

Optimized Neural Network Prediction Model for Potential Evapotranspiration Utilizing Ensemble Procedure

Ahmed El-Shafie · Ali Najah · Humod Mosad Alsulami · Heerbod Jahanbani

Received: 2 September 2012 / Accepted: 14 January 2014 /
Published online: 30 January 2014
© Springer Science+Business Media Dordrecht 2014

Abstract Potential evapotranspiration (ET_o) is an essential hydrologic parameter for having better understanding for hydrologic cycle in certain catchment area. In addition, ET_o is vital for calculating the agricultural demand. In fact, Penman-Monteith (PM) method is considered as reference method for estimating (ET_o), however, this method required a lot of data to be used which is not usually available in many catchment areas. Furthermore, there are several efforts that have been performed as competitor to reach accurate estimation of (ET_o) when there is lack of data to utilize (PM) method, but still required numerous research. Recently, methods based on Artificial Intelligence (AI) have been suggested to provide reliable prediction model for several application in engineering and especially for hydrological process. However, time series prediction based on Artificial Neural Network (ANN) learning algorithms is fundamentally difficult and faces problem. One of the major shortcomings is that the ANN model experiences over-fitting problem during training session and also occurs when a neural network loses its generalization. In this research a modification for the classical Multi Layer Preceptron- Artificial Neural Network (MLP-ANN) modeling namely; Ensemble Neural Network (ENN) is proposed and applied for predicting daily ET_o . The proposed model applied at two different region with two different climatic conditions, Rasht city located north part of Iran and Johor Bahru City, Johor, Malaysia using maximum and minimum daily temperature collected from 1975 to 2005. The result showed that the ENN outperformed the classical MLP-ANN method and successfully predict daily ET_o utilizing maximum and minimum temperature only with satisfactory level of accuracy. In addition, the proposed model could achieve accuracy level better than the traditional competitor methods for ET_o .

A. El-Shafie (✉)
Department of Civil and Structural Engineering, Faculty of Engineering, University Kebangsaan Malaysia,
Bangi, Malaysia
e-mail: elshafie@vlsi.eng.ukm.my

A. Najah
Department of Engineering Science, Universiti Malaysia Terengganu, Kuala Terengganu, Malaysia

H. M. Alsulami
Ministry of Higher Education, Kingdom of Saudi Arabia, Riyadh 11153, Saudi Arabia

H. Jahanbani
Civil Engineering Department, Faculty of Engineering, University Melbourne, Melbourne, Australia

Keywords Evapotranspiration · Neural network · Ensemble neural network · Over-fitting · Rasht City (Iran)

1 Introduction

1.1 Background

Evapotranspiration (ET_o), which is the combination of evaporation from soil and transpiration from plants, is a critical component of the energy budget and water cycle. The process is affected by many climatological parameters such as solar radiation, air temperature, air humidity and wind speed. Transpiration is vaporization of liquid water contented in plant tissue, which happens through plant stomata. Transpiration, like evaporation, depends on many factors: energy supply, vapor pressure gradient, wind, radiation, air temperature, air humidity, soil water content, the ability of the soil to conduct water to the roots, water logging, soil water salinity, crop characteristics, environmental aspects, and the cultivation practices. Evaporation and transpiration (ET) happen at the same time and there is no easy way to separate them. For example, when the crop is small the main process for losing water is evaporation, but once the crop completely developed and covered the soil, transpiration becomes the main process. Thus, an accurate estimation of ET_o will greatly improve the monitoring of changes in the water cycle, which is an important part of the on-going global climate change.

Numerous methods have been proposed for modeling evapotranspiration, as described by Brutsaert (1982). In the last decades, the neural networks approach has been applied to many branches of science and many studies have approved the successful application of ANN for estimating different phase of hydrological cycle such as rainfall runoff modeling (Hsu et al. 1995; Tokar and Johnson 1999); streamflow prediction (Imrie et al. 2000); reservoir inflow forecasting (Jain et al. 1999; El-Shafie et al. 2007, 2009; Sulaiman et al. 2011) and prediction of water quality parameters (Maier and Dandy 1999). Recently, artificial neural networks have been successfully used in modeling evapotranspiration. Trajkovic et al. (2003) applied a sequentially adaptive radial basis function network for the forecasting mean monthly ET_o . They concluded that ANN could predict monthly ET_o with high accuracy. Sudheer et al. (2003) checked the potential of ANN in estimating the monthly actual crop evapotranspiration (ET) from limited climatic data. Their research employed radial-basis function (RBF) type ANN and compared the result with measured lysimeter. Their study clearly showed that ANN could successfully estimate ET with limited climate data. Awchi (2008) estimated evapotranspiration by means of ANN and daily climatic data of temperature, relative humidity, sunshine hours, wind speed and rainfall, and compared them with Penman-Monteith (PM) results. He found that ANN could successfully be utilized for estimating ET_o . Zanetti et al. (2007) applied ANN for estimating daily ET_o as a function of maximum and minimum temperature, extraterrestrial radiation and the daylight. Wang et al. (2008) used ANN (FFBP) for estimation average decadal ET_o with maximum and minimum temperate in Burkina Faso. According to their results, output of ANN has a higher accuracy than Hargreaves (HGS) and Blaney-Criddle (BCR) (Blaney and Criddle 1962). Khoob (2008) compared method of converting pan evaporation (E_p) with ANN utilizing maximum and minimum temperature. He posited that ANN gives more accurate results than conversion method. Chauhan and Shrivastava (2008) evaluated the performance of ANN for estimating ET_o and showed that ANN models performs better than climatic models such as Blaney–Criddle, Radiation and Modified Penman. He reported that ANN model could provide estimating for ET_o relatively close to actual measures

from Pan evaporation. The result also suggested that ANN models could estimate ETo from maximum and minimum temperature.

One of the main issues in machine learning research is that of generalization. Generalization refers to the predictive ability of a base learner (or learning machine). The better a predictor performs on unseen data, the better it is said to possess the ability to generalize (Chandra and Yao 2006).

1.2 Problem Statement and Objective

Obviously, ANN provides a viable and effective approach for developing input–output prediction models in situations that do not require modeling of the whole and/or part of the internal parameter affecting ETo, although, those models have proved to be efficient, its convergence tends to be very slow, and yields sub-optimal solution. This may not be suitable for dynamic adaptive accurate forecasting purpose. In fact, the major objective of training an ANN for prediction is to generalize, i.e. to have the outputs of the network approximate target values given inputs that were not in the training set. One of the major shortcomings is that the ANN model experienced over-fitting problem during training session and occurs when a neural network loses its generalization. Based on the above assertion and according to Kumar et al. (2002), ETo depends on the interaction of many climatic parameters such as temperature air humidity, wind speed, solar radiation, and type and growth stage of the crop.

In our study the data are available in both the study areas and then the ETo could be calculated based PM method, which is considered as reference value of ETo. On the other hand, the ETo could be calculated utilizing the proposed method in this study with the limited data (maximum and minimum temperature) and also the other competitor method, and hence the comparison analysis could be carried out in order to evaluate and examine the performance of the proposed model in this study over the other competitor methods.

The main objective of this research is to introduce a modification for the classical neural network modeling, namely, Ensemble Neural Network (ENN) to overcome the over-fitting problem. In addition, this manuscript investigates the potential of utilizing the ENN model to estimate and predict the monthly time series of daily ETo at Rasht City, Iran and Johor City, Malaysia. The proposed ENN will use only the daily time series minimum and maximum temperature and solar radiation (T_{\min} , T_{\max} and R_s) as input pattern. Since there is no lysimeter installed in the case study area, PM method was accepted as reference ETo in order to evaluate the proposed model. The anticipated impact of this model is that it can predict the monthly time series of daily ETo similar to the level of accuracy of PM method without the need to explicitly consider the internal hydrologic or climatic parameters.

2 Material and Data Collection

In order to achieve the objective of this study, daily minimum and maximum temperature ($^{\circ}\text{C}$) have been collected. Two different data set that represent two different climate conditions humid and semi-humid in Johor, Malaysia and Rasht city, Iran. In fact, the idea behind utilizing such various data sets is to validate the applicability of the proposed model in predicting evapotranspiration.

There are several rivers in Johor State “located in south of Malaysia”, Johor River is considered as one of the main river not only in Johor state but also at Malaysian level. Johor River flows in a roughly north–south direction originating from Mount Gemuruh and then empties into the Strait of Johor. Johor River is the main water supply for several water

demands (domestic, industrial and agricultural) for Johor State and for some part in neighboring country Singapore. The major features of Johor River stream are 122.7 km long with a catchment of 2,636 km², on other words; it is almost 14 % of the Johor State of Peninsular Malaysia. There are several tributaries that connected with main stream of the river while the majors' one are Sayong, Linggui, Tiram and Lebam Rivers. Its banks are also known to be the location of past capitals of Johor. The Sungai Johor Bridge, officially opened in June 2011, is the first bridge to span the river and is currently the longest river bridge in Malaysia.

Syarikat Air Johor, SAJ (or Johor Water Company) and the Public Utilities Board of Singapore (PUB) each draws about 250,000 cubic metres/day of water from the Johor River near Kota Tinggi. Both water supply schemes have been operational since the mid-1960s. In addition, the Linggui Dam completed and impounded in 1993 also supplements the water supply to both Johor and Singapore.

For the Rasht station (latitude: 37 12 N, longitude: 49 39 E and elevation 36.7 m above sea level) were collected for the 30 years (01.01.1975 to 31.12.2005). North part of Iran, near Caspian Sea, has a special climate and plant cover, which is different other parts of Iran. It has cold and humid weather in winter and fall and hot and humid weather in summer. This region has the most precipitation, compared to other parts of Iran, and a maximum annual precipitation of 1,400 mm per year recorded in Anzali City in Gilan province. The principal agricultural product in that region is rice, which have a high crop water requirement. This region contains three provinces; Gilan, Mazandaran, and Golestan provinces. The case study is Rasht, the capital city of Gilan province, located west of the region. Geographical location of this province is from 36° 34' to 38° 27' north and from 48° 53' to 50° 34' south. Its area is 14,711 Km² with 2.4 million populations.

3 Methodology

3.1 Estimation of Reference Evapotranspiration

The Penman-Monteith (PM) equation for computation reference evapotranspiration proposed by Allen and Pruitt (1991) is as following:

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

where,

ET_o is reference evapotranspiration (mm day⁻¹); R_n is net radiation at the crop surface (MJ m⁻² day⁻¹); G is soil heat flux density (MJ m⁻² day⁻¹); T is mean daily air temperature at 2 m height (°C); u_2 is wind speed at 2 m height (m s⁻¹); e_s is saturation vapor pressure (kPa); e_a is actual vapor pressure (kPa); $e_s - e_a$ is saturation vapor pressure deficit (kPa); Δ is slope vapor pressure curve (kPa °C⁻¹); γ is psychrometric constant (kPa °C⁻¹).

In the case where only T_{\max} and T_{\min} are available, Hargreaves and Samani (1985) suggested that ET_o can be calculated as:

$$ET_o = C_o (T_{\max} - T_{\min})^{0.5} (T_{\text{mean}} + 17.8) R_a \quad (2)$$

Where, ET_o is reference evapotranspiration calculated by HGS method (mm/day); T_{\min} , T_{\max} and T_{mean} are minimum, maximum and mean temperature, respectively (°C); R_a is extraterrestrial radiation (mm/day).

The input data for training and simulation are T_{max} , T_{min} and R_s . In case of availability of T_{max} and T_{min} only, R_s can be calculated by Eq. (3):

$$R_s = K_{R_s} \sqrt{(T_{max} - T_{min})} R_a \tag{3}$$

where, R_s is solar radiation ($MJ\ m^{-2}\ d^{-1}$); K_{R_s} is adjustment coefficient (0.16–0.19); T_{max} and T_{min} are maximum and minimum air temperature, respectively ($^{\circ}C$); R_a extraterrestrial radiation ($(MJ\ m^{-2}\ d^{-1})$ Eq. 2).

3.2 Model Architecture

The actual data for T_{max} , T_{min} and R_s over the 30 years between 1975 and 2005 on daily basis at Rashat City, Iran is used in this study, while the ET_o was calculated based PM method. On the other hand, for similar period the same parameters were collected for Johor Bahru city, Malaysia, and these data were used to train, test and validate the proposed ANN model. In general, the model architecture is for predicting the average daily evapotranspiration. However, the nature of the data and its stochastic behavior is different from 1 month to another. In this context, the model is developed for average daily evapotranspiration but for each month.

In general, the ANN modeling method learnt from examples, it uses the preceding and the recent behavior of a system to predict its future changes with respect to the changes in the input parameter (Bishop 1996). In fact, the major advantage of ANN method is its capability to mimic the behavior of the feature of the input pattern and its mapping with the corresponding output. Furthermore, the ANN method has the potential to predict the behavior of systems without analytical prediction rules. One of the major shortcomings is the selection of ANN architecture in terms of input–output pattern that could provide the best result for desired output.

In this context, two different scenarios I and II for the model architecture are employed to predict the monthly ET_o at both case studies. Scenario I is organized such that the prediction of ET_o at particular month (t) and year (n) is based on the T_{max} , T_{min} and R_s , at month (t-1) year (n) as presented in the following equation

$$ET_o(t, n) = f\left(T_{(t-1, n)}^{max}, T_{(t-1, n)}^{min}, R_{(t-1, n)}^s\right) \tag{4}$$

On the other hand, Scenario II, is structured such that the prediction of ET_o at month (t) and year (n) is based on the T_{max} , T_{min} and R_s . of the same month for year (n-1)

$$ET_o(t, n) = f\left(T_{(t, n-1)}^{max}, T_{(t, n-1)}^{min}, R_{(t, n-1)}^s\right) \tag{5}$$

The architecture of the two scenarios applied in this research is a three-layer network including an input layer, an output layer and a hidden layer as shown in Fig. 1a, b.

The architecture of the network consists of an input layer of three neurons (corresponding to the monitored T_{max} , T_{min} and R_s), an output layer of one neuron (corresponding to the predicted ET_o) and a number of hidden layers of arbitrary number of neurons at each layer. In order to achieve the desirable forecasting accuracy, 12 ANN architectures were developed (one for each month). Daily T_{max} , T_{min} , R_s and ET_o for the period of 25 years, from 1975 to 2000, were utilized in order to train the 12 networks. The performance and the reliability of the ANN models were examined using the data monitored between 2000 and 2002. The capabilities of the developed ANN models were further verified for the period between years 2003 and 2005.

One classical problem of neural networks is called overfitting, which occurs especially with noisy data. It has been observed that excessive training results in decreased generalization.

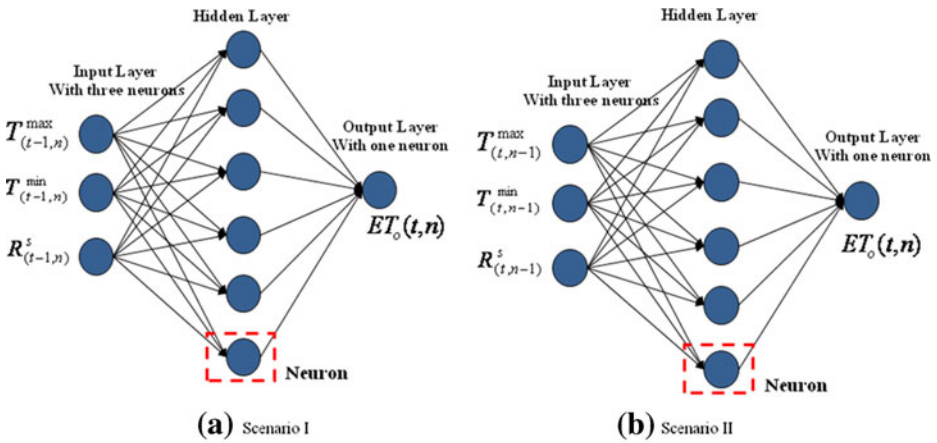


Fig. 1 Exact architecture of the ANN for the proposed two scenarios **a** scenario I **b** scenario II

Instead of finding general properties of the different input patterns that match a certain output, the training brings the network closer to each of the given input patterns. This results in less tolerance in dealing with new patterns. One Solution is using evaluation of the network performance as a different set of patterns than for the training. Hence, only networks that generate the ability to generalize are evaluated high.

3.3 Assembly Neural Network Procedure

In order to achieve this goal, we use a sequence of the previous behavior of the system as the training data then generate a sequence of inputs with proper length and their corresponding outputs for the first 90 % of 25 years (training data) and with respect to the size of the best period regarding the previous section. Subsequently, we construct a series of networks with an initial guess for the number of hidden layers neurons and initialize their parameters randomly. Finally for every network, the parameters vector will stop on a local minimum of its performance surface. Up to this point, all of the networks are over fitted on the training set. Afterward, a simulated annealing process is applied on each network. To do this, the model is modified to generate a set of vectors named the noise vectors. Length of each noise vector is equal to the length of each network parameters vector and its components are random numbers with uniform distribution between -0.05 and $+0.05$. By adding noise vectors to the network parameter vectors a new set of network parameters are obtained. This action makes for relatively minor changes in the location of each network in its state space.

Furthermore, the learning phase, a random vector of length N is generated where N is the length of the sequence of the first 25 years of time series values. This vector is called data noise vector and shown by following equation:

$$V_{dn} = M \times 10^{-2} \times z \times rand(1, N) \tag{6}$$

In this equation z is the number of networks added to the ensemble of neural networks before this step. $Rand(1, N)$ is a $1 \times N$ vector. The components of this vector are uniformly distributed random numbers between -0.05 and $+0.05$. Also, $M = \text{Max} - \text{Min}$ where Max and Min are the maximum and minimum values of time series of the system's behavior respectively. Once again, we select the network that has the best generalization on this new training

data sets. But this time the number of neurons in the hidden layers of networks is calculated using the following equations:

$$n'_1 = \left\lfloor x \times \frac{n_1}{n_2} \right\rfloor \tag{7}$$

$$n'_2 = x + n_2 \tag{8}$$

Where n_1 and n_2 are the initial number of neurons in the first and second hidden layers of the first set of networks and n'_1 and n'_2 are new values. The value of x is gained through the following equation:

$$x = \left\{ \left\lfloor \frac{|IN + 1|}{2} \right\rfloor \times \text{mod}((IN-1), 2) \right\} - \left\{ \left[\left\lfloor \frac{|IN + 1|}{2} \right\rfloor \times \left| \frac{1 + \text{sign}\left(n_2 - \left\lfloor \frac{|IN + 1|}{2} \right\rfloor\right)}{2} \right| + \left((2 \times n_2 - \left\lfloor \frac{|IN + 1|}{2} \right\rfloor) - 1 \right) \times \left| \frac{1 - \text{sign}\left(n_2 - \left\lfloor \frac{|IN + 1|}{2} \right\rfloor\right)}{2} \right| + 0.5 \right] \times \text{mod}(IN, 2) \right\} \tag{9}$$

In this equation, IN is the iteration number. Following the initial step, if IN is even the networks will be constructed with the previous structure but with more neurons. However if IN is odd then the number of neurons will be decreased until a zero limit is met. At this step we will continue the process by increasing the number of neurons in the hidden layer (n_1). This enables us to find more suitable number of neurons in the completion process of the ensemble if the initial guess was not accurate and networks need more (or less) number of neurons to achieve a good generalization.

Consider a single NN that has been trained on a given data set. Let x denote an input vector not seen before and let d denote the corresponding desired response; x and d represent realisation of the random vector X and random variable D , respectively. Let $F(x)$ denote the input–output function realised by the network. Networks are trained with these noisy parameters until another local minimum is achieved. Making noise vectors and training are repeated for a number of times and the outputs of these networks are compared on the followed 10 % of the 25 years which are not used during training steps. The winner has the best generalization amongst all and is selected as the first member of an ensemble of neural networks. Then, in light of the material on the bias/variance dilemma, we decompose the MSE between $F(x)$ and the conditional expectation,

The expectation, ED, is taken over the space D , defined as the space encompassing the distribution of all training sets (for example, inputs and target outputs) and the distribution of all initial conditions. There are different ways of individually training the networks and also different ways of combining their outputs. Here, we consider the situation where the networks have an identical configuration, but they are trained starting from different initial conditions by analogy with Eq. 9,

After finding the best network in each set, we compute the sum of absolute errors in the prediction of the last 10 % of data using the following equations:

$$e_{1=} = \frac{\sum_{i=1}^z \sum_{j=1}^k |T_s(j) - \text{Pr}(i, j)|}{z} \tag{10}$$

$$e_{2=} = \frac{\sum_{i=1}^{z+1} \sum_{j=1}^k |T_s(j) - \text{Pr}(i, j)|}{z + 1} \tag{11}$$

In the above equations, z is the number of networks that was added to the ensemble before this step. T_s is the sequence of the last 10 % events of time series in training data set and k is the size of T_s . $Pr(i,j)$ is the value that i^{th} member of ensemble predicts for the j^{th} event in the last 10 % of time series events. If $e_1 > e_2$, adding the best network of this step to the ensemble has caused improvements in the generalization of the ensemble totally. Otherwise, we don't add the selected network to the ensemble and repeat this step again with new noisy data sets and a new set of networks with different number of neurons in their hidden layers. The terminating condition is as follows: a predefined number of iterations (namely i_{tr}) are considered and at the end of these iterations the improvement of ensemble predictions (on the last 10 %) are measured. If this value is smaller than a predefined factor the termination condition is met. Otherwise this process will be repeated. Figure 2 illustrates the learning phase for the proposed ensemble ANN model.

3.4 Model Performance

In order to evaluate the proposed ANN model with ensemble procedure, various statistical indices that could be used. In this study, different methods for investigating model performance have been introduced. These were mean absolute relative error

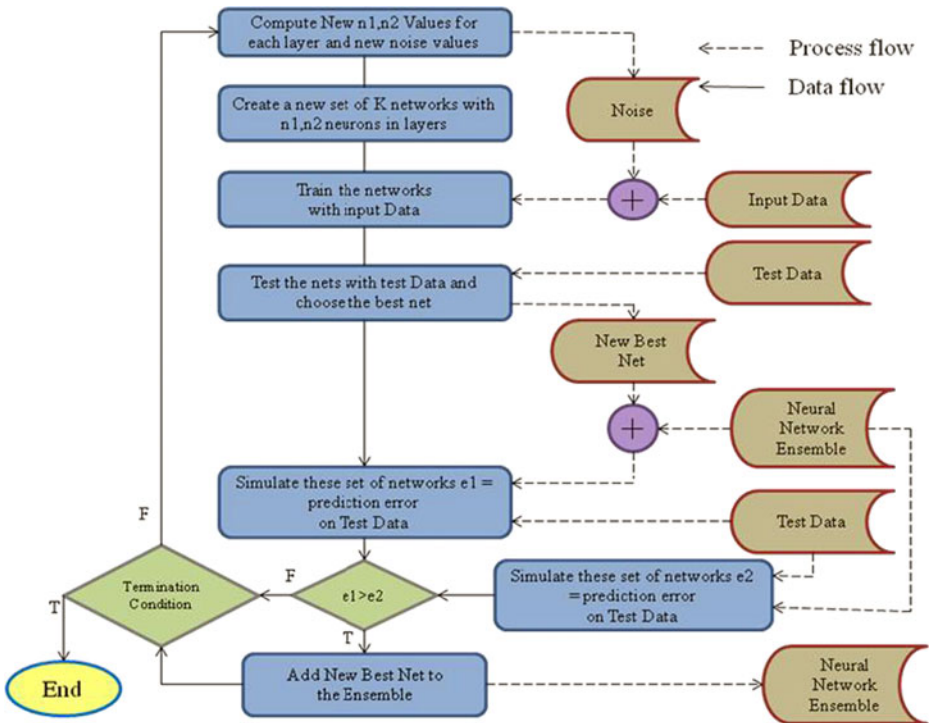


Fig. 2 Learning phase process for ensemble neural network

(MARE), mean absolute error (MAE), mean square error (MSE) and correlation coefficient (R^2). They evaluated models by the following equations:

$$MARE = \frac{1}{N} \sum \frac{|y - y'|}{y} * 100 \tag{12}$$

$$MAE = \frac{\sum |y - y'|}{N} * 100 \tag{13}$$

$$MSE = \frac{\sum (y - y')^2}{N} \tag{14}$$

Where, y is the ET_o calculated by PM, y' predicated ET_o by model and N is the total number of data. R^2 was resulted by linear correlation between actual and simulated data.

As the prediction accuracy of the peak and low ET_o events is of particular interest of the water resources management, especially in the field of irrigation, it is important to evaluate the model performance considering these extreme events. Therefore, the model performance during these events, another two statistical indices were recommended namely; Peak Value Criteria (PVC) and Low Value Criteria (LVC), were recommended to evaluate the model performance at the extreme values, which can be computed by Eqs. 15 and 16

$$PVC = \frac{\left(\sum_{i=1}^{T_p} (ET_{om} - ET_{op})^2 * (ET_{om})^2 \right)^{0.25}}{\left(\sum_{i=1}^{T_p} (ET_{om})^2 \right)^{0.5}} \tag{15}$$

$$LVC = \frac{\left(\sum_{i=1}^{T_l} (ET_{om} - ET_{ol})^2 * (ET_{om})^2 \right)^{0.25}}{\left(\sum_{i=1}^{T_l} (ET_{om})^2 \right)^{0.5}} \tag{16}$$

Where T_p =number of peak evapotranspiration greater than one-third of the mean peak ET_o observed; T_l =number of low ET_o lower than one-third of the mean low evapotranspiration observed. El-Shafie et al. (2009) reported that both PVC and LVC provide better performance indicators for assessment of the prediction model performance for the extreme inflow events. As the model can provide low PVC or LVC as the model represents better fit.

4 Results and Discussion

The daily PM was calculated by FAO ET_o calculator (version 3.1 January 2009). The time series input data of PM employed in the software and 30 years daily ET_o were generated and then sorted in to monthly time series. The 12 networks “one for each month” have successfully trained for the data between the period 1975 and 2000. Different Multi Layer Perceptron-

Artificial Neural Network (MLP-ANN) architectures (while keeping three neurons in the input layer and only one neuron in the output layer) were used to examine the best performance. In fact, there is no formal and/or mathematical method for determining the appropriate “optimal set” of the key parameters of Neural Network (number of hidden layers, number of neurons with each hidden layer and the type of transfer function between two consequence layers). Therefore, it was decided to perform this task utilizing trial and error method. Several sets were examined with maximum three hidden layers and maximum ten neurons within each layer. Table 1 shows the optimal number of neurons in the hidden layers for each network. It was decided to use the same number of neurons while implementing the model utilizing scenario II. Furthermore, the test session has been carried out for both data for Rasht and Johor cities utilizing the data for the years 2001 and 2005. Since the ET_o were accurately calculated using PM method over the years 2001 and 2005 period, the performance of the proposed ANN-based architecture can be examined and evaluated.

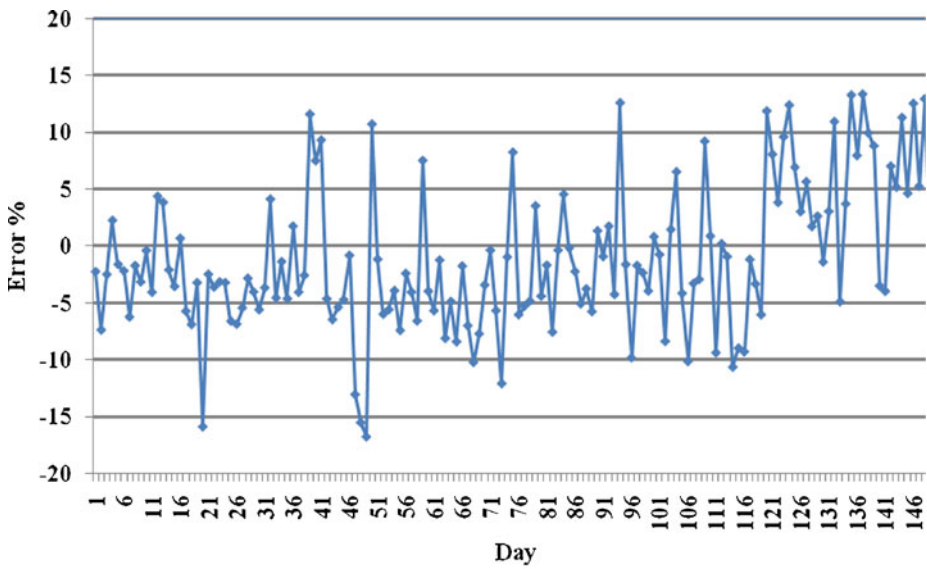
Before examining the model performance in details utilizing the statistical indices as presented in Sections 3–4, it is essential to pre-evaluate the model utilizing the error distribution. The error distribution equation is as follow;

$$Error = \frac{y - y'}{y} * 100 \quad (17)$$

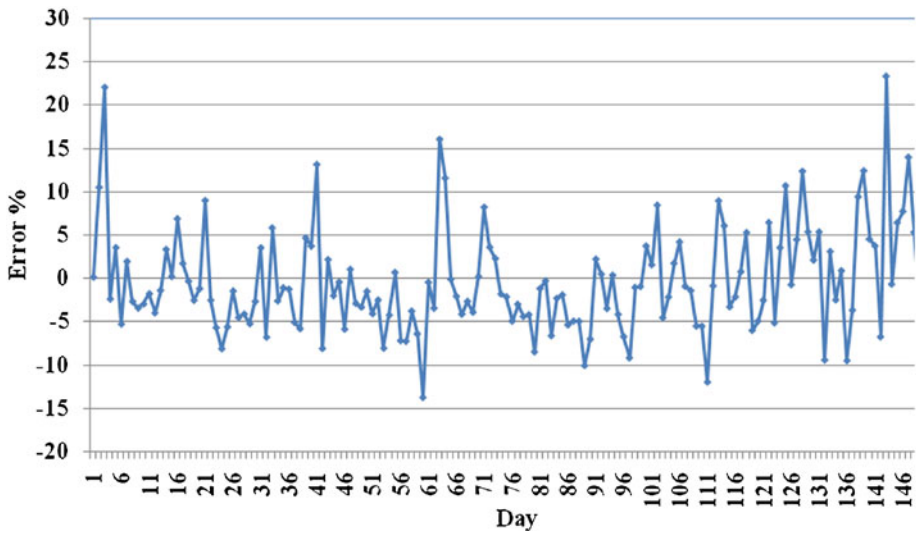
As an example for the model performance utilizing scenario I, Fig. 3a, b shows the result achieved for selected months April and September for Rasht city. It should be highlighted here that those two months present turning point for different season in terms of climate conditions for both cities. Figure 3 shows the error distribution for daily evapotranspiration as output from the proposed model versus the PM method for the testing period between 2001 and 2005, “5*30=150”. On the other hand, Fig. 4a, b demonstrates error distribution of the proposed ANN model for Johor city. In general, it can be noticed that ANN model could provide relatively reliable accuracy, however the maximum is relatively high which reflect a weakness on the proposed architecture of the model utilizing scenario I. For example, the maximum error for month of September for both cities is above 40 % in some days and more than 25 % for month of April. Such relatively high error is due to two major reasons, i) The model architecture and ii) Model overfitting.

Table 1 The ANN architecture for each month

Month	N1	Transfer function	Transfer function
Aug.	6	Tan_sigmoid	Pure-line
Sept.	6	Tan_sigmoid	Pure-line
Oct.	5	Log_sigmoid	Pure-line
Nov.	5	Log_sigmoid	Pure-line
Dec.	6	Tan_sigmoid	Pure-line
Jan.	6	Pure-line	Pure-line
Feb.	4	Pure-line	Pure-line
March	4	Pure-line	Pure-line
April	6	Pure-line	Pure-line
May	6	Pure-line	Pure-line
June	5	Tan_sigmoid	Pure-line
July	5	Tan_sigmoid	Pure-line



(a) April



(b) September

Fig. 3 Error distribution for ANN model scenario I for months April and September in Rasht City Iran

In order to improve the performance of the model architecture, scenario II has been implemented and examined as shown in Figs. 5 and 6 for the same selected months for both cities, respectively. Apparently, the results achieved from scenario II are relatively better than scenario I. It can be depicted from Fig. 3a, b that the maximum error decreased to 15 % for Rasht city for both months April and September, while it kept relatively high “30 %” for month of September for Johor city, as shown in Fig. 6, while, the maximum error value is reduced to 15 % for month of April for Johor City.

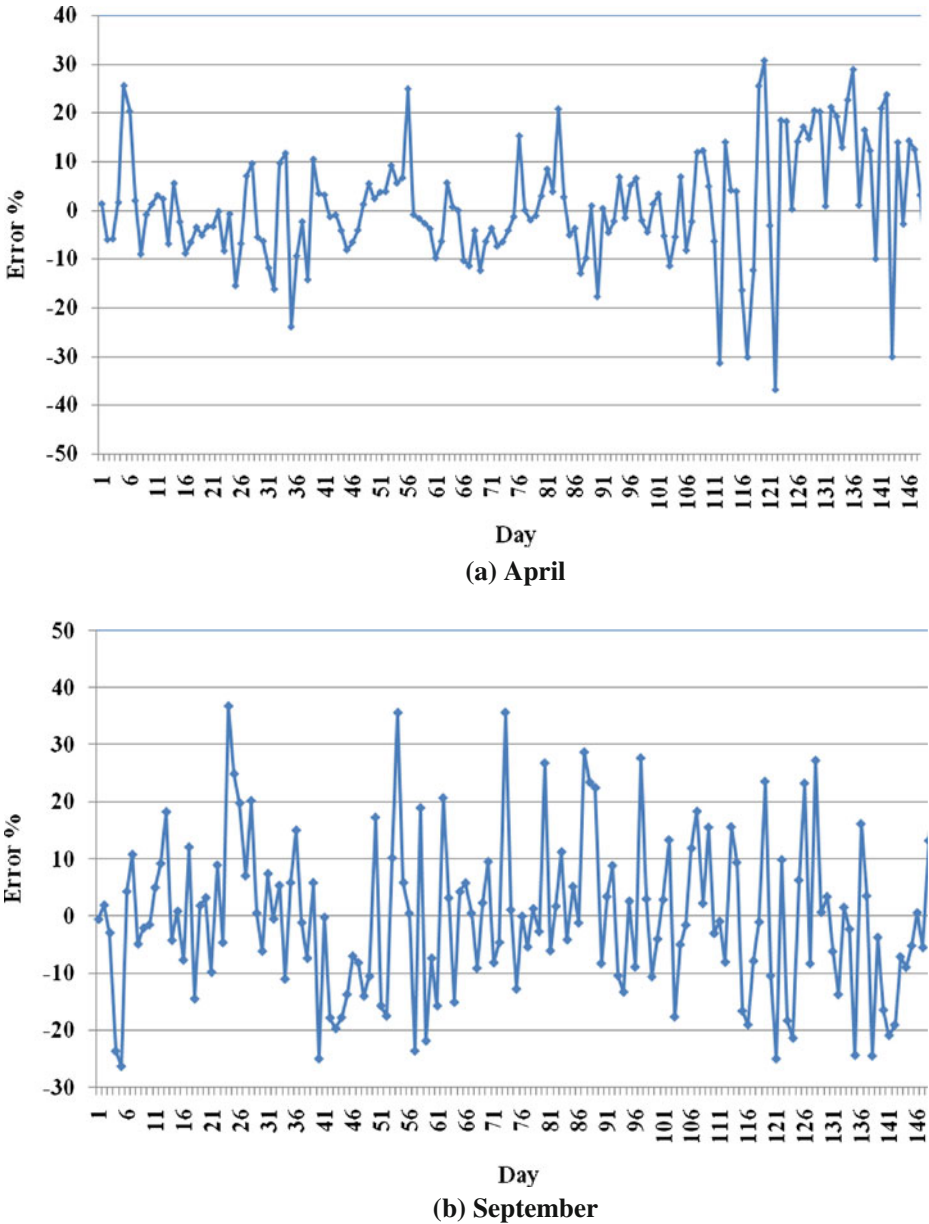


Fig. 4 Error distribution for ANN model scenario I for months April and September in Johor City Malaysia

In search for better understanding of the difference between the performance of both scenarios, further analysis to study the overall performance (not on monthly basis) of both scenarios have been carried out. Tables 2 and 3 show complete details for the selected statistical indices to evaluate the performance for both scenarios. In addition, these tables introduce the performance of ANN to PM method as a reference and to HGS as empirical

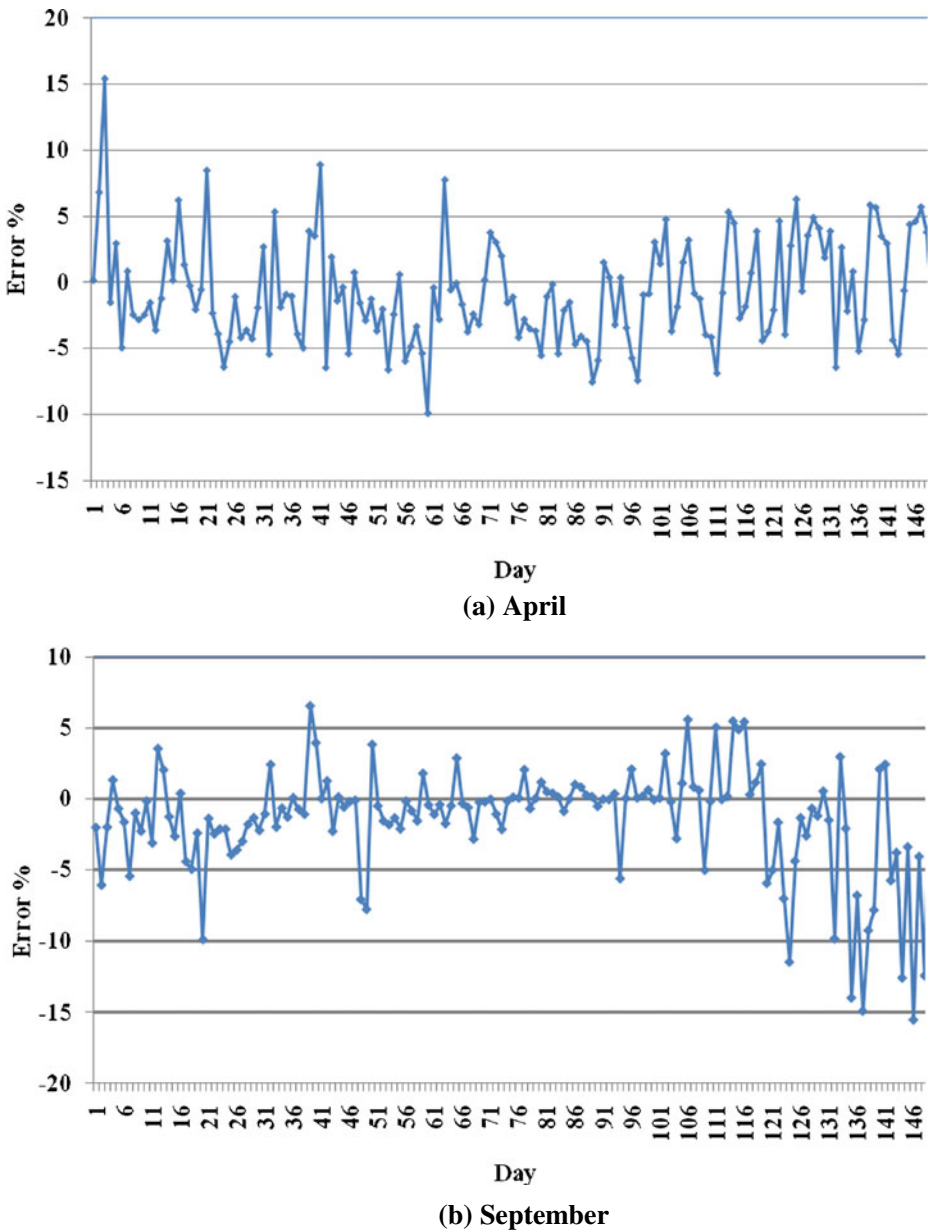


Fig. 5 Error distribution for ANN model Scenario II for months April and September in Rasht City Iran

method “competitor. It could be observed that ANN provides significant improvement on the level of accuracy in predicting the evapotranspiration if compared with HGS method. Furthermore, it could be found out that there is no remarkable change when comparing the results achieved for both classical ANN scenarios. In addition, the ANN model scenario I could provide better accuracy for 4 months January, March, April

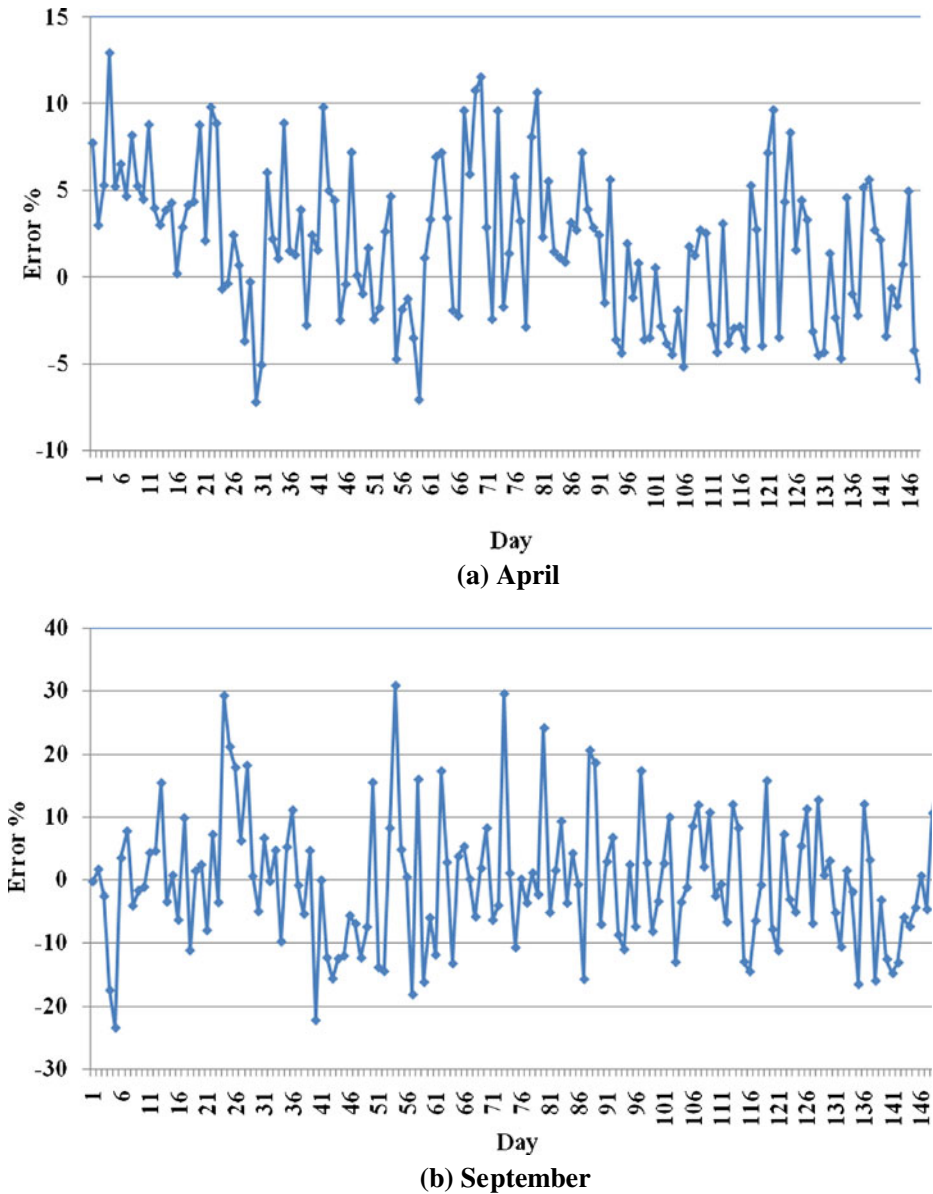


Fig. 6 Error distribution for ANN model Scenario II for months April and September in Johor City Malaysia

and September for Rasht City and only 1 month October for Johor City as highlighted with gray in the associated cell of the Tables 2 and 3, however, the ANN model scenario II could achieve better accuracy level than scenario I for the rest. Accordingly, in order to compare both ANN scenarios, averaging the whole statistical indices for all months. Tables 2 and 3 present the averaging values for MAE, MARE, MSE and R^2 in the last row. The results showed that scenario II outperformed scenario I and provides relatively higher performance. For further details, accuracy improvement (AI) index for each of

Table 2 MAE and MARE for ET_o prediction using HGS method, ANN scenario I and scenario II

Month	MAE						MARE					
	HGS		ANN				HGS		ANN			
			Scenario I		Scenario II				Scenario I		Scenario II	
	Rasht City	Johor City	Rasht City	Johor City	Rasht City	Johor City	Rasht City	Johor City	Rasht City	Johor City	Rasht City	Johor City
Jan	0.82	0.73	0.34	0.23	0.44	0.16	60.76	53.22	23.37	42.35	31.23	22.71
Feb	0.53	0.63	0.53	0.32	0.45	0.23	43.53	39.23	18.83	23.46	8.00	15.21
Mar	0.34	0.54	0.23	0.22	0.28	0.19	34.63	52.63	18.45	30.58	27.34	19.32
Apr	0.63	0.59	0.58	0.37	0.62	0.25	31.82	56.82	16.54	42.55	19.25	25.27
May	0.76	0.72	0.21	0.21	0.16	0.18	28.54	56.54	16.86	41.23	12.34	28.35
Jun	0.78	0.82	0.85	0.41	0.51	0.30	28.82	85.82	14.28	65.59	9.07	34.17
Jul	0.85	0.72	0.39	0.83	0.29	0.56	31.45	81.45	15.82	51.3	12.23	28.77
Aug	0.83	0.62	0.53	0.59	0.32	0.41	45.32	73.32	16.47	43.59	15.58	26.52
Sep	0.85	0.75	0.44	0.23	0.5	0.17	42.77	64.77	16.65	44.76	27.36	24.91
Oct	0.73	0.82	0.57	0.31	0.40	0.4	53.63	71.63	32.74	58.64	16.78	66.09
Nov	0.67	0.81	0.28	0.27	0.19	0.21	53.12	62.12	33.34	48.36	21.50	29.29
Dec	0.84	0.72	0.62	0.54	0.41	0.32	63.23	53.23	25.54	31.27	13.72	22.42
Average	0.72	0.71	0.46	0.38	0.38	0.28	43.13	62.56	20.74	43.64	17.87	28.59

Table 3 MSE and R² for ET_o prediction using HGS method, ANN scenario I and scenario II

Month	MSE						R ²					
	HGS		ANN				HGS		ANN			
			Scenario I		Scenario II				Scenario I		Scenario II	
	Rasht City	Johor City	Rasht City	Johor City	Rasht City	Johor City	Rasht City	Johor City	Rasht City	Johor City	Rasht City	Johor City
Jan	0.78	0.88	0.33	0.37	0.45	0.18	0.44	0.53	0.65	0.73	0.57	0.82
Feb	0.65	0.75	0.42	0.33	0.22	0.25	0.54	0.49	0.56	0.64	0.80	0.73
Mar	0.59	0.79	0.22	0.31	0.33	0.23	0.56	0.63	0.71	0.71	0.62	0.81
Apr	0.41	0.51	0.14	0.31	0.27	0.18	0.58	0.68	0.72	0.71	0.65	0.85
May	0.77	0.53	0.37	0.37	0.18	0.26	0.66	0.56	0.76	0.67	0.84	0.82
Jun	0.92	0.73	0.54	0.46	0.32	0.27	0.67	0.57	0.68	0.63	0.78	0.85
Jul	0.72	0.68	0.38	0.38	0.29	0.24	0.57	0.62	0.66	0.73	0.88	0.87
Aug	0.90	0.83	0.51	0.57	0.38	0.34	0.62	0.58	0.76	0.67	0.82	0.85
Sep	0.70	0.83	0.46	0.53	0.55	0.31	0.79	0.61	0.80	0.75	0.77	0.87
Oct	0.60	0.72	0.34	0.46	0.25	0.58	0.64	0.57	0.76	0.65	0.85	0.53
Nov	0.86	0.75	0.52	0.37	0.21	0.29	0.52	0.62	0.69	0.72	0.78	0.85
Dec	0.73	0.69	0.57	0.45	0.35	0.37	0.62	0.72	0.74	0.82	0.88	0.82
Average	0.72	0.72	0.37	0.38	0.30	0.27	0.60	0.60	0.72	0.72	0.78	0.81

Table 4 Accuracy Improvement AI% for MAE, MARE, MSE and R² associated with the output of both scenarios

Month	AI%							
	MAE		MARE		MSE		R ²	
	Rasht City	Johor City	Rasht City	Johor City	Rasht City	Johor City	Rasht City	Johor City
Jan	-29	30	-34	46	-36	51	12	-12
Feb	15	28	58	35	48	24	-43	-14
Mar	-22	14	-48	37	-50	26	13	-14
Apr	-7	32	-16	41	-93	42	10	-20
May	24	14	27	31	51	30	-11	-22
Jun	40	27	36	48	41	41	-15	-35
Jul	26	33	23	44	24	37	-33	-19
Aug	40	31	5	39	25	40	-8	-27
Sep	-14	26	-64	44	-20	42	4	-16
Oct	30	-29	49	-13	26	-26	-12	18
Nov	32	22	36	39	60	22	-13	-18
Dec	34	41	46	28	39	18	-19	0

the above statistical index to measure the significance of scenario II module over scenario I model could be expressed as follows

$$AI\% = \left(\frac{E_{scenarioI} - E_{scenarioII}}{E_{scenarioI}} \right) * 100 \tag{18}$$

Where $E_{scenarioI}$ is the value of the statistical error given by scenario I, while $E_{scenarioII}$ is the same statistical error given by the proposed scenario II. It should be noted that the negative

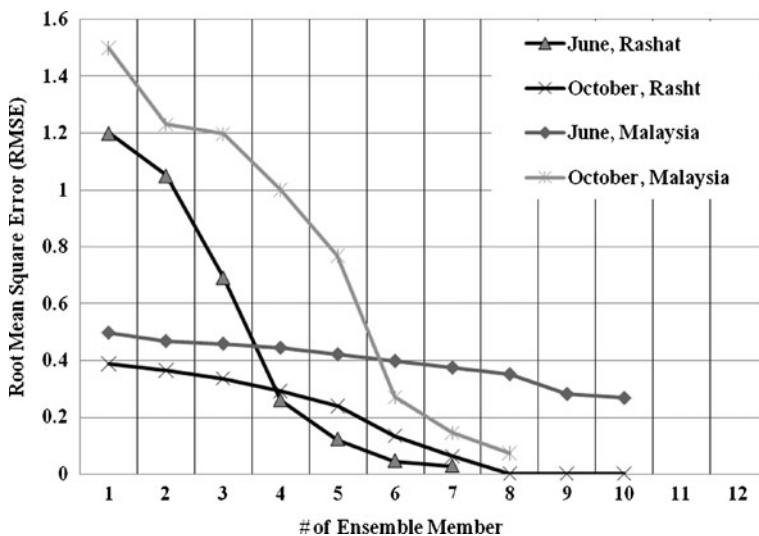
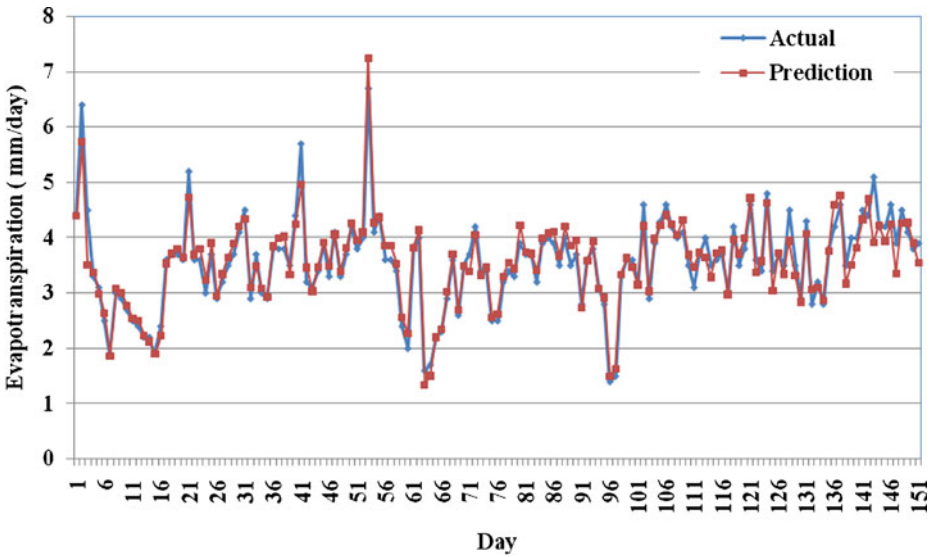


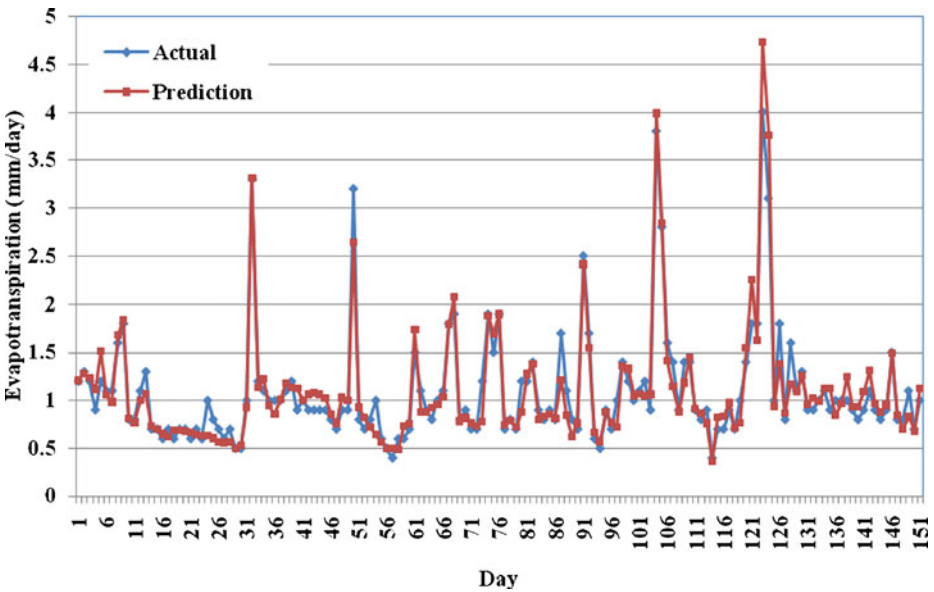
Fig. 7 The effect of increasing number of ensemble neural network on the mean square error for months July and December

Table 5 MAE, MARE, MSE and R² associated with the output of Ensemble ANN and its Accuracy Improvement AP% over classical ANN Scenario II

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Number of ensemble ANN	Rasht City	4	5	5	4	7	7	5	6	8	4	7
	Johor City	5	5	4	3	10	7	6	6	8	8	7
Ensemble ANN	MAE	Rasht City	0.15	0.15	0.08	0.15	0.21	0.19	0.22	0.15	0.10	0.12
	Johor City	0.14	0.13	0.13	0.15	0.13	0.20	0.26	0.21	0.15	0.10	0.13
MARE	Rasht City	8.90	7.5	6.26	5.44	4.36	5.07	4.59	5.95	6.05	6.74	11.27
	Johor City	7.26	12.81	8.36	7.13	4.36	4.72	6.68	6.53	4.79	6.08	12.43
MSE	Rasht City	0.04	0.04	0.01	0.04	0.06	0.05	0.06	0.08	0.04	0.03	0.04
	Johor City	0.04	0.03	0.03	0.04	0.03	0.06	0.08	0.06	0.04	0.02	0.04
R2	Rasht City	0.91	0.93	0.92	0.96	0.93	0.92	0.89	0.92	0.96	0.88	0.88
	Johor City	0.93	0.93	0.91	0.94	0.96	0.86	0.88	0.93	0.94	0.91	0.87
AP%	MAE	Rasht City	65.9	66.6	71.4	75.8	58.8	34.4	31.2	70.0	75.	70.7
	Johor City	6.25	43.4	31.5	40.0	27.7	33.3	53.7	48.7	11.7	47.6	59.3
MARE	Rasht City	71.5	0.12	77.1	71.7	64.6	44.1	62.4	61.8	77.8	48.6	17.8
	Johor City	68.0	15.7	56.7	71.7	84.6	86.1	76.7	75.3	80.7	66.1	44.5
MSE	Rasht City	91.1	81.8	96.9	85.1	66.6	84.3	79.3	78.9	92.7	90.4	88.5
	Johor City	77.7	88.0	86.9	77.7	88.4	77.7	66.6	82.3	87.1	93.1	89.1
R2	Rasht City	-59.6	-16.2	-48.3	-47.6	-10.7	-17.9	-1.14	-12.2	-24.6	-17.	0.00
	Johor City	-13.4	-27.4	-12.3	-10.5	-17.0	-1.18	-1.15	-9.41	-8.05	-71.7	0.00



(a) Rasht City



(b) Johor City

Fig. 8 Actual versus the prediction evapotranspiration utilizing ensemble ANN model, for month of April **a** Rasht **b** Johor Cities

values of MAE, MARE and MSE and positive values for R^2 mean that the scenario II provides better results over scenario I. Table 4 introduces the AI% calculated values for the above mentioned statistical indices. It could be observed from AI% that scenario II could significantly improve the accuracy level for most of the months.

Table 6 Classical ANN Scenario II and Ensemble ANN performance based on the peak and low flow error criteria

Month	Classical ANN Scenario II			Ensemble ANN		
	PVC (%)	LVC (%)	Average %	PVC (%)	LVC (%)	Average %
Jan	2.20	4.20	3.20	1.45	3.50	2.48
Feb	5.15	4.38	4.77	4.61	3.24	3.93
Mar	6.20	5.20	5.70	3.20	5.00	4.10
Apr	7.25	5.08	6.17	2.80	4.95	3.88
May	5.80	3.54	4.67	4.60	2.80	3.70
Jun	6.58	2.55	4.57	4.70	2.25	3.48
Jul	8.52	5.82	7.17	5.30	4.15	5.23
Aug	9.53	2.40	5.97	2.80	2.10	2.45
Sep	2.85	5.84	4.35	1.80	4.95	3.38
Oct	3.24	2.54	2.89	2.86	1.88	2.37
Nov	4.25	1.25	2.75	3.95	1.20	2.58
Dec	7.53	5.10	6.32	4.50	4.82	4.66

The ensemble procedure described in Section 3.3 was applied to overcome the overfitting and improve the generalization of the training process of the 12 networks with scenario II architecture. For example, the ensemble ANN model is applied for the same 2 months (June and October) which experienced relatively poor accuracy for testing data set because of the over fitting problem during training. Figure 7 shows that seven and eight networks are selected as ensemble members for the above mentioned months and ten and eight networks for Rasht city and Johor city, respectively, before the termination conditions met. As shown in Fig. 7, utilizing ensemble ANN method could reduce the MSE significantly if compared with the classical ANN method (over fitted network). The performance of the ensemble ANN model during the testing data (2001 and 2005) was examined and presented in Table 5. Table 5 summarizes the results of proposed ensemble ANN associated with each month for statistical indices for the testing period. In addition, Table 5 shows the AI% for all the statistical indices over the classical ANN model with scenario II. It can be depicted that the ensemble ANN model successfully provides better and consistent level of accuracy for all months for all statistical indices if compared with the classical ANN scenario II showed in Tables 2 and 3.

For further assessment, it is of important to examine the proposed model for predicting the extreme values of the evapotranspiration. This is due to the fact that the extreme value is essential for water resources managers. Figure 8 shows the actual values of evapotranspiration versus the prediction values as the output from the proposed model. Figure 8a shows the actual and prediction values for month of April for Rasht City, and Fig. 8b for the same month for Johor city. The selection for month of April is due to the fact that this month experienced the most fluctuated values for evapotranspiration. It could be noticed that the proposed model could accurately predict the extreme values. In addition, the ensemble ANN model is examined for the peak and low ET_o events using PVC and LVC statistics as discussed earlier in Section (4.3) and presented in Eqs. 15 and 16 are presented in Table 6. As presented in Table 6, it can be observed that the developed ensemble ANN model can perform the function of providing an accurate prediction value for the PVC ranging between 1.45 and 5.3 % for months January and July, respectively, while the LVC values experienced better prediction error if compared with the PVC ranging between 1.2 and 5 % for months November and

March, respectively. On the other hand, the PVC and LVC values achieved, utilizing the classical ANN model are relatively higher than ensemble ANN model ranging between 2.2 and 9.53 for PVC and 1.25 and 5.84 for LVC for January, August, November and September, respectively. As a result, utilizing the proposed ensemble ANN model outperforms the classical ANN model and significantly improves the prediction error for not only for the medium range of ET_o but also for the ET_o extreme values.

5 Conclusion

This study introduced a procedure for development of Artificial Neural Network (ANN) model for predicting daily evapotranspiration on monthly basis time series using only maximum, minimum range of temperature and solar radiation. The model was examined utilizing the actual data for two cities that experience two different climatic conditions namely Rasht in Iran and Johor in Malaysia. Two different scenarios for input–output pattern of the ANN model architecture in order to optimize the accuracy level were developed. In addition, a generalization technique is developed namely; Ensemble ANN to overcome the over-fitting problem that is experienced while training the ANN model. The proposed ANN model was evaluated and compared with the traditional model Hargreaves-Samani (HGS). The results showed that the classical ANN model outperformed the HGS model and achieved reliable performances including low MAE, MARE, MSE and high R^2 . In addition, the proposed ANN with ensemble procedure showed significant enhancement over the classical ANN. Furthermore, the ensemble ANN proved its reliability and high consistency level of accuracy when examined for predicting the extreme values of ET_o . The ability of the proposed ensemble ANN is extremely useful especially for predicting evapotranspiration which is highly needed for water resources managers while the availability of the data is limited.

References

- Allen RG, Pruitt WO (1991) FAO-24 reference evapotranspiration factors. *J Irrig Drain Eng* 117(5):758–773
- Awchi TA (2008) Application of radial basis function neural networks for reference evapotranspiration prediction. *Al-Rafidain Eng* 16(1):117–129
- Bishop CM (1996) *Neural networks for pattern recognition*, 1st edn. Oxford University Press, Oxford
- Blaney HF, Criddle WD (1962) Determining consumptive use and irrigation water requirements. *US Dept Agric Agric Res Serv Tech Bull* 1275, 59p
- Brutsaert WH (1982) *Evaporation into the atmosphere*. D. Reidel, Dordrecht
- Chandra A, Yao X (2006) Evolving hybrid ensembles of learning machines for better generalization. *Neurocomputing* 69:686–700
- Chauhan S, Shrivastava RK (2008) Performance evaluation of reference evapotranspiration estimation using climate based methods and artificial neural networks. *Water Resour Manag* 23:825–837
- El-Shafie A, Taha MR, Noureldin A (2007) A neuro-fuzzy model for inflow forecasting of the Nile river at Aswan high dam. *Water Resour Manag* 21(3):533–556
- El-Shafie A, Abdin AE, Noureldin A, Taha MR (2009) Enhancing inflow forecasting model at Aswan high dam utilizing radial basis neural network and upstream monitoring stations measurements. *Water Resour Manag* 23(11):2289–2315
- Hargreaves GH, Samani ZA (1985) Reference crop evapotranspiration from temperature. *Appl Eng Agric* 1(2): 96–99
- Hsu K, Gupta HV, Sorooshian S (1995) Artificial neural network modeling of the rainfall-runoff process. *Water Resour Res* 31(10):2517–2530
- Imrie CE, Durucan S, Korre A (2000) River flow prediction using artificial neural networks: Generalization beyond the calibration range. *J Hydrol* 233:138–153

- Jain SK, Das A, Srivastava DK (1999) Application of ANN for reservoir inflow prediction and operation. *J Water Resour Plan Manag* 125(5):263–271
- Khoob AR (2008) Comparative study of Hargreaves's and artificial neural network's methodologies in estimating reference evapotranspiration in a semiarid environment. *Irrig Sci* 26(3):253–259
- Kumar M, Raghuvanshi NS, Singh R, Wallender WW, Pruitt WO (2002) Estimating evapotranspiration using artificial neural network. *J Irrig Drain Eng* 128(4):224–233
- Maier HR, Dandy GC (1999) Empirical comparison of various methods for training feed-forward neural networks for salinity forecasting. *Water Resour Res* 35(8):2591–2596
- Sudheer KP, Gosain AK, Ramasastri KS (2003) Estimating actual evapotranspiration from limited climatic data using neural computing technique. *J Irrig Drain Eng* 129(3):214–218
- Sulaiman M, El-Shafie A, Karim O, Basri H (2011) Improved water level forecasting performance by using optimal steepness coefficients in an artificial neural network. *Water Resour Manag* 25(10):2525–2541
- Tokar AS, Johnson PA (1999) Rainfall-runoff modeling using artificial neural networks. *J Hydrologic Eng* 4(3): 232–239
- Trajkovic S, Todorovic B, Stankovic M (2003) Forecasting of reference evapotranspiration by artificial neural networks. *J Irrig Drain Eng* 129(6):454–457
- Wang YM, Traore S, Kerh T (2008) Neural network approach for estimating reference evapotranspiration from limited climatic data in Burkina Faso. *WSEAS Trans Comput* 7(6):704–713
- Zanetti SS, Sousa EF, Oliveira VPS, Almeida FT, Bernardo S (2007) Estimating evapotranspiration using artificial neural network and minimum climatological data. *J Irrig Drain Eng* 133(2):83–89