Evaluation of different data management scenarios for estimating daily reference evapotranspiration

Jalal Shiri, Ali Ashraf Sadraddini, Amir Hossein Nazemi, Ozgur Kisi, Pau Marti, Ahmad Fakheri Fard and Gorka Landeras

ABSTRACT

Temperature and solar radiation-based modeling procedures are reported in this study for estimating daily reference evapotranspiration (ET_0) by using gene expression programming (GEP) and adaptive neuro-fuzzy inference system (ANFIS). A comparison is also made among these techniques and the corresponding traditional temperature/radiation-based ET_0 estimation equations. Two data management scenarios were evaluated for estimating ET_0: (1) the models were trained and tested using the local data of each studied weather station; and (2) the models were trained using the pooled data from all the stations and tested in each individual station. The GEP and ANFIS models were found to be better than the Hargreaves–Samani, Makkink and Turc ET_0 equations in the first scenario. Comparison of GEP and ANFIS models trained with pooled data and tested for each station showed that the ANFIS models generally performed better than the GEP models. However, the comparison of GEP and ANFIS models trained and tested with pooled data revealed that the GEP models performed better than the ANFIS models in the second scenario.

Key words | evapotranspiration, gene expression programming, local training, neuro-fuzzy, regional training

Jalal Shiri (corresponding author)
Ali Ashraf Sadraddini
Amir Hossein Nazemi
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,
Architecture and Engineering Faculty,
Canik Basari University, Samsun,
Turkey

Pau Marti
Department of Rural and Agrifood Engineering,
Universidad Politécnica de Valencia, Valencia,
Spain

Gorka Landeras
Basque Country Research Institute for Agricultural Development, Alava,
Spain

INTRODUCTION

An accurate estimation of evapotranspiration (ET) is needed for the computation of crop water requirement, water resources planning and 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 evapotranspiration' (ET_0) was introduced by the Food and Agriculture Organization of United Nation (FAO) for computing crop ET (Doorenbos & Pruitt 1977). The interdependence of the factors affecting ET makes the study of the evaporative demand of the atmosphere difficult regardless of crop type, its stage of development and management level. In the recent past, the Penman–Monteith equation has been adopted as a reference equation for estimating ET_0 and calibrating other ET_0 equations (Allen *et al.* 1998). The adapted Penman–Monteith model (which will be referred to as FAO56-PM model in this paper) has two important advantages (Landeras *et al.* 2008): (1) it can be applied in a great variety of environments and climate scenarios without local calibration; and (2) 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 model. The development and validation of models relying on fewer climatic data is therefore of critical importance for the regions where the measured climatic data are limited.

Notable research has been carried out to estimate ET_0 using weather parameters (e.g. Turc 1961; Makkink 1957; Hargreaves & Samani 1985; Allen *et al.* 1998). Nevertheless, in the recent past the application of artificial intelligence

doi: 10.2166/nh.2013.154

(AI) based models (e.g. artificial neural networks or ANN, adaptive neuro-fuzzy inference system or ANFIS and genetic programming or GP) has become viable, leading to numerous publications in hydrology and water resources engineering. A complete review of such applications is beyond the scope of the present paper, and only the most relevant studies will be discussed here.

ANNs have been widely applied for estimating pan-evaporation as well as ET_0 values (e.g. Trajkovic 2005; Kim & Kim 2008; Landeras *et al.* 2008; Kisi 2009; Kumar *et al.* 2011; Kim *et al.* 2012).

ANFIS is a combination of an adaptive neural network and a fuzzy inference system (FIS). The adaptive neural network is a superset of all kinds of feed-forward neural networks (Jang 1993). The parameters of the FIS are determined by the ANN learning algorithms. Since this system is based on the FIS, reflecting extensive knowledge, an important aspect is that the system should always be interpretable in terms of fuzzy IF-THEN rules. ANFIS is capable of approximating any real continuous function on a compact set (Jang *et al.* 1997).

Kisi (2006) investigated the ability of the ANFIS technique to improve the accuracy of daily evaporation estimation. Dogan (2009) examined the capability of the ANFIS technique for estimating ET_0 by performing a sensitivity analysis of the applied input parameters. Shiri & Kisi (2010) introduced a new hybrid wavelet-neuro-fuzzy conjunction model for predicting short-term and long-term streamflow. Shiri *et al.* (2011) compared ANFIS to ANN when modeling evaporation using climatic data from Illinois, USA, and found ANFIS to be better than ANN in this instance. Kisi & Shiri (2012a) introduced a new wavelet-ANFIS conjunction model for predicting short-term groundwater level variations. Pour Ali Baba *et al.* (2013) applied ANFIS model for estimating daily ET_0 values using available and estimated weather data.

GP, firstly proposed by Koza (1992) as a generalization of genetic algorithms (GA, Goldberg 1989), employs a 'parse tree' structure for the determination of solutions. This technique has the ability to derive a set of explicit formulations that rule the phenomenon to describe the relationship between the independent and dependent variables using various operators. Gene expression programming (GEP) is comparable to GP, yet evolves computer programs of different sizes and shapes encoded in linear chromosomes of fixed lengths. The chromosomes are composed of multiple genes, each gene encoding a smaller subprogram. Furthermore, the structural and functional organization of the linear chromosomes allows the unconstrained operation of important genetic operators such as mutation, transposition and recombination.

Babovic *et al.* (2001) employed GP for modeling risk in water supply. Shiri & Kisi (2011a) applied AI techniques to estimate daily pan-evaporation by using available and estimated climatic data in Iran. Shiri & Kisi (2011b) compared ANFIS with GP to forecast groundwater depth fluctuations. Kisi & Shiri (2011) introduced a new hybrid wavelet-GEP model for predicting daily rainfall values. Shiri *et al.* (2012) applied GEP for estimating daily ET_0 values in the Basque Country (northern Spain). Kisi & Shiri (2012b) applied GEP for estimating daily suspended sediment load by climatic variables implications. Landeras *et al.* (2012) applied GEP for estimating daily incoming solar radiation.

So far, only a few attempts have been made to assess the definition criterion of the different datasets required for the application of the mentioned AI techniques, mostly considering a single dataset assignment. The present paper aims to evaluate different data management scenarios when dealing with GEP and ANFIS models for ET_0 estimation. Specifically, we investigate whether the performance of the previous AI techniques improves when they are trained with local patterns (i.e. considering only training patterns of the test station), if exogenous patterns from other stations are added to the existing training set (pooled training). Further, these approaches are compared with existing empirical equations based on the same input combinations.

MATERIALS AND METHODS

Study area

In the present study, daily data from five weather stations (Bam, Esfahan, Kerman, Semnan and Shahrood) in the central part of Iran were used. The available dataset covers a period of 5 years (1 January 2004–31 December 2008), and consists of daily maximum air temperature T_{max}, minimum temperature T_{min}, mean air temperature T_{mean}, solar

radiation R_S, relative humidity R_H and wind speed at 2 m above ground surface u_2. Figure 1 displays the locations of the studied weather stations. The geographical positions of the studied weather stations are summarized in Table 1.

In this study, the aridity index I_A (UNEP 1992) was employed as a numerical indicator of the degree of dryness of the climate. According to the World Atlas of Desertification (UNEP 1992, 1997), dry lands have an aridity index of less than 0.65 and precipitation of less than 600 mm per year. In Table 1, it is possible to find the values of these indices along with the average annual precipitation values for each studied weather station. It can be seen from Table 1 that all the studied weather stations are categorized as dry climate; the lack of easily accessible (fresh) water means that accurate estimation of ET_0 is crucial. The aridity index I_A is assessed on the basis of climate variables using the ratio of annual average precipitation to potential evapotranspiration:

$$I_A = \frac{P}{\mathrm{PET}} \tag{1}$$

where PET is annual potential evapotranspiration (assumed to be the same as ET_0, mm) and P (mm) is the average annual precipitation (UNEP 1992). Table 2 lists the statistical parameters of the data used during the study period. In the table, X_{mean}, X_{max}, X_{min}, S_X, C_V and C_{SX} denote the mean, maximum, minimum, standard deviation, coefficient of variation and coefficient of skewness, respectively. This table shows that the wind speed has a high skewed distribution for all stations.

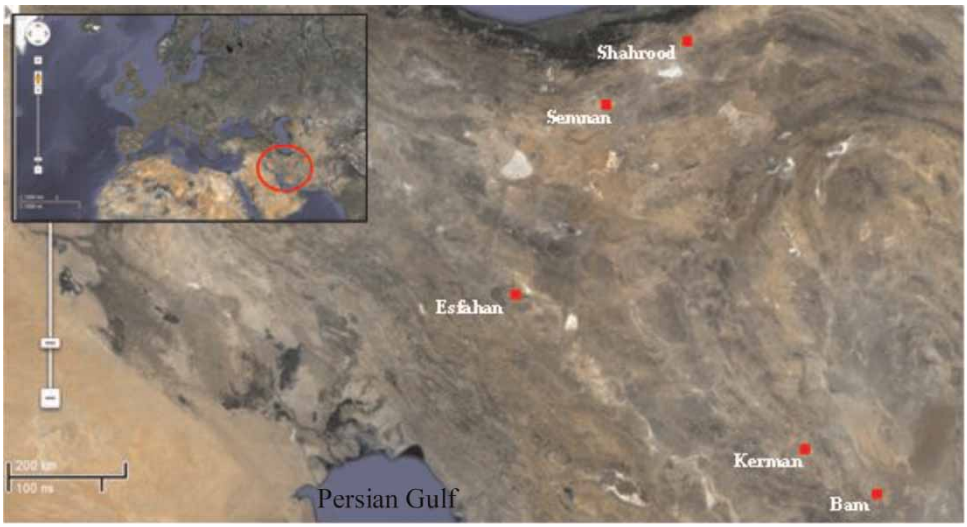

Figure 1 | Locations of the studied weather stations in Iran.

Table 1 | Summary of the weather station sites used in this study

		UTM coordinates					
Station	**Station code**	**Latitude (°N)**	**Longitude (°E)**	**Height (m)**	***N***	I_A	***P* (mm)**
Bam	1	29.06	58.21	1,066.9	1,829	0.063	116
Esfahan	2	32.37	51.40	1,550.4	1,829	0.097	159
Kerman	3	30.15	56.58	1,753.8	1,829	0.098	173
Semnan	4	35.35	53.33	1,130.8	1,829	0.131	179
Shahrood	5	36.25	54.57	1,345.3	1,829	0.125	143

Table 2 | The statistical parameters of the daily dataset

Data set	Unit	X_{max}	X_{min}	X_{mean}	S_X	C_V	C_{SX}
Bam							
T_{max}	°C	47.6	0	30	9.3	0.31	−0.56
T_{min}	°C	34.8	−4.4	17.8	8.6	0.48	−0.41
T_{mean}	°C	40.5	−1.1	24.3	8.9	0.36	−0.53
R_S	MJ m^{-2} day	28.9	3.9	17.7	4.8	0.27	−0.23
R_H	%	95.3	17	68.3	7.9	0.11	−2.4
u_2	m s^{-1}	19.8	0	6.1	2.5	0.42	2.14
Esfahan							
T_{max}	°C	41.6	−3.4	24.1	10.4	0.43	−0.29
T_{min}	°C	28.8	−11	9.7	8.8	0.92	−0.16
T_{mean}	°C	34.8	−5.9	16.9	9.5	0.56	−0.21
R_S	MJ m^{-2} day	30.4	3.5	18.8	6.2	0.33	−0.25
R_H	%	92	22	63	12.5	0.21	−0.21
u_2	m s^{-1}	11.4	0	1.9	1.5	0.79	1.32
Kerman							
T_{max}	°C	41.6	−2.8	25.2	9.1	0.36	−0.45
T_{min}	°C	26.6	−15.2	7.7	8.1	1.04	−0.17
T_{mean}	°C	39.4	−7.3	16.8	8.7	0.51	−0.11
R_S	MJ m^{-2} day	33.1	4.3	20.1	6.3	0.29	−0.34
R_H	%	93.9	11.6	55.9	10.9	0.19	−0.01
u_2	m s^{-1}	15.7	0	2.6	2.1	0.79	3.58
Semnan							
T_{max}	°C	43	−8.2	23.4	11.5	0.49	−0.32
T_{min}	°C	31.4	−20.2	11.4	10.3	0.89	−0.17
T_{mean}	°C	36.8	−13.9	17.4	10.8	0.62	−0.24
R_S	MJ m^{-2} day	24.5	3.3	16.7	6.2	0.37	−0.12
R_H	%	94.4	29.4	68.2	6.6	0.09	0.43
u_2	m s^{-1}	17.1	0	3.7	2.3	0.62	1.28
Shahrood							
T_{max}	°C	40.6	−6	21.3	10.6	0.49	−0.34
T_{min}	°C	27.6	−11.8	10.4	8.9	0.86	−0.21
T_{mean}	°C	34.1	−8.9	15.8	9.7	0.61	−0.28
R_S	MJ m^{-2} day	28.5	3.4	15.8	6.1	0.38	−0.14
R_H	%	93.9	48.5	70.4	6.2	0.08	0.85
u_2	m s^{-1}	8.5	0	2.3	1.5	0.67	0.81

Adaptive neuro-fuzzy inference system

There are two approaches for FISs, namely the approach of Mamdani (Mamdani & Assilian 1975) and that of Sugeno (Takagi & Sugeno 1985). The neuro-fuzzy model used in this study implements the Sugeno fuzzy approach to obtain the values of the output variable from the input variables. In the implementation of fuzzy logic, several types of

membership functions (MFs) can be used. However, recent studies have shown that the type of MF does not affect the results fundamentally (Vernieuwe *et al.* 2005). In the present study, triangular MFs were used as normal in practical applications (Russel & Campbell 1996). The number of MFs was determined iteratively. Figure 2 illustrates a schematic representation of the ANFIS model for ET_0 estimation using weather data. Detailed information about ANFIS may be found in, for example, Jang (1993).

In the present study, the grid partitioning identification methods of the Sugeno FIS models are applied for mapping the non-linear relationship among the input–output variables. Here the input variables of the ANFIS models are the weather data and the output layer corresponds to the ET_0 values. The grid partitioning method proposes independent partitions of each antecedent variable by defining the MFs of all antecedent variables. The number of MFs of each input variable (triangular functions in this case) is selected iteratively. A large number of MFs of input variables should be avoided to reduce computational cost (Keskin *et al.* 2004).

Gene expression programming

The procedure starts by random generation of chromosomes of the certain program (initial population). The generated chromosomes are then expressed and the performance of each individual program is evaluated against a set of performance cases (Ferreira 2006). The programs are then selected according to their own performance in that particular environment. The process is repeated until a good solution can be found for the studied phenomenon.

The application of GEP involves the selection of the following functions and operators:

1. The fitness function.
2. The set of terminals T and the set of functions F to create the chromosomes.
3. The chromosomal architecture.
4. The linking function.
5. The genetic operators.

In the present work the GeneXpro program was applied for modeling daily evaporation. Figure 3

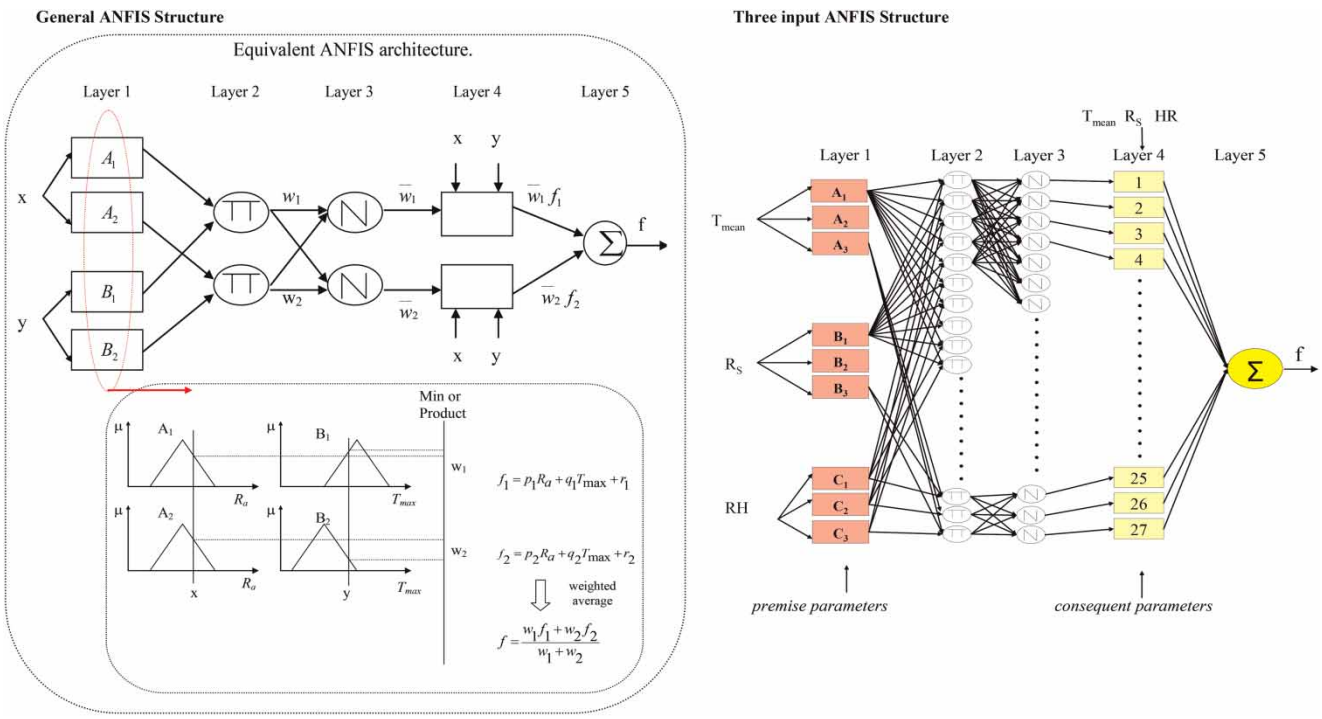

Figure 2 | ANFIS architecture.

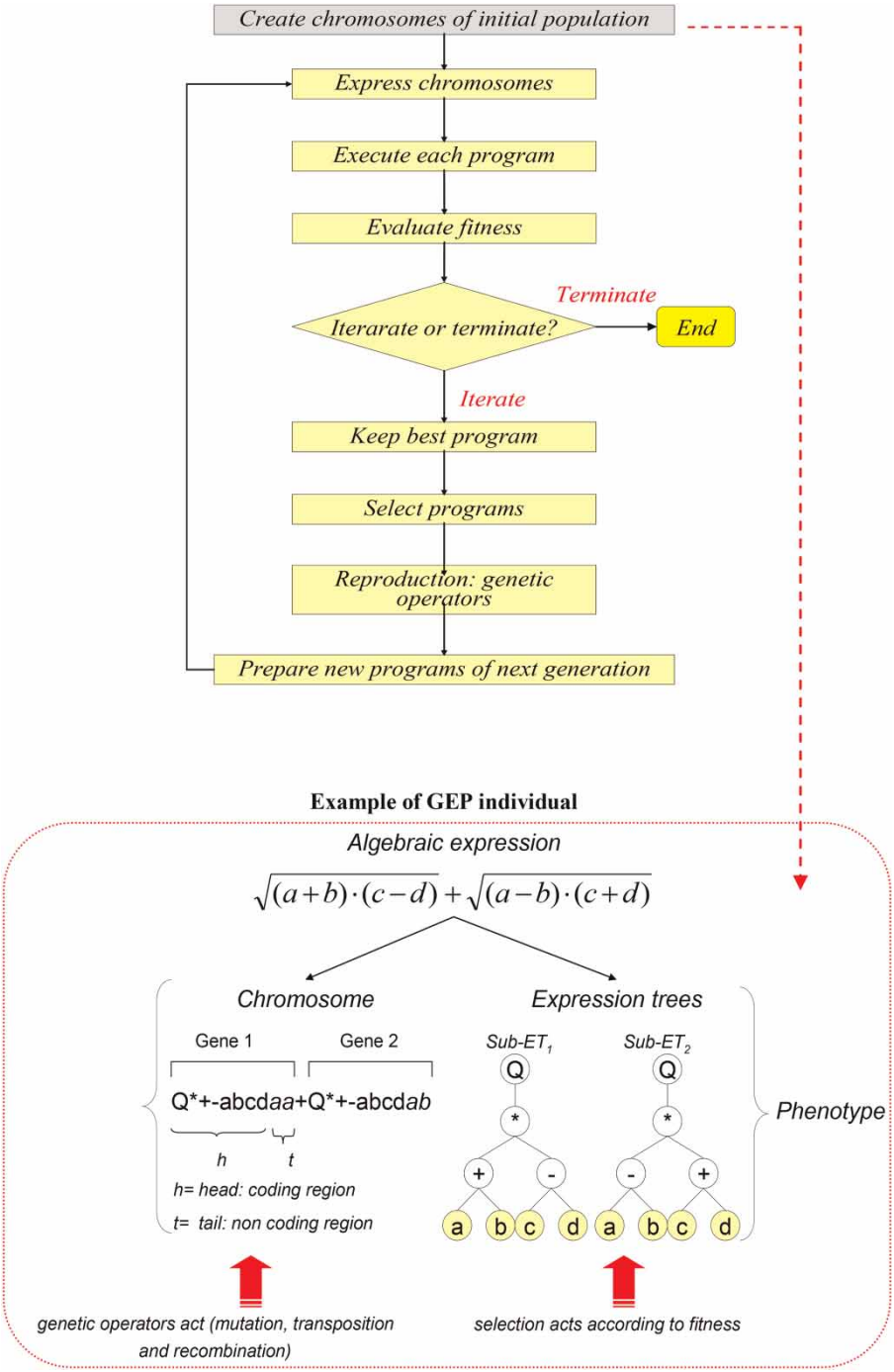

Figure 3 | General GEP model implementation and general structure.

shows the GEP model implementation steps. Detailed information about the aforementioned procedure for modeling ET can be found in, for example, Shiri *et al.* (2012).

Input selection and dataset assignment

The applied ET_0 equations are summarized in Table 3. ET_0 values produced by the FAO56-PM equation were considered

Table 3 | Applied ET_0 equations (where Δ is slope of the saturation vapor pressure function, kPa °C^{-1}; γ is psychometric constant, kPa °C^{-1}; R_n is net radiation, MJ m^{-2} day^{-1}; G is soil heat flux density, MJ m^{-2} day^{-1}; e_a is saturation vapor pressure, kPa; e_d is actual vapor pressure; $\alpha = 1.26$; λ is latent heat of the evaporation, MJ kg^{-1}; and R_a is extraterrestrial radiation, mm day^{-1})

ET_0 equation	Meteorological inputs	Equation
Hargreaves & Samani (1985)	T_{mean}, T_{max}, T_{min}	$ET_0 = 0.0023 R_a (T_{mean} + 17.8)\sqrt{T_{max} - T_{min}}$
Makkink (1957)	T_{mean}, T_{max}, T_{min}, R_S	$ET_0 = 0.61 \frac{\Delta}{\Delta + \gamma} \frac{R_S}{\lambda} - 0.12$
Turc (1961)	T_{mean}, R_S, R_H	$ET_0 = a_T 0.013 \frac{T_{mean}}{T_{mean} + 15} \frac{23.8856 R_S + 50}{\lambda}$ $R_H \geq 50 \rightarrow a_T = 1$ $R_H < 50 \rightarrow a_T = 1 + \frac{50 - R_H}{70}$
FAO56-PM (Allen *et al.* 1998)	T_{mean}, R_S, u, R_H	$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T_{mean} + 273} u_2 (e_a - e_d)}{\Delta + \gamma(1 + 0.34 u_2)}$

as the benchmark values for calibrating all the applied models. The Hargreaves–Samani model is recommended by Allen *et al.* (1998) in the presence of only air temperature records. This model requires only air temperature (as meteorological input, which is easily measured at each weather station) and extraterrestrial radiation (Droogers & Allen 2002). This model is generally more accurate for windy locations with large daily temperature amplitude, as well as for locations with light wind conditions, where the daily air temperature amplitude is low to moderate (Gavilán *et al.* 2006).

The Makkink (1957) equation is a semi-empirical model for ET_0 calculation, which has been developed in the Netherlands by comparing the Penman (1948) ET values with lysimetric values. The Turc (1961) model is also a simplification of the Makkink model and requires air temperature, solar radiation and relative humidity as input variables. The consideration of temperature-based solar-radiation-based models depends on the availability of meteorological sensors in the studied weather station. The presence of temperature sensors is more usual than the presence of solar radiation sensors in weather stations.

According to the applied equations, the input combinations evaluated in the current work are summarized in Table 4. Input combinations (1)–(3) were considered to allow for a valid comparison with the Hargreaves–Samani, Makkink and Turc equations, respectively.

For each applied model, the data from 1 January 2004 to 31 December 2006 were used for training, while the remaining data were reserved for testing. Two main data management scenarios were evaluated: (1) at each station; and (2) pooled data. In scenario 1, the performance of each applied model was evaluated at each individual station. Data from the period 2004–2006 were used as training data for all locations: Bam, Esfahan, Kerman, Shahrood and Semnan, with 1,095 patterns for each. Data from the period 2007–2008 were used for independent testing of the models at each station. In scenario 2, the performance of the models was analyzed globally using the pooled data of the five studied weather stations. In this case, data from the period 2004–2006 were used for training (5,475 patterns) and data from the period 2007–2008 were used for independent validation of the models. The performance results of all AI models were referred to the global test set, considering all stations together, and split up per test station.

Table 4 | Inputs used for implementation of AI models

	AI1	AI2	AI3
T_{max}	–	–	
T_{min}	–	–	
T_{mean}	–	–	–
R_a	–		
R_S		–	–
R_H			–

Statistical parameters

Four statistical parameters were used to assess model performance, namely: the coefficient of determination (R^2), scatter index (SI), mean absolute error (MAE) and

Nash–Sutcliffe coefficient (NS), defined as follows:

$$R^2 = \left(\frac{\sum_{i=1}^{n} (ET_{i0} - ET_{0,mean})(ET_{iM} - ET_{M,mean})}{\sqrt{\sum_{i=1}^{n} (ET_{i0} - ET_{0,mean})^2 \sum_{i=1}^{n} (ET_{iM} - ET_{M,mean})^2}} \right)^2 \quad (2)$$

$$SI = \frac{RMSE}{ET_{0,mean}} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n} (ET_{iM} - ET_{i0})^2}}{ET_{0,mean}} \quad (3)$$

$$MAE = \frac{\sum_{i=1}^{n} abs(ET_{i0} - ET_{iM})}{n} \quad (4)$$

$$NS = 1 - \frac{\sum_{i=1}^{n} (ET_{iM} - ET_{i0})^2}{\sum_{i=1}^{n} (ET_{i0} - ET_{0,mean})} \quad (5)$$

where ET_M and ET_0 denote the values generated by different models and by the FAO56-PM ET_0 equation at the ith time step, respectively; n is number of time steps; and $ET_{0,mean}$ and $ET_{M,mean}$ represent the mean ET values estimated by FAO56-PM and applied models, respectively.

RESULTS AND DISCUSSION

Individual station scenario

In this part of the study, the performance of GEP, ANFIS and empirical models was evaluated locally at each individual station. Table 5 sums up the accuracy parameters of the applied models considering the individual station

Table 5 | Error statistics of GEP and ANFIS models (scenario 1) during the test period

		Test station				
Model	**Index**	**Bam**	**Esfahan**	**Kerman**	**Semnan**	**Shahrood**
GEP1	R^2	0.959	0.926	0.907	0.920	0.956
	SI	0.11	0.14	0.14	0.16	0.12
	MAE (mm)	0.452	0.491	0.428	0.480	0.301
	NS (–)	93.1	92.4	90.5	91.5	95.5
ANFIS1	R^2	0.978	0.923	0.907	0.921	0.968
	SI	0.07	0.14	0.15	0.16	0.11
	MAE (mm)	0.320	0.482	0.343	0.467	0.252
	NS (–)	96.6	92	89.6	91.5	96.6
GEP2	R^2	0.969	0.929	0.838	0.909	0.886
	SI	0.08	0.14	0.21	0.17	0.21
	MAE (mm)	0.354	0.489	0.399	0.511	0.512
	NS (–)	95.8	92.7	79.1	90.1	87.4
ANFIS2	R^2	0.981	0.924	0.810	0.883	0.967
	SI	0.07	0.14	0.22	0.19	0.11
	MAE (mm)	0.304	0.483	0.375	0.471	0.254
	NS (–)	97	92	77.5	87.7	96.5
GEP3	R^2	0.973	0.933	0.911	0.877	0.936
	SI	0.08	0.13	0.14	0.19	0.15
	MAE (mm)	0.350	0.477	0.466	0.585	0.371
	NS (–)	96.3	93.1	90.8	87.4	93.3
ANFIS3	R^2	0.980	0.935	0.763	0.926	0.953
	SI	0.07	0.14	0.26	0.15	0.13
	MAE (mm)	0.309	0.454	0.424	0.448	0.294
	NS (–)	96.8	92.8	68.3	92.3	95.1

approach. Figure 4 represents the scatter plots of the applied GEP and ANFIS models in each station corresponding to this scenario. The obtained results clearly show the superiority of the GEP and ANFIS models over the corresponding ET_0 equations. A comparison between input combinations reveals that the input combination (1) (i.e. AI1 models) produces more accurate results than the other applied models for ET_0 modeling. A comparison of GEP and ANFIS models indicates that the latter show better accuracy than the corresponding GEP models for the first and second input combinations. For the third input combination however, the GEP3 model (relying on mean air temperature, solar radiation and relative humidity) performs better than the ANFIS3 model. From the presented values in Table 5, it is seen that the GEP and ANFIS models for all input combinations perform better than the corresponding ET_0 equations. The Hargreaves–Samani model shows higher accuracy than the other equations and its error statistics are closer to the corresponding GEP and ANFIS results for the whole studied stations.

In the case of first input combination (AI1 models), the GEP1 and ANFIS1 models almost have the same results based on SI statistics in Esfahan, Kerman and Shahrood. In Bam and Semnan, however, the ANFIS1 model performs more accurately than the GEP1 model. For the second input combination (AI2 models), GEP2 shows better accuracy than the ANFIS2 model in Esfahan, Kerman and Shahrood. In Bam and Semnan, the ANFIS2 model has a higher accuracy than the GEP2 model. For the third input combination (AI3 models), the ANFIS3 model performs better than the GEP3 model in Shahrood and Semnan, while the GEP3 model has a higher accuracy in Kerman. In Bam and Esfahan, the models demonstrate similar accuracies.

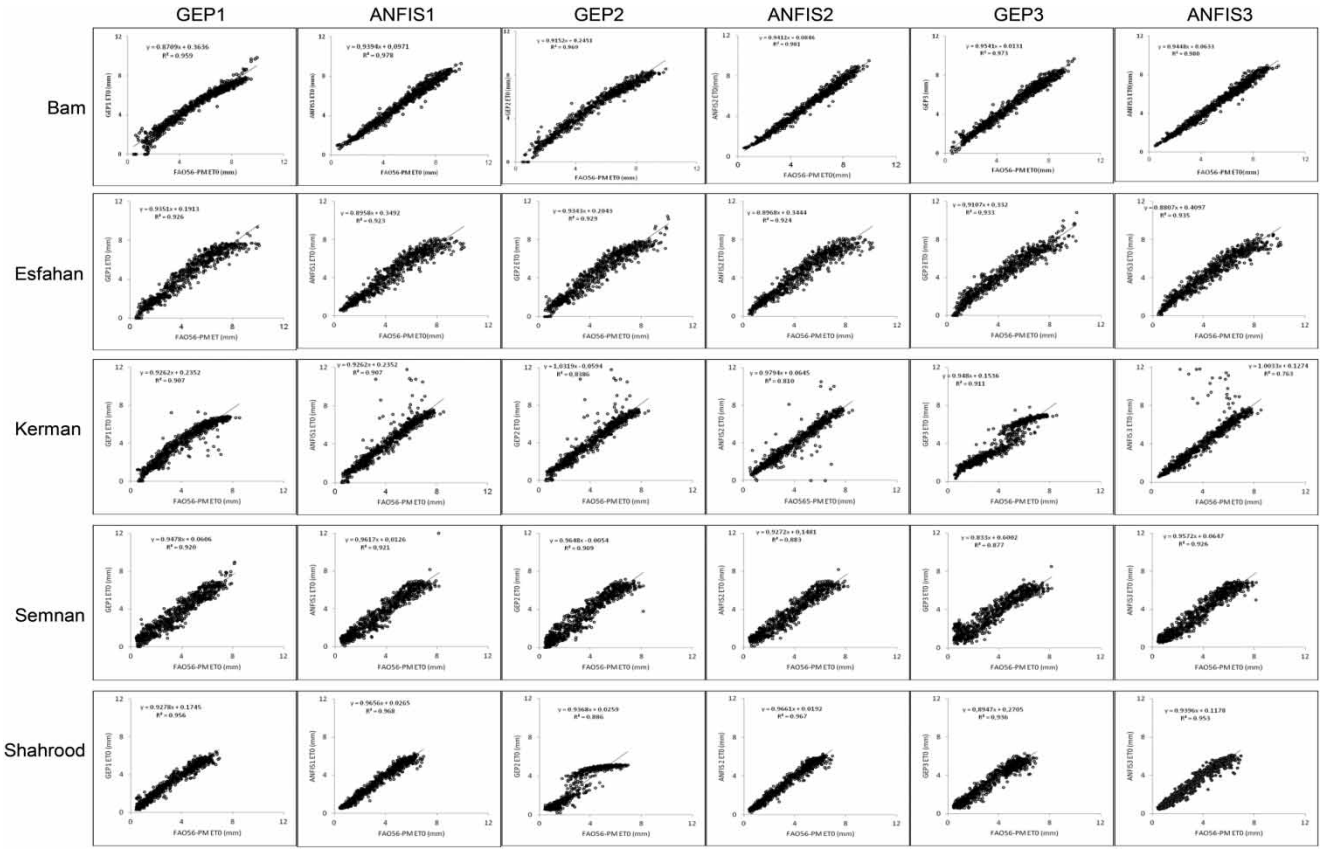

Figure 4 | Scatter plots of the applied models for individual station approach (scenario 1).

Pooled approach

The comparison of different GEP and ANFIS models trained with pooled data and tested in each station are presented in Table 6 and depicted in Figure 5. At Esfahan and Semnan stations, the GEP2 and GEP3 models showed almost the same accuracy. For the first input combination (AI1 models), the GEP1 and ANFIS1 models present similar results in Bam. The GEP1 model performs better than the ANFIS1 model in Kerman, while the ANFIS1 model has a higher accuracy in Esfahan, Semnan and Shahrood. For the second input combination (AI2 models), the GEP2 and ANFIS2 models demonstrate similar results in Esfahan and Semnan. The ANFIS2 model performs better than the GEP2 model in Bam, Kerman and Shahrood. For the third input combination (AI3 models), the ANFIS3 model performs better than the GEP3 model in Esfahan and Kerman. In Bam, Semnan and Shahrood, the models show similar accuracy.

Table 7 sums up the global indicators/parameters of the GEP and ANFIS models trained with pooled individual data. In the case of first input combination (AI1 models), pooled training decreases the accuracy of the GEP1 model while it increases the accuracy of the ANFIS1 model. The accuracy of the GEP1 and ANFIS2 models decreases by pooled training for the first input combination (AI1 models). In the case of second input

Table 6 | Error statistics of GEP and ANFIS models (trained with pooled data-tested at each station)

		Test station				
Model	**Index**	**Bam**	**Esfahan**	**Kerman**	**Semnan**	**Shahrood**
GEP1	R^2	0.958	0.895	0.913	0.953	0.943
	SI	0.10	0.22	0.14	0.14	0.21
	MAE (mm)	0.419	0.745	0.417	0.454	0.531
	NS (–)	94.5	82	91.3	92.9	88.3
ANFIS1	R^2	0.970	0.923	0.867	0.959	0.956
	SI	0.11	0.18	0.19	0.12	0.16
	MAE (mm)	0.451	0.595	0.405	0.355	0.417
	NS (–)	93.9	88.8	86.9	96.8	96.5
GEP2	R^2	0.964	0.917	0.864	0.957	0.943
	SI	0.13	0.19	0.21	0.12	0.21
	MAE (mm)	0.551	0.635	0.478	0.358	0.544
	NS (–)	90.7	86.6	81.1	95	88
ANFIS2	R^2	0.971	0.923	0.910	0.958	0.958
	SI	0.11	0.18	0.15	0.13	0.16
	MAE (mm)	0.449	0.592	0.352	0.357	0.409
	NS (–)	93.9	88.9	91.9	96.8	96.6
GEP3	R^2	0.972	0.913	0.807	0.956	0.943
	SI	0.11	0.19	0.25	0.12	0.16
	MAE (mm)	0.488	0.642	0.559	0.348	0.413
	NS (–)	92.5	86.8	72.6	95.2	92.1
ANFIS3	R^2	0.981	0.929	0.836	0.959	0.949
	SI	0.11	0.16	0.22	0.12	0.15
	MAE (mm)	0.455	0.544	0.478	0.343	0.384
	NS (–)	94.1	90.5	83.1	97	96.9

Figure 5 | Scatter plots of the applied models for pooled data approach (scenario 2).

Table 7 | Global assessment of the applied models (using two scenarios) during the test period

Data management criterion	Model	R^2	SI	MAE (mm)	NS
Individual training	GEP1	0.933	0.136	0.430	92.6
	ANFIS1	0.939	0.130	0.373	93.3
Pooled training	GEP1	0.932	0.160	0.513	89.3
	ANFIS1	0.935	0.150	0.444	92.6
–	Hargreaves–Samani	0.936	0.162	0.513	91.8
Individual training	GEP2	0.906	0.166	0.453	89.2
	ANFIS2	0.912	0.150	0.377	90.1
Pooled training	GEP2	0.928	0.170	0.513	88.3
	ANFIS2	0.944	0.140	0.431	93.6
–	Makkink	0.927	0.327	0.137	78.1
Individual training	GEP3	0.926	0.133	0.449	92.2
	ANFIS3	0.912	0.153	0.386	89.1
Pooled training	GEP3	0.918	0.160	0.490	87.8
	ANFIS3	0.930	0.150	0.441	92.3
–	Turc	0.900	0.748	2.788	69.2

combination (AI2 models), pooled training decreases the accuracy of the GEP2 model while it increases the accuracy of the ANFIS1. The accuracy of the GEP3 model decreases while the accuracy of the ANFIS3 model increases slightly by pooled training for the third input combination. It can be said that pooled training generally increases ANFIS model accuracies while it decreases GEP model accuracies. The comparison of GEP and ANFIS models highlights that the ANFIS models generally perform better than the corresponding GEP models. The reason for poor performance of AI3 models (relative to AI1 and AI2 models) may be the weather conditions of the studied stations, which is a dry climate meaning that relative humidity cannot affect ET_0 values as much as in more humid regions.

The results described here, derived from a single chronological dataset, should be confirmed in further studies through dataset scanning approaches (leave-one-out procedures). Further research should also deal with models

fed with more climatic inputs, as well as under different climatic conditions.

CONCLUSIONS

In this modeling study, we have evaluated different data management scenarios when developing GEP and ANFIS models for estimating daily ET_0 by using weather data from five stations in Iran. Two main training scenarios were evaluated: in the first scenario the GEP and ANFIS models were trained and tested using the local data of each studied weather station and in the second the models were developed using the combined data from the five weather stations. AI techniques always demonstrated higher accuracy than the corresponding empirical approaches for the three input combinations.

In the first scenario, the comparison of GEP and ANFIS models revealed that the ANFIS models generally performed better than the GEP models. In the second scenario, comparison of different GEP and ANFIS models trained and tested with pooled data indicated that the GEP models performed slightly better than the ANFIS models. The comparison of the two scenarios revealed that pooled training (scenario 1) might increase the accuracy of ANFIS models while it might decrease the accuracy of GEP models. Accordingly, the consideration of additional exogenous patterns does not drastically improve the performance of locally trained models relying on temperature and solar radiation information in the studied region. The encountered relationships based on these inputs might be site specific. These inputs therefore might not allow performance improvement through wider training ranges. These conclusions might not apply for other input combinations and climatic conditions.

ACKNOWLEDGEMENTS

This work was prepared as part of Jalal Shiri's PhD thesis.

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.

Babovic, V., Kanizares, R., Jenson, H. R. & Klinting, A. 2001 Neural networks as routine for error updating of numerical models. *Journal of Hydrological Engineering* **127** (3), 181–193.

Dogan, E. 2009 Reference evapotranspiration estimation using adaptive neuro-fuzzy inference system. *Journal of Irrigation and Drainage Engineering* **58**, 617–628.

Doorenbos, J. & Pruitt, W. O. 1977 Crop water requirements. Irrigation Drainage paper No. 24, FAO, Rome, Italy.

Droogers, P. & Allen, R. G. 2002 Estimating reference evapotranspiration under inaccurate data conditions. *Irrigation and Drainage Systems* **16** (1), 33–45.

Ferreira, C. 2006 *Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence*. Springer, Berlin, Heidelberg, New York.

Gavilán, P., Lorite, I. J., Tornero, S. & Berengena, J. 2006 Regional calibration of Hargreaves equation for estimating reference ET in a semi arid environment. *Agricultural Water Management* **81**, 257–281.

Goldberg, D. E. 1989 *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, MA.

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

Jang, J. S. R. 1993 ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics* **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.

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

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

Kim, S., Shiri, J. & Kisi, O. 2012 Pan evaporation modeling using neural computing approach for different climatic zones. *Water Resources Management* **26** (11), 3231–3249.

Kisi, O. 2006 Daily pan evaporation modeling using a neuro-fuzzy computing technique. *Journal of Hydrology* **329**, 636–646.

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

Kisi, O. & Shiri, J. 2011 Precipitation forecasting using wavelet-genetic programming and wavelet-neuro-fuzzy conjunction models. *Water Resources Management* **25** (13), 3135–3152.

Kisi, O. & Shiri, J. 2012a Wavelet and neuro-fuzzy conjunction model for predicting water table depth fluctuations. *Hydrology Research* **43** (3), 286–300.

Kisi, O. & Shiri, J. 2012b River suspended sediment estimation by climatic variables implication: comparative study among soft computing techniques. *Computers and Geosciences* **43**, 73–82.

Koza, J. R. 1992 *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. The MIT Press, Cambridge, MA.

Kumar, M., Raghuwanshi, N. S. & Singh, R. 2011 Artificial neural networks in evapotranspiration modelling: a review. *Irrigation Science* **1**, 11–25.

Landeras, G., Lopez, J. J., Kisi, O. & Shiri, J. 2012 Comparison of gene expression programming with neuro-fuzzy and neural network computing techniques in estimating daily incoming solar radiation in the Basque Country (Northern Spain). *Energy Conversion and Management* **62**, 1–13.

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). *Agricultural Water Management* **95**, 553–565.

Makkink, G. F. 1957 Testing the Penman formula by means of lysimeters. *Journal of the Institution of Water Engineers and Scientists* **11** (3), 277–288.

Mamdani, E. H. & Assilian, S. 1975 An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies* **7** (1), 1–13.

Penman, H. L. 1948 Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society London, A* **193** (1032), 120–145.

Pour Ali Baba, A., Shiri, J., Kisi, O., Fakheri Fard, A., Kim, S. & Amini, R. 2013 Estimating daily reference evapotranspiration using available and estimated climatic data by adaptive neuro-fuzzy inference system (ANFIS) and artificial neural networks (ANN). *Hydrology Research* **44** (1), 131–146.

Russel, S. O. & Campbell, P. F. 1996 Reservoir operating rules with fuzzy programming. *Journal of Water Resources Planning and Management* **122** (3), 165–170.

Shiri, J. & Kisi, O. 2010 Short term and long term streamflow forecasting using a wavelet and neuro-fuzzy conjunction model. *Journal of Hydrology* **394**, 486–493.

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). *Journal of Irrigation and Drainage Engineering* **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. *Computers and Geoscience* **37** (10), 1692–1701.

Shiri, J., Dierickx, W., Pour-Ali Baba, A., Neamati, S. & Ghorbani, M. A. 2011 Estimating daily pan evaporation from climatic data of the State of Illinois, USA using adaptive neuro-fuzzy inference system (ANFIS) and artificial neural networks (ANN). *Hydrology Research* **42** (6), 491–502.

Shiri, J., Kisi, O., Landeras, G., Lopez, J. J., Nazemi, A. H. & Stuyt, L. C. P. M. 2012 Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country (Northwestern Spain). *Journal of Hydrology* **414–415**, 302–316.

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

Trajkovic, S. 2005 Temperature-based approaches for estimating reference evapotranspiration. *Journal of Irrigation and Drainage Engineering* **131** (4), 316–323.

Turc, L. 1961 Evaluation des besoins en eau d'irrigation, évapotranspiration potentielle. *Annales Agronomiques* **12** (1), 13–49.

UNEP (United Nations Environment Programme) 1992 *World Atlas of Desertification*. Edward Arnold, London.

UNEP (United Nations Environmental Programme) 1997 *World Atlas of Desertification*. Editorial commentary by N. Middleton and D.S.G. Thomas. Edward Arnold, London.

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. *Journal of Hydrology* **302** (1–4), 173–186.

First received 22 September 2012; accepted in revised form 15 December 2012. Available online 18 February 2013