Research Journal of Applied Sciences, Engineering and Technology 14(10): 386-398, 2017
DOI:10.19026/rjaset.14.5131
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: May 15, 2017 Accepted: June 23, 2017 Published: October 15, 2017

Research Article

A Committee Machine with Intelligent Systems for Estimating Monthly Mean Reference Evapotranspiration in an Arid Region

[1]Ali H. Al-Aboodi, [2]Alaa M. Al-Abadi and [1]Husham T. Ibrahim
[1]Department of Civil Engineering, College of Engineering, University of Basrah, Basrah, Iraq
[2]Department of Geology, College of Sciences, University of Basrah, Basrah, Iraq

Abstract: The aim of this research is to estimate the reference evapotranspiration ET_o as given by FAO-56 PM equation in Basrah city, southern Iraq by using several climatic inputs data including maximum monthly mean air temperature, minimum monthly mean air temperature, monthly mean relative humidity and monthly mean wind speed. Three artificial intelligent systems (generalized regression neural network GRNN, multi-layer perceptron MLP and adaptive neuro-fuzzy inference systems ANFIS) were used for predicting reference evapotranspiration. Root mean squared error and coefficient of determination were used as comparison criteria for evaluation of performance of all the developed models. The results shown that the models performances of multi-layer perceptron models are better than adaptive neuro-fuzzy inference systems models and slightly better than generalized regression neural network models with different inputs combination. A Committee Machine with Intelligent Systems (CMIS) was constructed for estimation of ET_o by integrating the results of predicting ET_o from GRNN, MLP and ANFIS, each of them has a weight factor representing its contribution in overall estimation. The results illustrated that the performance of committee machine with intelligent systems is better than any one of the individual artificial intelligent systems for predicting ET_o.

Keywords: Adaptive neuro-fuzzy inference systems, artificial neural network, basrah, committee machine, evapotranspiration, Iraq

INTRODUCTION

Evapotranspiration (ET) is a term used to denote all processes that converting the existing water on the surface into water steam. ET is an essential component in global water energy and carbon cycles and thus provides a link between the atmosphere and the Earth's surface (Tang *et al*., 2014). The accurate estimation of ET is important for studying hydrological water balance, design of irrigation systems, simulation of crop yield and even efficient planning of water resources projects (Kumar *et al*., 2011). However, ET is a complex process because it depends on different factors such as weather data and growth stage of the crop (Trajkovic and Kolakovic, 2009). To avoid the need to calibrate a separate ET equation for each crop and stage of growth, the concept of reference evapotranspiration (ET_0) was introduced by Allen *et al*. (1998). ET_0 is defined as the rate of ET from a hypothetical crop with an assumed crop height of (0.12 m), a fixed surface resistance of (70 sec/m) and an albedo of (0.23), which would closely resemble ET from an extensive surface of green grass of uniform height, activity growing, well-watered and completely shading the ground.

The importance of ET_0 in hydrological and agricultural studies leads to develop different instruments and methodologies to estimate it. The ET_0 can be directly measured using either lysimeter field instrument, water balance approach, or estimated indirectly using the climatological data (Kumar *et al*., 2011). Unfortunately, the available lysimeter data are very limited or sometimes non-existent in developing countries. Because of these difficulties in estimating ET_0, the indirect ET_0 estimation methodologic which are essentially depending on an easy to capture meteorological data become more popular. In recent few decades, numerous methodologies, classified as temperature-based, radiation-based, pan-evaporation-based and combination-type, have been developed for estimating ET_0 (Trajkovic and Kolakovic, 2009). One of the methodologies that are widely used to estimate ET_0 is the FAO Penman-Monteith method.

Corresponding Author: Ali H. Al-Aboodi, Department of Civil Engineering, College of Engineering, University of Basrah, Basrah, Iraq

The Penman-Monteith method is an accurate method for estimating evapotranspiration and can be used in different regimes (Kumar *et al*., 2002). The effectiveness of this method for estimating ET_0 and for evaluating other equations have been indicated by many studies (Pereira and Pruitt, 2004; López-Urrea *et al*., 2006; Gavilān *et al*., 2006). The main advantages of Penman-Monteith equation are (Landeras *et al*., 2008):

- It is applicable in different environments under different climatic scenarios without local calibration
- The adapted equation has been validated using lysimeters data under a wide range of climatic conditions.

The main disadvantage of this method is that it requires a large number of climatic variables such as air temperature, relative humidity, solar radiation and wind speed to compute ET_0 which are not always available in meteorological stations or at least missing for a certain period. To fill this gap, many researchers attempted to use artificial intelligent techniques such as Artificial Neural Networks (ANNs), Adaptive Neuro-Fuzzy Inference System (ANFIS) and Genetic Programming to estimate ET_0 with promising and successful results.

Most of the previous studies mainly focused on one or more techniques for estimating ET_0, independently. A Committee Machine (CM), or committee neural network, has a parallel architecture that produces a final output by combining the results of individual experts (Haykin, 1991). The experts may be neural networks, empirical formulas, or other algorithms (Chen and Lin, 2006). The main advantage of CM technique is that it can lead to significant improvements in the performance on new data, with little extra computational effort. In fact, the combined response of the CM performs the best to those of its constituent experts. The efficacy of CM for estimating ET_0 is not investigated yet; therefore, the objective of this study is to use three intelligent systems namely Generalized Regression Neural Network (GRNN), Multi-Layer Perceptron (MLP) and Adaptive Neuro-Fuzzy Inference System (ANFIS) along with CM to develop more accurate model for estimating ET_0 from available meteorological data in an arid region. The Basra City in southern Iraq has been selected to demonstrate the adapted methodology. The optimum weights for CM is optimally computed for the first time using Pattern Search (PS) optimization technique.

MODELINGTECHNIQUSE

Penman-monteith method: The FAO-56 PM method is recommended as highly accurate method for determining ET_0. This method is a physically based approach and requires measurements of air temperature, relative humidity, solar radiation and wind speed as input to estimate ET_0. In this study the FAO-56 PM method was used as a reference model for assessing the performance of the usedapproaches.FAO-56 PM equation which given by (Allen *et al*., 1998) is shown in Eq. (1):

$$ET_o = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

where,

ET_o : Reference evapotranspiration [mm/day]
Rn : Net radiation at the crop surface [MJ/m^2/day]
G : Soil heat flux density [MJ/m^2/day^2]
T : Mean daily air temperature at 2 m height [°C]
u_2 : Wind speed at 2 m height [m/s]
e_s : Saturation vapor pressure [kPa]
e_a : Actual vapour pressure [kPa]
e_s-e_a : Saturation vapour pressure deficit [kPa]
Δ : Slope vapour pressure curve [kPa°C^{-1}]
γ : Psychrometric constant [kPa°C^{-1}]

Generalized Regression Neural Network (GRNN): GRNN is a variation of radial basis neural networks, which is designed for function approximation and regression (Alilou and Yaghmaee, 2015). GRNN is a universal approximation for smooth function, allowing it to solve any function approximation and estimate any continuous variables when giving enough data (Disorntetiwat, 2001). GRNN is a one-pass learning algorithm with a highly parallel structure (Specht, 1991). Basically, GRNN consists of four layer (Fig. 1); the input layer, the pattern layer, the summation layer and the output layer (Barzegar *et al*., 2016). The number of input units in input layer depends on the total number of the observation parameters (Hannan *et al*., 2010). The first layer feeds the inputs to the second layer, where each unit represents a training input pattern. In the second layer (pattern layer) the Euclidean distance and activation function are calculated. The pattern layer is connected with the weights of the summation layer to the two neurons in the summation layer. Summation layer has two subparts namely Numerator and Denominator parts. Summation of the multiplication of training data and activation function continued within Numerator part, while the summation of all activation function implement in Numerator part. The summation layer feeds both the Numerator and Denominator to the next output layer. Equations (2 and 3) show the output layer which dividing the Numerator part by that of each

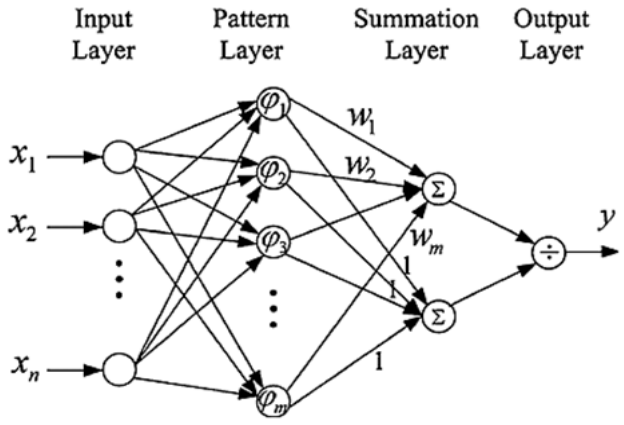

Fig. 1: Block diagram of GRNN architecture

Denominator part yielding the predicted values of an unknown input vector x (Specht, 1991):

$$y_i = \frac{\sum_{i=1}^{n} W_i \exp\left(-D\left(x, x_i\right)\right)}{\sum_{i=1}^{n} \exp\left(-D\left(x, x_i\right)\right)} \quad (2)$$

$$D\left(x, x_i\right) = \sum_{k=1}^{m}\left(\frac{x_i - x_{ik}}{\sigma}\right)^2 \quad (3)$$

where,

W_i : The weight connection between the i^{th} neuron in the pattern layer and summation neuron
n : The number of the training patterns
D : The Gaussian function
m : The number of elements of an input vector
x_k,x_{ik} : The j^{th} element of x and x_i, respectively
σ : The spread parameter, whose optimal value is determined experimentally

During the training process, the error is measured by the Means Squared Error (MSE). The training process is repeated for several times with different spread factors until the network is optimized according to the minimum amount of MSE or a pre-defined threshold value (Kisi *et al*., 2015).

Multi-Layer Perceptron (MLP): The limitations of single layer artificial neural network have led to development of multi-layer feed-forward networks with one or more hidden layers, called Multi-Layer Perceptron (MLP) networks. MLP networks overcome many of the limitations of single layer perceptrons. Multi-Layer Perceptron (MLP) is artificial neural network, the computation in MLP is performed using a set of many simple units with weighted connections between them. MLP is a feed-forward artificial neural network model that maps sets of input data onto a set of appropriate outputs. MLP consists of multiple layers of nodes; each layer is fully connected with another one. Node is called a neuron (or processing element) with a nonlinear activation function. MLP utilizes a supervised learning technique called back-propagation for training the network (Rumelhart and McClelland, 1986).

Learning occurs in the perceptron by changing connection weights after each piece of data is processed, based on the amount of error in the output compared to the expected result (target). Figure 2 show the two-layered feed forward neural networks with sigmoid hidden neurons and linear output neurons. This network includes a nonlinear activation function. The important point to emphasize here is that the smoothly nonlinearity (i.e., differentiable everywhere), as opposed to the hard limiting used in Rosenblatt's perceptron. A commonly used form of nonlinearity that satisfies this requirement is a sigmoid nonlinearity defined by the logistic function which shown in Eq. (4):

$$y_i = \frac{1}{1 + \exp\left(-v_j\right)} \quad (4)$$

where,

y_i : The output of the neuron;
v_j : The induced local field of neuron j (i.e., the weighted sum of all synaptic inputs plus the bias).

The explicit expression for an output value of MLP as shown in Eq. (5) (Nourani and Babakhani, 2013):

$$y_k = f_\circ\left[\sum_{i=1}^{M_N} W_{kj} \cdot f_h\left(\sum_{i=1}^{N_N} W_{ji} X_i + W_{j\circ}\right) + W_{k\circ}\right] \quad (5)$$

where,

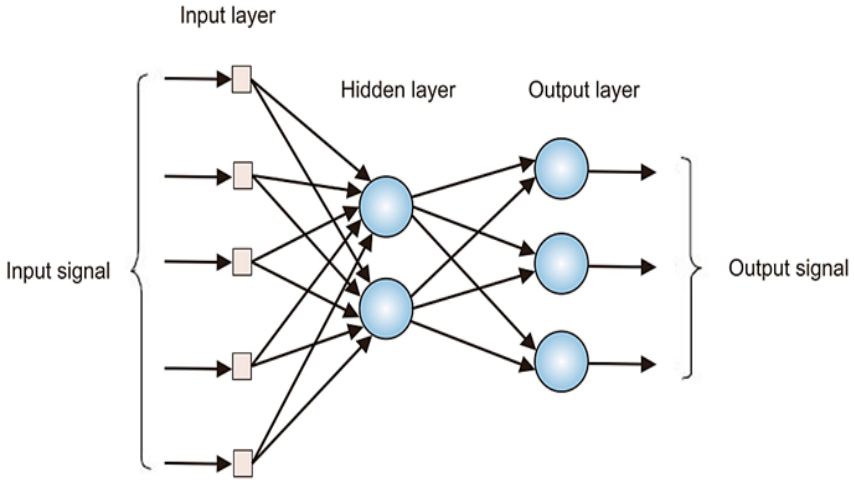

Fig. 2: Two-layered feed forward neural networks

W_{ji} : A weight in the hidden layer connecting the i^{th}neuron in the input layer and the *j*th neuron in the hiddenlayer
W_{jo} : The bias for the j^{th} hidden neuron
f_h : The activation function of the hidden neuron
W_{kj} : A weight in the output layer connecting the j^{th} neuron in the hidden layer and the k^{th} neuron in the output layer
W_{ko} : The bias for the k^{th} output neuron
f_o : The activation function for the output neuron
X_i : i^{th} input variable for input layer
y_k : Computed output variable
M_N&N_N : The number of the neurons in the input and hidden layers, respectively

MLP has been applied successfully to solve difficult problems in different cases with a highly popular algorithm known as the error back-propagation algorithm. This algorithm is based on the error-connection learning rule.

Adaptive Neuro-Fuzzy Inference Systems (ANFIS): Adaptive Neuro Fuzzy Inference System (ANFIS) is a fuzzy mapping algorithm that is based on Takagi-Sugeno fuzzy inference system. It integrates both neural networks and fuzzy logic principles (Loukas, 2001). The parameters associated with the membership functions changes through the learning process. The computation of these parameters (or their adjustment) is facilitated by a gradient vector. This gradient vector provides a measure of how well the fuzzy inference system is modeling the input/output data for a given set of parameters. When the gradient vector is obtained, any of several optimization routines can be applied in order to adjust the parameters to reduce some error measure. This error measure is usually defined by the sum of the squared difference between actual and desired outputs. The shape of membership functions is obtained in neuro-fuzzy by training them with input/output data rather than specifying them manually. The ANFIS consists of five layers (Fig. 3), the basic functions of each layer are the input, fuzzification, rule inference, normalization and defuzzification.

ANFIS can be represented as a linear arrangement of input variables and a constant term as described by Eq. (6) (Hossen *et al*.,2013):

$$Rulei: If\, x_{k1} is F_{i1} AND x_{k2} is F_{i2} \ldots AND x_{km} is F_{im} THEN y_i = c_{i0} + c_{i1}x_{k1} + \cdots + c_{im}x_{km}$$
$$y = \frac{\sum_{i=1}^{N} w_i y_i}{\sum_{i=1}^{N} w_i}, w_i = \prod_{j=1}^{m} F_{ij}(x_{kj}) \quad (6)$$

where,
$Rulei\ (i = 1,2, \ldots, N)$: The i^{th} fuzzy rule
$x_{kj} (j = 1,2, \ldots, m)$: The j^{th} input variable of the k^{th} pattern vector
F_{ij} : A fuzzy variable of the j^{th} input variable in the i^{th} rule
$\Pi_{j=1}^{m}$: A fuzzy T-norm operator
w_i : A rule firing-strength of the i^{th} rule
y_i : The i^{th} rule output
y : The overall output

The clustering algorithm is used in this research, the clustering algorithm is a method which is usually employed to discover a cluster center and inform the position of heart (center) of each cluster (Stoffel *et al*., 2012). It provides a method that shows how to group data points that populate some multidimensional space into a specific number of different clusters (Elleithy, 2010).

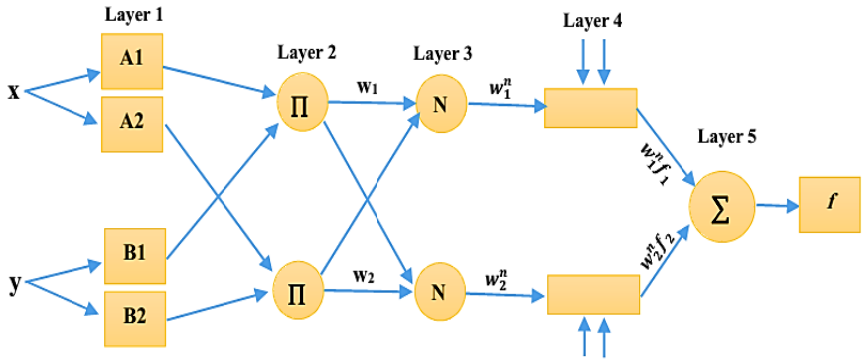

Fig. 3: ANFIS architecture

Committee Machine with Intelligent Systems (CMIS): The standard practices with artificial neural networks or any other intelligent systems to train many different candidate networks or systems, then the best one is selected according to the comparison criteria, such as statistical criteria (for example, the selection is based on the performance of an independent validation test) while discarding the rest. There are two disadvantages in this method:

- All the effort involved in training of the discarded networks is wasted
- Randomness of the noise in the data, which have the best validation set performance, will not necessarily have the best test set performance.

These disadvantages can be overcome by combining the intelligent systems together to form a Committee Machine with Intelligent Systems (CMIS); the importance of this procedure is that lead to significant improvements in the performance on new data, with little extra computational effort. The committee often does better than the best single constituent network in isolation (Bagheripour and Asoodeh, 2014).

The idea behind the committee machine is to fuse the knowledge acquired by experts in order to arrive at an overall decision that is superior to that of any of the individual experts acting alone (Haykin, 1991). The assumption is that, there are *N* trained intelligent systems with output vector O_i, which are used to predict target vector *T*. The prediction error could be written asshown below (Bhatt and Helle, 2002; Lim, 2005; Chen and Lin, 2006; Kadkhodaie-Ilkhchi *et al*., 2009; Ghiasi-Freez *et al*., 2012):

$$e_i = O_i - T \ (i = 1, \dots, N) \tag{7}$$

The expectation of the squared error for the *i*th intelligent system (O_i) is:

$$E_i = \xi[(O_i - T)^2] = \xi[e_i^2] \tag{8}$$

where,
$\xi[.]$: The expectation

The average error for each intelligent systems is shown in Eq. (9):

$$E_{avg} = 1/N \sum_{i=1}^{N} E_i = (\tfrac{1}{N}) \sum_{i=1}^{N} \xi[e_i^2] \tag{9}$$

Applying the averaging method, output vector O_i of the CMIS is shown in Eq. (10):

$$O_{CMIS} = 1/N \sum_{i=1}^{N} O_i \tag{10}$$

The prediction-squared error of CMIS is shown in Eq. (11):

$$E_{CMIS} = \xi[(O_i - T)^2] = \xi[(1/N \sum_{i=1}^{N} O_i - T)^2] = \xi[(1/N \sum_{i=1}^{N} e_i)^2] \tag{11}$$

Considering Cauchy's inequality:

$$(a_1 b_1 + a_2 b_2 + \cdots + a_n b_n)^2 \le (a_1^2 + a_2^2 + \cdots + a_n^2) \times (b_1^2 + b_2^2 + \cdots + b_n^2) \tag{12}$$

Equations (9) and (11) can be extended as below:

$$E_{CMIS} = \xi\left[\left(\frac{1}{N \sum_{i=1}^{N} e_i}\right)^2\right] = \frac{\xi}{N^2}(e_1 \times 1 + e_2 \times 1 + \cdots + e_N \times 1)^2 \tag{13}$$

$$E_{avg} = \frac{1}{N} \sum_{i=1}^{N} \xi[e_i^2] = \frac{\xi}{N}(e_1^2 + e_2^2 + \cdots + e_N^2) \times (1 + 1 + \cdots + 1) = \frac{\xi}{N}(e_1^2 + e_2^2 + \cdots + e_N^2) \times (N) \tag{14}$$

Cauchy's inequality can be applied to Eq. (13) and (14):

$$\frac{\xi}{N^2}(e_1 \times 1 + e_2 \times 1 + \cdots + e_N \times 1)^2 \le \frac{\xi}{N}(e_1^2 + e_2^2 + \cdots + e_N^2) \times (N) \tag{15}$$

By a simple substitution of two sides of Eq. (14), it will be concluded that:

$$E_{CMIS} \le E_{avg} \tag{16}$$

The error of CMIS is less than or equal to the average of all the intelligent systems (Lim, 2005). A

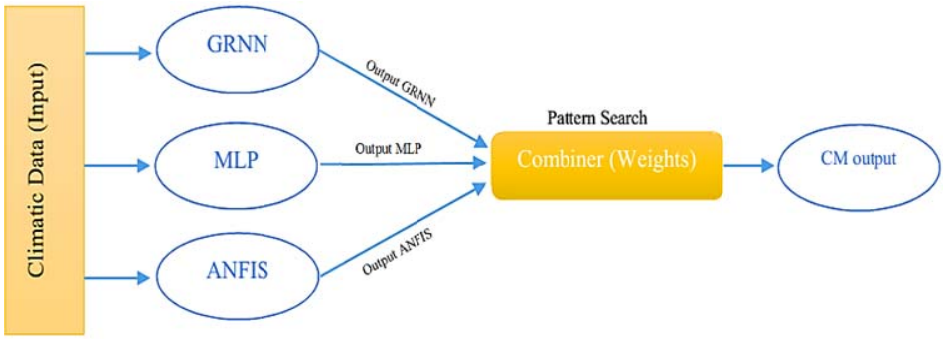

Fig. 4: Schematic diagram of CM model

schematic diagram of the developed CMIS is shown in (Fig. 4). There are different methods of combining the intelligent systems outputs in the combiner. The simple ensemble averaging method is the popular one (Chen and Lin, 2006). In the proposed CM in this research, each of the intelligent models has a weight factor defining its contribution in the overall estimation. A pattern search optimization technique is used in this study to derive the optimal combination of the weights. PS is a family of numerical optimization methods that does not require any information about the gradient of the objective function. Unlike more traditional optimization methods that use information about the gradient or higher derivatives to search for an optimal point, a PS algorithm searches a set of points around the current point, looking for one where the value of the objective function is lower than the value at the current point. It converges using the theory of positive bases (Yu, 1979).

STUDYAREAAND DATAPREPARATION

Basrah City is located on Shatt Al-Arab River in southern Iraq and has borders with Kuwait, Saudi

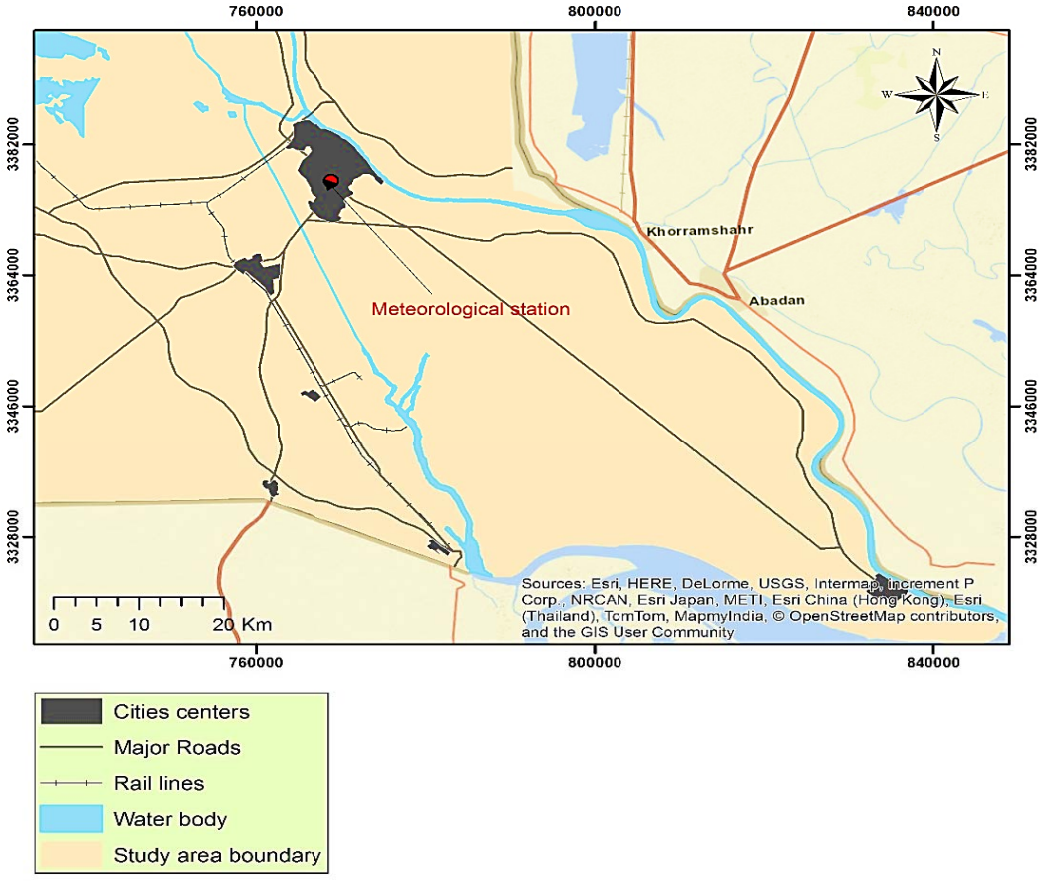

Fig. 5: Location of the study area

Table 1: The monthly statistical summary of data set used in this study

Variable	X_{mean}	X_{max}	X_{min}	SD	C_v	K	Correlation with ET_0
T (°C)	26.41	40.20	9.40	9.49	35.94	-0.14	0.95
RH (%)	41.29	80.00	17.00	17.38	42.08	0.43	-0.93
U_2	4.15	7.70	1.70	1.13	27.20	0.70	0.73
ET_0	8.02	17.78	1.80	4.47	55.56	0.29	1.00

Arabia and Islamic Republic of Iran. It is located between longitude line (47° 30'-48° 30') and latitude line (30°00'-30° 30') as shown in (Fig. 5). Basrah has a hot desert climate, like the rest of the surrounding region, though it receives slightly more precipitation than inland locations due to its location near the Arabian Gulf. During the summer months, from June to August, Basrah is consistently one of the hottest cities on the world, with temperatures regularly exceeding 50°C in July and August. In winter, Basrah experiences mild weather with average high temperatures around 20°C. At some winter nights, minimum temperatures may be reaching to 0°C. The City experience high humidity, sometimes exceeding 90%, due to its location close to the Arabian Gulf. Basrah is relatively an agricultural area where palm trees, fruit and vegetables are planted. Basrah is also known for planting tomatoes in Safwan-Al Zubair area (south west of center city) in winter season, which supplies the tomatoes demands of other Iraqi Provinces.

The climate information used in this research was obtained from the meteorological recording station in Hi Al-Hussain at the center of the Basrah City (Fig. 5).The samples data which consist of 22 years (1991-2012) monthly records of maximum monthly mean air temperature (T_{max}), minimum monthly mean air temperature (T_{min}), monthly mean Relative Humidity (RH) and monthly mean wind speed at 2 m above the ground surface (U_2). A statistical summary of these variables with obtained $ET_0$by using Penman-Monteith equation (FAO-56 PM) is presented in Table 1. The RH shows low variation if comparing with T and U_2. On the other hand, U_2 have the lowest correlation with ET_0 and have high skewed and distribution. All variables seem to be effective parameters on ET_0 with respect to correlation values. The inputs T, RH and U_2 and output ET_0 values were used for the constructing intelligent models. Three models for each intelligent system with different inputs combination are employed. The information and input variables for these models are shown in Table 2.

Two statistical errors namely, Root Mean Squared Error (RMSE) and Coefficient of Determination (R^2) are used to evaluate the performance of the developed models. The RMSE and $R^2$are computed as shown in Eq. (17 and 18):

$$RMSE = \left(\frac{\sum_{j=1}^{n}(Y_j - \hat{Y}_j)^2}{n}\right)^{1/2} \quad (17)$$

Table 2: Different inputs combination of models with their intelligent systems

Model no.	Input variables	Intelligent system
1	T_{min}, T_{max}	GRNN
2	T_{min}, T_{max}, RH	
3	T_{min}, T_{max}, RH, U_2	
4	T_{min}, T_{max}	MLP
5	T_{min}, T_{max}, RH	
6	T_{min}, T_{max}, RH, U_2	
7	T_{min}, T_{max}	ANFIS
8	T_{min}, T_{max}, RH	
9	T_{min}, T_{max}, RH, U_2	

Table 3: The GRNN spread values for different combination inputs

Model no.	Input variables	Spread value
1	T_{min}, T_{max}	0.02
2	T_{min}, T_{max}, RH	0.02
3	T_{min}, T_{max}, RH, U_2	0.04

$$R^2 = 1 - \frac{\sum_{j=1}^{n}(Y_j - \hat{Y}_j)^2}{\sum_{j=1}^{n}(Y_j - \bar{Y})^2} \quad (18)$$

where,

Y&$\hat{Y}$: The observed and estimated values respectively

n : The number of observations

$\bar{Y}$ and $\hat{Y}$: The mean of observed and estimated values

The RMSE shows the goodness of fit relevant to high values. The R^2 shows the degree to which two variables are linearly related (Karunanidhi *et al*.,1994). In case of GRNN, the output value is estimated using weighted average of the training dataset, where the weight is calculated using the Euclidean distance between the training and testing data. If the weight or distance is large, then the weight will be very less and if the distance is small, it will put more weight to the output. The decision that is required for each of the models inputs is the selection of the appropriate smoothing factors to be applied. For different input combinations, the optimum spread for the GRNN model was determined according to the MSE criterion. The determined spread values for different combination inputs are shown in Table 3.

A two-layer feed forward network with sigmoid hidden neurons and linear output neurons is used in the present research. The network is trained with Levenberg-Marquardt back propagation algorithm. Many researchers employed the Levenberg-Marquardt algorithm which is an approximation to Newton's method for adjusting the weights of the ANN model because it is more powerful than the conventional gradient descent techniques (Kişi, 2007). The optimal

number of neurons in the hidden layer is determined using trial and error method and found to be (20). The model is evaluated by the testing data set which is not used during the training phase. The total number of observations is 264 samples; these observations are divided into three parts. 60% (158 samples) for training, these are presented to the network during training and the network is adjusted according to its error. 20% (53 samples) is used for validating part; this set of data is used for measuring the generalization of network and to halt training when generalization stops improving. Also, the testing part is taken as the same percentage of validation test (20%, 53 samples), these have no effect on training and so provide an independent measure of network performance during and after training.

In this research, a subtractive clustering method is used for extraction of clusters and fuzzy if-then rules for ANFIS model. The subtractive clustering algorithm is an attractive approach to the synthesis of ANFIS networks, which estimates the cluster number and its cluster location automatically. By using this method, each sample point is seen as a potential cluster center. Computation time in this method becomes linearly proportional to data size, but independent of the dimension problem under consideration. The effective and important parameter in subtractive clustering which controls number of clusters and fuzzy if-then rules is clustering radius. This parameter is ranged of (0, 1). The training error can be controlled by adjusting clustering radius. Specifying a smaller cluster radius usually yields smaller clusters and more rules, a large cluster radius when approaching to one yields few large cluster in the data and few rules. Optimum clustering radius is determined by performing subtractive clustering network for several times, with changing radius value between (0, 1), leads to different number of if-then rules that could be established. According to the RMSE, the best fuzzy model is selected. The observations are divided into two statistically parts. 80% (211 samples) for training, these are presented to the network during training and the network is adjusted according to its error. 20% (53 samples) is used for checking part. The checking data is used for both checking and testing the fuzzy inference system parameters. Here, chkRMSE is the root mean square error of the system generated by the checking data. Table 4 show that the best value of clustering radius which equal to (0.3) is associated with lowest value of chkRMSE which equal to (1.1228) for model No. (7). By the same way, Table 5 and 6 show the optimum value of clustering radius with lowest value chkRMSE for model No. (8) and model No. (9), respectively.

A Gaussian membership function (mf) is selected to the extracted input clusters. The normal distribution of input data is carried out by using Gaussian function *f(x)* as shown in Eq. (19):

Table 4: Clustering radius with root mean square error generated by the checking data of model No. (7)

Clustering radius	No. of rules	chkRMSE
1	2	1.1899
0.9	2	1.1837
0.8	2	1.1847
0.7	2	1.1740
0.6	3	1.1908
0.5	3	1.1734
0.4	4	1.1582
0.3	5	1.1228
0.2	10	1.2279
0.1	26	1.3915

Table 5: Clustering radius with root mean square error generated by the checking data of model no. (8)

Clustering radius	No. of rules	chkRMSE
1	2	1.1567
0.9	2	1.1594
0.8	2	1.1594
0.7	3	1.0951
0.6	3	1.1247
0.5	3	1.1318
0.4	4	1.1459
0.3	5	1.0826
0.2	10	1.2402
0.1	42	1.4865

Table 6: Clustering radius with root mean square error generated by the checking data of model no. (9)

Clustering radius	No. of rules	chkRMSE
1	2	0.2497
0.9	2	0.2786
0.8	3	0.2141
0.7	3	0.1737
0.6	4	0.1689
0.5	5	0.1816
0.4	6	0.2149
0.3	12	0.2074
0.2	21	0.3503
0.1	116	0.9572

$$f(x) = \frac{e^{-(x-\mu)^2/\sigma^2}}{\sigma\sqrt{2\pi}} \quad (19)$$

where,

$\mu\&\sigma$: The parameter of normal distribution showing the mean and standard deviation of data, respectively.

The mean represents the cluster center, while, the standard deviation is calculated by the following function:

$$\sigma = (radii \times (\max(data) - \min(data)))/sqrt(8.0) \quad (20)$$

The Gaussian membership function parameters for the models of ANFIS are shown in Table 7.

Output of each Membership Function (mf) is linear equation consist of equation parameters multiply by input variable. For example, output mf1 in model No. (7), which is the consequent of rule no. 1 is constructed from two climatic inputs T_{min} and T_{max} as shown in Eq. (21):

Table 7: The gaussian membership function parameters for the models of intelligent system (ANFIS)

	Model No. (7)							
	Min. Temperature		Max. Temperature					
Inputs parameters	σ	μ	σ	μ				
Input mf No.								
mf1	2.896	9.3	3.986	20.2				
mf2	2.896	14.5	3.986	26.3				
mf3	2.896	20.4	3.986	35.3				
mf4	2.896	25.8	3.986	42.1				
mf5	2.896	28	3.986	45.3				
	Model No. (8)							
	Min. Temperature		Max. Temperature		Relative humidity			
Inputs parameters	σ	μ	σ	μ	σ	μ		
Input mf No.								
mf1	2.896	9.3	3.638	20.2	6.682	22		
mf2	2.896	13.3	3.638	26.5	6.682	27		
mf3	2.896	20	3.638	34.1	6.682	41		
mf4	2.896	25.1	3.638	42.7	6.682	53		
mf5	2.896	29	3.638	45.4	6.682	67		
	Model No. (9)							
	Min. Temperature		Max. Temperature		Relative humidity		Wind speed	
Inputs parameters	σ	μ	σ	μ	σ	μ	σ	μ
Input mf No.								
mf1	5.791	9.628	7.276	21.7	13.36	24	1.273	3.6
mf2	5.791	20	7.276	34.1	13.36	25	1.273	3.8
mf3	5.791	25.9	7.276	43.3	13.36	41	1.273	4.7
mf4	5.791	29.6	7.276	46.1	13.36	60	1.273	6.8

Table 8: Membership functions parameters

	Model No. (7)				
Output mf No.	C1	C2	C3	C4	C5
mf1	-0.090	0.204	-0.458		
mf2	-1.744	-1.093	122.894		
mf3	-0.060	0.251	-1.1		
mf4	-0.156	-0.204	17.137		
mf5	-1.121	-1.42	89.065		
Model No. (8)					
mf1	-0.960	-0.241	0.009	54.861	
mf2	0.006	0.101	-0.061	4.740	
mf3	-0.664	-0.098	0.0	29.183	
mf4	0.292	-0.178	-0.104	10.528	
mf5	0.002	-0.144	-0.011	12.390	
Model No. (9)					
mf1	0.122	0.1223	-0.08297	1.421	-0.9129
mf2	0.06974	0.06057	-0.06034	0.3155	3.791
mf3	0.105	0.1037	-0.05619	1.06	-0.2413
mf4	-0.05656	0.4549	-0.1661	1.343	-8.69

$$outputmf1 = c1 \times Tmin + c2 \times Tmax + c3 \quad (21)$$

where,

$c1$ &$c2$: The coefficients corresponding to T_{min} and T_{max} inputs, respectively

$c3$: The constant

Parameters in the above Eq. (21) are determined by linear least squares estimation. The other parameters of ANFIS models corresponding with different inputs combination are shown in Table 8.

Pattern search method was used to determine optimal combination of the weights for construction CMIS. The fitness function for PS can be expressed as follows:

$$MSE_{CMIS} = \sum_{i=1}^{n} 1/n(w_1 o_{1i} + w_2 o_{2i} + w_{31} o_{3i} - T_i)^2 \quad (22)$$

where,

MSE_{CMIS} : The mean square error of CMIS

$w_1, w_2 \text{and} w_3$: The weight factors corresponding to GRNN (o_{1i}), MLP (o_{2i}) and ANFIS (o_{3i}) predictions, respectively.

T_i : The ET_o value (as given by the FAO-56 PM equation)

n : The number of training data (158 samples).

RESULTSAND DISCUSSION

Table 9 show the results of applying three intelligent models used in this study with prediction accuracy. Results illustrated that the MLP models are better than ANFIS models and slightly better than GRNN models. Add RH as inputs in addition to the T_{min} and T_{max} increases the model's performances by reducing RMSE 35.07%, 36.84% and 3.58% and increasing R^2 by 2.96%, 1.33% and 0.62% for models No. (2, 5 and 8), respectively. Adding U_2 to the inputs combinations (T_{min}, T_{max}, RH) increases the model's performance by reducing RMSE 61.54%, 72.74% and11.58% and increasing R^2 by 0.76%, 1.76% and 1.91% for models No. (3, 6 and 9) respectively.

Output of the best performance models from the previous steps, were then used an input for constructing CM model for the overall estimation of ET_0. Each intelligent model has a weigh factor representing its contribution in overall estimation. At the first part, CMIS is constructed to obtain optimum combination of the weights with models that have only two inputs climatic data T_{min} and T_{max}. Then, the weights obtained from pattern search method as shown below:

$$ET_{oCMIS} = 0.401 \times ET_{oModel\ No.1} + 0.445 \times ET_{oModel\ No.4} + 0.154 \times ET_{oModel\ No.7} \quad (23)$$

At the second part, CMIS is constructed to obtain optimum combination of the weights with models that have only three inputs climatic data (Tmin, Tmax, RH).Then, the weights obtained from pattern search method as shown below:

$$ET_{oCMIS} = 0.398 \times ET_{oModel\ No.2} + 0.451 \times ET_{oModel\ No.5} + 0.151 \times ET_{oModel\ No.8} \quad (24)$$

Table 9: R^2 and RMSE of intelligent system models in testing period

Model No.	Intelligent system	R^2	RMSE
1	GRNN	0.9513	1.0075
2		0.9795	0.6542
3		0.9870	0.2516
4	MLP	0.9589	0.9481
5		0.9717	0.5988
6		0.9888	0.1632
7	ANFIS	0.9396	1.1228
8		0.9454	1.0826
9		0.9635	0.9572

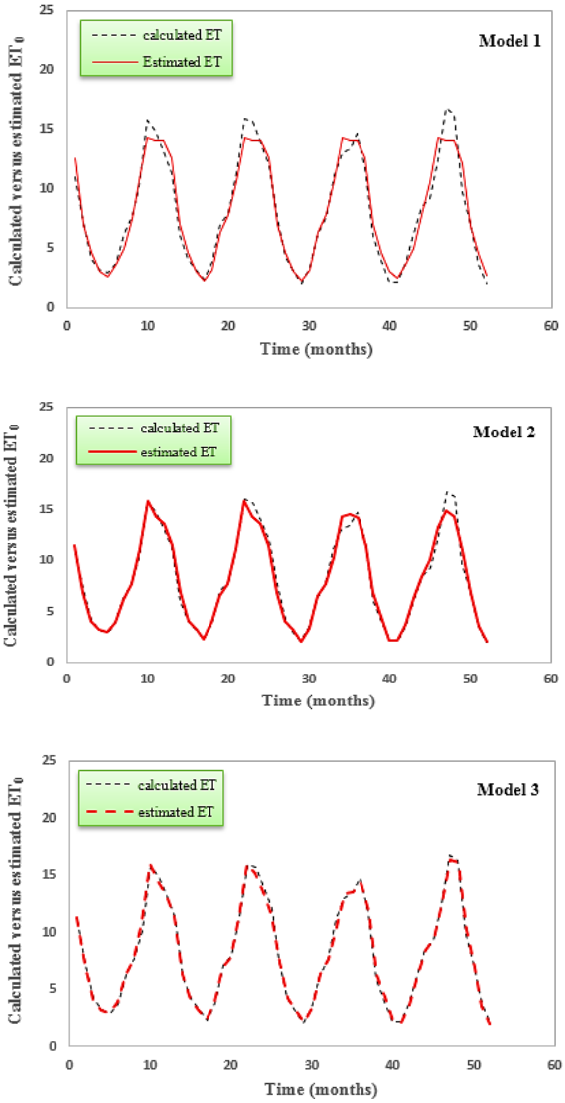

Fig. 6: Comparative plot in the testing period for ET_o given by the FAO-56 PM equation and ET_o estimated by artificial intelligent models for models 1, 2 and 3

RMSE of predicted ET_o from CMIS with models that have only two inputs climatic data is 0.8321 and R^2 is 0.9623. There is a significant improvement in the models' performance when comparison with best model of intelligent system (MLP), by reducing RMSEby12.23% and increasing R^2 by 0.35%. Also, when applying CMIS with models that have three inputs climatic data, the values of RMSE and R^2 are 0.4251 and 0.9823 respectively. Reducing RMSE by 29.01% and increasing R^2 by 1.09% comparison with best model of intelligent system (MLP). Figure 6 to 8 presents the details of the target and estimated monthly reference evapotranspiration by artificial intelligent models, while Fig. 9 presents the details of the target and estimated ET_o by CMIS with models have two inputs climatic data (a) and models have three inputs climatic data (b).

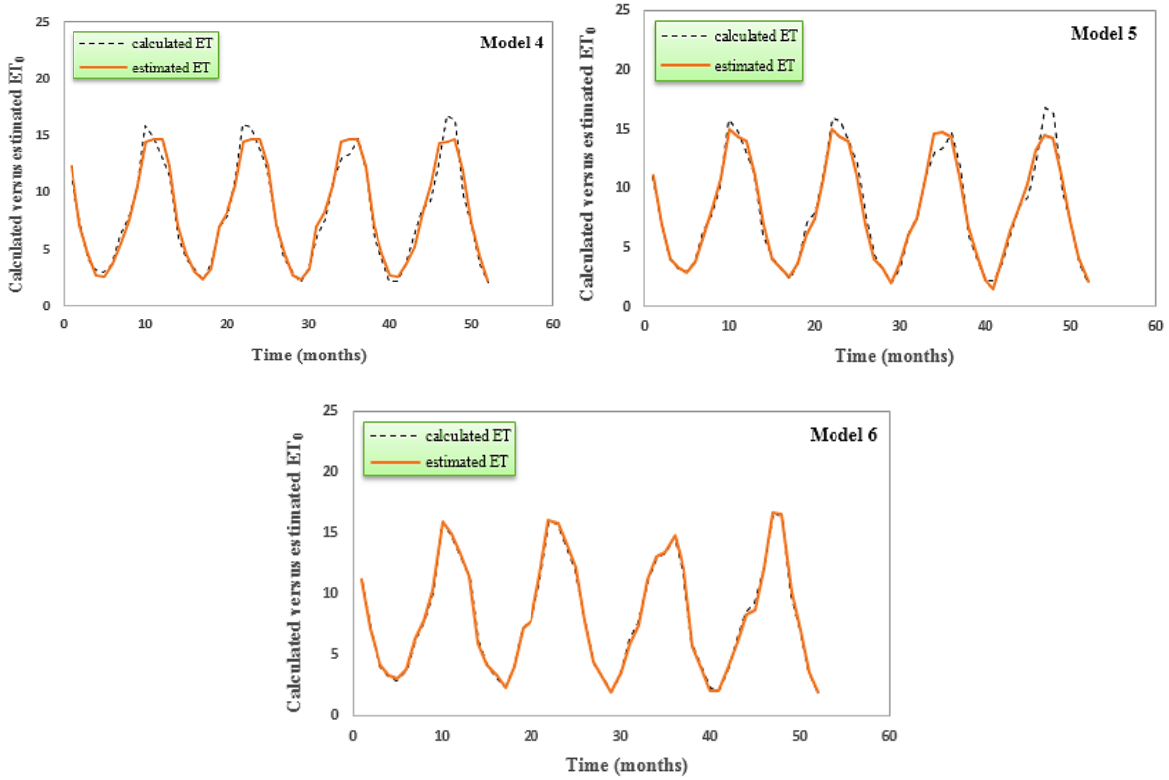

Fig. 7: Comparative plot in the testing period for ET_o given by the FAO-56 PM equation and ET_o estimated by artificial intelligent models for models 4, 5 and 6

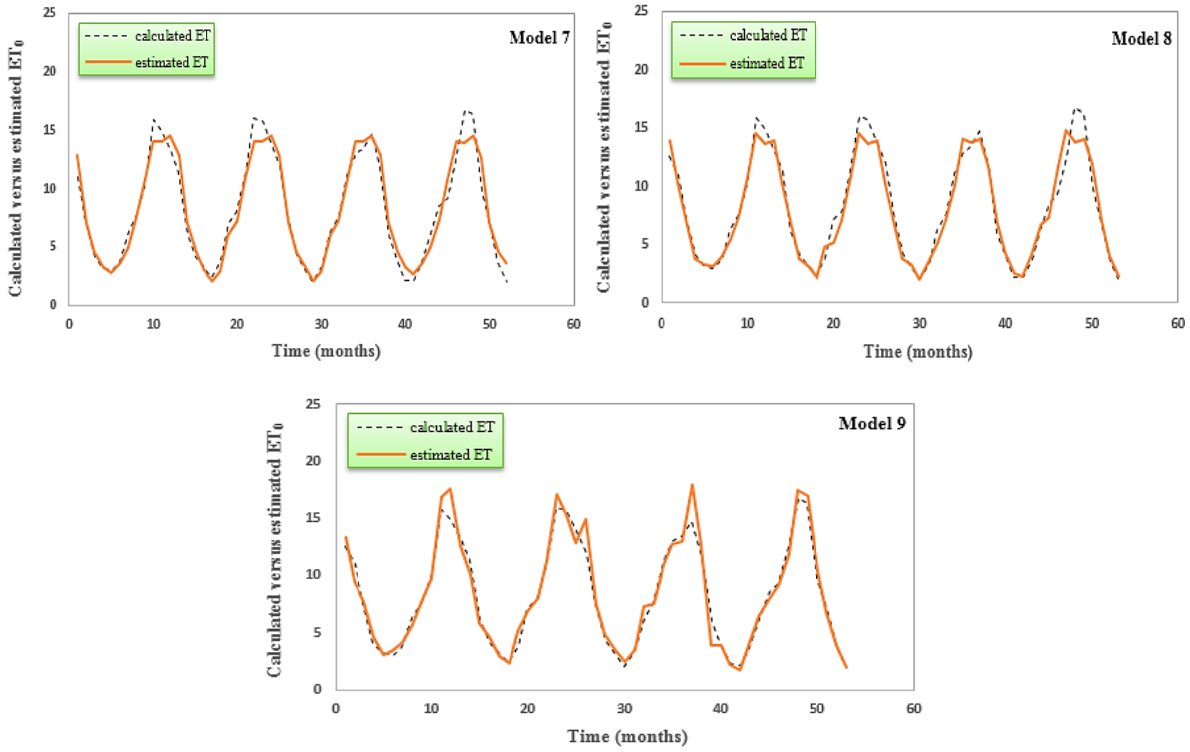

Fig. 8: Comparative plot in the testing period for ET_o given by the FAO-56 PM equation and ET_o estimated by artificial intelligent models for models 7, 8 and 9

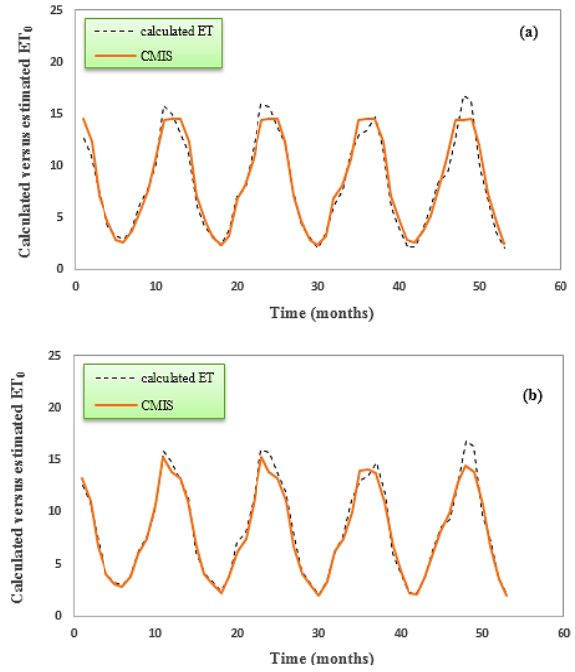

Fig. 9: Comparative plot in the testing period for ET_o given by the FAO-56 PM equation and ET_o estimated by A Committee Machine with Intelligent Systems (CMIS), (a): Models have two inputs climatic data; (b): Models have three inputs climatic data

CONCLUSION

In this study, three artificial intelligent systems (GRNN, MLP and ANFIS) were used for predicting ET_o in Basrah city, southern Iraq. Root Mean Squared Error (RMSE) and coefficient of determination (R^2) were used as evaluation criteria. At general, the models' performances of MLP models were better than ANFIS models and slightly better than GRNN models with different inputs com2bination. The Relative humidity is added to the models with two inputs combination (T_{min}, T_{max}), the models' performances is increased by reducing RMSE by 35.07%, 36.84% and 3.58% andincreasing R^2 by 2.96%, 1.33% and 0.62% for models No. (2, 5 and 8) respectively. In similar way, wind speed (U_2) is added also to other inputs combination to study the effect of each parameter on ET_o. The models' performances with this case is increased by reducing RMSE by 61.54%, 72.74% and 11.58% and increasing R^2 by 0.76%, 1.76% and 1.91% for models No. (3, 6 and 9) respectively. Wind speed has more effective than relative humidity. In other part of study, a CMIS was constructed for the overall estimation of ET_o by integrating the results of predicting ET_o from GRNN, MLP and ANFIS, each of them has a weight factor representing its contribution in overall estimation. The values of RMSE and R^2 when using the concept of CMIS for predicting ET_o with models that have only two inputs climatic data (T_{min}, T_{max}) are 0.8321 and 0.9623 respectively. There is a significant improvement in the models' performance when comparison with best model of artificial intelligent system (MLP) by reducing RMSE by 12.23% and increasing R^2 by 0.35%. When applying CMIS with models have three inputs climatic data (T_{max}, T_{min}, RH), the values of RMSE and R^2 are 0.4251 and 0.9823 respectively. As can be seen, that RMSE is reduced by 29.01% and R^2 is increased by 1.09% comparison with best model of intelligent system (MLP). This indicates that CMIS has a significant improvement for estimation of ET_o.

REFERENCES

Alilou, V.K. and F. Yaghmaee, 2015. Application of GRNN neural network in non-texture image inpainting and restoration. Pattern Recogn. Lett., 62: 24-31.

Allen, R.G., L.S. Pereira, D. Raes and M. Smith, 1998. Crop evapotranspiration: Guidelines for computing crop water requirements. Irrigation and Drainage Paper No. 56, FAO, Rome, Italy.

Bagheripour, P. and M. Asoodeh, 2014. Committee machine reaping of three well-known models: Established between saturation pressure and gas chromatography data. Appl. Petrochem. Res., 4(3): 313-316.

Barzegar, R., M. Sattarpour, M.R. Nikudel and A.A. Moghaddam, 2016. Comparative evaluation of artificial intelligence models for prediction of uniaxial compressive strength of travertine rocks, case study: Asarshahr area, NW Iran. Model Earth Syst. Environ., 2: 76.

Bhatt, A. and H.B. Helle, 2002. Committee neural networks for porosity and permeability prediction from well logs. Geophys. Prospect., 50(6): 645-660.

Chen, C.H. and Z.S. Lin, 2006. A committee machine with empirical formulas for permeability prediction. Comput. Geosci., 32(4): 485-496.

Disorntetiwat, P., 2001. Global stock index forecasting using multiple generalized regression neural networks with a gating network. Ph.D. Thesis, University of Missouri Rolla, Missouri, United States.

Elleithy, K., 2010. Advanced Techniques in Computing Sciences and Software Engineering. Springer, Netherlands, pp: 599.

Gavilān, P., I.J. Lorite, S. Tornero and J. Berengena, 2006. Regional calibration of Hargreaves equation for estimating reference ET in a semiarid environment. Agr. Water Manage., 81(3): 257-281.

Ghiasi-Freez, J., A. Kadkhodaie-Ilkhchi and M. Ziaii, 2012. Improving the accuracy of flow units prediction through two committee machine models: An example from the South Pars Gas Field, Persian Gulf Basin, Iran. Comput. Geosci., 46: 10-23.

Hannan, S.A., R.R. Manza and R.J. Ramteke, 2010. Generalized regression neural network and radial basis function for heart disease diagnosis. Int. J. Comput. Appl., 7(13): 7-13.

Haykin, S., 1991. Neural Networks: A Comprehensive Foundation. Prentice-Hall, Englewood Cliffs, pp: 842.

Hossen, J., S. Sayeed, I. Yusof and S.M.A. Kalaiarasi, 2013. A framework of modified adaptive fuzzy inference engine (MAFIE) and its application. Int. J. Comput. Inform. Syst. Ind. Manage. Appl., 5: 662-670.

Kadkhodaie-Ilkhchi, A., M.R. Rezaee and H. Rahimpour-Bonab, 2009. A committee neural network for prediction of normalized oil content from well log data: An example from South Pars Gas Field, Persian Gulf. J. Petrol. Sci. Eng., 65(1-2): 23-32.

Karunanidhi, N., W.J. Grenney, D. Whitley and K. Bovee, 1994. Neural networks for river flow prediction. J. Comput. Civil Eng., 8(2): 201-219.

Kişi, O., 2007. Streamflow forecasting using different artificial neural network algorithms. J. Hydrol. Eng., 12(5): 532-539.

Kisi, O., M. Tombul and M.Z. Kermani, 2015. Modeling soil temperatures at different depths by using three different neural computing techniques. Theor. Appl. Climatol., 121(1-2): 377-387.

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

Kumar, R., V. Shankar and M. Kumar, 2011. Modeling of crop reference evapotranspiration: A review. Univer. J. Environ. Res. Technol., 1(3): 239-246.

Landeras, G., A. Ortiz-Barredo and J.J. López, 2008. Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). Agr. Water Manage., 95(5): 553-565.

Lim, J.S., 2005. Reservoir properties determination using fuzzy logic and neural networks from well data in offshore Korea. J. Petrol. Sci. Eng., 49(3-4): 182-192.

López-Urrea, R., F. Martin de Santa Olalla, C. Fabeiro and A. Moratalla, 2006. Testing evapotranspiration equations using lysimeter observations in a semiarid climate. Agr. Water Manage., 85(1-2): 15-26.

Loukas, Y.L., 2001. Adaptive neuro-fuzzy inference system: An instant and architecture-free predictor for improved QSAR studies. J. Med. Chem., 44(17): 2772-2783.

Nourani, V. and A. Babakhani, 2013. Integration of artificial neural networks with radial basis function interpolation in earthfill dam seepage modeling. J. Comput. Civil Eng., 27(2): 183-195.

Pereira, A.R. and W.O. Pruitt, 2004. Adaptation of the Thornthwaite scheme for estimating daily reference evapotranspiration. Agr. Water Manage., 66(3): 251-257.

Rumelhart, D.E. and J.L. McClelland, 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition. MIT Press, Cambridge, Mass.

Specht, D.F., 1991. A general regression neural network. IEEE T. Neural Networ., 2(6): 568-576.

Stoffel, K., P. Cotofrei and D. Han, 2012. Fuzzy clustering based methodology for multidimensional data analysis in computational forensic domain. Int. J. Comput. Inform. Syst. Ind. Manage. Appl., 4: 400-410.

Tang, X.G., H.P. Li, A.R. Desai, Z. Nagy, J.H. Luo, T.E. Kolb, A. Olioso, X.B. Xu, L. Yao, W. Kutsch, K. Pilegaard, B. Köstner and C. Ammann, 2014. How is water-use efficiency of terrestrial ecosystems distributed and changing on Earth? Sci. Rep-UK, Vol. 4, Article No. 7483.

Trajkovic, S. and S. Kolakovic, 2009. Evaluation of reference evapotranspiration equations under humid conditions. Water Resour. Manag., 23: 3057-3067.

Yu, W.C., 1979. Positive basis and a class of direct search techniques. Sci. Sin. Math., 9(1): 53-67.