Desalination and Water Treatment
www.deswater.com

doi:10.5004/dwt.2018.22249

111 (2018) 183–191
April

Evolutionary support vector machine for evapotranspiration estimation (case study: Haji Abad region, Hormozgan province)

Omolbani Mohammadrezapour[a,*], Abolfath Moradi[b], Ozgur Kisi[c], Salman Sharifazari[d]

[a]*Associated Professor, Department of Water Engineering, University of Zabol,. P. O. Box 98615-538, Zabol, Iran, email: omohammadrezapour@gmail.com, mohammadrezapour@uoz.ac.ir*
[b]*Soil and Water Research Department, Fars Agricultural and Natural Resources Researchand Education Center, AREEO, Shiraz, Iran, email: Moradi1373@gmail.com*
[c]*Ilia State University, Faculty of Natural Sciences and Engineering, Tbilisi, Georgia, email: ozgur.kisi@iliauni.edu.ge*
[d]*Department of Water Engineering, University of Zabol. P. O. Box 98615-538, Zabol, Iran.s.sharifazari@gmail.com*

Received 22 November 2017; Accepted 24 March 2018

ABSTRACT

Accurate estimation of evapotranspiration (ET) values is of crucial importance in hydrology, agriculture and agro-meteorology issues. The objective of this research was to evaluate the use of evolutionary support vector machine (ESVM) to model daily ET using limited climatic data. For this aim, the most common evolutionary method, genetic algorithm (GA), was used for optimization of SVM variables. For the ESVM, four input combinations of maximum air temperature (T_{max}), minimum air temperature (T_{min}), wind speed (U_2), daily solar radiation (Rs), relative humidity (Rh_{mean}) and mean temperature (T_{mean}) were tried. Climatic data covering 3-year period of October 2004–October 2007 were obtained from the extremely arid and hot region of Haji Abad located in the northern region of Hormozgan province, Iran. Artificial Neural Network (ANN) as a base model was also applied for evaluating modeling accuracy of the ESVM in estimating ET. The results of the ESVM and ANN models were evaluated by comparing their estimates with the measured lysimetric data. The root mean square error (RMSE), coefficient of efficiency (CE) and the coefficient of determination (R^2) were used as comparison criteria. According to the results obtained, the ESVM2 whose input variables are T_{mean} and Rh_{mean} was selected as the best model in estimating ET.

Keywords: Evapotranspiration; Lysimetric data; Estimation; SVM; Genetic algorithm; ANN

1. Introduction

Evapotranspiration is one of the most important components of hydrological cycle, so its precise estimation seems quite essential for many research projects, especially in irrigation and drainage schemes and water resources. Evapotranspiration is the process by which water is lost from soil surface by evaporation and from plant by both evaporation and transpiration to the atmosphere [1]. Evaporation is the process by which liquid water changes to vapor water, and transpiration is the process of water movement through areal parts of plant mainly leaves as vapor to the surrounding atmosphere [2]. Since evaporation and transpiration occur simultaneously at the field scale and cannot be easily separated, the term evapotranspiration has been defined as the sum of evaporation and transpiration to describe crop water requirement [3]. This is because more than 98% of plant water uptake is lost through evapotranspiration. Crop evapotranspiration is governed by weather parameters including air temperature and relative humidity, solar radiation and wind speed, crop characteristics and the availability of soil water. Due to the variability of these factors over time and space, crop evapotranspiration changes spatially and temporally. Therefore, in order

*Corresponding author.

to precise planning of irrigation projects such as irrigation scheduling, evapotranspiration needs to be determined accurately in any location and over time during the growth season of crops. Of course, it requires costly tools that are hard to be used everywhere. Therefore, researchers have always sought applied cheap, and accurate relationships and methods for correct estimation of this parameter. Several methods such as the empirical equations of Priestley-Taylor, Blaney–Criddle, McGuinness-Bradley, and data-axis methods, like neural-artificial networks, support vector machine, and tree models have been developed for accurate estimation of evapotranspiration worldwide. One of the main problems of the empirical equations is their low flexibility in dealing with various data, which leads to weak results of the equations. Due to their higher flexibility, data-driven methods have a good ability in modeling different situations. Many studies have been conducted on the use of data-based methods for estimating the evapotranspiration of reference plants, which are referred in this article. Using the Penman-Monteith method, Kumar et al. [4] identified the use of artificial neural network in predicting the evapotranspiration of the superior reference plant. The studies on prediction of runoff associated with atmospheric precipitation in 21 areas in Canada carried out by Cannon and Whitfield [5]. They recognized that the artificial neural network has a higher efficiency than the linear regression model. Neural networks trained with Levenberg-Marquardt algorithm in prediction of daily reference evapotranspiration used by Kisi [6]. For making a comparison between the network results, he also used conjugate gradient (CG) algorithm. Comparison of the artificial neural network results with those of Penman and Hargreaves models showed that the artificial neural network has a higher accuracy to predict reference evapotranspiration using meteorological data.

Neural network with minimum meteorological data to estimate evapotranspiration in the city of Rio de Janeiro, Brazil also used by Zanetti et al. [7]. they used the meteorological data for the period of September, 1996–August, 2002 to train the network. The results showed that the artificial neural network with the maximum and minimum air temperatures inputs can successfully predict evapotranspiration. The prediction of reference evapotranspiration by self-organizing neural networks evaluated by Adeloye et al. [8]. In this study, the synoptic daily data (from 1811 to 1915) obtained from the meteorological station of the University of Edinburgh, England were used as inputs of the self-organizing map neural network and the experimental models. The results showed the high ability of the neural networks and the Hargreaves empirical equation in predicting the proper amount of the potential evapotranspiration. Daily reference evapotranspiration (ET_0) in the north of Algeria using generalized regression neural networks (GRNN) and radial basis function neural networks (RBFNN) was modeled by Ladlani et al [9]. Also the empirical Hargreaves-Samani (HG) and Priestley-Taylor (PT) equations are also considered for the comparison. The result showed that the GRNN was better than the RBFNN, Priestley-Taylor and Hargreaves-Samani models.The accuracy of nine solar radiation estimating models (Rs) using data gathered from eight weather stations in Canada to calculate evapotranspiration was evaluated by Aladenola et al. [10]. The results showed high accuracy of Hargreaves-Samani (H-S) model and low accuracy of Mahmoud-Hobart (M-H) model. The support vector machine (SVM) and artificial neural network (ANN) to model the evapotranspiration in an area in California were used by Kisi and Cimen [11]. In this study, the ET_0 data obtained from the standard Penman-Monteith FAO 56 equation were used as the SVM and ANN inputs. The results revealed that the SVM at each level of input data in the model has better results than the ANN.

For modeling daily ET_0 with limited climatic data the support-vector-machine used by Wen et al. [12]. The results of SVM models were evaluated by comparing the output with the ET_0 calculated using Penman–Monteith FAO 56 equation (PMF-56) and Artificial Neural Network (ANN). They found that three climatic parameters, T_{max}, T_{min}, and R_s were enough to predict the daily ET_0 satisfactorily. 3 hybrid models of support vector regression (SVR) consisting of wavelet transforms (WT), singular spectrum analysis (SSA), and a chaotic approach (CA) were applied to the input time series, evaluated for predicting river flow of Kizilirmak River in Turkey by Baydaroglu et al. [13]. The result showed that the SVR-WT had the highest coefficient of determination and the lowest mean absolute error than the other models. There have been many studies (Trajkovic et al. [14], Shiri et al. [15], Marti et al. [16], Shiri et al. [17], Shiri et al. [18], Kim et al. [19], Landeras et al. [209], Shiri et al. [21] and Shiri et al. [22]) in the field of evapotranspiration estimation. Giving all the related studies is beyond the scope of this study and therefore some of them were only presented in this section. In the previous studies, reference evapotranspirations calculated by Penman Monteith FAO 56 were generally used for modeling amount of evapotranspiration by artificial intelligence models such as (ANN, SVM, GRNN,etc). It is evident from the related literature, there is not any work which investigates the applicability of SVM (SVM) and ANN in estimation of ET measured by lysimeter. Also, for better performance of the SVM model, the control parameters were optimized using genetic algorithm. Therefore, the main objective of this study was to investigate the accuracy of evolutionary support vector machine (ESVM) models for estimating daily ET measured by lysimeter using various combinations of daily meteorological data including maximum air temperature (T_{max}, °C), minimum air temperature (T_{min}, °C), mean air temperature (T_{mean}, °C), wind speed (U, km/month), solar radiation (R_s) and relative humidity (Rh, %) in a humidity area. In addition, the performance of the ESVM models was compared with those of the ANN.

2. Materials and methods

The purpose of this study was to estimate evapotranspiration by using ESVM and ANN models. Thus, MATLAB software was used to train models by using the two-thirds of the data, and the remaining one- third was used for testing each model.

3.1. Area of study and data

This research was carried out in Hajiabad area with the coordinates of 27°13′ N 55° 22′ E elevated by 920 m in the northern region of Hormozgan province for 3 years (Fig. 1).

In this research, first, a drainage lysimeter, with a dimension of 3*3*1.9 m, was installed in the center of a 2-ha land plot. Then, the Bermudagras cultivar was planted in the lysimeter and its surrounding with an area of 900 m^2. After complete covering was done on the ground and the height of the grass reached about 8 cm, its weekly evapotranspiration was calculated for three years (2004-2007) using the water balance method (Eq. (1)).

$$ET = I + P + \Delta SW\text{-}D \quad (1)$$

where ET: Evapotranspiration of the reference plant (mm), I: irrigation water depth (mm) measured by the watermeter installed in the inlet of irrigation water to lysimeter, P: rainfall (mm) measured by rain gauge, D: amount of water drained from the lysimeter (mm) measured at the end of the weekly period by measuring the amount of water collected in the open-air slab along the Lysimeter range, ΔSW: soil moisture changes in the lysimeter at the beginning and end of the weekly period (mm) and equal to SW_1–SW_2, SW_1: amount of water in the soil at the beginning of the period, SW_2: amount of water in the soil at the end of the period.

The climatological parameters such as relative temperature, relative humidity, wind speed, and solar radiation, were measured at the meteorological station of the experimental site and were also used for weekly periods. The actual evapotranspiration in this study were measured from Drainage lysimeterfrom October 2005- October 2008. Table 1 shows the mean of the variables used for modeling daily evapotranspiration during the statistical period for the study area.

The dailystatistical parameters of the climatic data are given in Table 1; X_{min}, X_{max}, X_{mean}, $S_x$and C_v symbolize the maximum, minimum, mean, standard deviation, skewness and variation coefficient, respectively.

Fig. 1. The studied area in Hormozgan province.

3.2. Model definition

3.2.1. Support vector machine (SVM)

The support machine is an efficient learning system based on the compact optimization theory, which uses the principle of deployment of a structural error minimization and leads to an overall optimal response.

In the SVM regression model, a function associated with the dependent variable Y is estimated which is a function of several independent variables X by itself. Like other regression issues, it is assumed that the relationship between independent and dependent variables is determined by an algebraic function such as $f(x)$ plus some noise (admissible error ε).

$$f(x) = W^T.Q(x) + b \quad (2)$$

$$y = f(x) + noise \quad (3)$$

If W is the vector of the coefficients, b is the characteristic of the regression function, and Q is also a kernel function, the purpose is to find a functional form for $f(x)$. This is accomplished by calibrating the SVM model which is obtained with a set of samples (the calibration set).

This process involves the sequential optimization of the error function. Depending on the definition of this error function, two SVM models can be defined: SVM regression type I (known as ε-svm regression) and SVM regression type II (known as ε-SVM regression). It is worth noting that in this research, the regression model ε-SVM has been used due to its extensive application in regression studies.

Therefore, in order to calculate W and b, the error function (Eq. (4)) in the ε-SVM model must be optimized by considering the conditions (constraints) in Eq. (5).

$$\frac{1}{2}W^T.w + c\sum_{i=1}^{N}\varepsilon_i + c\sum_{i=1}^{N}\varepsilon_i^* \quad (4)$$

$$y_i - W^T.Q(x_i) - b \le \varepsilon + \varepsilon_i w^T.Q(x_i) + b - y_i \le \varepsilon + \varepsilon_i^*, \varepsilon_i^* \ge 0, i = 1,2,\ldots N \quad (5)$$

In the above equations, c is a positive integer that determines the fine when calibration error occurs. Q is the kernel function, N is the number of samples, and the two attributes ε_i, ε_i^* are slack variables which determine the upper and lower limits of the admissible error ε.

In problems, it is predicted that the data will fall within the boundary interval ε. However, if the data is outside the interval ε, there will be an error equal to ε_i^* and ε_i. It

Table 1
Statistical characteristics of the data used in the study

	X_{mean}	X_{max}	X_{min}	S_x	C_v
Lysimeter evapotranspiration (mm/day)	7.24	15.20	1.22	4.037	0.55
Mean temperature (C°)	23.17	35.16	8.72	8.011	0.34
Relative humidity (%)	42.83	95.86	9.29	12.93	0.30
Wind speed (km/h)	3.81	7.8	1.32	1.16	0.30
Solar radiation (R_s)	8.68	11.75	3.95	2.052	0.23

is also worth mentioning that the SVM model solves problems caused by under fitting and ultrafine estimate through simultaneously minimizing both the two semesters of WT·W/2 and the training error, i.e., $c\sum_{i=1}^{N}\left(\varepsilon_i+\varepsilon_i^*\right)$ in Eq. (3). Thus, with the introduction of two Lagrange coefficients, a_i^* and a_i the optimization problem with the numerical maximization of the following quadratic function (Eq. (6)) will be solved under the conditions of Eq. (7).

$$\sum_{i=1}^{N}y_i\left(a_i-a_i^*\right)-\varepsilon\sum_{i=1}^{N}\left(a_i+a_i^*\right)-0.5\sum_{i,j=1}^{N}\left(a_i-a_i^*\right)\left(a_j-a_j^*\right)Q\left(x_i\right)^T.Q\left(x_j\right) \quad (6)$$

$$\sum_{i=1}^{N}\left(a_i+a_i^*\right)=0 \quad 0\le a_i\le C \quad 0\le a_i^*\le C \quad i=1,2,\ldots,N \quad (7)$$

The objective function of Eq. (6) is a convex function; thus, the solution of Eq. (6) is unique and optimal. After defining the Lagrange coefficients in Eq. (4), the characteristics of W and b in the SVM regression model are calculated using the Karush-kuhn-tucker theory conditions [23], where. $W=\sum_{i=1}^{N}Q\left(x_i\right)\left(a_i+a_i^*\right)$

As a result, the regression for the SVM model will be:

$$W=\sum_{i=1}^{N}\left(a_i+a_i^*\right)Q\left(x_i\right)^T+b \quad (8)$$

It should be noted that Langrage terms (a_i+a_i*) can be zero or nonzero. Therefore, only the data sets whose coefficients $|a_i|$ are non-zero are entered in the final regression equation, so this set of data is considered as support vectors. Simply, the support vectors of those data whose $\bar{a}_i$ values are less than C are called marginal support vectors. When the a_i value of support vectors is equal to C, it is called an error support vector or a bounded support vector. Marginal support vectors are found at the insensitive marginal boundary, while vectors supporting the error are out of the range. Finally, the regression SVM function can be rewritten in the form of Eq. (9):

$$f\left(x\right)=\sum_{i=1}^{N}\bar{a}_i.Q\left(x_i\right)^T.Q\left(x\right)+b \quad (9)$$

In the above equation, $\bar{a}_i$ is the mean value of the Langrage coefficients. Calculating $Q(x)$ in its characteristic space may be very complicated. To solve this problem, the common way in the regression SVM model is choosing kernel function. A support vector machine is usually used with three radial basis functions (RBF), and several linear and degree polynomials [24,25].

Given the fact that the most commonly used kernel functions are linear, kernel, radial base, and polynomial functions [26–28], the radial basis kernel function is also used in this study and its function is mentioned in Eq. (10).

$$k\left(x,x_i\right)=\exp\left(-\left\|x-x_i\right\|^2\Big/Q^2\right) \quad (10)$$

Two important SVM parameters, C and gamma (γ) values, should be selected appropriately. The C parameter represents how much outliers are taken into account. The choice of the value of C affects the classification output. In other words, if the value of C is too large, the rate of classification accuracy in the training phase will be very high, whereas at the testing phase, the rate of classification accuracy sharply reduces, specifically, too much fitting occurs. On the other hand, if the value of C is too small, it will reduce the accuracy of the classification, which is also inappropriate. The gamma (γ) parameter has a greater effect than the C parameter on the output of the classification, since its value affects the resulting feature vector. A very large amount of this parameter results in excessive fitting and its very small amount reduces accuracy. The objective function in optimizing the support vector model using the genetic algorithm involves minimizing the amount of evapotranspiration error calculated by the support vector model and the measured values using the Lysimeter. As this parameter reaches its lowest value, the parameters of SVM model (C, γ) will reach their optimal values. For this purpose, the genetic algorithm (GA) was used. In ESVM modeling, there are three parameters (C, γ and ε) whose variation makes changes in the model performance and affects the convergence speed and the quality of the answers.

3.2.2. Genetic algorithm (GA)

Genetic algorithm, inspired by genetic science and the Darwin's natural selection principles, is based on the survival of superiors or natural selection. This algorithm is also derived from nature and based on the evolutionary principle of "Survival of the fittest". Although proposed after the evolutionary strategy algorithm, the most popular method is evolutionary algorithms. In a genetic algorithm, a population of individuals will survive in their desirability in the environment. People with superior abilities will find more chances of marriage and reproduction. So, after a few generations, better-off children will be created. In the genetic algorithm, each person in the population, as a chromosome, will become more perfect over generations. In each generation, chromosomes are evaluated, and in proportion to their value, they can survive and multiply. Generation production is discussed in the genetic algorithm with cross over and mutation functions. Superior parents are selected on the basis of a fitness function. Likewise, superior chromosomes are selected by considering the amount of fitness to be presented in the production process of the new generation. More fitted chromosomes have a higher chance of getting chosen with methods such as roulettewheeland tournament selection. The new population is produced regarding the elite preservation strategy, and the evolution process continues until the exit condition is fulfilled.

3.2.3. Objective function

The objective function of the optimization problem is defined as the result of dividing the value of RMSE by the Nash coefficient value (CE). Thus, by subtracting the numerator RMSE and adding its denominator (CE coefficient), the objective function will tend toward its minimum value, and the performance of the support vector model (SVM) will change for more accurate estimation. The decision variables of this objective are the SVM parameters

whose values will be determined optimally by the genetic algorithm (GA). The objective function is presented in the following equation:

$$\text{minimize } f = \frac{\sqrt{\sum_{i=1}^{n}\left(O_i - P_i\right)^2}}{1 - \frac{\sum_{i=1}^{n}\left(O_i - P_i\right)^2}{\sum_{i=1}^{n}\left(O_i - \bar{O}\right)^2}} \quad (11)$$

In the above equation, O_i and P_i are respectively the observed and the estimated values of evapotranspiration, $\bar{O}$ is the mean of the observed values, and n is the number of data.

3.3. Artificial Neural Network (ANN) with Levenberg-Marquardt training

Artificial neural network (ANN) is a distributed information processing system whose performance characteristics are similar to the neural structure of the human brain and its expansion is based on the following rules:

- Processing information is done in single elements called nodes, and these nodes are arranged in bunches called layers.
- Signals are transmitted between the nodes through the connections.
- Each connection has a weight that indicates its binding force.
- Each node converts a series of the weighing of signals to an output signal by applying a nonlinear transformation, which is called the activation function [29]. In other words, 90% of the neural network models used in water resources problems are of the feed forward neural networks among which the most important is the MLP model with Levenberg-Marquardt (LM) training algorithm. The typical architecture of this network consists of three sections: input layer, hidden layer, and output layer. The number of neurons in the input and output layers depends on the type of problem, but the number of hidden layer nodes is calculated by trial and error.

In the feed- forward neural networks, there are connections between nodes in different layers. Moreover, the inputs are displayed in the network input layer, and the stimulation of the input starts from the input to the output. Fig. 2 illustrates the overall structure of the artificial neural network.

In Fig. 2, X (x_1,x_2,…x_n) is the input vector, W_{ij} is the connection weight of the i^{th} node of the I^{th} node in the next layer, b_j is the bias of the j^{th} node, yi is the output of each layer, and y is the final output of the ANN. Also, f is the stimulus function of each neuron that maps the input information of the previous layer's neurons to the next layer's. In feed forward neural networks, there are interconnections between nodes in different layers. Moreover, the inputs are displayed in the network input layer and the stimulation.

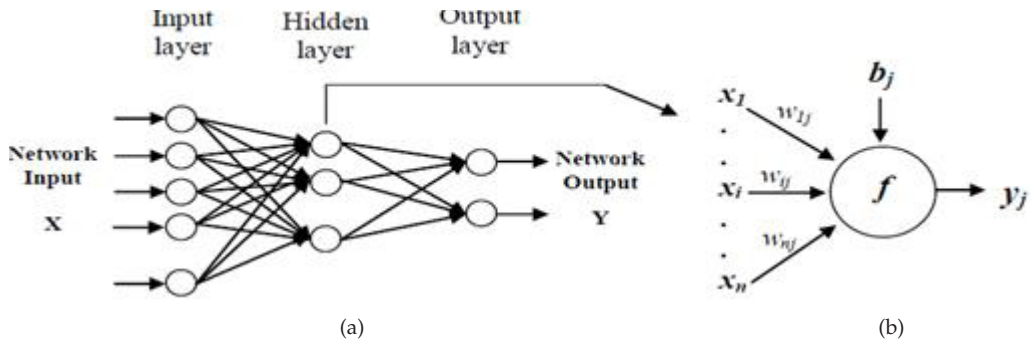
3.4. Input variables of models

In this study, according to input data, including average temperature, solar radiation, wind speed, and relative humidity per month, different models were offered to select four different patterns. Pattern 1 includes minimum and maximum air temperature (Model 1 = f(T_{min}, T_{max}) in the same month, pattern 2 includes average air temperature and average relative humidity (Model 2 = f(T_{mean}, Rh_{mean})), and wind speed in the same month, pattern 3 includes average air temperature, average relative humidity and R_s (Model 3 = f(T_{mean}, Rh_{mean}, R_s)) in the same month, pattern 4 includes average air temperature, relative humidity, wind speed, and, R_s, Model 4 = f(T_{mean}, Rh_{mean}, R_s, U_2) in the same month. Here T: air temperature (°C), U: wind speed (km/month), R_s: solar radiation and Rh: relative humidity (%).

To model the evapotranspiration with ANN and ESVM, MATLAB software was used. The models were trained up to reach a determined value of square error. After reaching goal error, the training was stopped and models were tested. Thus, we tried to cope with overfitting of the applied models. Various numbers of control parameters (The cross over rates from 0.65 to 0.8 and the mutation rates between 0.02 and 0.05) were tried using GA to get optimal ESVM models. Two-thirds of data were used for training the models and the remaining one-third was used to test the models.

3.5. Evaluation criteria

The statistical measures including the coefficient of determination (R^2) which expresses the coordination of data



(a) (b)

Fig. 2. (a) General overview of the three-layer feed forward neural networks; (b) the J^{th} node structure of the network.

predicted by the models and the computing data as well as the root mean square error (RMSE) that states the error value which gives more weight to large errors. They were used to check and evaluate the accuracy of the models. The mentioned parameters can be calculated using Eqs. (12) and (13)

$$R^2 = \left(\frac{\sum_{i=1}^{N} \left(T_i^O - \overline{T}^O\right)\left(T_i^E - \overline{T}^E\right)}{\sqrt{\sum_{i=1}^{N} \left(T_i^O - \overline{T}^O\right)^2 \sum_{i=1}^{N} \left(T_i^E - \overline{T}^E\right)^2}} \right)^2 \quad (12)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left(T_i^E - T_i^O\right)^2}{N}} \quad (13)$$

where T_i^O and T_i^E respectively are the observed and predicted values at the time step i, and $\overline{T}^O$ and $\overline{T}^E$ are the means of the observed and predicted values. N also represents the number of data.

Another criterion that used in this study is the coefficient of efficiency (CE). It is calculated by Eq. (14):

$$CE = 1 - \frac{\frac{1}{N}\sum_{1}^{N}(O-P)^2}{\frac{1}{N}\sum_{1}^{N}(O-\ddot{O})^2} \quad (14)$$

where CE: the coefficient of efficiency, N: the number of data points, O: observed value and P: predicted value. The range of the Nash–Sutcliffe efficiencies is from (–∞) to 1. An efficiency of 1 (CE = 1) means a perfect match of model estimates to the observed data.

The main evaluation criterion is selected as RMSE. By using this criterion, the best model was selected.

4. Results and discussion

To improve the performance of ESVM, numerous trial and errors were made to get the best value for C, γ and ε. The performance of the ESVM models in estimating lysimeter ET is given in Table 2. The values for the parameters C, γ, and ε of the optimum ESVM are also provided in this table. The number of population in the genetic algorithm for all ESVM models is 50. The cross over rate ranges from 0.65 to 0.8 and the mutation rate varies between 0.02 and 0.05 and the selection function is Roulette wheel. Fig. 3 shows the objective function for the genetic algorithm with 50 iterations. As can be seen from the figure, the ESVM model easily fit to data with a low number of iterations.

It is obvious from Table 2 that all the applied ESVM models can perform similarly in training and test stages, since the values of RMSE, R^2 and CE do not change significantly. During the testing period, it was observed that the ESVM2 model was better than the other modes for lysimeter ET estimation. Based on the performance statistics, ESVM2, whose input combination is T_{mean}, Rh_{mean}, has the smallest value of the RMSE (0.766 mm/day). It also has higher values of the CE (0.923) and R^2 (0.929) than the other models in the testing period. Therefore, it was selected as the best fit model for estimating the lysimeter ET. ESVM4 model whose inputs include (T_{mean}, Rh_{mean}, R_s, U_2) with RMSE of 0.780 (mm/ day), CE (0.920) and R^2 of 0.942 which provides the second best lysimeter ET estimation. However, these two models are almost equal accuracy but the ESVM2 is preferred because it has only two inputs and simpler than the ESVM4.

The comparison of the ET values measured by the lysimeter and the values estimated by the ESVM2 and ESVM4 models are shown in the form of line graphs and scatter plots in Fig. 4.

The ET values which are estimated by the ESVM models are closely followed the corresponding measured lysimeter ET values and both have the same trend. There is a slight difference between these two models. Such consistency reveals that both two models show good estimation accuracy. In some cases, ESSVM4 performs better than the two-parameter model (ESSVM2).

For evaluating the ability of ESVM model relative to ANN model, four ANN models were developed using the same input combinations for ET modeling. The optimal architectures and the performance statistics of the ANN models are shown in Table 2. According to the testing period results, the ANN4 model with the input combination (T_{mean}, Rh_{mean}, Rs, U_2), which has the lowest RMSE (0.793 mm/day) and the highest R^2 (0.927) and the nearest value of CE to unit, performed the best. ANN3 model whose inputs are the T_{mean}, Rh_{mean}, R_s, with a RMSE of 0.861 mm/day, R^2 of 0.9116 and CE of 0.873 was also ranked as the second best in ET estimation. From the comparison of Tables 2 and 3, it can be seen that the best ESVM model (ESVM2) increased the accuracy of the best ANN model (ANN4) by 0.2, 3.4 and 2%with respect to R^2, RMSE and CE, respectively.

The comparison result of the performance criteria given in Table 2 and 3 showed that all the ESVM models could perform better than the corresponding ANN models in estimating lysimeter ET. The comparison of the best ANN model (ANN4) with the lysimeter ET values is also represented in Fig. 5 in the form of scatterplot and time variation

Table 2
The best structures and the performance statistics of the ESVM models during training and testing periods

Models	Input	Parameter			Training period			Testing period		
		γ	C	ε	R^2	RMSE	CE	R^2	RMSE	CE
ESVM1	(T_{min}, T_{max})	1.081	49.24	0.08	0.946	0.934	0.946	0.921	0.801	0.916
ESVM2	(T_{mean}, Rh_{mean})	2.01	28.43	0.164	0.951	0.894	0.950	0.929	0.766	0.923
ESVM3	(T_{mean}, Rh_{mean} , R_s)	3.99	13.24	0.199	0.948	0.913	0.948	0.916	0.864	0.902
ESVM4	(T_{mean}, Rh_{mean} , R_s, U_2)	0.335	61.75	0.106	0.951	0.898	0.950	0.942	0.780	0.920

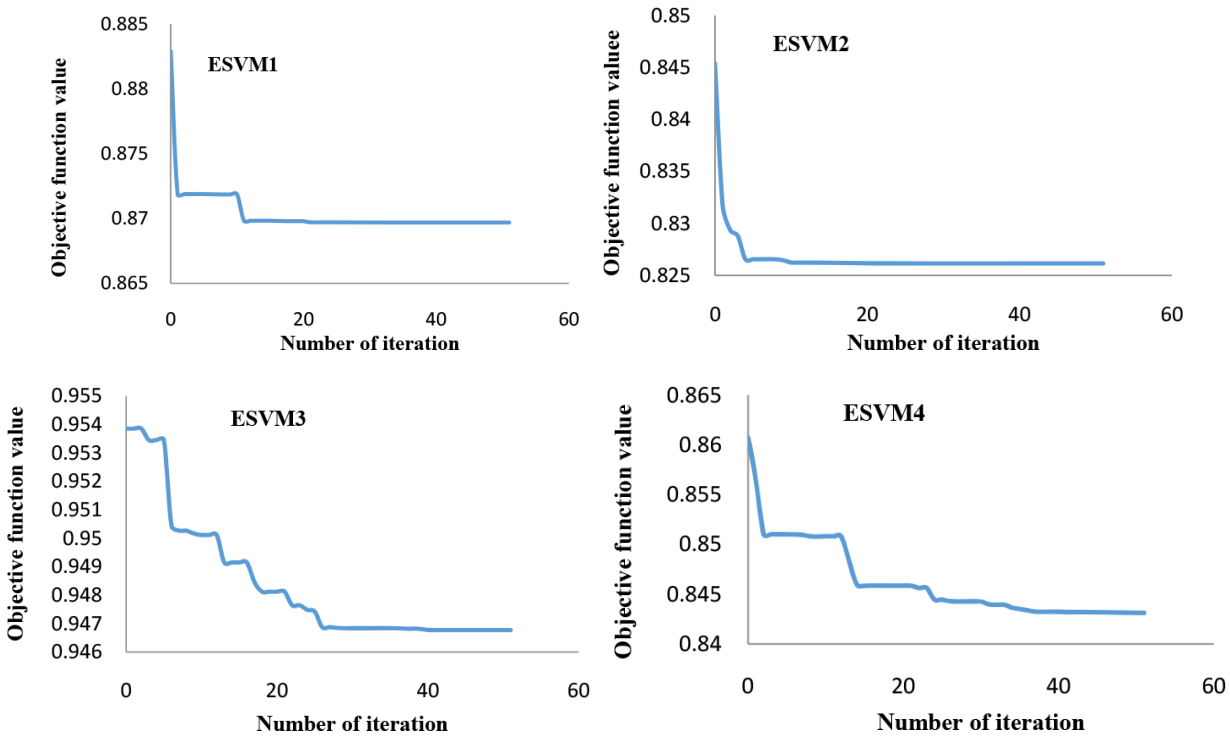


Fig. 3. Convergence trajectories of genetic algorithm.

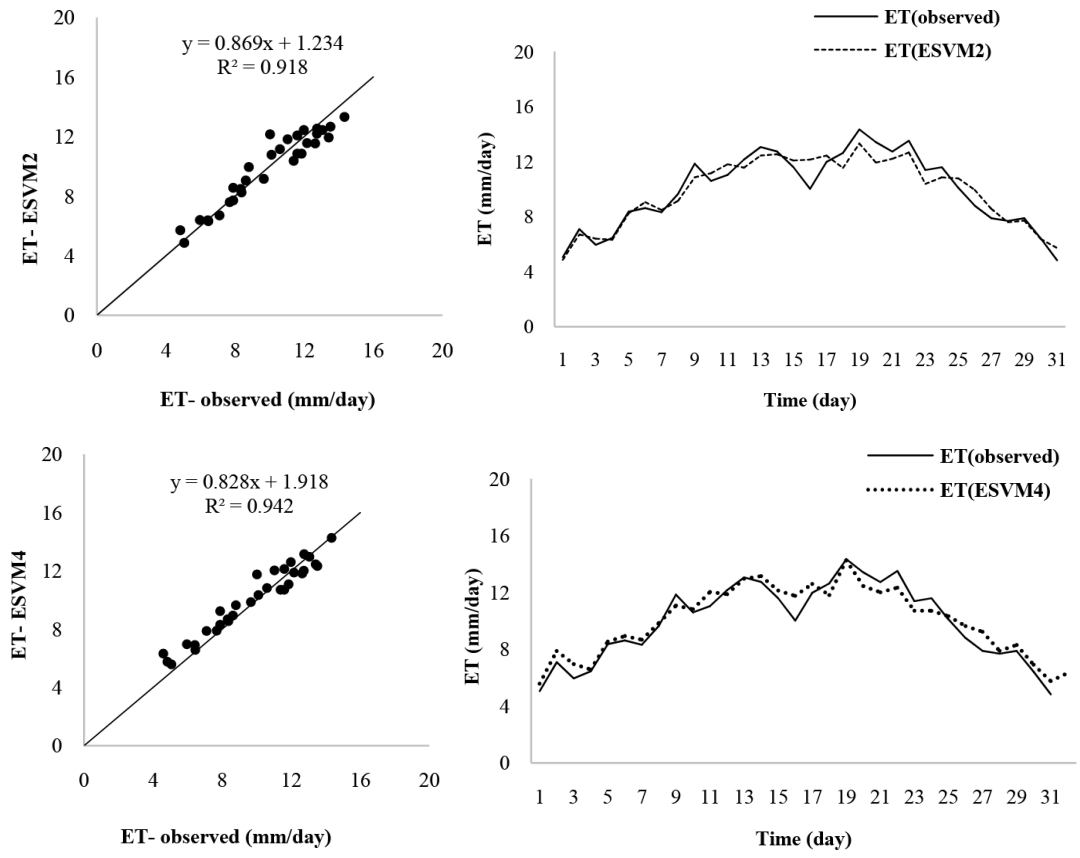


Fig. 4. Comparison of the ET values estimated by ESVM2, ESVM4 and lysimeter ET during testing period.

Table 3
The best structures and the performance statistics of the ANN models during training and testing periods

Models	Input	Input-hidden-output nodes	Training period			Testing period		
			R^2	RMSE	CE	R^2	RMSE	CE
ANN1	(T_{min}, T_{max})	2-5-1	0.830	1.121	0.801	0.824	1.119	0.784
ANN2	(T_{mean}, Rh_{mean})	3-4-1	0865	1.124	0.863	0.864	1.026	0.852
ANN3	(T_{mean}, Rh_{mean} , R_s)	4-2-1	0.909	0.872	0.868	0.912	0.861	0.873
ANN4	(T_{mean}, Rh_{mean} , R_s, U_2)	3-3-1	0.921	0.826	0.913	0.927	0.793	0.906

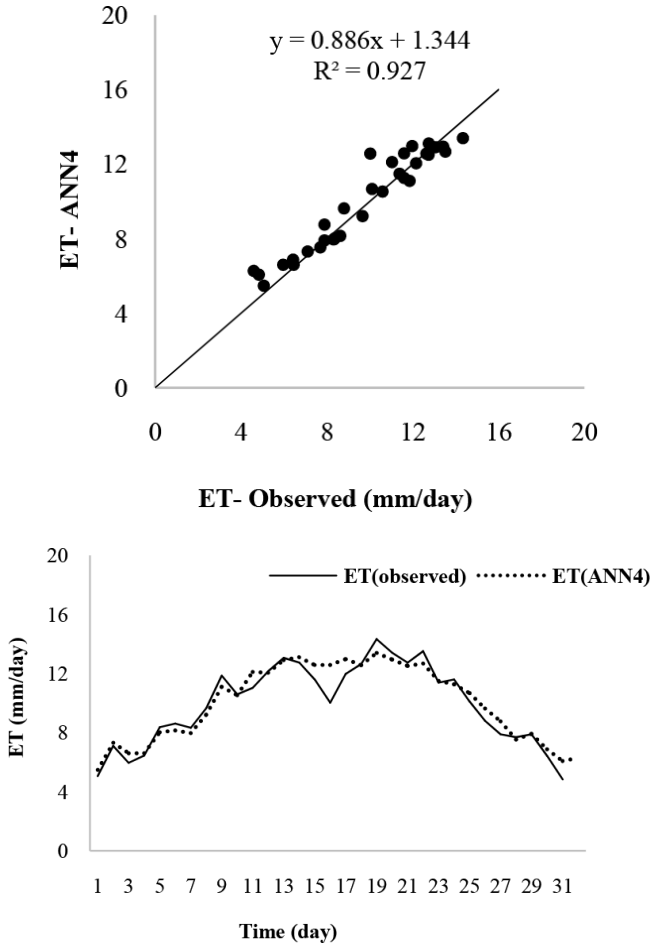


Fig. 5. Comparison of the ET values estimated by the best ANN model and lysimeter ET during testing period.

graph. Figs. 4-5 reveals that the ANN model has more scattered estimates and inferior results compared to the ESVM models.

5. Conclusion

In this study, usability of evolutionary SVM method in estimation of evapotranspiration amount has been studied. Four models were developed using different combinations of six daily climatic data including T_{mean}, T_{max}, T_{min}, Rh_{mean}, R_s and U_2. The ESVM models were tested using the ET observed by drainage lysimeter. The results demonstrated that the ESVM could be successfully applied to establish accurate and reliable lysimeter ET modeling. Particularly, ESVM model whose inputs included T_{mean}, Rh_{mean} provided good ET estimate. This is very important especially for the developing areas where reliable weather data sets are limited. The ESVM models were also compared with those of the ANN models using same input combinations. Based on the comparison, the ESVM models gave more accurate results than the ANN models in estimation of lysimeter ET.

References

[1] G. Rana, N. Katerji, M. Mastrorilli, Environmental and soil-plant parameters for modelling actual crop evapotranspiration under water stress conditions, Ecol. Model, 101 (1997) 363–371.
[2] R.G. Allen, L.S. Pereira, D. Raes, M. Smith, Crop Evapotranspiration: Guidelines for computing crop water requirements. Rome, Italy, FAO Irrigation Drainage Paper, (1998) No. 56, 300p.
[3] J. Doorenbos, W.O. Pruitt, Crop Water Requirements, Irrigation Drainage Paper, No. 24. FAO. (1977). Rome, Italy.
[4] M. Kumar, N. Raghuwanshi, R. Singh, W. Wallender, W. Pruitt, Estimating evapotranspiration using artificial neural network, J. Irrig. Drain. E, 128(4) (2002) 224–233.
[5] A.J. Cannon, P.H. Whitfield, Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models, J. Hydrol., 259 (2002) 136–151.
[6] O. Kişi, Evapotranspiration estimation using feed-forward neural networks, J. Hydrol. Res., 37(3) (2006) 247–260.
[7] S. Zanetti, E. Sousa, V. Oliveira, F. Almeida, S. Bernardo, Estimating evapotranspiration using artificial neural network and minimum climatological data, J. Irrig. Drain. E., 133(2) (2007) 83–89.
[8] A.J. Adeloye, R. Rustum, I.D. Kariyama, Neural computing modeling of the reference crop evapotranspiration, Environ. Model. Softw., 29(1) (2012) 61–73.
[9] Ladlani, L. Houichi, L. Djemili, H. Salim, B. Khaled, Modeling daily reference evapotranspiration (ET0) in the north of Algeria using generalized regression neural networks (GRNN) and radial basis function neural networks (RBFNN): a comparative study, Meteorol. Atmos. Phys., 118 (2012) 163–178.
[10] O.O. Aladenola, C.A. Madramootoo, Evaluation of solar radiation estimation methods for reference evapotranspiration estimation in Canada, Theor. Appl. Climatol., 118(3) (2014) 377–385.
[11] O. Kisi, M. Cimen, Evapotranspiration modelling using support vector machines, Hydrolog. Sci. J., 54(5) (2009) 918–928.
[12] X. Wen, J. Si , Z. He, J. Wu, H. Shao, H. Yu, Support-vector-machine-based models or modeling daily reference evapotranspiration with limited climatic data in extreme arid regions, Water Resour Manag., 29 (2015) 3195–3209.
[13] O. Baydaroğlu, K. Koçak, K. Duran, River flow prediction using hybrid models of support vector regression with the wavelet

transform, singular spectrum analysis and chaotic approach, Meteorol. Atmos. Phys., (2017).
[14] S. Trajkovic, M. Stankovic, B. Todorovic, Estimation of FAO Blaney-Criddle b Factor by RBF Network, J. Irrig. Drain. E, ASCE, 126(4) (2000) 268–27.
[15] J. Shiri, P. Marti, A.H. Nazemi, A.A. Sadraddini, O. Kisi, G. Landeras, A. Fakheri Fard, Local vs. external training of neuro-fuzzy and neural networks models for estimating reference evapotranspiration assessed through k-fold testing, J. Hydrol. Res., 46(1) (2015) 72–88.
[16] P. Marti, P. González-Altozano, R. López-Urrea, L.A. Mancha, J. Shiri, Modeling reference evapotranspiration with calculated targets: Assessment and implications, Agric. Water Manag., 149 (2015) 81–90.
[17] J. Shiri, A.A Sadraddini,A.H. Nazemi, P. Marti, A. Fakheri Fard, O. Kisi, G. Landeras, Independent testing for assessing the calibration of the Hargreaves–Samani equation: New heuristic alternatives for Iran, Comput. Electron. Agr., 117 (2015) 70–80.
[18] J. Shiri, A.H. Nazemi, A.A. Sadraddini, G. Landeras, O. Kisi, A. Fakheri Fard, P. Marti, Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran, Comput. Electron. Agr., 108 (2014) 230–241.
[19] S. Kim, J. Shiri, V.P. Singh, O. Kisi, G. Landeras, Predicting daily pan evaporation by soft computing models with limited climatic data, Hydrolog. Sci. J., 60(6) (2015) 1120–1136.
[20] G. Landeras, E. Bekoe, J. Ampofo, F. Logah, M. Diop, M. Cisse, J. Shiri, New alternatives for reference evapotranspiration estimation in West Africa using limited weather data and ancillary data supply strategies, Theor. Appl. Climatol., (2017) DOI: 10.1007/s00704-017-2120-y.
[21] J. Shiri, Evaluation of FAO56-PM, empirical, semi-empirical and gene expression programming approaches for estimating daily reference evapotranspiration in hyper-arid regions of Iran, Agric. Water Manag., 188 (2017) 101–114.
[22] J. Shiri, P. Marti, V.P. Singh, Evaluation of gene expression programming approaches for estimating daily evaporation through spatial and temporal data scanning, Hydrol. Process., 28(3) (2014) 1215–1225.
[23] R. Fletcher, (2013) Practical Methods of Optimization: John Wiley & Sons, Chichester, West Sussex England.
[24] B. Guo, S.R. Gunn, R.I. Damper, J.D. Nelso, Customizing kernel functions for SVM-based hyperspectral image classification, IEEE. Trans. Image. Process, 17(4) (2008) 622–629.
[25] T. Kavzoglu, I. Colkesen, A kernel functions analysis for support vector machines for land cover classification, Int. J. Appl. Earth Obs. Geoinf., 11(5) (2009) 352–359.
[26] D. Basak, S. Pal, D.C. Patranabis, Support vector regression, Natl. Westm. Bank Q R., 11(10) (2007) 203–224.
[27] G.Q. Liu, (2011) Comparison of regression and ARIMA models with neural network models to forecast the daily streamflow of White Clay Creek, Dissertation, University of Delaware.
[28] V.N. Vapnik, A.J. Chervonenkis, The necessary and sufficient conditions for consistency of the method of empirical risk, Pattern Recog. Image Anal., 1(3) (1991) 284–305.
[29] R.S Govindaraju, Artificial neural networks in hydrology, I: Preliminary concepts, J. hydro. Engin., 5(2) (2000) 115–123.
[30] C.M. Zealand, D.H. Burn, S.P. Simonovic, Short term streamflow forecasting using artificial neural networks, J. Hydro., 214(1) (1999) 32–48.