Estimating daily reference evapotranspiration using available and estimated climatic data by adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN)

Ana Pour-Ali Baba, Jalal Shiri, Ozgur Kisi, Ahmad Fakheri Fard, Sungwon Kim and Rouhallah Amini

ABSTRACT

Daily reference evapotranspiration (ET_0), as a dependent variable, was estimated for two weather stations in South Korea, using 8 years (1985–1992) of measurements of independent variables of air temperature, sunshine hours, wind speed and relative humidity. The model uses the adaptive neuro-fuzzy inference system (ANFIS) and artificial neural networks (ANNs) for estimating daily ET_0. In the first part of the study, the applied models were trained, tested and validated using various combinations of the recorded independent variables, which corresponded to the Hargreaves–Samani, Priestly–Taylor and FAO56-PM equations. The goodness of fit for the models was evaluated in terms of the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE) and Nash–Sutcliffe coefficient (NS). In the second part of the study, the estimated solar radiation data were applied as input parameters (for the same input combinations, as the first part), instead of recorded sunshine values. The results indicated that the both applied ANFIS and ANN models performed quite well in ET processes from the available climatic data. The results also showed that the application of estimated solar radiation data instead of the recorded sunshine values decreases the models' accuracy.

Key words | empirical equations, evapotranspiration, neural networks, neuro-fuzzy

Ana Pour-Ali Baba
Department of Agronomy, Miyaneh Branch,
Islamic Azad University, Miyaneh,
Iran

Jalal Shiri (corresponding author)
Ahmad Fakheri Fard
Water Engineering Department,
Faculty of Agriculture,
University of Tabriz, Tabriz,
Iran
E-mail: *j_shiri2005@yahoo.com*

Ozgur Kisi
Civil Engineering Department,
Architectural and Engineering Faculty,
Canik Basari University,
Samsun,
Turkey

Sungwon Kim
Department of Railroad and Civil Engineering,
Dongyang University, Yeongju,
Republic of Korea

Rouhallah Amini
Agronomy Department, Faculty of Agriculture,
University of Tabriz, Tabriz,
Iran

INTRODUCTION

Accurate estimation of evapotranspiration (ET) (the process of water loss to the atmosphere by the combined processes of evaporation and transpiration) is needed for computation of crop water requirement, water resources management, and determination of the water budget, especially under arid conditions where water resources are scarce and fresh water is a limited resource.

The term reference ET (ET_0) was introduced by the United Nations Food and Agriculture Organization (FAO) as a methodology for computing crop ET (Doorenbos & Pruitt 1977), because the interdependence of the factors affecting the ET makes the study of the evaporative demand of the atmosphere regardless of crop type, its stage of development and its management difficult. Recently, the Penman–Monteith equation has been adopted as a reference equation for estimating the reference ET (ET_0) and calibrating other ET_0 equations (Allen *et al.* 1998). The adapted Penman–Monteith equation (which will also be referred to as FAO56-PM model in short) has two important advantages (Landeras *et al.* 2008): (i) it can be applied in a great variety of environments and climate scenarios without local

doi: 10.2166/nh.2012.074

calibration, and (ii) it has been validated using lysimeters under a wide range of climatic conditions. On the other hand, the need for a large number of climatic variables (e.g. air temperature, relative humidity, solar radiation and wind speed) is a major disadvantage of the FAO56-PM equation.

In recent years, the application of artificial intelligence (AI) in water resources engineering has become viable. In this context, the application of artificial neural networks (ANNs) and adaptive neuro-fuzzy inference systems (ANFIS) is widening, leading to the publication of numerous research papers. The review of all the papers regarding the application of these techniques is beyond the scope of the present paper and we shall review only some of the most relevant studies. Odhiambo *et al.* (2001) showed that ET_0 estimates based on an optimized fuzzy-neural model were comparable to those based on FAO56-PM equation. Kumar *et al.* (2002) applied a multi-layer neural network (NN) using a back propagation training algorithm for the estimation of ET_0. Sudheer *et al.* (2003) applied a radial basis function NNs with limited climate data to model ET_0. Trajkovic (2005) modeled FAO56-PM ET_0 using a temperature-based radial basis NN. Kisi (2006a, b) applied generalized regression NN and feed forward NN to the ET_0 estimation. Kisi (2007) modeled ET_0 from climatic data using a neural computing technique. Landeras *et al.* (2008) applied ANNs for the estimation of daily ET_0 in Northern Spain, and compared the results with those obtained by using empirical equations. Kumar *et al.* (2008) compared the conventional and ANN-based ET_0 estimation models and concluded that ANNs are better tools for estimating ET_0 than empirical methods. Kim & Kim (2008) applied a NN-genetic algorithm approach for nonlinear evaporation and ET modeling. Kisi (2009) employed multilayer perceptrons and radial basis NNs for estimating monthly pan evaporation values. Wang *et al.* (2010) applied ANN with feed forward back-propagation algorithm for modeling ET_0 in arid regions of Africa. Trajkovic (2010) applied ANNs to estimate hourly ET_0 values and tested the obtained values as well as FAO56-PM ET_0 values against the lysimetric data and found the ANNs results to be promising. Kim (2011) applied a stochastic and NN approach for nonlinear hydrologic modeling. Guo *et al.* (2011) predicted daily crop ET_0 by least square support vector machines.

Kisi (2006c) investigated the ability of the ANFIS technique to improve the accuracy of daily evaporation estimation. Keskin *et al.* (2004) used fuzzy models to estimate daily pan evaporation in western Turkey. Kisi & Ozturk (2007) used the ANFIS computing technique for ET_0 estimation. Moghaddamnia *et al.* (2009) applied NN and ANFIS techniques for evaporation estimation in a hot and dry climate in Iran. Dogan (2009) examined the capability of ANFIS technique for estimating ET_0 by making some sensitivity analysis about the applied input parameters. Shiri & Kisi (2011a) applied AI techniques (i.e. ANFIS, ANN and genetic programming) to estimate daily pan evaporation by using available and estimated climatic data in Iran. Shiri & Kisi (2011b) compared ANFIS with genetic programming to forecast groundwater depth fluctuations. Shiri *et al.* (2011) compared ANFIS to ANNs for modeling daily pan evaporation values and found ANFIS to be better than ANNs. The objective of this paper is to demonstrate the feasibility of the neuro-fuzzy and NN techniques in modeling daily reference ET using climatic parameters in the Yeongsan River catchment, South Korea. The neuro-fuzzy and NN techniques are compared with commonly used Hargreaves–Samani and Priestly–Taylor empirical equations.

MATERIALS AND METHODS

Used data

The daily climatic data of two weather stations, Gwangju and Haenam in South Korea, were used in the present study. The locations of the both stations are shown in Figure 1. The climatic data consisted of 8 years (January 1985–December 1992) of daily records of air temperature (T_A), sunshine hours (S), wind speed (W_S) and relative humidity (R_H). Daily ET_0 values calculated by the FAO56-PM equation were applied as training patterns for the development of the ANFIS and ANN models. For each station, the first 4 years of data (50% of whole data) were applied for training the models, 2 years for testing and the remaining 2 years are applied for models' validation. Such a manner (data division in three parts) is much better than the data

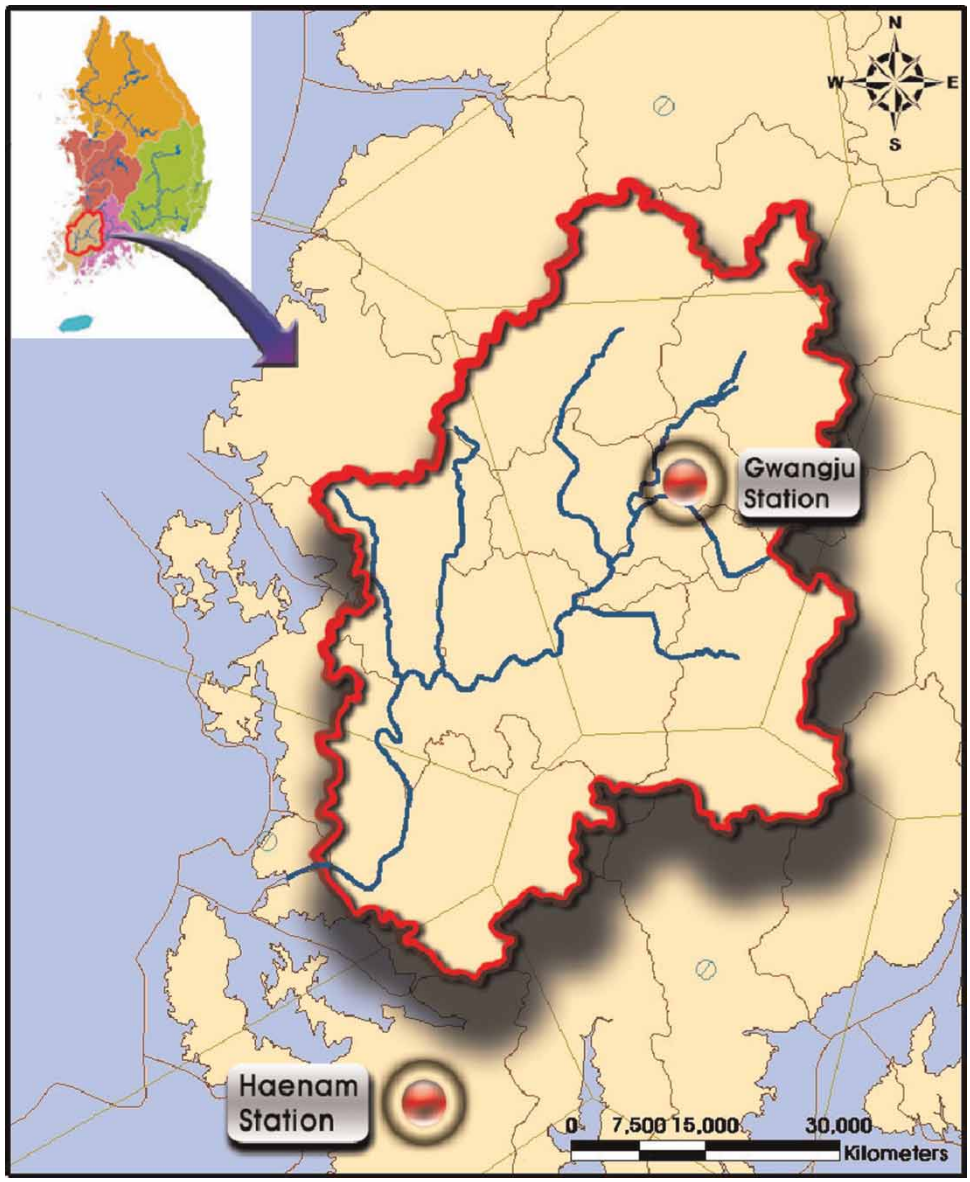

Figure 1 | The location of the Gwangju and Haenam stations.

division in two parts. First, one can obtain models' parameters by using training data and then choose the optimal model according to their testing performances. Finally, the evaluation and comparison of the optimal models can be achieved by using different data (validation) sets which are not used for model development stages (training and testing). The description of the studied weather stations as well as the mean weather data are presented in Table 1. Figure 2 displays the time series of the ET_0 values for the both stations during the study period.

ET_0 equations

FAO56-PM equation

The Penman–Monteith equation modified by Allen *et al.* (1998) reads:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T_{mean} + 273} W_s(e_a - e_d)}{\Delta + \gamma(1 + 0.34U_2)} \quad (1)$$

Table 1 | Summary of the studied weather stations and the mean values of the applied weather parameters

				Weather parameters				
	Longitude (°W)	Latitude (°N)	Altitude (m)	T_A (°c)	S (hours)	W_S (m/s)	R_H (%)	ET_0 (FAO56-PM) (mm)
Gwangju	126.89	35.17	74.5	13.6	5.95	2.27	72.39	3.29
Haenam	126.56	34.55	4.6	13.4	6.49	2.17	76.18	3.35

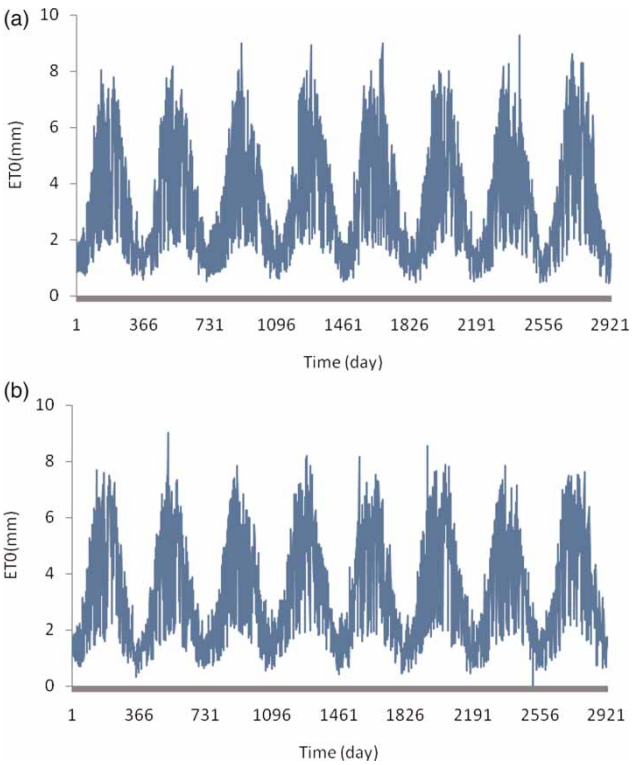

Figure 2 | Daily ET_0 values during the study period for: (a) Gwangju station and (b) Haenam station.

where ET_0 = reference evapotranspiration (mm/day), Δ = slope of the saturation vapor pressure function (kPa/°C), γ = psychometric constant (kPa/°C), R_n = net radiation (MJ/m^{2}day), G = soil heat flux density (MJ/m^{2}day), T_{mean} = mean air temperature (°C), W_s = average 24 h wind speed at 2 m height (m/s), e_a = saturation vapor pressure (kPa), and e_d = actual vapor pressure. The ET_0 is then multiplied by a crop coefficient (K_c) to calculate actual ET from a reference ET_0. The crop coefficient is acquired with respect to plant type, the plant's maturity and local factors such as soil type (Jensen *et al.* 1990).

Hargreaves–Samani equation

The Hargreaves model (Hargreaves & Samani 1985) can be expressed as follows:

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

where ET_0 = reference evapotranspiration (mm/day), R_a = extraterrestrial radiation (mm/day), T_{max} = maximum air temperature (°C) and T_{min} = minimum air temperature (°C).

Priestly–Taylor equation

The Priestley & Taylor (1972) equation for computing ET_0 value is:

$$ET_0 = \frac{\alpha}{\lambda}\frac{\Delta}{\Delta + \gamma}(R_n - G) \quad (3)$$

where ET_0 = reference evapotranspiration (mm/day), α = 1.26, λ = latent heat of the evaporation (MJ/kg) and the other applied parameters were introduced before (Gibson *et al.* 1994).

Adaptive neuro-fuzzy inference system (ANFIS)

An ANFIS is a combination of an ANN and a fuzzy inference system. The parameters of the fuzzy inference system are determined by the NN learning algorithms. Since this system is based on the fuzzy inference system, reflecting amazing knowledge, an important aspect is that the system should be always interpretable in terms of fuzzy IF-THEN rules. ANFIS is capable of approximating any real continuous function on a compact set of parameters to any degree of accuracy (Jang *et al.* 1997). ANFIS identifies a set of parameters through a hybrid learning rule combining back

propagation gradient descent error digestion and a least squared error method. There are mainly two approaches for fuzzy inference systems, namely the approaches of Mamdani (Mamdani & Assilian 1975) and Sugeno (Takagi & Sugeno 1985). The differences between the two approaches arise from the consequent part where Mamdani's approach uses fuzzy membership functions (MFs), while linear or constant functions are used in Sugeno's approach. In this study, the Sugeno method was applied for estimating daily reference ET values. The use of fuzzy theory together with ANNs combines the ability of fuzzy set to represent understandable knowledge of human with learning capability of ANNs (Jin 2003).

As a simple example, assume a fuzzy inference system having two inputs x and y and one output f. The first-order Sugeno fuzzy model, a typical rule set with two fuzzy IF-THEN rules can be given as:

$$\text{Rule 1: If } x \text{ is } A_1 \text{ and } y \text{ is } B_1, \text{ then } f_1 = p_1x + q_1y + r_1 \quad (4)$$

$$\text{Rule 2: If } x \text{ is } A_2 \text{ and } y \text{ is } B_2, \text{ then } f_2 = p_2x + q_2y + r_2 \quad (5)$$

where A_1, A_2 and B_1, B_2 are the MFs for inputs x and y, respectively, p_1, q_1, r_1 and p_2, q_2, r_2 are the parameters of the output function. Here the output f is the weighted average of the individual rule outputs and is itself a crisp value. The corresponding equivalent ANFIS architecture is represented in Figure 3. Detailed information for ANFIS can be found in Jang (1993).

Artificial neural networks (ANNs)

ANNs are basically parallel information-processing systems. The internal architecture of ANNs is similar to the structure of a biological brain with a number of layers of fully interconnected nodes or neurons. Each neuron is connected to other neurons by means of direct communication links, each with an associated weight. The NN usually has two or more layers of neurons in order to process nonlinear signals. The input layer admits the incoming information, which is processed by the hidden layer(s), and the output layer presents the network result. During the learning process, the weights (W) of the interconnections and the neural biases are adjusted in trial and error procedures, to minimize the errors. Three-layer feed-forward networks were employed in this study, with a sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. Figure 4 illustrates the schematic architecture of the applied NNs, consisting of three layers (i, j and k) with the interconnection weights W_{ij} and W_{jk}. The hidden-layer-node numbers of each model were determined after an iterative process, because there is not yet a definite theoretical background for determining the interconnections of neurons. The basic details and concepts of the working of an ANN can be found in Bishop (1995) or Haykin (1999).

Goodness of fit of model performance

Four statistical evaluation criteria were used to assess the models' performances: the correlation coefficient (R^2):

$$R^2 = \left(\frac{\sum_{i=1}^{n} (\mathrm{ET_0} - \text{mean } \mathrm{ET_0})(\mathrm{ET_M} - \text{mean } \mathrm{ET_M})}{\sqrt{\sum_{i=1}^{n} (\mathrm{ET_0} - \text{mean } \mathrm{ET_0})^2 \sum_{i=1}^{n} (\mathrm{ET_M} - \text{mean } \mathrm{ET_M})^2}} \right)^2 \quad (6)$$

The Pearson correlation coefficient (R) term and the coefficient of determination (R^2) provide information for

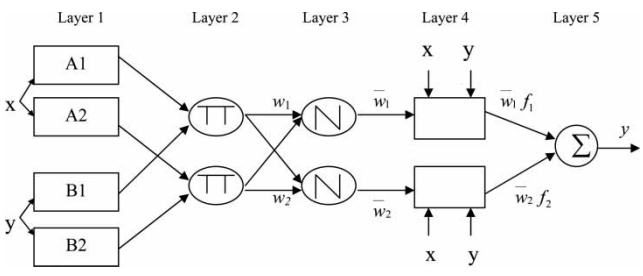

Figure 3 | Equivalent ANFIS architecture.

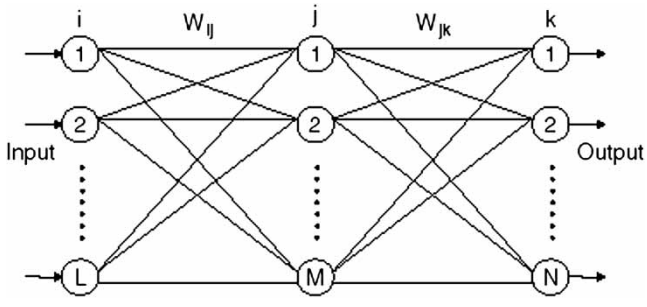

Figure 4 | A three layered feed-forward NN.

linear dependence between observations and corresponding estimates and should not be applied alone as performance indicators, as has been discussed by Legates & McCabe (1999). Therefore, other statistical measures such as mean absolute error, MAE (which is a linear scouring rule and describes only the average magnitude of the errors, ignoring their direction) and root mean square error, RMSE (which describes the average magnitude of the errors by giving more weight on large errors) should be applied to evaluate the models' performance. Also the Nash–Sutcliffe coefficient (NS) can provide good insight about the applied model. The optimal value of NS is 1, representing the perfect fit. The mentioned scours can be defined as:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\mathrm{ET_M} - \mathrm{ET_0})^2}{n}} \tag{7}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} \mathrm{abs}(\mathrm{ET_0} - \mathrm{ET_M})}{n} \tag{8}$$

$$\mathrm{NS} = 1 - \frac{\sum_{i=1}^{n} (\mathrm{ET_M} - \mathrm{ET_0})}{\sum_{i=1}^{n} (\mathrm{ET_0} - \mathrm{mean\ ET_0})} \tag{9}$$

where $\mathrm{ET_M}$ and $\mathrm{ET_0}$ denote the values generated by different models and FAO56-PM $\mathrm{ET_0}$ equation, respectively. Also mean $\mathrm{ET_0}$ and mean $\mathrm{ET_M}$ represent the mean ET values, estimated by FAO56-PM and applied models, respectively.

RESULTS AND DISCUSSION

The applied models, namely ANFIS, ANNs and empirical models, were implemented using the recorded weather data at Gwangju and Haenam stations in South Korea. The study examined various combinations of these data (as inputs for ANFIS and ANN models) to evaluate the degree of the effect of each of these variables on ET as well as inter-comparison of the obtained results. Also, two empirical models, namely Hargreaves–Samani and Priestly–Taylor (which will be also referred to as HS and PT models in short, respectively), were considered for comparison with data-driven approaches in the study. It is relevant to note that the present study consisted of examining two approaches: approach 1 in which various combinations of T_A, S, W_S and R_H were applied as inputs for ANFIS and ANN models; approach 2 in which the results produced by using the same input combinations as part 1, but by applying estimated solar radiation (R_S) values, instead of sunshine hours (S).

According to Allen *et al.* (1998), in the case of the absence of recorded solar radiation data, one may estimate R_S values by using the temperature difference method, as follows:

$$R_S = K_{rs}\sqrt{T_{max} - T_{min}}R_a \tag{10}$$

where R_a is the extraterrestrial radiation (MJ/m^2/d), T_{max} and T_{min} are maximum and minimum air temperatures ($^\circ$C) and K_{rs} is the adjustment coefficient ($^\circ\mathrm{C}^{-0.5}$). K_{rs} coefficient is empirical and differs in interior ($K_{rs} = 0.16\ ^\circ\mathrm{C}^{-0.5}$) and coastal ($K_{rs} = 0.19\ ^\circ\mathrm{C}^{-0.5}$) regions (Hargreaves & Samani 1982). There are some empirical models for estimating incoming solar radiation data, but the results of investigations (e.g. Bandyopadhyay *et al.* 2008) show that the choice of method for estimating R_S has relatively small effect on estimating $\mathrm{ET_0}$. Therefore the commonly used temperature difference method (Equation (10)) was applied to estimate R_S values.

In the first part of the study (i.e. approach 1), various ANFIS and ANN models were developed using the recorded weather parameters. In this way, the evaluated input combinations were:

(i) T_A (ANFIS1, ANN1);
(ii) T_A and S (ANFIS2, ANN2);
(iii) T_A, S, W_S and R_H (ANFIS3, ANN3).

It is noted that the aforementioned input combinations corresponded to the required input variables for the HS model (single input model), PT model (double-input model) and FAO56-PM model (quadruple-input model) for making a valid comparison between the results of data driven models and $\mathrm{ET_0}$ equations. The HS and PT models are selected for the comparison because these are commonly used equations in the literature. The calculated

FAO56-PM ET_0 values were used as the reference values for the comparisons and evaluation of other applied models.

Models implementation

Implementation of ET_0 equations

The FAO56-PM equation was selected as reference equation for the comparison of other applied models. The HS and PT were the ET_0 equations applied in the present study. These equations were also trained (locally calibrated) by using the data from 1985 to 1988, and then tested and validated during the period 1989–1990 and 1991–1992, respectively. Allen *et al.* (1994) recommended that the ET equations be calibrated by using the standard FAO56-PM method as follows:

$$ET_0 = a + bET_M \quad (11)$$

where the ET_0 denotes the ET value calculated by the standard FAO56-PM model and the ET_M denotes the ET value estimated by one of the empirical or semi-empirical equations and a and b are the linear regression coefficients. The weather data applied for training the ANFIS and ANN models (the period from 1985 to 1988) were applied for calibrating the HS and PT equations. Table 2 represents the

Table 2 | Statistical analysis of the ET_0 equations for the period of the study

Study period	ET_0 equation	R^2	RMSE (mm)	MAE (mm)	Calibration parameters	
					a	*b*
Gwangju station						
Training period (1985–1988)	Hargreaves–Samani	0.758	1.021	0.833	–	–
	Priestly–Taylor	0.703	1.220	0.907	–	–
Testing period (1989–1990)	Hargreaves–Samani	0.772	1.040	0.826	–	–
	Priestly–Taylor	0.748	1.264	0.905	–	–
Validation period (1991–1992)	Hargreaves–Samani	0.769	1.089	0.856	–	–
	Priestly–Taylor	0.731	1.313	0.955	–	–
Haenam station						
Training period (1985–1988)	Hargreaves–Samani	0.777	1.007	0.797	–	–
	Priestly–Taylor	0.647	1.258	0.938	–	–
Testing period (1989–1990)	Hargreaves–Samani	0.761	1.137	0.888	–	–
	Priestly–Taylor	0.627	1.405	1.039	–	–
Validation period (1991–1992)	Hargreaves–Samani	0.781	0.784	0.437	–	–
	Priestly–Taylor	0.640	0.989	0.538	–	–
Gwangju station						
Training period (1985–1988)	Hargreaves–Samani_calibrated	0.758	0.915	0.730	0.311	1.049
	Priestly–Taylor_calibrated	0.703	1.015	0.781	–1.054	1.519
Testing period (1989–1990)	Hargreaves–Samani_calibrated	0.772	0.935	0.740	–	–
	Priestly–Taylor_calibrated	0.748	0.986	0.734	–	–
Validation period (1991–1992)	Hargreaves–Samani_calibrated	0.769	0.965	0.750	–	–
	Priestly–Taylor_calibrated	0.731	1.039	0.783	–	–
Haenam station						
Training period (1985–1988)	Hargreaves–Samani_calibrated	0.777	0.902	0.701	0.215	1.102
	Priestly–Taylor_calibrated	0.647	2.143	1.928	–0.924	1.471
Testing period (1989–1990)	Hargreaves–Samani_calibrated	0.761	1.012	0.785	–	–
	Priestly–Taylor_calibrated	0.627	2.011	1.783	–	–
Validation period (1991–1992)	Hargreaves–Samani_calibrated	0.781	0.911	0.691	–	–
	Priestly–Taylor_calibrated	0.640	2.152	1.941	–	–

error statics of each applied calibrated and un-calibrated ET_0 equation during the training period. The calibration parameters are also provided in this table. It is clear from Table 2 that the calibration of the empirical models increases their accuracy except the PT model for the Haenam station. This may be due to the fact that the correlation between the un-calibrated PT model and FAO56-PM model is low (0.647).

Implementation of ANFIS and ANN models

Table 3 gives the final architectures of ANFIS and ANN models for both the stations. The second column indicates the number of MFs of each input variable of ANFIS models. The third column shows the number of input, hidden and output nodes in each of the ANN models. For instance, in the input combination (iii), the ANFIS model has 2, 3, 3 and 3 triangular MFs for the input variables, T_A, S, W_S and R_H, respectively. Fuzzy MFs can take many forms, but triangular MFs are often selected for practical applications (Russel & Campbell 1996), although the type of MF does not affect the results significantly (Vernieuwe *et al.* 2005). So far, there is not any basic rule available to determine the number of MFs, so they usually determined iteratively. However, a modeler should avoid using a large number of MFs or parameters to save time and calculation effort (Keskin *et al.* 2004). Also, the selection of the ANNs architectures was based on the method proposed by Kumar *et al.* (2002). Based on this method, the hidden layer may have $k+1$ nodes in which, k represents the nodes of input layer (the number of applied input parameters).

Table 4 gives the R^2, RMSE and MAE values for ANFIS and ANN models during the training period. From the table it is clearly seen that the quadruple-input ANFIS and ANN models whose inputs are the T_A, S, W_S and R_H, give the best results among all the models considered. The considered ANFIS and ANN models show good performance for approximating FAO56-PM ET_0 values from the available climatic data. In the second part of the study, the computed solar radiation (R_S) data were applied instead of the sunshine hour data to estimate ET_0 values. The statistical performances of the models during the training period are also given in Table 4. It can be seen from the table that the R_S-based ANFIS and ANN models are inferior to those developed by using recorded sunshine hour values. The foregoing sections will discuss the models' differences precisely.

Table 3 | Final architecture of ANFIS and ANN models

Models input	ANFIS structure (the number of MFs of each variable)	ANN structure (the number of input-hidden-output nodes)
T_A	4	1–2–1
T_A,S	3 and 3	2–3–1
T_A, S, W_S, R_H	2, 3, 3 and 3	4–5–1

Model testing

The ET_0 values were calculated by the calibrated and un-calibrated equations for the testing period (1989–1990) and the obtained results were compared to the results obtained by the standard FAO56-PM model as reference equation. The error statistics of the ET_0 equations during the test period are given in Table 2. The testing results of the ANFIS and ANN models are represented in Table 5 for both stations. In the case of application of sunshine hour data as one of the models' inputs, the quadruple input ANFIS model (corresponded to the FAO56-PM equation) produces the best results with a R^2 value of 0.948, RMSE value of 0.446 mm and MAE value of 0.349 mm (for Gwangju station) and an R^2 value of 0.915, RMSE value of 0.559 mm and MAE value of 0.430 mm (for Haenam station). The ANN model can be ranked as the second model with R^2 value of 0.939, RMSE value of 0.479 mm and MAE value of 0.305 mm (for Gwangju station) and a R^2 value of 0.893, RMSE value of 0.624 mm and MAE value of 0.388 mm (for Haenam station). From this table, it can be concluded that introducing air temperature as the unique input parameter (corresponding to the HS equation) cannot give good results for both stations. Introducing sunshine hours as input variable along with the air temperature (corresponding to the PT equation), improves the model accuracy. Comparison of the results of the applied models during the testing period (i.e. Tables 2 and 5) reveals that the performances of the PT equation are inferior to ANFIS and ANN models in the both stations, but the HS

Table 4 | Error statistics of ANFIS and ANN models during the training period

	ANFIS model			ANN model		
Input combination	R^2	**RMSE (mm)**	**MAE (mm)**	R^2	**RMSE (mm)**	**MAE (mm)**
Models with actual sunshine hours data						
Gwangju station						
T_A	0.434	1.390	1.081	0.447	1.38	1.07
T_A,S	0.904	0.574	0.440	0.919	0.528	0.43
T_A, S, W_S, R_H	0.943	0.441	0.348	0.935	0.47	0.31
Haenam station						
T_A	0.486	1.312	1.032	0.498	1.3	1.29
T_A,S	0.899	0.578	0.457	0.886	0.622	0.45
T_A, S, W_S, R_H	0.937	0.451	0.354	0.918	0.522	0.356
Models with estimated solar radiation data						
Gwangju station						
T_A,R_S	0.794	0.842	0.651	0.793	0.849	0.626
T_A, S, W_S, R_H	0.879	0.646	0.476	0.876	0.656	0.462
Haenam station						
T_A,R_S	0.793	0.830	0.628	0.787	0.844	0.608
T_A, S, W_S, R_H	0.896	0.587	0.434	0.889	0.609	0.417

Table 5 | Summary of the testing process of the applied models

	ANFIS model			ANN model		
Input combination	R^2	**RMSE (mm)**	**MAE (mm)**	R^2	**RMSE (mm)**	**MAE (mm)**
Models with sunshine hour values						
Gwangju station						
T_A	0.412	1.498	1.218	0.447	1.452	1.173
T_A,S	0.911	0.588	0.469	0.888	0.649	0.446
T_A, S, W_S, R_H	0.948	0.446	0.349	0.939	0.479	0.305
Haenam station						
T_A	0.481	1.366	1.110	0.470	1.38	1.1
T_A,S	0.889	0.063	0.506	0.869	0.687	0.506
T_A, S, W_S, R_H	0.915	0.559	0.430	0.893	0.624	0.388
Models with R_S values						
Gwangju station						
T_A,R_S	0.816	0.837	0.641	0.802	0.872	0.616
T_A, R_S, W_S, R_H	0.875	0.688	0.466	0.872	0.707	0.432
Haenam station						
T_A,R_S	0.793	0.871	0.676	0.782	0.885	0.651
T_A, R_S, W_S, R_H	0.889	0.630	0.459	0.858	0.717	0.514

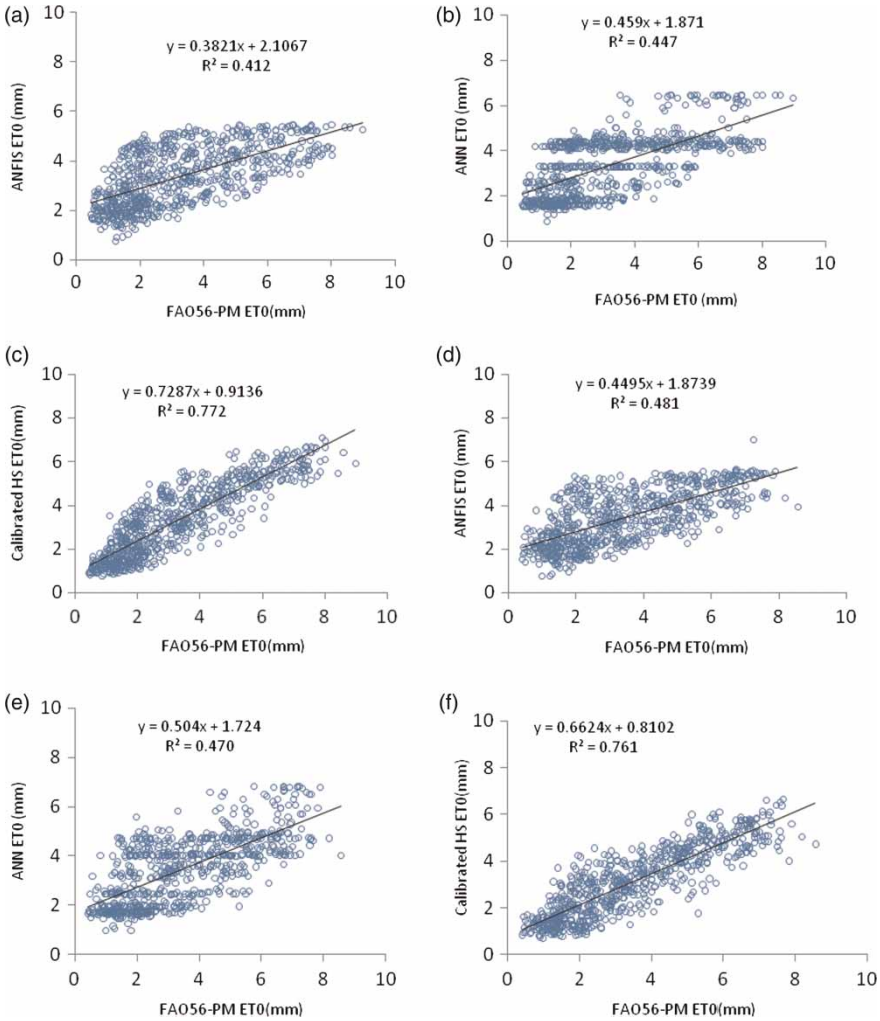

Figure 5 | FAO56-PM and estimated ET_0 values of single-input models for Gwangju (a, b, c) and Haenam (d, e, f) stations during the test period.

equation produces better results than ANFIS and ANN models. Figure 5 displays the FAO56-PM and estimated ET_0 values for the single-input ANFIS and ANN models as well as for the HS model, in the both stations during the test period. It is clear that the HS equation can produce the ET_0 values better than the corresponded ANFIS and ANN models. The testing values of each double-input model versus the FAO56-PM ET_0 values are presented through the scatterplots in Figure 6. The comparison of the scatterplots reveals that the PT equation cannot surpass the corresponded ANFIS and ANN models in estimating ET_0 values. Finally, Figure 7 displays the FAO56-PM ET_0 values versus the estimated ET_0 by using quadruple-input ANFIS and ANN models during the test period. The figure clearly demonstrated that the quadruple-input ANFIS and ANN models (corresponded to the FAO56-PM equation) produce promising results in estimating daily reference ET values. A comparison of the scatterplots presented through Figures 4–6 reveals that the FAO56-PM corresponded ANFIS and ANN models surpasses the other two input combinations.

In the second part of the study, the computed solar radiation (R_S) data were applied instead of the sunshine hour data to estimate ET values. Therefore the double- and quadruple-input ANFIS and ANN models were re-constructed and tested in both stations. The statistical performances of these models during the test period are given in the lower part of Table 5. It can be seen from

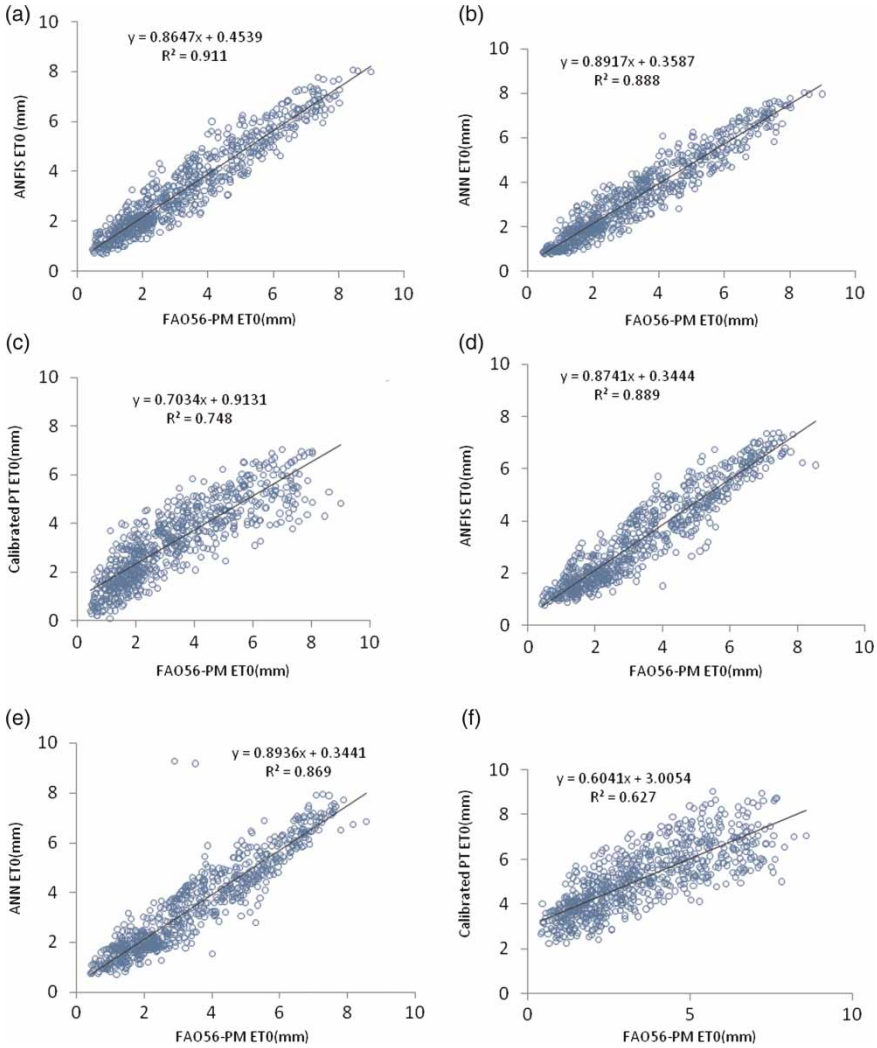

Figure 6 | FAO56-PM and estimated ET_0 values of double-input models for Gwangju (a, b, c) and Haenam (d, e, f) stations during the test period.

the table that, similar to the models which apply sunshine hours, the FAO56-PM corresponded ANFIS and ANN model gives the best results among all the models considered with R^2 value of 0.875, RMSE value of 0.688 mm and MAE value of 0.466 mm (for Gwangju station) and R^2 value of 0.889, RMSE value of 0.630 mm and MAE value of 0.459 mm (for the Haenam station). The ANN model can be ranked as the second best model. Comparing the results presented in Table 5 demonstrated that the application of sunshine hour data as input variables for ANFIS and ANN models gives better results than those produced by the models with application of estimated R_S data.

Model validation

Similar to the models' testing procedure, the ET_0 equations which have been calibrated by using the weather data from training period were validated by using the weather data from the period 1991–1992. The lower part of Table 2 represents the statistical analysis of the validation period for ET_0 equations.

In order to demonstrate the model capabilities for estimating daily reference ET using daily recorded (and estimated) weather data, the present discussion will deal with the validation process of the optimal models. As mentioned in the earlier section, the period from 1991 to 1992 has been reserved for validation purposes, so the same

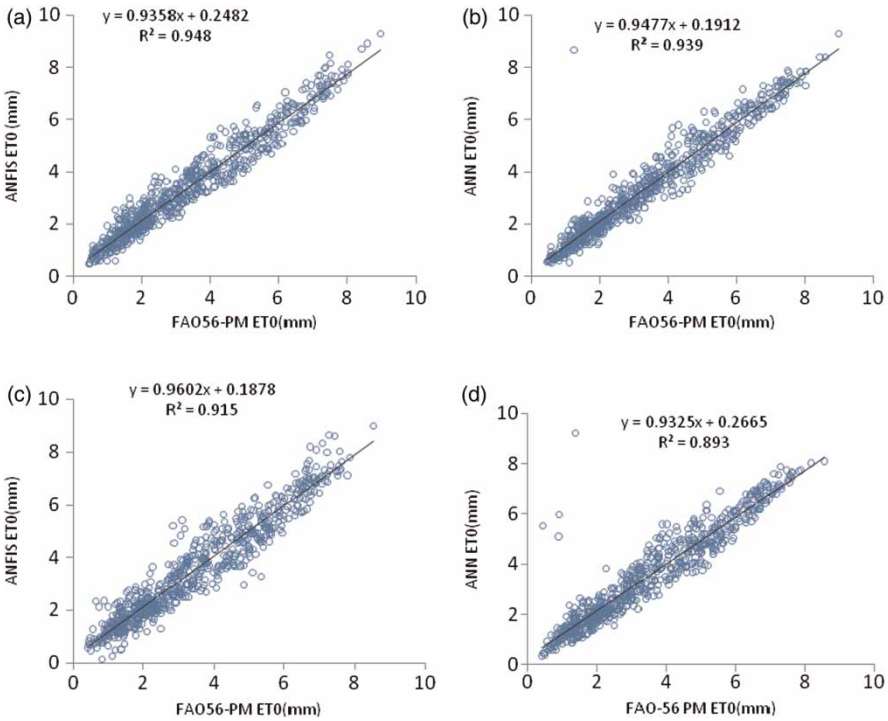

Figure 7 | FAO56-PM and estimated ET_0 values of quadruple-input models for Gwangju (a, b) and Haenam (c, d) stations during the test period.

statistical measures as those used for the training and testing period will be applied here to also show the feasibility and accuracy of employed ANFIS and ANN models in the reference ET estimation. In this way, the double- and quadruple-input ANFIS and ANN models will be compared through the statistical measures. Table 6 presents a summary of the

Table 6 | Summary of the validation process of the optimal models

	ANFIS model				ANN model			
Input combination	R^2	**RMSE (mm)**	**MAE (mm)**	**NS (–)**	R^2	**RMSE (mm)**	**MAE (mm)**	**NS (–)**
Models with sunshine hour values								
Gwangju station								
T_A, S	0.903	0.636	0.507	0.897	0.919	0.574	0.442	0.916
T_A, S, W_S, R_H	0.943	0.474	0.370	0.943	0.959	0.400	0.307	0.959
Haenam station								
T_A, S	0.917	0.546	0.434	0.911	0.877	0.645	0.467	0.884
T_A, S, W_S, R_H	0.926	0.501	0.391	0.926	0.902	0.577	0.394	0.901
Models with R_S values								
Gwangju station								
T_A, R_S	0.820	0.851	0.654	0.817	0.765	0.965	0.659	0.764
T_A, R_S, W_S, R_H	0.899	0.634	0.470	0.898	0.858	0.765	0.477	0.852
Haenam station								
T_A, R_S	0.808	0.805	0.615	0.808	0.794	0.834	0.618	0.764
T_A, R_S, W_S, R_H	0.886	0.626	0.452	0.84	0.845	0.732	0.447	0.841

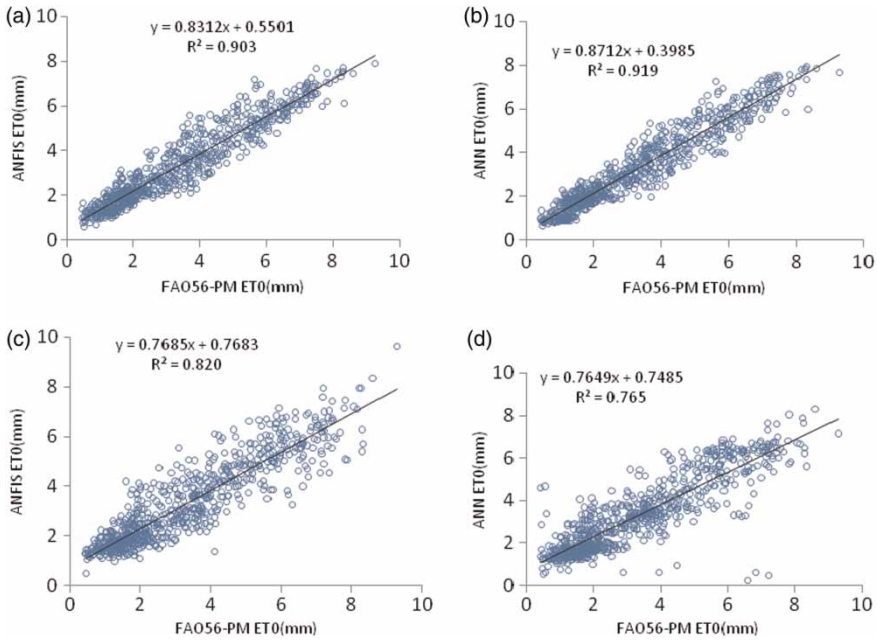

Figure 8 | FAO56-PM and estimated ET_0 values of double-input models for Gwangju station during the validation period: (a, b) with sunshine values; (c, d) with estimated R_S values.

validation process for the optimal ANFIS and ANN models. From the table, it can be concluded that the ANFIS and ANN models corresponded to the FAO56-PM equation give the most accurate results. Also, the comparison of the validation results of the models constructed by using the recorded sunshine hour data with those developed by using

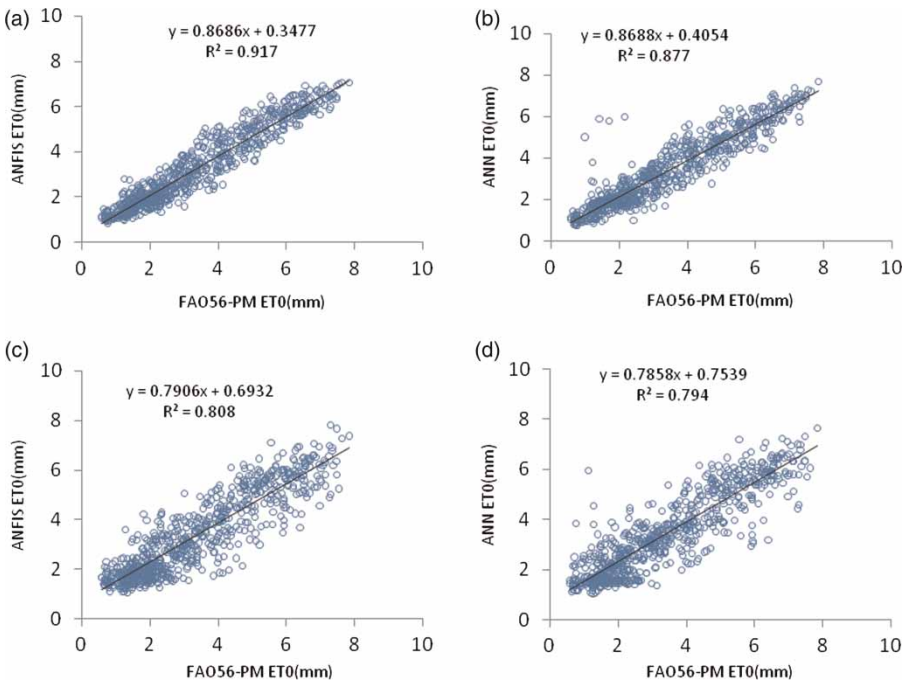

Figure 9 | FAO56-PM and estimated ET_0 values of double-input models for Haenam station during the validation period: (a, b) with sunshine values; (c, d) with estimated R_S values.

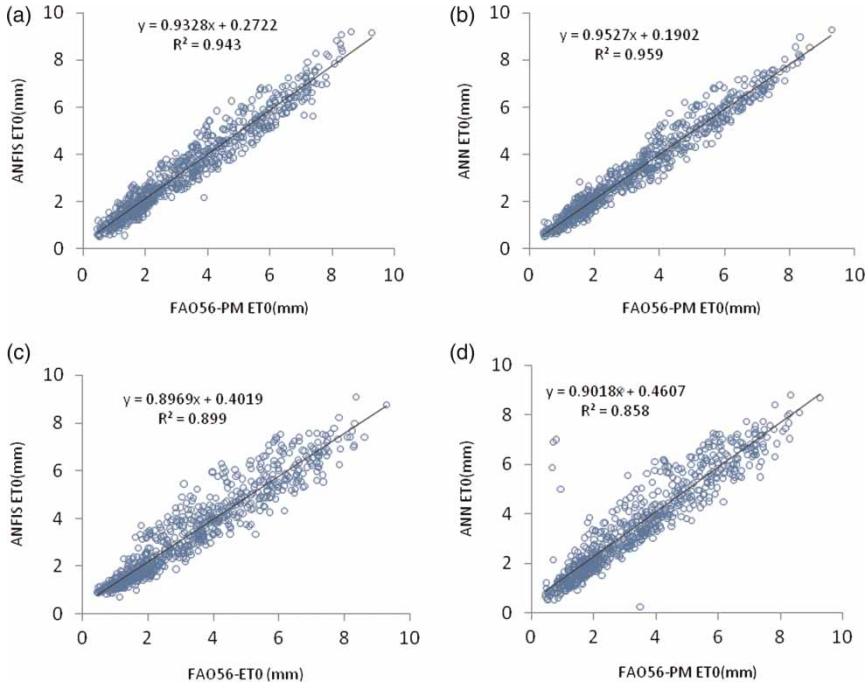

Figure 10 | FAO56-PM and estimated ET_0 values of quadruple-input models for Gwangju station during the validation period: (a, b) with sunshine values; (c, d) with estimated R_S values.

estimated solar radiation values demonstrated that applying the actual sunshine hour data gives much better results than the approach which applies estimated data. The estimation of total ET was also considered for comparison due to its importance for irrigation scheduling. In cases of the models with applying sunshine hour data, the double-input ANFIS model (ANFIS2) estimated the total ET of Gwangju station as 2,436 mm, compared to the calculated FAO56-PM value of 2,446 mm with an underestimation of 0.4% during the validation period, the ANN2, ANN3 and ANFIS3 models resulted in 2,423, 2,470 and 2,481 mm with an underestimation of 0.9% and overestimations of 0.010 and 0.014%, respectively. Also the calibrated HS and PT equations produce total ET_0 values of 2,425 and 2,414 mm, respectively, during the validation period with underestimations of 0.8 and 1.3%. For Haenam station, the total ET estimates of the ANFIS3, ANN3, ANFIS2 and ANN2 models are 0.2% lower and 0.9, 1.1 and 2.8% higher than the FAO56-PM ET_0 value (2,465 mm) during the validation period. The calibrated HS and PT equations produce total ET_0 values of 2,284 and 3,821 mm, respectively, during the validation period with underestimations of 7.3 and 55%. Also, in the case of applying estimated solar radiation data instead of recorded sunshine hour data, while the ANFIS2 model estimated the total ET value for Gwangju station as 2,443 mm, compared to the calculated FAO56-PM value of 2,446 mm with an underestimation of 0.1%, the ANN2, ANFIS3 and ANN3 models resulted in 2,419, 2,488 and 2,543 mm, with an underestimation of 1.1% and overestimations of 1.7 and 3.9%, respectively. For the Haenam station, the ANFIS2 model gives the total ET value (during the validation period) as 2,452 mm with an underestimation of 0.5%. The ANN2, ANFIS4 and ANN4 models gives the total ET_0 value with 0.6, 1.3 and 1.4% overestimations, respectively. Figures 8 and 9 display the FAO56-PM and double-input ANFIS and ANN models estimates for the both stations during the validation period. From the scatterplots it can be observed that the ANFIS and ANN models have good capabilities for estimating ET_0 values by using limited climatic data. Although the models which apply the recorded sunshine data surpass the models which apply estimated R_S values, the comparison of the whole scatterplots of the studied stations shows that whenever the sunshine records are missing, the estimated R_S values can be applied instead. This would be of great use, especially in areas where the advanced instruments for detecting other climatic data are not available and T_A is the

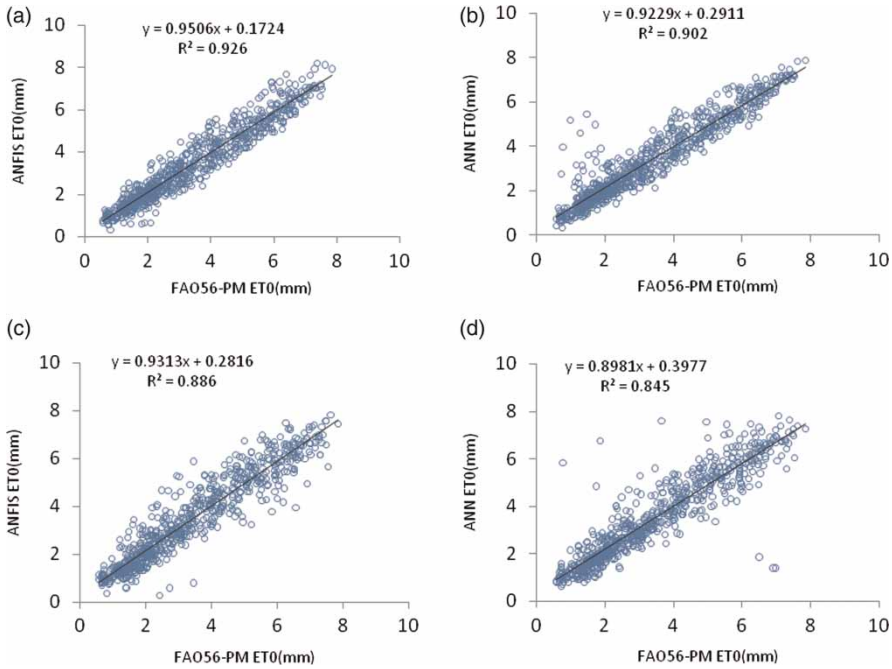

Figure 11 | FAO56-PM and estimated ET_0 values of quadruple-input models for Haenam station during the validation period: (a, b) with sunshine values; (c, d) with estimated R_S values.

unique recorded weather parameter. Figures 10 and 11 show the same results for FAO56-PM corresponded ANFIS and ANN models.

CONCLUSIONS

A modeling study is reported here that develops ANFIS and ANN models for simulating daily reference ET from daily climatic data for two gauging stations in the Yeongsan River catchment in South Korea. The ANFIS and ANN models were compared with those of the calibrated and un-calibrated Hargreaves-Samani and Priestly-Taylor equations. The comparison of the results proved that the ANFIS and ANN could successfully be applied to modeling reference ET from available climatic data. It was found that the calibrated empirical equations generally increased the accuracy of their un-calibrated versions. The estimated solar radiation data were applied as input parameters to the ANFIS and ANN models, instead of recorded sunshine values. The results indicated that both the applied ANFIS and ANN models performed quite well in ET processes from the available climatic data. The results also showed that the applying estimated solar radiation values instead of the recorded sunshine data decreases the models' accuracy.

REFERENCES

Allen, R. G., Pereira, L. S., Raes, D. & Smith, M. 1998 Crop Evapotranspiration – Guidelines for Computing Crop Evapotranspiration. Irrigation Drainage Paper 56, FAO, Rome, Italy.

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

Bandyopadhyay, A., Bhadra, A., Raghuwanshi, N. S. & Singh, R. 2008 Estimation of monthly solar radiation from measured air temperature extremes. *Agric. Forest Meteorol.* **148**, 1707–1718.

Bishop, C. M. 1995 *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford, 504 pp.

Dogan, E. 2009 Reference evapotranspiration estimation using adaptive neuro-fuzzy inference system. *Irrig. Drain. J.* **58**, 617–628.

Doorenbos, J. & Pruitt, W. O. 1977 Crop Water Requirements. Irrigation Drainage paper 24, FAO, Rome, Italy.

Gibson, J. J., Prowse, T. D. & Edwards, T. W. D. 1994 Evaporation from a small lake in the continental arctic using multiple methods. *Hydrol. Res.* **27** (1–2), 1–24.

Guo, X., Sun, X. & Ma, J. 2011 Prediction of daily crop reference evapotranspiration (ET_0) values through a least-squares support vector machine model. *Hydrol. Res.* **42** (4), 268–274.

Hargreaves, G. L. & Samani, Z. A. 1982 Estimating potential evapotranspiration. *ASCE J. Irrig. Drain. Eng.* **108** (3), 225–230.

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

Haykin, S. 1999 *Neural Networks: a Comprehensive Foundation*. Prentice-Hall, Upper Saddle River, NJ, pp. 842.

Jang, J. S. R. 1993 ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans. Sys. Man. Cyber.* **23** (3), 665–685.

Jang, J. S. R., Sun, C. T. & Mizutani, E. 1997 *Neurofuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Prentice-Hall, New Jersey.

Jensen, M. E., Burman, R. D. & Allen, R. G. 1990 *Evapotranspiration and Irrigation Water Requirements*. ASCE Manuals and Reports on Engineering Practices no. 70. American Society of Civil Engineers, NY.

Jin, Y. 2003 *Advanced Fuzzy Systems Design and Applications*. Physica, New York.

Keskin, M. E., Terzi, O. & Taylan, D. 2004 Fuzzy logic model approaches to daily pan evaporation estimation in Western Turkey. *Hydrol. Sci. J.* **49** (6), 1001–1010.

Kim, S. 2011 Nonlinear hydrologic modeling using the stochastic and neural networks approach. *Disaster Adv.* **4** (1), 53–63.

Kim, S. & Kim, H. S. 2008 Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling. *J. Hydrol.* **351** (3–4), 299–317.

Kisi, O. 2006a Generalized regression neural networks for evapotranspiration modeling. *Hydrol. Sci. J.* **51** (6), 1092–1105.

Kisi, O. 2006b Evapotranspiration estimation using feed forward neural networks. *Nor. Hydrol.* **37** (3), 247–260.

Kisi, O. 2006c Daily pan evaporation modeling using a neuro-fuzzy computing technique. *J. Hydrol.* **329**, 636–646.

Kisi, O. 2007 Evapotranspiration modeling from climate data using a neural computing technique. *Hydrol. Process.* **21** (6), 1925–1934.

Kisi, O. 2009 Modeling monthly evaporation using two different neural computing techniques. *Irrig. Sci.* **27**, 417–430.

Kisi, O. & Ozturk, O. 2007 Adaptive neurofuzzy computing technique for evapotranspiration estimation. *ASCE J. Irrig. Drain. Eng.* **133** (4), 368–379.

Kumar, M., Bandyopadhyay, A., Raghuwanshi, N. S. & Singh, R. 2008 Comparative study of conventional and artificial neural network based ET_0 estimation models. *Irrig. Sci.* **26** (6), 531–545.

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

Landeras, G., Ortiz-Barredo, A. & Lopez, J. J. 2008 Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). *Agric. Water Manage.* **95**, 553–565.

Legates, D. R. & McCabe, G. J. 1999 Evaluating the use of goodness-of-fit measures in hydrologic and hydroclimatic validation. *Water Resour. Res.* **35** (1), 233–241.

Mamdani, E. H. & Assilian, S. 1975 An experiment in linguistic synthesis with a fuzzy logic controller. *Int. J. Man. Mach. Stud.* **7** (1), 1–13.

Moghaddamnia, A., Ghafari Gousheh, M., Piri, J., Amin, S. & Han, D. 2009 Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. *Adv. Water Resour.* **32**, 88–97.

Odhiambo, L. O., Yoder, R. E., Yoder, D. C. & Hines, J. W. 2001 Optimization of fuzzy evapotranspiration model through neural training with input-output examples. *Trans. ASAE* **44**, 1625–1633.

Priestley, C. H. B. & Taylor, R. J. 1972 On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon. Weather Rev.* **100** (2), 81–92.

Russel, S. O. & Campbell, P. F. 1996 Reservoir operating rules with fuzzy programming. *J. Water Resour. Plan. Manage.* **122** (3), 165–170.

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

Shiri, J. & Kisi, O. 2011a Application of artificial intelligence to estimate daily pan evaporation using available and estimated climatic data in the Khozestan Province (South Western Iran). *ASCE J. Irrig. Drain. Eng.* **137** (7), 412–425.

Shiri, J. & Kisi, O. 2011b Comparison of genetic programming with neuro-fuzzy systems for predicting short-term water table depth fluctuations. *Comput. Geosci.* **37** (10), 1692–1701.

Shiri, J., Dierickx, W., Pour-Ali Baba, A., Nemati, S. & Ghorbani, M. A. 2011 Estimating daily pan evaporation from climatic data of the state of Illinois, USA using adaptive neuro-fuzzy inference system and artificial neural network. *Hydrol. Res.* **42** (6), 491–502.

Takagi, T. & Sugeno, M. 1985 Fuzzy identification of systems and its application to modeling and control. *IEEE Trans. Sys. Man. Cyber.* **15** (1), 116–132.

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

Trajkovic, S. 2010 Testing hourly reference evapotranspiration approaches using lysimeter measures in a semi arid climate. *Hydrol. Res.* **41** (1), 38–49.

Vernieuwe, H., Georgieva, O., De Baets, B., Pauwels, V. R. N., Verhoest, N. E. C. & De Troch, F. P. 2005 Comparison of data-driven Takagi–Sugeno models of rainfall-discharge dynamics. *J. Hydrol.* **302** (1–4), 173–186.

Wang, Y. M., Traore, S., Kerh, T. & Leu, J. M. 2010 Modeling reference evapotranspiration using feed forward backpropagation algorithm in arid regions of Africa. *Irrig. Drain. J.* **60** (3), 404–417.

First received 3 May 2011; accepted in revised form 30 November 2011. Available online 30 July 2012