Contents lists available at SCCE

Journal of Soft Computing in Civil Engineering

Journal homepage: www.jsoftcivil.com

Reference Evapotranspiration Estimation Using ANN, LSSVM, and M5 Tree Models (Case Study: of Babolsar and Ramsar Regions, Iran)

Yashar Dadrasajirlou[1*], Hamidreza Ghazvinian[1], Salim Heddam[2], Mariam Ganji[3]

1. Faculty of Civil Engineering, Semnan University, Semnan, Iran
2. Professor, Faculty of Science, Agronomy Department, Hydraulics Division, Laboratory of Research in Biodiversity Interaction Ecosystem and Biotechnology, University 20 Août 1955, Skikda, Algeria
3. Faculty of Natural Resources and Environment, Islamic Azad University Science and Research Branch, Tehran, Iran

Corresponding author: Dadras_yashar@semnan.ac.ir

https://doi.org/10.22115/SCCE.2022.342290.1434

ARTICLE INFO

Article history:
Received: 14 May 2022
Revised: 02 October 2022
Accepted: 03 October 2022

Keywords:
Reference evapotranspiration;
LSSVM;
ANN;
M5 Tree;
Babolsar;
Ramsar.

ABSTRACT

Evapotranspiration is a non-linear and complex phenomenon requiring different climatic variables for accurate estimation. In this study, the performance of several artificial intelligence models in estimating the amount of monthly reference evapotranspiration was investigated. Babolsar and Ramsa regions located in the north of Iran were selected as case study models proposed in this study: artificial neural network (ANN), least square support vector machines (LSSVM), and M5 tree models. The data used in this study was gathered between 2009 till 2019 (11 consecutive years). In the present study, 70% of the data were used for the training stage, and 30% of the data were reserved for testing the proposed models. Models' performances were evaluated using several evaluation criteria, i.e., the coefficient of determination (R2), the root mean square error (RMSE), and the mean absolute error (MAE). The results for Babolsar and Ramsar stations showed that all three models have a relatively good performance in estimating the rate of reference evapotranspiration. However, the LSSVM model performed better than the other models. The R^2, MAE, and RMSE for the LSSVM model in the test stage were 0.982, 0.366 mm, 0.425 mm, 0.937, 0.018 mm, and 0.350 mm for Babolsar and Ramsar stations, respectively.

How to cite this article: Dadrasajirlou Y, Ghazvinian H, Heddam S, Ganji M. Reference evapotranspiration estimation using ANN, LSSVM, and M5 tree models (case study: of Babolsar and Ramsar regions, Iran). J Soft Comput Civ Eng 2022;6(3):101–118. https://doi.org/10.22115/scce.2022.342290.1434

1. Introduction

Reference evapotranspiration (ET0) is a variable that is used in irrigation planning, water resources management and hydrological studies, and several other applications, especially for estimating the water demand of crops in large irrigation areas [1]. ET0 is a non-linear and complex phenomenon that requires different climatic variables [2]. Factors affecting ET0 include relative humidity, solar radiation, wind speed, temperature, water solute concentration, and atmospheric pressure. With decreasing relative humidity, the concentration of solutes in water, atmospheric pressure, and increasing solar radiation, wind speed, and temperature, the rate of ET0 increases. The ET0 rate is measured based on height, and its unit is millimeters, centimeters, or inches [3–6]. The simplest way to measure evaporation is to use trays placed on the ground or floating on water [7,8]. Calculating actual evapotranspiration is a difficult task and requires extensive research. Therefore, evapotranspiration studies are raised in two methods: a) estimation of actual evapotranspiration (the amount of water that evaporates under the current conditions) and b) estimation of potential evapotranspiration (i.e., the amount of water that evaporates, if resources are limited No water) [9]. Excessive estimation of plant water consumption leads to wastage of irrigation water, wetting lands, and contamination of groundwater resources. Also, a lower estimation of required water will cause drought stress on the plant and reduce yield [10,11]. Researchers have conducted many studies to obtain evapotranspiration.

Many researchers worldwide have estimated reference evapotranspiration in different climates under different conditions and scenarios. They have evaluated various experimental and intelligent models and methods with minimum and maximum input data. This is due to the importance of evapotranspiration in the water cycle and related issues.

SamadianFard and Panahi [12] estimated daily ET0 using support vector regression (SVR) and M5Tree models and compared their results with Hargreaves and Torrent White experimental models. Their study used meteorological variables from 1971 to 1994 in Tabriz synoptic station. In addition, the Penman-Monteith-FAO method's output was considered the basis. Then 17 scenarios made of one to six meteorological data were examined. The results showed that M5Tree and SVR models, considering all meteorological variables, showed better results in estimating than Hargreaves and Torrent White model. Goodarzi et al. (2015) [13] calculate the rate of evapotranspiration concerning climate change (B1, A2, A1B, and B2) in the watershed of Lake Urmia using statistical microscale models LARS-WG and SDSM and model output HadCM3 public circulation covered the next three time periods (2030-2011, 2065-2046 and 2099-2080). The rate of evapotranspiration was calculated at a monthly and seasonally time scale using Priestley-Taylor and Hargreaves-Samani methods. The results showed that, on average, in the long run at the basin level, the minimum temperature will increase between 0.2 to 3.4 degrees, and the maximum temperature varies between 0.9 to 2.9 degrees in future periods compared to the base period (1990-1961). Also, under the influence of temperature, the estimated evapotranspiration rate increases in all months and seasons in future periods.

Mohtarami et al. (2015) [14] improved the accuracy of the Hargreaves-Samani method in estimating ET0 using the correction factor using the ANN model and the M5Tree between 2004

and 2013 at Farkushhar station and Shahrekord airport. The results of this study indicated that the ANN and the M5Tree models have good performance in correction factor modeling. The ANN model, on the other hand, performed more accurately. The accuracy of the Hargreaves-Samani model before using the correction coefficient compared to the Penman-Monteith-FAO method had an RMSE equal to 0.900; the value after applying the correction coefficient reached RMSE of 0.69 and RMSE of 0.72 using ANN and M5Tree models respectively. The results show that the performance of the Hargreaves-Samani method has improved after applying correction coefficients. Zortipour et al. (2017) [15] simulated and compared potential evapotranspiration daily using ANN, ANFIS (Adaptive Neuro Fuzzy Inference Systems), and M5 decision tree methods at Shiraz synoptic station. ANFIS and ANN models had acceptable accuracy due to the high correlation coefficient (R) and low error rate. Also, considering that the values of R, RMSE, and MAE for the M5 model were estimated to be 0.7170, 0.1088 mm, and 0.0877 mm, the results indicate good performance of the M5 Tree model in evaluating ET0.

The primary purpose of Ferreira et al. (2019) [1] study was to calculate daily ET0 with limited data. They sought to investigate the performance of the MARS model when confronted with limited data. After comparing the results of the Penman-Monteith equation and the MARS model, he observed that the intelligent model performed better. Meanwhile, solar radiation scenario, relative humidity, and wind speed were more effective in estimating ET0. Temperature-based procedures had the highest performance. Using climatic data, Wu et al. (2019) [2] evaluated the ET0. The results showed that the M5Tree model can successfully estimate ET0.

Babu and Thomas (2022) [16] modeled evaporation data from pans in South India from 2000 to 2019 with the help of three models: decision tree regression, random forest and regression and Gradient Boosting Regression Trees. The input data in the models included dry bulb temperature, wet bulb temperature, maximum and minimum air temperature, vapor pressure, relative humidity, wind speed, wind direction, sunshine hours, and rainfall. The random forest model had the best performance.

Batra and Gandhi (2022) [17] in a study on pan evaporation for Rajendranagar, Hyderabad, a part of India with the help of maximum and minimum temperature data, wind speed, relative humidity, precipitation and sunshine hours for the year 1993 Until 2008, different ANN models were used, and the least error was reported for the MLP model.

Considering that relatively limited studies have been done in the field of evaporation and transpiration with the ANN, LSSVM, M5 tree model, in the northern region of Iran, In the present study, three data mining methods namely M5 tree, ANN and LSSVM in Babolsar and Ramsarcites have been used to estimate the evapotranspiration of plants.

Studies related to the combination of artificial intelligence and evaporation studies in these areas have not been combined with each other, and the innovation of this research is to combine these studies. In this research, first, the meteorological data of the study stations were collected. After checking the data, they were separated for the testing and training stages and then they were simulated by intelligent algorithms considering different scenarios. In the following, using the error criteria, a comparison was made between the performance of the algorithms to evaluate

their simulation capability in evaporation. Finally, the obtained results were expressed and suggestions for future researches were mentioned.

2. Methods

2.1. Study area and available data

Mazandaran province is affected by latitude, Alborz Mountains, altitude, proximity to the sea, local and regional winds, displacement of northern and western air masses, and even dense forest cover. Mazandaran region has a unique climatic diversity. Figure 1 shows the geographical location of Mazandaran province, and Table 1 shows the location of the studied stations.

Two major currents play an important role in the climate of Mazandaran province: one is the north and northeast air current that travels from Siberia and the North Pole to the south and southwest, causing cold weather, frost, and snow and rain [18]. This air mass has little effect on the climate of Mazandaran. The other is the westerly winds that cross the Mediterranean Sea and the Black Sea in winter, and after entering Iran, cause heavy and continuous rains [19]. In the summer months, the rainmaking power of these winds decreases and only increases the humidity and sultry weather, leading to unfavorable living conditions [7].

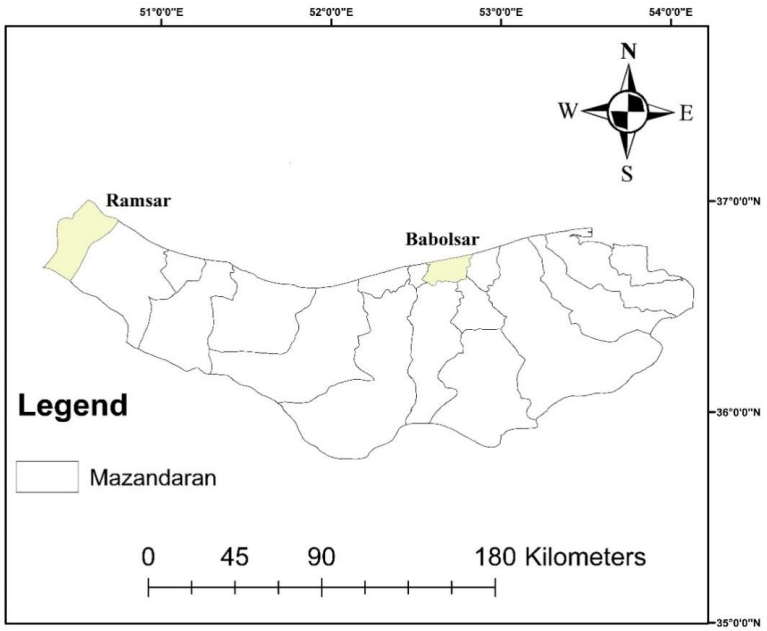


Fig. 1. Mazandaran Province [20].

Table 1
Synoptic stations of Mazandaran province.

Synoptic station	Longitude Degree Minute	Latitude Degree Minute	Altitude
Babolsar	52 48	36 25	16
Ramsar	53 00	36 20	12

In this study, the statistical data of two synoptic stations of Mazandaran province namely Babolsar and Ramsar have been collected in 11 consecutive years between 2009 and 2019 and used for modelling the ET_0. The meteorological data used are: air temperature, sunshine hours, relative humidity, solar radiation and wind speed. Tables 2 and 3 show the monthly average of some meteorological data of Babolsar and Ramsar stations To evaluate the simulated results, the statistical indices of coefficient of determination (R^2)[21], root mean square error (RMSE) [21] and mean absolute error (MAE)[8] were calculated [22–26]. The values of the mentioned indicators are calculated from one to three relations.

In Equations (1) to (3), O_i is the observed evaporation value in a month, P_i is the predicted evaporation value of the same month, $\overline{O}$ is the average of the observed evaporation values, and $\overline{P}$ is the similar average for the predicted values [27].

Table 2
Monthly average of Babolsar synoptic station.

Month	Average maximum temperature (°C)	Average minimum temperature (°C)	Average temperature (°C)	Monthly rainfall (mm)	Maximum rainfall in a day (mm)	Relative humidity (percentage)	Number of hours of sunshine	Maximum wind speed (meters per second)
April	18.2	11.1	14.2	91.4	28.2	73	134	15
May	24.1	14.1	21.1	34.1	13.2	71	165	19
June	29.1	20.5	22.1	60.3	21.2	72	247.4	5
July	31.2	33.5	25.8	38.3	23.2	80	233.2	7
August	33.0	21.9	28.1	30.3	10.3	71	248.3	7
September	32.4	20.1	23.3	61.2	51	72	222.7	34
October	26.4	17.1	22.2	123.1	35.6	70	172.1	13
November	10.1	11.1	13.2	61	22.4	81	91.7	16
December	12.2	3.9	8.3	57	25.5	83	134.2	17
Januray	12.1	1.9	7.4	3.8	3.1	75	129.1	10
Februrary	11.5	3.1	5.3	131.2	22.7	76	106.8	14
March	14.1	5.1	11.4	9	3.4	84	162.8	8

$$MAE = \frac{\sum_{i=1}^{N} |P_i - O_i|}{N} \tag{1}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (O_i - P_i)^2}{N}} \tag{2}$$

$$R^2 = \left(\frac{\sum_{i=1}^{N} (O_i - \overline{O})(P_i - \overline{P})}{\sqrt{\sum_{i=1}^{N} (O_i - \overline{O})^2 \sum_{i=1}^{N} (P_i - \overline{P})^2}} \right)^2 \tag{3}$$

2.2. Network configuration

In this study, the ratio of data used in the testing phase to the training phase was 70%/30%. Temperature, relative humidity, solar radiation, wind speed, and sunny hours were used as input for the models, and the output of the Penman-Monteith-FAO model was used as the objective function in the innovative models. In this study, intelligent models, namely ANN, M5Tree, and LSSVM models, were used. The experimental method used is the Penman-Monteith-FAO method (based on mass transfer). MATLAB software was used to calculate innovative models [28,29].

2.3. Artificial neural network model (ANN)

Artificial neural networks (ANNs) are one of the methods of artificial intelligence techniques that are inspired by the functional system of the human brain and cannot be compared with natural systems [30]. Conquering strategic principles is the approach of the computational model of ANNs, which is the basis of the brain process for answering questions and using them in computer systems. ANN models must be able to store knowledge in similar ways and process it accordingly [31]. The large number of simple computational units that can work together is called a parallel structure [24]. One of the features of ANN is that it is three-layered. This three-layer structure uses input, hidden, and output layers [7, 27]. The main application of this structure is to establish communication between the mentioned layers to explain the relationships between input and output data [28]. This structure is introduced by considering three components (i, j, k), where i, j, and k represent the nodes of the input layer, the number of hidden layers, and the number of output layers, respectively (figure 2) [32].

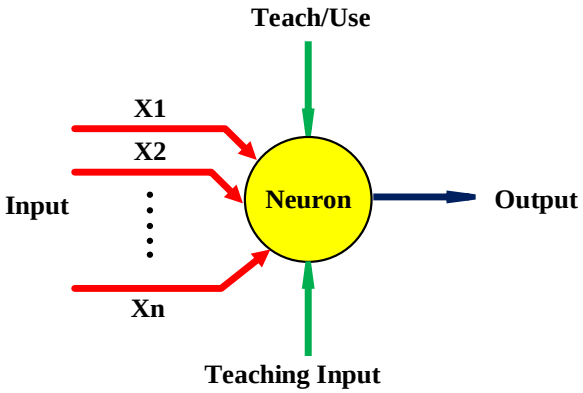


Fig. 2. Schematic structure of artificial neural network model.

2.4. M5 decision tree model

In 1992, Quinlan was the first researcher to propose the idea of the M5 tree. He based his method presented in the middle based on three factors: binary decision making, linear regression functions, and communication between input and output data [29]. Another reason why researchers welcome the use of this intelligent method is its multidimensional space and sub-layer [33]. The M5 can learn well and take on substantial tasks [34]. The M5 tree model is created in two stages: the growth stage and tree pruning. In the tree growth stage, also known as the division stage, the input space will be divided into several classes using linear regression

models that minimize errors between predicted and actual values. Finally, a decision tree is constructed using the information obtained [29]. In the M5 tree model, the division of the idea follows the decision tree, but with the difference that it can be used in quantitative data [33]. Standard deviation is used to select the best feature to divide the data set in each node [34]. The M5 is obtained using a common deviation reduction calculated according to equation (4):

$$SDR = sd(E) - \sum_{i} \frac{|Ei|}{|E|} sd(Ei) \tag{4}$$

In above equation, E is a set of instances that reach the node, and E_i is a subset of the input data to the parent node. These steps are completed until a suitable tree structure is formed. The way to deal with this problem is to prune the tree in the previous step.

2.5. Least square support vector machine (LSSVM)

SVM is an efficient learning system based on the theory of constrained optimization. This model uses the inductive principle of structural error depreciation, leading to an optimal solution. Unlike SVM, which uses quadratic programming to solve problems, LSSVM uses linear equations to solve problems. For this reason, it has higher computational accuracy than the SVM models [35]. The following regression equation is used in the LSSVM model to estimate various problems (Equation 5) [35–37]:

$$Y\left(X_i\right) = W^T.\Phi(X_i) + b \tag{5}$$

In this equation, $\Phi(X_i)$is called nonlinear drawing of inputs in the feature space with high dimensions, and b and W are the deviation of the regression function and the values of the weights, respectively, which are calculated using the minimization of the objective function according to equation (6):

$$\min_{w,e,b} j(w,e) = \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{i=1}^{N} e_i^2 \tag{6}$$

With considering the equation (7) as limits for above equation:

$$y_i = w^T \Phi(x_i) + b + e_i \quad \text{i=1,2,3,...,N} \tag{7}$$

In the above relations, e_i is the error of the training data and x is the parameter regulating the error section. Finally, the LSSVM model estimation function is defined as Equation (8):

$$y(x) = \sum_{i=1}^{N} a_i K(x_i, x_j) + b \tag{8}$$

In the above relation, x is a kernel function and is expressed as a function with internal multiplication in the property space according to the equation (9):

$$K(x_i, x_j) = \Phi(x_i)\Phi(x_j) \quad \text{i,j=1,2,3,...,N} \tag{9}$$

Compared to conventional models such as neural networks (ANN) and partial least squares regression, LSSVM is more preferred for signal processing, pattern recognition, and nonlinear regression estimation because it takes less time to compute [38].

2.6. Penman-monteith-FAO model

The Penman-Monteith-FAO model is the reference model for calculating the evapotranspiration rate of the reference plant, which was introduced by the FAO in 1998 in FAO Journal No. 24 Irrigation and Drainage, and is considered by many researchers around the world [39]. The reference plant is grass with a height of 8 to 15 cm, radiation coefficient (albedo) of about 0.23, shading resistance is 70 seconds per meter [39,40].

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma[\frac{900}{(T+273}]U_2(e_a - e_d)}{\Delta + \gamma(1 + 0.34U_2)} \quad (10)$$

In the equation 10, ET_0 is Evapotranspiration of reference plant (mm d-1), T is air temperature ($^{\circ}$c), U_2 stands for wind speed at a height of 2 meters (ms-1), R_n refers to the net surface radiation (MJ m-2 d-1), e_a-e_d is the lack of saturated vapor pressure (Kpa), Δ stands for the Slope of saturated steam pressure curve with temperature (KpaC-1), and γ is the Constant psychrometers (KpaC-1).

3. Results

For Babolsar station, model M5 recorded an average R^2 of 0.9069 and a coefficient of variation of 0.0538. In addition, the average MAE and RMSE for the Babolsar station were 0.0411 mm and 0.3133 mm, respectively, and the coefficient of variation was 0.9546 and 0.1548. The LSSVM model recorded an average R^2 of 0.9821 and a coefficient of variation of 0.0051. In addition, the average MAE and RMSE for this station were 0.0425 mm and 0.3366 mm, respectively, and the coefficients of variation were 0.0745 and 0.8121, respectively.

The ANN model recorded an average R^2 of 0.9254 and a coefficient of variation of 0.0236. Also, the average MAE and RMSE for this station were 0.0658 mm and 0.3432 mm, respectively, and the coefficients of variation were 0.7232 and 0.0432, respectively. Among these, the M5 model had the lowest average for R^2, and the LSSVM model had the highest standard for R^2.

Table 3 shows the modeling results for Babolsar station with the three mentioned models for the test phase. Figures 3 to 5 compare experimental and simulation results of reference transpiration evaporation for Babolsar station.

Table 3
Babolsar station modeling results.

Model	R^2	MAE (mm)	RMSE (mm)
M5	0.9069	0.0411	0.3313
LSSVM	0.9821	0.0425	0.3366
ANN	0.9254	0.0658	0.3432

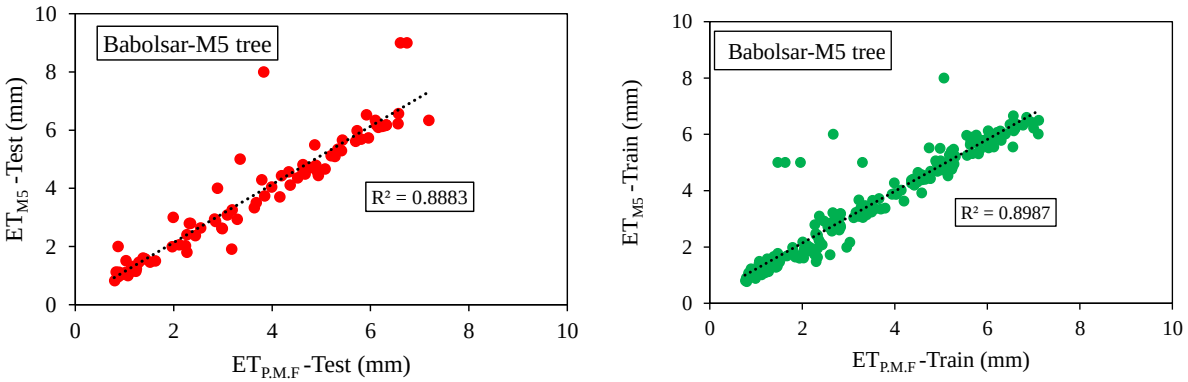


Fig. 3. R^2 for M5 model for training and test data in Babolsar station.

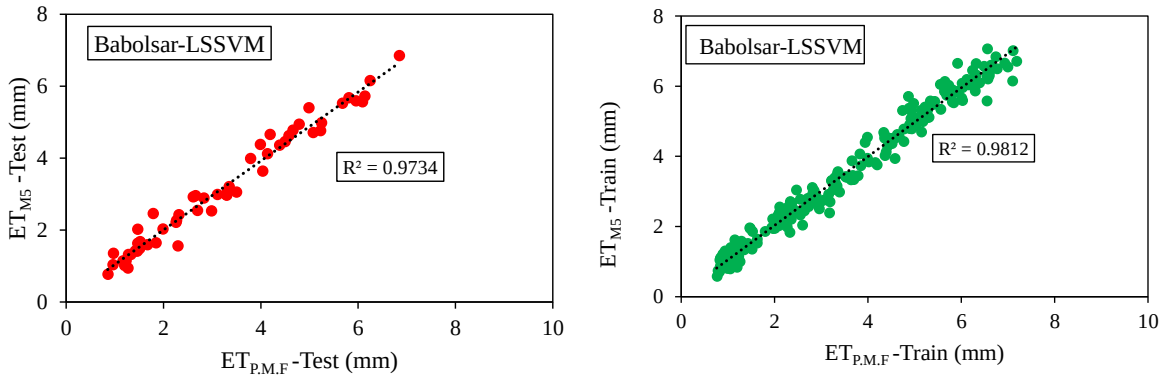


Fig. 4. R^2 for LSSVM model for training and test data in Babolsar station.

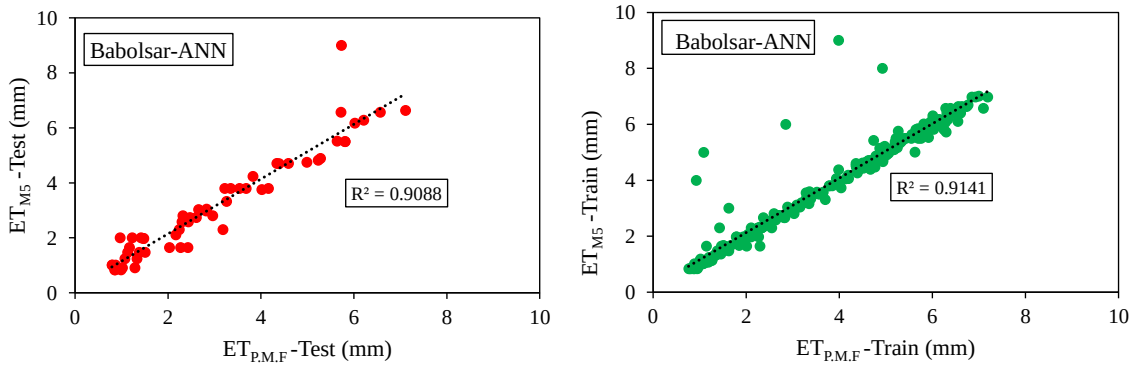


Fig. 5. R^2 for ANN model for training and test data in Babolsar station.

For Ramsar station, M5 recorded an average R^2 of 0.9023. In addition, the average MAE and RMSE for Ramsar stations were recorded as 0.0321 mm and 0.4535 mm, respectively. The LSSVM model recorded an average $R^2$of 0.9737. In addition, the average MAE and RMSE for this station were recorded as 0.0318 mm and 0.3504 mm, respectively. The ANN model recorded an average R^2 of 0.9476. In addition, this station's average MAE and RMSE were recorded as 0.0672 mm and 0.4543 mm, respectively. Among these, the M5 model had the lowest average for R^2, and the LSSVM model had the highest standard for R^2. Table 4 shows the modeling results for the Ramsar station with the three models for the test phase. Figures 6 to 8 show a comparison diagram of the experimental results and the simulation results of the reference transpiration evaporation for the Ramsar station.

Table 4
Ramsar station modeling results.

Model	R^2	MAE (mm)	RMSE (mm)
M5	0.9023	0.0321	0.4535
LSSVM	0.9657	0.0318	0.3504
ANN	0.9476	0.0672	0.4543

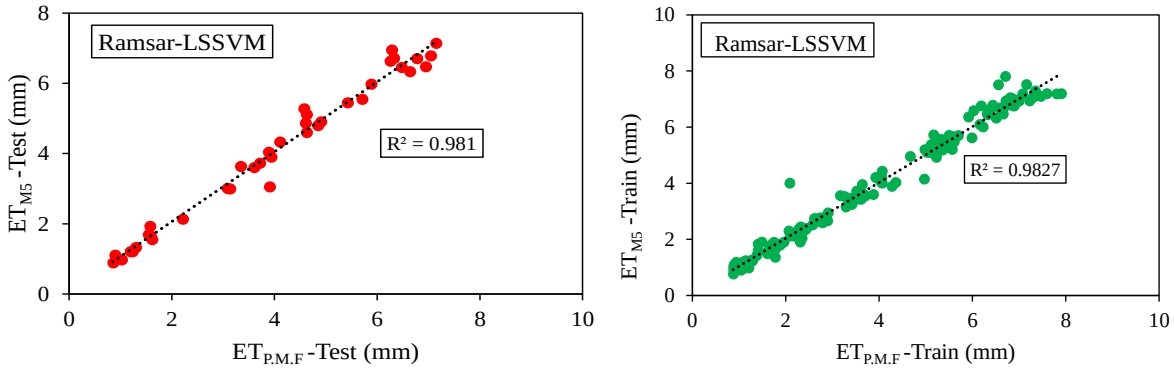


Fig. 6. R^2 for LSSVM model for training and test data in Ramsar station.

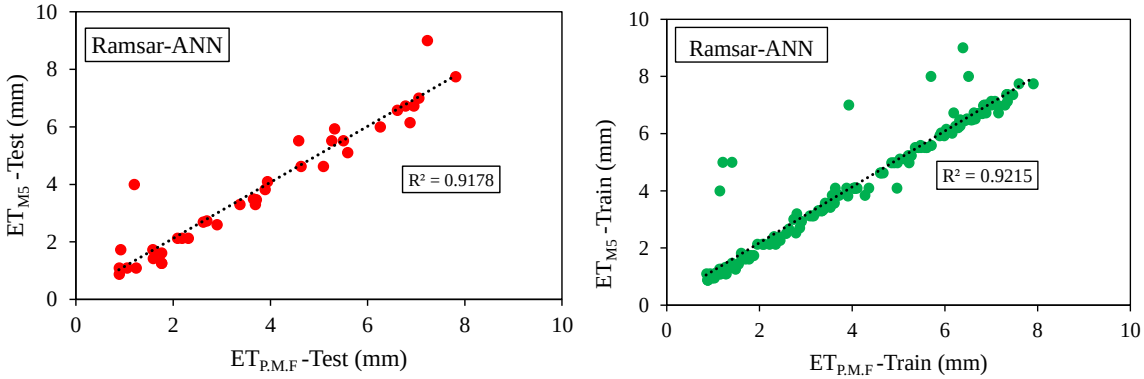


Fig. 7. R^2 for ANN model for training and test data in Ramsar station.

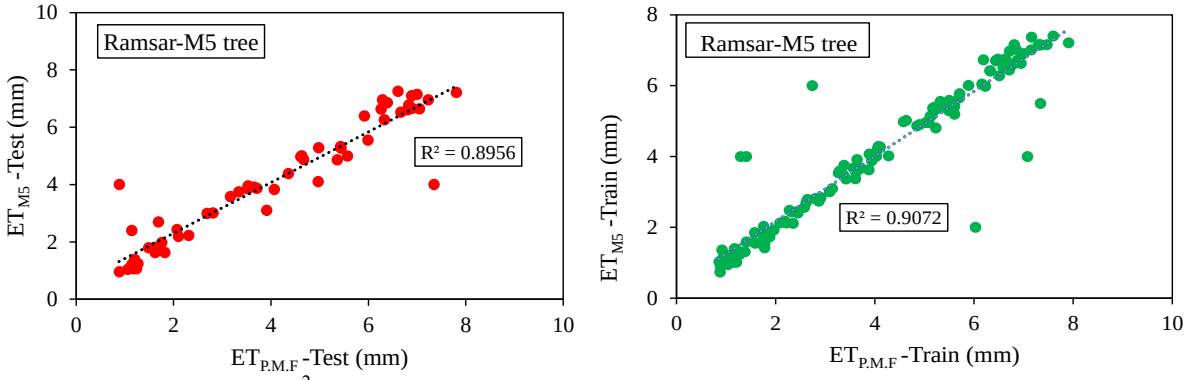


Fig. 8. R^2 for M5 model for training and test data in Ramsar station.

Also, Figures 9 and 10 give the scatter plot or scatter curve of the met values versus the Penman-Monteith-FAO (observational) values for the test step. These figures show that all three models perform relatively well in modeling, and the LSSVM model performs better than the other two models.

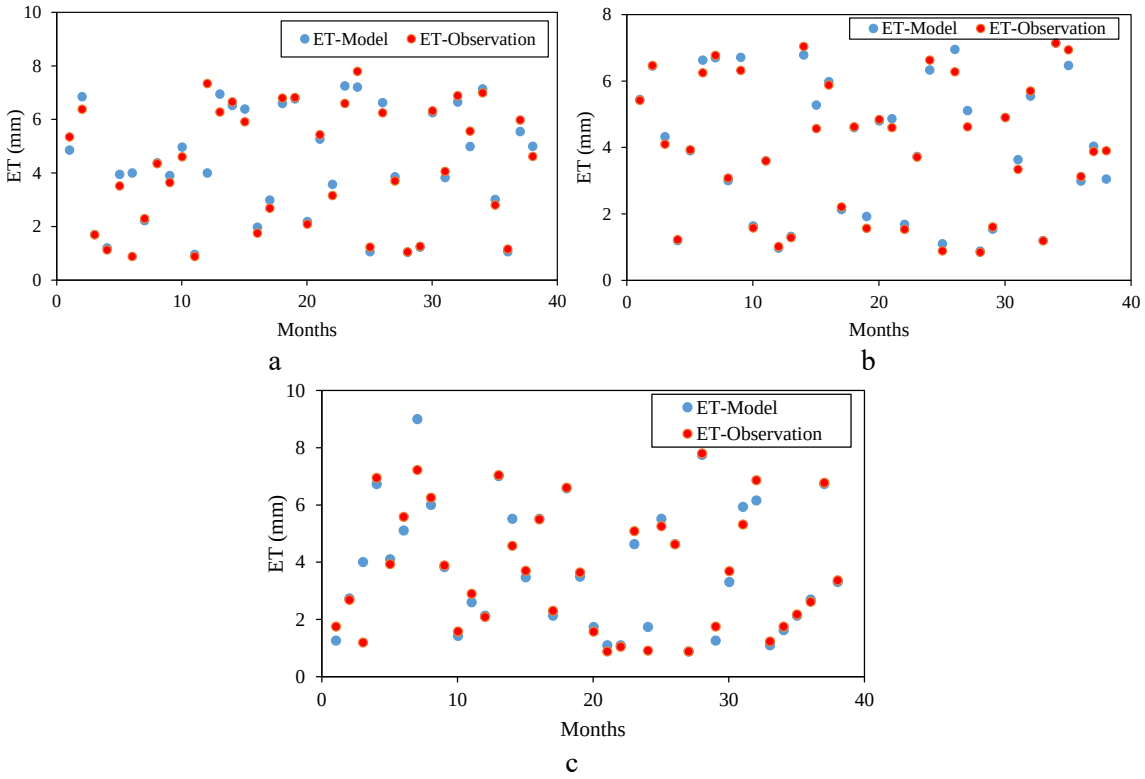


Fig. 9. Scattering of estimated data with observational data in the test phase for Ramsar, a) M5 TREE, b) LSSVM and c) ANN.

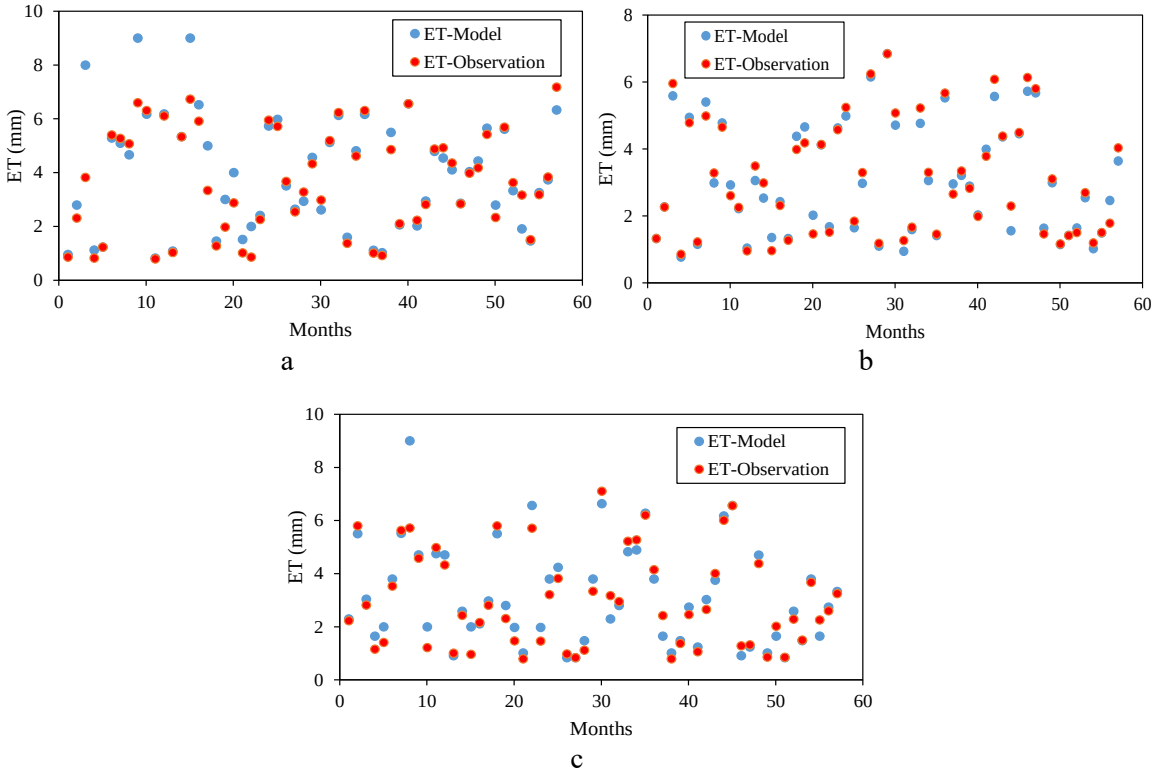


Fig. 10. Scattering of estimated data with observational data in the test phase for Babolsar, a) M5 TREE, b) LSSVM and c) ANN.

Figures 11 and 12 show a histogram of the difference between the Penman-Monteith-FAO (observational) and metered data in the three models for Ramsar and Babolsar in the test phase. The slightest difference in the LSSVM model is for both cities.

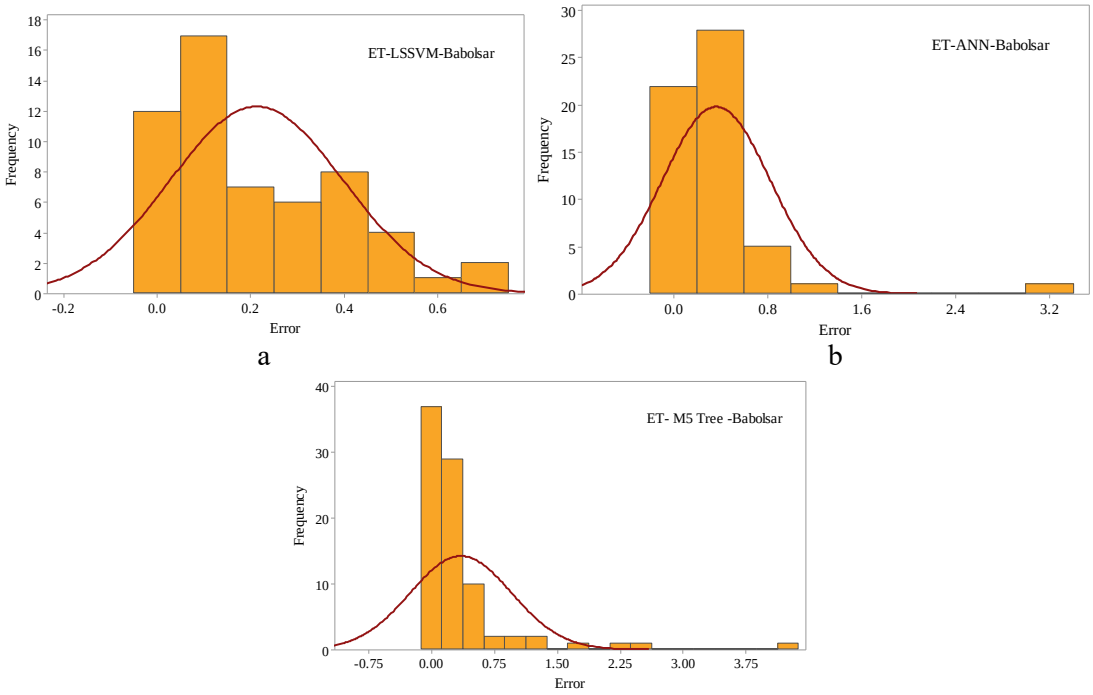


Fig. 11. Histogram of the difference between the observed data and the estimated data for Babolsar: a) ANN, b) LSSVM and c) M5 tree.

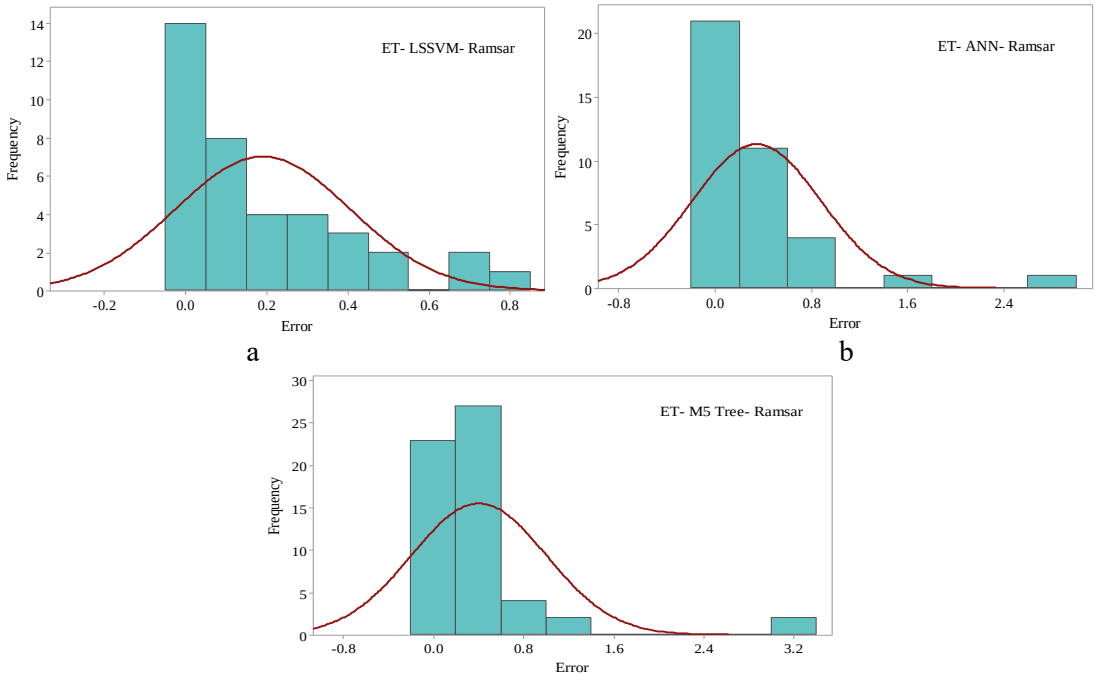
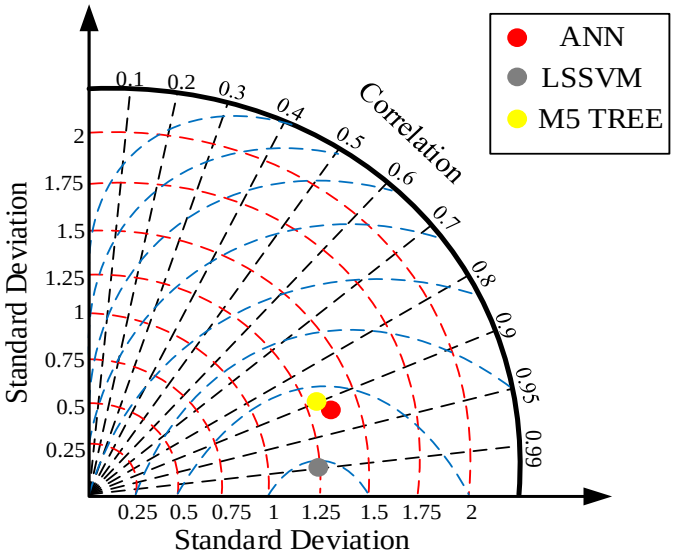


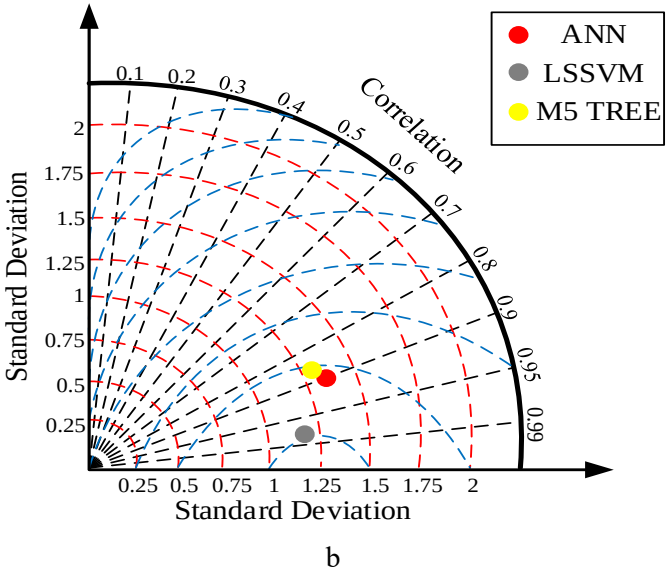
Fig. 12. Histogram of the difference between the observed data and the estimated data for Ramsar: a) ANN, b) LSSVM and c) M5 tree.

The Taylor diagram [41] is a valuable tool for evaluating the outcomes of various methods and has recently been used extensively in water engineering studies. This diagram is presented in two forms, semicircle, and quadrilateral. In both cases, the correlation coefficient values are plotted as the circle's radius on its arc, the standard deviation values are plotted as concentric circles relative to the circle's center, and the RMSE values are plotted as concentric circles close to the reference point.

The evaluation method in this diagram is that the estimated data based on three statistical criteria (RMSE, correlation coefficient between observational data and calculated data, and standard deviation) are plotted on the diagram. Based on Figure 13, the results show that the performance of all three models for both cities is relatively high.



a



b

Fig. 13. Taylor diagram of the studied models, a: Ramsar and b: Babolsar.

Abed et al. (2021) [42] used minimum temperature, maximum temperature, average temperature, wind speed, relative humidity, and solar radiation data as input to estimate the evaporation rate using Extreme Gradient Boosting (EGB), ElasticNet Linear Regression, and LSTM models. Also, Majhi et al. (2019) [43] have used minimum and maximum temperature, morning and afternoon relative humidity, wind speed, and solar radiation data as input for LSTM and Multilayer Perceptron(MLP) models. The best results for the above pieces of research are obtained when all the input data are considered in the estimation process. Adding to that, the results of Qasem et al. (2019) [44] also showed that the higher the number and types of inputs, the higher the accuracy of the results by ANN, Wavelet-Artificial Neural Network (WANN), SVR, and Wavelet-Support Vector Regression (WSVR) models. The above-cited articles' results support scenario No. 4, where all data were considered input.

Alsumaiei (2020) [45] used the ANN for the dry area of Kuwait International Airport (KIA) to model the evaporation rate, the results of which are consistent with the obtained results in the current study. It was stated that the MLP model could have a relatively good performance in predicting evaporation in arid and very arid regions, which is confirmed by the results obtained in the present study.

Arid and semi-arid climates have unique climate regimes characterized by "scarce water resources", "bare vegetation" and "high evaporation rates." According to the report of the Food and Agriculture Organization of the United Nations, extremely dry climates are defined as areas where annual precipitation does not exceed 3% of annual evaporation. It is necessary to compare the performance of smart models with other experimental and practical models, by addressing the issue of the contribution of the pan body in heat exchange. In the future research, we can mention this issue, considering that the temporal behavior of daily evaporation is unstable, it is better to investigate and compare other intelligent methods for modeling.

4. Conclusions

Evapotranspiration of reference plant (ET0) is a variable used in irrigation planning, water resources management, and hydrological studies. Its other application is to estimate the water requirement of crops in large irrigation areas. Evapotranspiration is a nonlinear and complex phenomenon that requires different climatic parameters. Calculation of evapotranspiration helps to save water and agricultural issues. This study used three intelligent models: the M5 decision tree model, least square support vector machine (LSSVM), and artificial neural network (ANN). This study used two synoptic stations of Mazandaran province with humid climates named Ramsar station and Babolsar station, located in the north of Iran. At Babolsar station, the LSSVM model had the best coefficient of determination and the lowest error rate, and the ANN model was in the next rank. In Ramsar station, the highest coefficient of perseverance and the lowest error rate belongs to the LSSVM model. The necessary suggestions for continuing this research are presented as follows:

1) Investigation of intelligent and experimental models in the desired stations in limited data conditions under different scenarios

2) Investigation of climate change in the studied stations

3) Using TOPSIS and AHP methods to weight the criteria as well as ranking the models

Funding

This research received no external funding.

Conflicts of interest

The authors declare no conflict of interest.

Authors contribution statement

YD, HG: Conceptualization; SA, HG: Data curation; SH, HG: Formal analysis; YA, HG, SH: Investigation; YD, SH: Methodology; HG, YD: Project administration; HG, YD, SH: Resources; YD, HG: Software; SH: Supervision; HG, YD: Validation; HG, YD: Visualization; HG, YD, SH: Roles/Writing – original draft; HG, YD, SH: Writing – review & editing.

References

[1] Ferreira LB, França F, Oliveira RA De, Inácio E, Filho F. Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM ; a new approach. J Hydrol 2019. https://doi.org/10.1016/j.jhydrol.2019.03.028.

[2] Wu L, Zhou H, Ma X, Fan J, Zhang F. Daily reference evapotranspiration prediction based on hybridized extreme learning machine model with bio-inspired optimization algorithms: Application in contrasting climates of China. J Hydrol 2019;577:123960. https://doi.org/10.1016/j.jhydrol.2019.123960.

[3] Ghazvinian H, Karami H, Farzin S, Mousavi SF. Effect of MDF-Cover for Water Reservoir Evaporation Reduction, Experimental, and Soft Computing Approaches. J Soft Comput Civ Eng 2020;4:98–110. https://doi.org/10.22115/scce.2020.213617.1156.

[4] Karami H, Ghazvinian H, Dehghanipour M, Ferdosian M. Investigating the Performance of Neural Network Based Group Method of Data Handling to Pan's Daily Evaporation Estimation (Case Study: Garmsar City). J Soft Comput Civ Eng 2021;5:1–18. https://doi.org/10.22115/scce.2021.274484.1282.

[5] Ghazvinian H, Karami H, Farzin S, Mousavi SF. Experimental Study of Evaporation Reduction Using Polystyrene Coating, Wood and Wax and its Estimation by Intelligent Algorithms. Irrig Water Eng 2020;11:147–65. https://doi.org/10.22125/iwe.2020.120727.

[6] Ghazvinian H, Farzin S, Karami H, Mousavi SF. Investigating the Effect of using Polystyrene sheets on Evaporation Reduction from Water-storage Reservoirs in Arid and Semiarid Regions (Case study: Semnan city). J Water Sustain Dev 2020;7:45–52. https://doi.org/10.22067/jwsd.v7i2.81748.

[7] Dehghanipour MH, Karami H, Ghazvinian H, Kalantari Z, Dehghanipour AH. Two Comprehensive and Practical Methods for Simulating Pan Evaporation under Different Climatic Conditions in Iran. Water 2021;13:2814. https://doi.org/10.3390/w13202814.

[8] Ghazvinian H, Karami H, Farzin S, Mousavi S-F. Introducing affordable and accessible physical covers to reduce evaporation from agricultural water reservoirs and pools (field study, statistics, and intelligent methods). Arab J Geosci 2021;14:2543. https://doi.org/10.1007/s12517-021-08735-3.

[9] Karamouz M, Nazif S, Falahi M. Hydrology and hydroclimatology: principles and applications. 6000 Broken Sound Parkway NW, Suite 300: CRC Press; 2012.

[10] Dehghanipour MH, Ghazvinian H, Dehghanipour A. Evaluation of the efficiency of artificial intelligence models for simulating evaporation in arid, semi-arid, and very-wet climates of Iran. Iran-Water Resour Res 2021.

[11] Babamiri O, Dinpazhoh Y. Comparison and Evaluation of Twenty Methods for Estimating Reference Evapotranspiration Based on Three General Categories: Air Temperature, Solar Radiation and Mass Transfer in the Basin of Lake Urmia. JSTNAR 2016;20:145–61. https://doi.org/10.18869/acadpub.jstnar.20.77.145.

[12] Samadianfard S, Panahi S. Estimating Daily Reference Evapotranspiration using Data Mining Methods of Support Vector Regression and M5 Model Tree. Jwmr 2019;9:157–67. https://doi.org/10.29252/jwmr.9.18.157.

[13] Goudarzi M, Salahi B, Hosseini SA. Estimation of Evapotranspiration Rate Due to Climate Change in the Urmia Lake Basin. Ijwmse 2018;12:1–12.

[14] Mohtarami O, Hosseini MR, Fattahi R. Hargreaves Method Improves Accuracy in Estimating Reference Evapotranspiration Adjustment Weight With the Help of Artificial Neural Network and Decision Tree. Water Eng 2018;6:112–23.

[15] Zorati pur E, Neisi L, Golabi M, Bazaz A, Zoratipur A. Simulation and Comparison of Potential Evapotranspiration by Artificial Neural Networks, ANFIS (Fuzzy Neural Network) and Decision Making M5 (Case Study; Synaptic Station of Shiraz). Iran-Water Resour Res 2019;15:365–71.

[16] Babu S, Thomas B. Modeling Daily Pan Evaporation Using Tree-Based Regression Methods, 2022, p. 605–14. https://doi.org/10.1007/978-981-19-0475-2_53.

[17] Batra K, Gandhi P. Artificial Neural Network-Based Model for the Prediction of Evaporation in Agriculture, 2022, p. 561–9. https://doi.org/10.1007/978-981-16-8542-2_46.

[18] Solaimani K. Seasonal relationship between climatic variables and evaporation based on Bayesian quantile regression method in southern Caspian region. Arab J Geosci 2022;15:1000. https://doi.org/10.1007/s12517-022-10263-7.

[19] Jannar Fereidouni F, Nahavandian Esfahani S, Mahmoudi N. Seasonal variations of the water column structure and estimation of the mixed layer depth based on the temperature using threshold method in Babolsar and Ramsar regions. J Earth Sp Phys 2020;46:159–74. https://doi.org/10.22059/jesphys.2020.286089.1007154.

[20] Moghaddam AS, Massoud J, Mahmoodi M, Mahvi AH, Periago MV, Artigas P, et al. Human and animal fascioliasis in Mazandaran province, northern Iran. Parasitol Res 2004;94. https://doi.org/10.1007/s00436-004-1169-6.

[21] Karami H, DadrasAjirlou Y, Jun C, Bateni SM, Band SS, Mosavi A, et al. A Novel Approach for Estimation of Sediment Load in Dam Reservoir With Hybrid Intelligent Algorithms. Front Environ Sci 2022;10. https://doi.org/10.3389/fenvs.2022.821079.

[22] Ghazvinian H, Mousavi S-F, Karami H, Farzin S, Ehteram M, Hossain MS, et al. Integrated support vector regression and an improved particle swarm optimization-based model for solar radiation prediction. PLoS One 2019;14:e0217634. https://doi.org/10.1371/journal.pone.0217634.

[23] Ghazvinian H, Bahrami H, Ghazvinian H, Heddam S. Simulation of Monthly Precipitation in Semnan City Using ANN Artificial Intelligence Model. J Soft Comput Civ Eng 2020;4:36–46. https://doi.org/10.22115/scce.2020.242813.1251.

[24] Naderpour H, Rafiean AH, Fakharian P. Compressive strength prediction of environmentally friendly concrete using artificial neural networks. J Build Eng 2018;16:213–9.

[25] Naderpour H, Rezazadeh Eidgahee D, Fakharian P, Rafiean AH, Kalantari SM. A new proposed approach for moment capacity estimation of ferrocement members using Group Method of Data Handling. Eng Sci Technol an Int J 2020;23:382–91. https://doi.org/10.1016/j.jestch.2019.05.013.

[26] Naderpour H, Nagai K, Fakharian P, Haji M. Innovative models for prediction of compressive strength of FRP-confined circular reinforced concrete columns using soft computing methods. Compos Struct 2019;215:69–84. https://doi.org/10.1016/j.compstruct.2019.02.048.

[27] Dadrasajirlou Y. Quantitative flood mitigation in urban basins with optimal low-impact development and best management practices designs under climate change conditions. Semnan Universtiy, 2021.

[28] Eftekhari M, Eslaminezhad SA, Akbari M, DadrasAjirlou Y, Haji Elyasi A. Assessment of the Potential of Groundwater Quality Indicators by Geostatistical Methods in Semi-arid Regions. J Chinese Soil Water Conserv 2021;52:158–67. https://doi.org/10.29417/JCSWC.202109_52(3).0004.

[29] Javadi S, Maghami A, Hosseini SM. A deep learning approach based on a data-driven tool for classification and prediction of thermoelastic wave's band structures for phononic crystals. Mech Adv Mater Struct 2021:1–14. https://doi.org/10.1080/15376494.2021.1983088.

[30] Heddam S, Kisi O. Modelling daily dissolved oxygen concentration using least square support vector machine, multivariate adaptive regression splines and M5 model tree. J Hydrol 2018;559:499–509. https://doi.org/10.1016/j.jhydrol.2018.02.061.

[31] Rezazadeh Eidgahee D, Jahangir H, Solatifar N, Fakharian P, Rezaeemanesh M. Data-driven estimation models of asphalt mixtures dynamic modulus using ANN, GP and combinatorial GMDH approaches. Neural Comput Appl 2022;34:17289–314. https://doi.org/10.1007/s00521-022-07382-3.

[32] Naderpour H, Parsa P, Mirrashid M. Innovative Approach for Moment Capacity Estimation of Spirally Reinforced Concrete Columns Using Swarm Intelligence–Based Algorithms and Neural Network. Pract Period Struct Des Constr 2021;26. https://doi.org/10.1061/(ASCE)SC.1943-5576.0000612.

[33] Wang L, Niu Z, Kisi O, Li C, Yu D. Pan evaporation modeling using four different heuristic approaches. Comput Electron Agric 2017;140:203–13. https://doi.org/10.1016/j.compag.2017.05.036.

[34] Rezaie-Balf M, Kim S, Fallah H, Alaghmand S. Daily river flow forecasting using ensemble empirical mode decomposition based heuristic regression models: Application on the perennial rivers in Iran and South Korea. J Hydrol 2019;572:470–85. https://doi.org/10.1016/j.jhydrol.2019.03.046.

[35] Nikpour MR, Sani Khani H, Mahmodi Babelan S, Mohammadi A. Application of LS-SVM, ANN, WNN and GEP in rainfall- runoff modeling of Kiyav-Chay River. Iran J Ecohydrol 2017;4:627–39. https://doi.org/10.22059/ije.2017.61501.

[36] Hu Z, Karami H, Rezaei A, DadrasAjirlou Y, Piran MJ, Band SS, et al. Using soft computing and machine learning algorithms to predict the discharge coefficient of curved labyrinth overflows. Eng Appl Comput Fluid Mech 2021;15:1002–15. https://doi.org/10.1080/19942060.2021.1934546.

[37] Naderpour H, Mirrashid M, Parsa P. Failure mode prediction of reinforced concrete columns using machine learning methods. Eng Struct 2021;248:113263. https://doi.org/10.1016/j.engstruct.2021.113263.

[38] Barzegar R, Ghasri M, Qi Z, Quilty J, Adamowski J. Using bootstrap ELM and LSSVM models to estimate river ice thickness in the Mackenzie River Basin in the Northwest Territories, Canada. J Hydrol 2019;577:123903.

[39] Cunha AC, Gabriel Filho LRA, Tanaka AA, Goes BC, Putti FF. Influence Of The Estimated Global Solar Radiation On The Reference Evapotranspiration Obtained Through The Penman-Monteith Fao 56 Method. Agric Water Manag 2021;243:106491.

[40] Córdova M, Carrillo-Rojas G, Crespo P, Wilcox B, Célleri R. Evaluation of the Penman-Monteith (FAO 56 PM) method for calculating reference evapotranspiration using limited data. Mt Res Dev 2015;35:230–9.

[41] Taylor KE. Summarizing multiple aspects of model performance in a single diagram. J Geophys Res Atmos 2001;106:7183–92. https://doi.org/10.1029/2000JD900719.

[42] Abed M, Imteaz MA, Ahmed AN, Huang YF. Application of long short-term memory neural network technique for predicting monthly pan evaporation. Sci Rep 2021;11:1–19.

[43] Majhi B, Naidu D, Mishra AP, Satapathy SC. Improved prediction of daily pan evaporation using Deep-LSTM model. Neural Comput Appl 2020;32:7823–38. https://doi.org/10.1007/s00521-019-04127-7.

[44] Qasem SN, Samadianfard S, Kheshtgar S, Jarhan S, Kisi O, Shamshirband S, et al. Modeling monthly pan evaporation using wavelet support vector regression and wavelet artificial neural networks in arid and humid climates. Eng Appl Comput Fluid Mech 2019;13:177–87. https://doi.org/10.1080/19942060.2018.1564702.

[45] Alsumaiei AA. Utility of Artificial Neural Networks in Modeling Pan Evaporation in Hyper-Arid Climates. Water 2020;12:1508. https://doi.org/10.3390/w12051508.