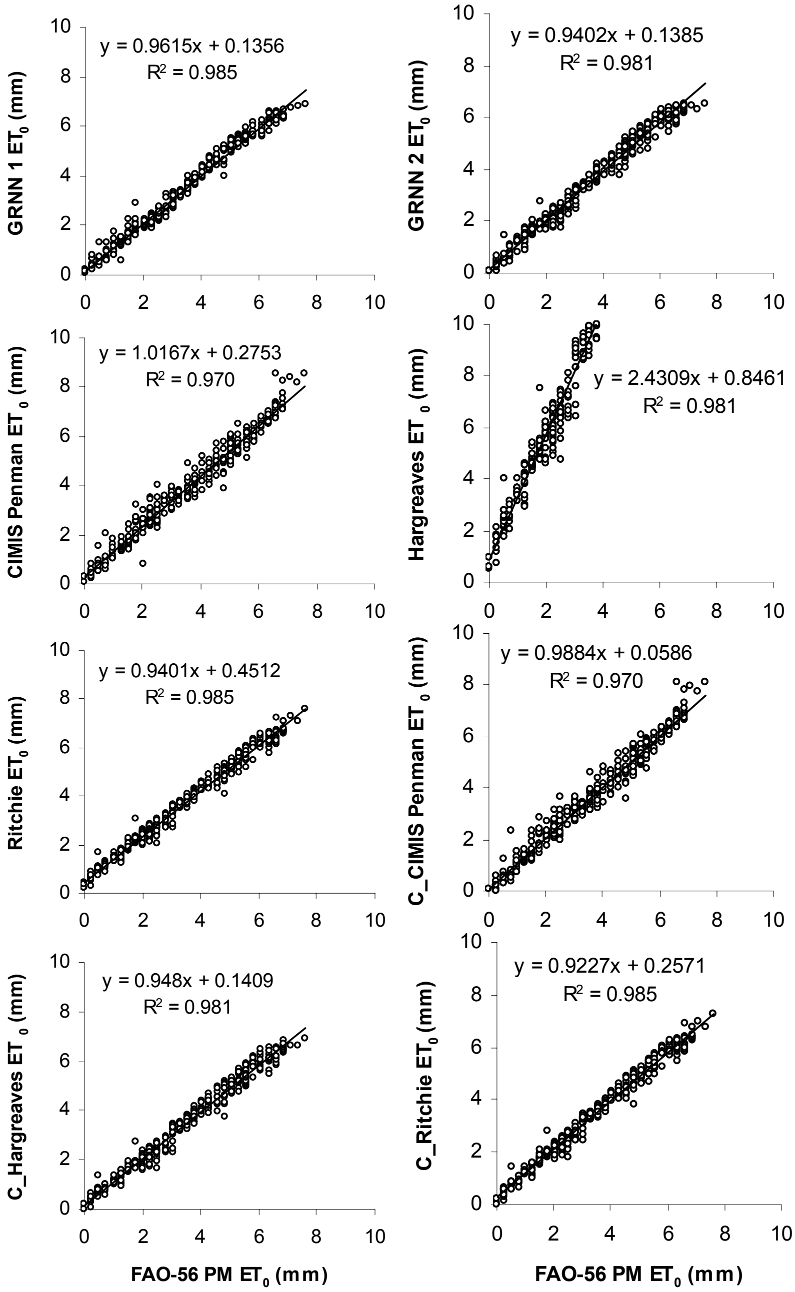

Fig. 2 The FAO-56 PM and estimated ET_0 values of the Pomona Station.

calibrated empirical methods were also considered for the comparison in the present study.

The GRNN1 and GRNN2 models (input combinations (ii), (iv)) are compared with the CIMIS Penman, Hargreaves, Ritchie methods and their calibrated versions denoted as C_CIMIS Penman, C_Hargreaves and C_Ritchie, in respect of the MSE, MAE and R^2 statistics, for the Pomona Station in Table 3. The input variables used for each model are also given in this table. The GRNN2, Hargreaves and Ritchie models use the same input variables. It can be seen from Table 3 that the GRNN1 model outperforms all other models in terms of various performance criteria. It can be clearly seen that the calibration of the empirical models significantly increases the estimation accuracy. The C_Hargreaves and C_Ritchie models seem to be slightly better than the GRNN2. Among the empirical methods, the C_Hargreaves and C_Ritchie models have similar results and they perform better than the CIMIS Penman. The ET_0 estimates of each model are represented in Fig. 2 in the form of a scatterplot. It is seen from the scatter-plots that the GRNN1 estimates are closer to the corresponding observed ET_0 values than those of the other models. The Hargreaves model overestimated ET_0 obtained by the FAO-56 PM method, as reported by Allen *et al.* (1994). However, the calibration considerably increased the performance of the Hargreaves model (see Table 3 and Fig. 2). The estimation of total ET_0 was also considered for comparison due to its importance in irrigation management. The GRNN1, GRNN2, C_Hargreaves and C_Ritchie methods computed the total FAO-56 PM ET_0 of 1288.80 mm as 1288.76, 1262, 1273 and 1283 mm, respectively, with underestimations of 0.003%, 2%, 1.2% and 0.4%, while the CIMIS Penman, Hargreaves, Ritchie and C_CIMIS Penman methods resulted in 1411, 3443, 1377 and 1295 mm, with overestimations of 9.5%, 167%, 6.8% and 0.5%, respectively. The GRNN1 estimate is almost the same as the total FAO-56 PM ET_0 value. The GRNN2 is ranked as the second best. The calibration significantly increases the performance of the Hargreaves and Ritchie methods in total ET_0 estimation.

For the Santa Monica Station, the performances of the GRNN models in the test period are summarized in Table 4. Here also, the trend of the results is the same as that of the Pomona Station. The best performance criteria (MSE = 0.032 $mm^2\ day^{-2}$, MAE = 0.127 mm day^{-1} and R^2 = 0.986) are obtained by the GRNN model whose inputs are R_s, T, RH and U_2. In contrast to the Pomona Station, RH seems to be more effective than T in estimation of ET_0 (the reduction in the MSE (85%), MAE (65%) and the increment in R^2 (16%) are the largest; see input combination (iii) in Table 4).

The comparison of GRNN estimates with those of the empirical models is made for the Santa Monica Station in Table 5. Results from this table clearly indicate that the GRNN1 provides the most accurate ET_0 estimates, as found for the Pomona Station. For this station, the GRNN2 model also performs better than the empirical models. Unlike the Pomona Station, the C_Hargreaves and C_Ritchie methods perform worse

Table 4 The performance statistics of the GRNN models in test period—Santa Monica Station.

Input combinations		MSE ($mm^2\ day^{-2}$)	MAE ($mm\ day^{-1}$)	R^2
(i)	R_s	0.429	0.474	0.793
(ii)	R_s and T	0.301	0.392	0.854
(iii)	R_s, T and RH	0.044	0.137	0.980
(iv)	R_s, T, RH and U_2	0.032	0.127	0.986

Table 5 The performance statistics of the models in test period—Santa Monica Station.

Models	Model inputs	MSE (mm^2 day^{-2})	MAE (mm day^{-1})	R^2
GRNN 1	R_s, T, RH and U_2	0.032	0.127	0.986
GRNN 2	R_s and T	0.301	0.392	0.854
CIMIS Penman	R_s, T, RH and U_2	0.410	0.392	0.830
Hargreaves (1985)	R_s and T	23.57	4.35	0.818
Ritchie (1990)	R_s and T	0.641	0.650	0.780
C_CIMIS Penman	R_s, T, RH and U_2	0.371	0.415	0.830
C_Hargreaves (1985)	R_s and T	0.381	0.471	0.818
C_Ritchie (1990)	R_s and T	0.456	0.474	0.780

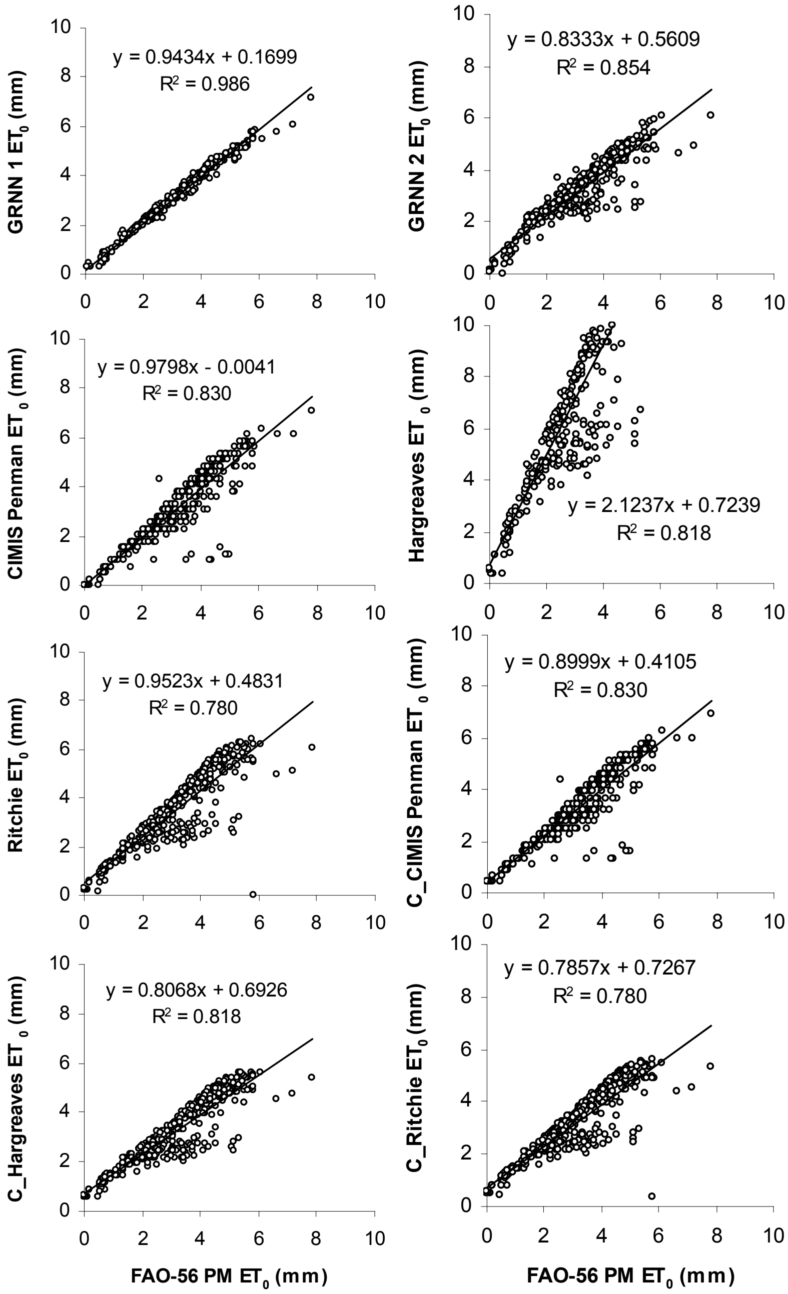

Fig. 3 The FAO-56 PM and estimated ET_0 values of the Santa Monica Station.

than the GRNN2 model. The estimates of each model are plotted in Fig. 3 in terms of scatter diagrams. As seen from the figure, the GRNN1 has less scattered estimates than the other models. Here also, overestimations are obviously seen for the Hargreaves model. While the GRNN1 and CIMIS Penman estimated the total FAO-56 PM ET_0 as 1176 and 1155 mm, compared to the computed FAO-56 PM ET_0 of 1181 mm, with underestimations of 0.4% and 2.1%, the GRNN2, Hargreaves, Ritchie, C_CIMIS Penman, C_Hargreaves and C_Ritchie models resulted in 1189, 2772, 1301, 1213, 1206 and 1194 mm, respectively, with overestimations of 0.7%, 135%, 10%, 2.7%, 2.1% and 1.1%. The GRNN estimates are closest to the observed one. Again, the calibration significantly increased the performance of the temperature-based models, Hargreaves and Ritchie, in total ET_0 estimation, as found in the preceding application.

Estimation of the Pomona Station ET_0 data using the data of Santa Monica

The ET_0 estimation using nearby station data is an important issue since the data of some stations are missed. To solve this problem, the regression techniques are frequently used. This part of study focused on the investigation of GRNN and empirical models performances to solve this problem. The R_s, T, RH and U_2 data of the Santa Monica Station are used to estimate the FAO-56 PM ET_0 of the Pomona Station. In this application three years (2001–2003) of data are again used for the calibration of the GRNN and empirical models and the remaining one year data are used for testing. The optimum spread parameter was found as 0.06 and 0.03 for the GRNN1 and GRNN2 models, respectively. The empirical models and multilinear regression (MLR) are also applied to the same data. The performance statistics of the models are given in Table 6. The C_Hargreaves model has the lowest MSE (0.308 mm^2 day^{-2}) and MAE (0.478 mm day^{-1}) and the highest R^2 statistic (0.977). The Ritchie model is ranked as the second best. The increment in the Hargreaves performance after calibration is obviously seen. However, the calibration decreases the estimation performances of the CIMIS Penman and Ritchie models. The MLR model estimates are not accurate. The FAO-56 PM ET_0 and estimated ET_0 of the Pomona Station using the data of Santa Monica Station are shown in Fig. 4. It can be seen from the fit line equations (assume that the equation is $y = a_0x + a_1$) in Fig. 4 that the a_0 and a_1 coefficients for the C_Hargreaves model are respectively closer to the 1 and 0 with a higher R^2 of 0.977 than those of the GRNN, MLR and other empirical models. This confirms the MSE and MAE statistics evaluated in Table 6. The Hargreaves, C_ Hargreaves and Ritchie estimated the total FAO-56 PM ET_0 1289 mm as 2815, 1424 and 1301 mm, with overestimations of 118%, 11% and 1%, while the GRNN1, GRNN2, CIMIS Penman, C_CIMIS Penman, C_Ritchie, MLR1 and MLR2 resulted in 1154, 1155, 1183, 1121, 1119, 1099 and 1133 mm, with underestimations of 10.5%, 10.4%, 8.2%, 13%, 13.2%, 14.7% and 12.1%, respectively. The Ritchie method seems to be best in estimation of total ET_0 value of the Santa Monica Station. The CIMIS Penman, C_CIMIS Penman, C_Hargreaves, C_Ritchie, MLR2 gave the total ET_0 estimates close to each other. Unlike the preceding applications, the GRNN models did not provide accurate results in cross-station validation. The results imply that the C_Hargreaves and Ritchie empirical methods may be used instead of GRNN models in cross-station applications.

Table 6 The performance statistics of the models in ET_0 estimation using nearby station's data - Pomona Station.

Models	Model inputs	MSE (mm^2 day^{-2})	MAE (mm day^{-1})	R^2
GRNN 1	R_s, T, RH and U_2	0.726	0.664	0.861
GRNN 2	R_s and T	0.745	0.679	0.848
CIMIS Penman	R_s, T, RH and U_2	0.946	1.299	0.678
Hargreaves (1985)	R_s and T	26.47	4.471	0.977
Ritchie (1990)	R_s and T	0.640	0.649	0.815
C_CIMIS Penman	R_s, T, RH and U_2	1.460	0.999	0.678
C_Hargreaves (1985)	R_s and T	0.308	0.478	0.977
C_Ritchie (1990)	R_s and T	0.972	0.740	0.815
MLR 1	R_s, T, RH and U_2	1.051	0.826	0.828
MLR 2	R_s and T	1.129	0.850	0.811

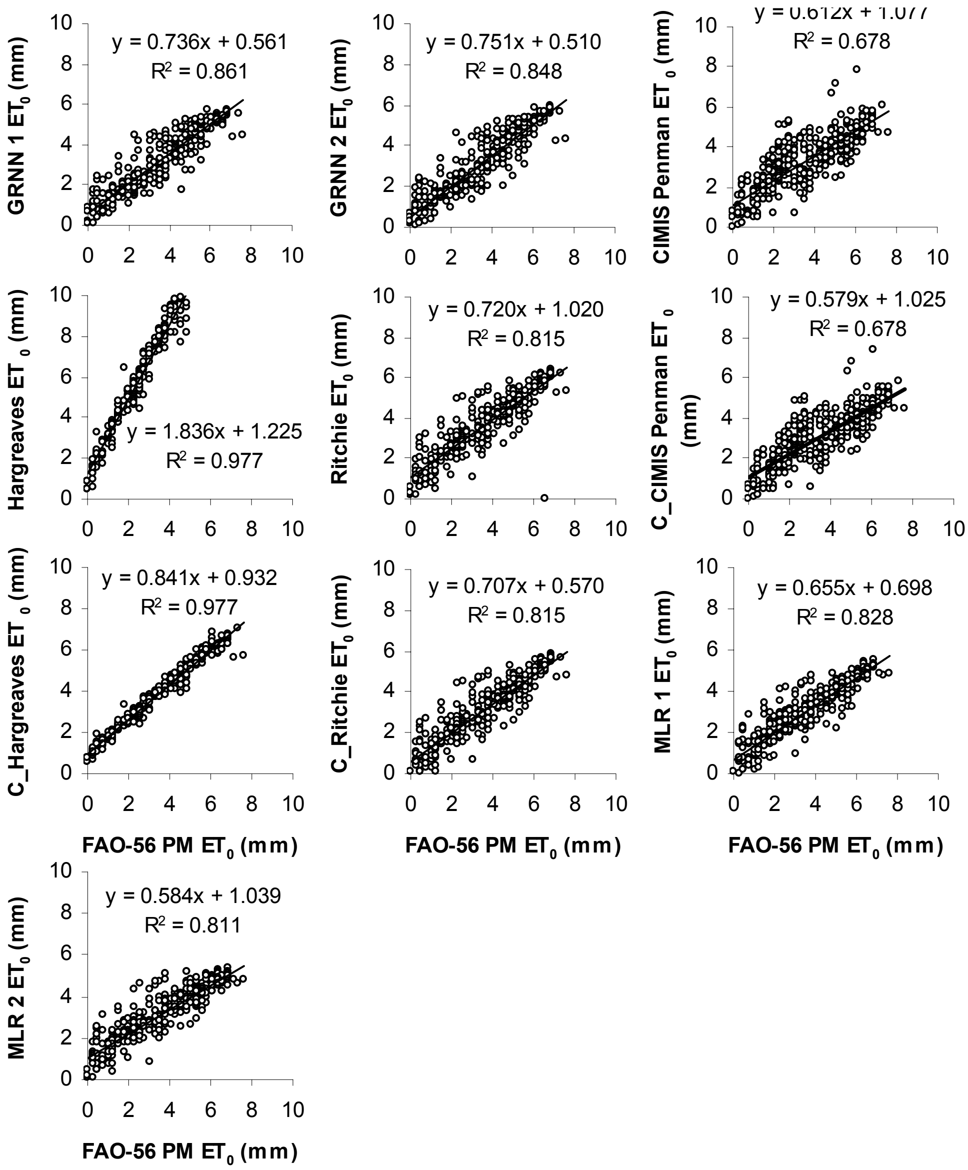

Fig. 4 The FAO-56 PM and estimated ET_0 values of the Pomona Station using the data of Santa Monica Station.

CONCLUDING REMARKS

The following conclusions could be drawn from this study:

– The potential of GRNN models for the estimation of evapotranspiration using climatic variables has been illustrated. The study demonstrated that modelling of daily reference evapotranspiration is possible through the use of the GRNN technique.
– The GRNN model, whose inputs are solar radiation (R_s), air temperature (T), relative humidity (RH) and wind speed (U_2), was found to perform better than the Penman, Hargreaves and Turc methods in estimation of FAO-56 PM ET_0.
– The calibrated temperature-based models, C_Hargreaves and C_Ritchie, performed slightly better than the GRNN model whose inputs are solar radiation and air temperature for the Pomona station. It was found that the calibration significantly increased the performance of the temperature-based models in estimation of ET_0. The GRNN2 model can be used in estimation of FAO-56 PM ET_0 where there exist only solar radiation and air temperature data. In such a situation, the Hargreaves and Ritchie models should be used after calibration.
– The air temperature and relative humidity were found to be more effective than the wind speed in modelling ET_0.
– The GRNN model whose inputs are the wind speed, solar radiation, relative humidity and air temperature performs the best among the input combinations tried in the study. This indicates that all these variables are needed for better evapotranspiration modelling.
– The C_Hargreaves method was found to be more adequate than the GRNN, MLR and other empirical models in estimation of ET_0 using nearby station data.

The GRNN technique could be of use in water budgeting of basins, design of reservoirs and various other hydrological analyses where other models may be inappropriate. The study used data from only two stations and further studies using more data from various areas may be required to reinforce the conclusions drawn from this study.

REFERENCES

Allen, R. G., Smith, M., Perrier, A. & Pereira, L. S. (1994) An update for the calculation of reference evapotranspiration. *ICID Bull.* **43**(2), 35–92.

Allen, R. G., Pereira, L. S., Raes, D. & Smith, M. (1998) Crop evapotranspiration guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper no. 56, FAO, Rome, Italy.

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

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

Cigizoğlu, H. K. (2005) Application of generalized regression neural networks to intermittent flow forecasting and estimation. *J. Hydrol. Engng ASCE* **10**(4), 336–341.

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

Disorntetiwat, P. (2001) Global stock index forecasting using multiple generalized regression neural networks with a gating network. PhD Thesis, University of Missouri-Rolla, USA.

Hargreaves, G. H. & Samani, Z. A. (1985) Reference crop evapotranspiration from temperature. *Appl. Engng. in Agric*. **1**(2), 96–99.

Hidalgo, H. G., Cayan, D. R. & Dettinger, M. D. (2005) Sources of variability of evapotranspiration in California. *J. Hydroet.* **6**, 3–19.

Jain, A. & Indurthy, S. K. V. P. (2003) Comparative analysis of event based rainfall–runoff modeling techniques—deterministic, statistical, and artificial neural networks. *J. Hydrol. Engng ASCE* **8**(2), 1–6.

Jensen, M. E., Burman, R. D. & Allen, R. G. (1990) Evapotranspiration and irrigation water requirements. ASCE Manuals and Reports on Engineering Practices no. 70., ASCE, New York, USA.

Jones, J. W. & Ritchie, J. T. (1990) Crop growth models. In: *Management of Farm irrigation Systems* (ed. by G. J. Hoffman, T. A. Howel & K. H. Solomon), 63–89. ASAE Monograph no. 9, ASAE, St Joseph, Michigan, USA.

Karunanithi, N., Grenney, W. J., Whitley, D. & Bovee, K. (1994) Neural networks for river flow prediction. *J. Comput. Civil Engng ASCE* **8**(2), 201–220.

Kisi, O. (2005) Suspended sediment estimation using neuro-fuzzy and neural network approaches. *Hydrol. Sci. J.* **50**(4), 683–696.

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

Minns, A. W. & Hall, M. J. (1996) Artificial neural networks as rainfall–runoff models. *Hydrol. Sci. J.* **41**(3), 399–416.

Naoum, S. & Tsanis, I. K. (2003) Hydroinformatics in evapotranspiration estimation. *Environ. Modell. Software* **18**, 261–271.

Pruitt, W. O. & Doorenbos, J. (1977) Empirical calibration, a requisite for evapotranspiration formulae based on daily or longer mean climatic data. In: Proc. Int. Round Table Conf. on Evapotranspiration,, Budapest, Hungary,.

Sahoo, G. B. & Ray, C. (2006) Flow forecasting for a Hawaii stream using rating curves and neural networks. *J. Hydrol.* **317**, 63–80.

Singh, V. P. (1988) *Hydrologic System Rainfall–Runoff Modelling*, vol. 1. Prentice Hall, Englewood Cliffs, New Jersey, USA.

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

Snyder, R. & Pruitt, W. (1985) Estimating reference evapotranspiration with hourly data. Ch. VII in: California Irrigation Management Information System Final Report (ed. by R. Snyder *et al.*). Univ. of California-Davis. Land, Air and Water Resources Paper no. 10013, California, USA.

Specht, D. F. (1991) A general regression neural network. *IEEE Trans. Neural Networks* **2**(6), 568–576.

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

Tayfur, G. (2002) Artificial neural networks for sheet sediment transport. *Hydrol. Sci. J.* **47**(6), 879–892.

Trajkovic, S. (2005) Temperature-based approaches for estimating reference evapotranspiration. *J. Irrig. Drain. Engng ASCE* **131**(4), 316–323.

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

Wasserman, P. D. (1993) *Advanced Methods in Neural Computing*, 155–161. Van Nostrand Reinhold, New York, USA.

Received 26 July 2005; accepted 12 July 2006