Comparative analysis of data-driven methods for daily reference evapotranspiration estimation of Southern Caspian Sea

Bahram Bakhtiari | Alireza Mohebbi-Dehaghani | Kourosh Qaderi

Abstract
Reference evapotranspiration (ET₀) as a component in the hydrological cycle is calculated using many methods. In this study, the capability of four data-driven methods including artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), support vector machine (SVM), and M5 tree model has been evaluated for daily ET₀ estimation at the south of the Caspian Sea. Different combinations of climatic data, solar radiation (Rs), mean air temperature (Tmean), mean relative humidity (RHmean), and wind speed (U) during 1991–2020 were used as input variables. The data were divided into training and test data. The values generated from the methods were compared with those of the FAO-56 Penman-Monteith as a standard method. The results indicated that the accuracy of ANFIS was increased for estimating ET₀, especially in validation phase, when all climatic variables were used as inputs in the synoptic stations. Totally, based on the evaluation of the performances, it was created that the ANFIS with Tmean, RHmean, Rs, and U variables had the best accuracy, while the ANN, SVM, and M5 with only one input of U had the worst performance. The ANFIS with Tmean and Rs was recommended for modelling ET₀ if there are fewer climatic variables in these regions.

Keywords
adaptive neuro-fuzzy, climate variables, data-driven methods, evapotranspiration modelling, neural networks, Penman-Monteith

1 | INTRODUCTION
Reference evapotranspiration (ET₀) is one of the main variables for the control of reference crop growth. Accuracy estimation of ET₀ is necessary for many studies such as water resources management, drought management (Jiang et al., 2022), irrigation system design, water budget estimation, and meteorological and climatological studies (Raza et al., 2020). This water cycle component is a complex non-linear phenomenon that has a great deal of dependence on climatological factors. Therefore, research in this field is very important. Evapotranspiration is achieved by direct and indirect methods. It can be measured directly by some instrument such as lysimeters. However, construction of this equipment is difficult and expensive (Ma et al., 2015). In an indirect method, usually...
empirical, semi-empirical, or physical equations have been used to estimate ET₀ based on meteorological or climatological data (Lopez-Urrea et al., 2006). The Food and Agriculture Organization (FAO) recommends the use of the standardized FAO Penman-Monteith (PMF-56) equation for ET₀ estimation (Allen et al., 1998). Considering the definition of the reference crop (Doorenbos, 1977), the accuracy of this method has been confirmed in different climates and time periods (Allen et al., 2005). Therefore, in most studies, this method is known as a standard method for comparing other empirical or soft computing methods (Xu et al., 2015). The Penman-Monteith equation needs many meteorological factors, such as air temperature, relative humidity, solar radiation, vapour pressure deficit, and wind speed, to calculate ET₀. The application and evaluation of this method in some areas is a challenge due to lack of dense weather stations and incomplete meteorological data (Valipour, 2017). Nonetheless, a large number of researchers compared the standard Penman-Monteith with other empirical methods for estimating ET₀ (Celestin et al., 2020).

Over the past few years, many soft computing approaches and data-driven methods like artificial neural networks (ANN), support vector machines (SVM), adaptive networks based on fuzzy inference system (ANFIS) and M5 model have been used in the field of various environmental sciences, such as water engineering, hydrological process, and meteorological and climatological studies (Pal & Deswal, 2009; Sekertekin et al., 2021).

Since the 1990s, ANN, based on the realization of the nervous systems, has been used extensively in hydrological prediction and simulation (Yang et al., 1996), optimization (Wen & Lee, 1998), water resources, and meteorological applications (Narvekar & Fargose, 2015). ASCE Task Committee went over the application of this method in hydrology problems (ASCE 2000a, 2000b). Kisi (2008) investigated the accuracy of three different ANN methods, including the multi-layer perceptron (MLPs), radial basis neural networks (RBNNs), and generalized regression neural networks (GRNNs), in modelling ET₀. The results demonstrated that the MLP and RBNN models could be used successfully in ET₀ estimation. In another study, Chauhan and Shrivastava (2009) evaluated the performance of four climate-based methods and ANN for ET₀ estimation in India. The results demonstrated that the ANN models performed better than the others. For the past 2 years, the use of ANN in ET₀ estimation has still been used by researchers (Majhi & Naidu, 2021; Walls et al., 2020).

The ANFIS method was first developed by Jang (1993) and has been used in many respects, for example, hydrology and water resources problems (Alquraish et al., 2021). Pulido-Calvo and Gutierrez-Estrada (2009) discussed the potential of ANFIS, ANN, and multivariate regression in ET₀ estimation in southern Spain. They found that the ANFIS performed better than the other methods. Ladlani et al. (2014) modelled the daily ET₀ in Mediterranean region of Algeria using ANFIS and multiple linear regression (MLR). They drew the conclusion that ANFIS simulated ET₀ better than the MLR. Keshtegar et al. (2017) studied the accuracy of subset ANFIS in estimating the daily ET₀ and compared it with M5, ANN, and ANFIS. They found that the accuracy of studied models was increased using subset ANFIS. The use of ANFIS is now recommended in modelling ET₀ in different climates (Roy et al., 2021).

The SVM is based on the concept of decision planes that define decision boundaries. This technique has a good distribution capability and has been proven to be a strong algorithm in many fields (Vapnik, 1998). Lately, the use of SVM has been considered in hydrology engineering, such as lake water level prediction (Wu et al., 2008) and stream flow simulations (Lin et al., 2006). Eslamian et al. (2009) investigated the ability of the SVM and ANN methods to model the ET₀ in a controlled greenhouse environment. They concluded that the SVM model was superior to ANN in modelling evapotranspiration. Rahimikhoob (2014) evaluated the performance of ANN and M5 methods with standardized PMF-56 method in an arid climate. The results showed that the ANN computed ET₀ better than the M5, but both methods performed well for this climate.

The M5 tree model that was developed by Quinlan (1992) is a logical learning technique that deals with continuous numerical values, and its learning algorithm is called M5. This algorithm is the most popular categorization handled in the family of decision-making tree models. This model has been applied to study many problems in hydrology, such as flood forecasting (Solomatine & Xue, 2004) and meteorology (Kisi et al., 2021).

However, simultaneous use and comparison of these four data-driven methods for daily ET₀ estimation in different climates are limited in the literature (Rahimikhoob, 2014). Since ET₀ is a non-linear and stochastic component of the hydrologic cycle and has a great deal of dependence on meteorological factors; therefore, the study and comparison of many data-driven methods are necessary for estimating this component. From the literature review, it was considered that data-driven methods have been applied to modelling ET₀ in different regions and climates; however, the application of these methods has been studied finitely (Murat, 2011; Shiri et al., 2013). Because of the dynamic nature or even non-stationary of evapotranspiration, it is difficult to estimate ET₀ data accurately, especially in areas with limited meteorological data.

Therefore, the aim of this study is to peruse the usability of four data-driven methods (ANN, ANFIS,
SVM, and M5 tree model) with different characteristics for estimating \( \text{ET}_o \) in daily time step using different climatic input combinations of solar radiation \((R_s)\), mean air temperature \((T_{\text{mean}})\), mean relative humidity \((\text{RH}_{\text{mean}})\), and wind speed at 2-m height \((U)\) at the southern climates of the Caspian Sea. The gathered data of the three synoptic stations (from 1991 to 2020) were used for training and validation of the applied methods. The performances of these methods were compared with the standardized PMF-56 method. Some statistical criteria have been used for the model’s performance evaluation. The best one with a smaller error and highest coefficient of determination was selected for future estimation of daily \( \text{ET}_o \). This inspection will help to find a correlation between meteorological variables and daily \( \text{ET}_o \) and to recognize the most substantial climatic factors that effect on daily \( \text{ET}_o \).

2 | MATERIALS AND METHODS

2.1 | Study area and data

The research was carried out in the coastal zone of the southern part of the Caspian Sea (North of Iran). The Caspian Sea is the largest interior body of water on Earth, which has an area of 360,000 km\(^2\). Similar to many lakes, it does not feed into an ocean, but it looks like a sea in terms of its depth and size. That is why the lake is mentioned as ‘Sea’ in the literature. The Caspian Sea has about a third of the mean salinity of seawater, thus it is not actually a freshwater lake (Ghadiri et al., 2006). Because this sea has no outlets, the main factor in the loss of water in this area is the evaporation phenomenon.

The coastline in the south of the Caspian Sea, expanding over 800 km, contains nearly 18% of the whole Caspian coastline. The Caspian Sea climate is mostly affected by its geographical status and the prevailing regional atmospheric circulation. Most of the precipitation occurs during December to March, and most of evapotranspiration occurs during June to September; therefore, there is a serious water scarcity during the growing season of crops due to the evapotranspiration phenomenon in this region. Evapotranspiration varies depending on the climatic conditions, the surface water availability, and the crop pattern in this area. Thus, accurate estimation of \( \text{ET}_o \) is an important component in irrigation management in this coastal area. For this purpose, three synoptic weather stations, namely Gorgan, Babolsar, and Rasht, were selected in this study. These stations are located in the range of 36°54’ to 37°19’ N (latitude), from 49°37’ to 54°24’ E (longitude), and from –8.6 to 13.3 m above sea level elevation, covering three climate zones based on De Martonne Aridity index (Rahimi et al., 2013). These stations are on or near the coast of the Caspian Sea and are supposed to illustrate the general state of the southern Caspian Sea. The distances between Gorgan to Babolsar and Babolsar to Rasht synoptic weather stations are about 184 and 313 km, respectively.

Table 1 shows the detailed geographical and climatological characteristics of these synoptic stations. The locations of these stations are shown in Figure 1. The long-term available daily meteorological records of solar radiation \((R_s)\), mean air temperature \((T_{\text{mean}})\), mean relative humidity \((\text{RH}_{\text{mean}})\), and wind speed \((U)\) were collected from IRIMO (Islamic Republic of Iran Meteorological Organization, 2021) during 1991–2020 were applied to calculate \( \text{ET}_o \). The quality control tests of raw data were accomplished according to Allen (1996). The statistical analysis was carried out for the data before using for the estimation of \( \text{ET}_o \). Moreover, the main parameters of the statistical analysis are illustrated in Table 1. The standard deviation value for \( \text{RH}_{\text{mean}} \) is higher as compared with the other variables at the three stations. The daily \( \text{ET}_o \) data for this study were divided into two parts: training and validation data sets. The data sets from January 1991 to December 2013 were used for training (80% of the data), meanwhile those from 1 January 2014 to December 2020 (20% of the data) were used for validation of the applied methods.

2.2 | Explanation of chosen methods

A physical base method (PMF-56) and four data-driven methods (ANN, ANFIS, SVM, and M5 tree model) were used in this study to estimate \( \text{ET}_o \). So, a summary of these methods is provided below.

2.2.1 | FAO-56 Penman-Monteith method

As there is no lysimeter for measuring \( \text{ET}_o \) in the studied stations, daily \( \text{ET}_o \) data were calculated by the FAO-56 Penman-Monteith (PMF-56) method. This method has been suggested as the standard equation for defining \( \text{ET}_o \) by FAO (Doorenbos & Pruitt, 1977) and was applied as the benchmark model. According to Smith et al. (1992), this physical method estimates the \( \text{ET}_o \) closer to the lysimetric values than other empirical methods. The PMF-56 can be used for daily or shorter time steps. For daily scales, Equation (1) is proposed:

\[
\text{ET}_o = \frac{0.408(\Delta R_s - G) + 0.900}{\Delta + 0.34 u_2} T_{\text{mean}} + 273 (e_s - e_a),
\]
| Station/climate     | Long. (E) | Lat. (N) | Elevation (m) | Variables                  | Mean | Standard deviation | Min  | Max  | Range |
|---------------------|-----------|----------|---------------|----------------------------|------|--------------------|------|------|-------|
| Gorgan (Mediterranean) |           |          |               | $T_{\text{mean}}$ (°C)     | 17.8 | 7.9                | -4.4 | 33.6 | 38.0  |
|                     |           |          |               | $\text{RH}_{\text{mean}}$ (%) | 73.3 | 10.8               | 25.1 | 98.4 | 73.3  |
|                     | 54° 24'   | 36° 54'  | 13.3          | $R_s$ (MJ m⁻² day⁻¹)        | 9.6  | 4.4                | 2.7  | 19.7 | 17.0  |
|                     |           |          |               | $U$ (m s⁻¹)                 | 1.2  | 1.1                | 0.0  | 5.1  | 5.1   |
|                     |           |          |               | $\text{ET}_o$ (mm day⁻¹)    | 2.4  | 1.62               | 0.12 | 8.81 | 8.69  |
| Babolsar (Humid)    |           |          |               | $T_{\text{mean}}$ (°C)     | 17.0 | 7.2                | -0.4 | 31.0 | 31.4  |
|                     |           |          |               | $\text{RH}_{\text{mean}}$ (%) | 82.2 | 6.9                | 38.1 | 99.8 | 61.7  |
|                     | 52° 39'   | 36° 43'  | -21.0         | $R_s$ (MJ m⁻² day⁻¹)        | 9.3  | 4.5                | 2.7  | 18.8 | 16.1  |
|                     |           |          |               | $U$ (m s⁻¹)                 | 1.5  | 1.1                | 0.0  | 5.1  | 5.1   |
|                     |           |          |               | $\text{ET}_o$ (mm day⁻¹)    | 2.2  | 1.37               | 0.10 | 6.02 | 5.92  |
| Rasht (Per-Humid)   |           |          |               | $T_{\text{mean}}$ (°C)     | 16.4 | 7.5                | -5.2 | 32.1 | 37.3  |
|                     |           |          |               | $\text{RH}_{\text{mean}}$ (%) | 84.5 | 8.9                | 50.0 | 100 | 50.0  |
|                     | 49° 37'   | 37° 19'  | -8.6          | $R_s$ (MJ m⁻² day⁻¹)        | 8.7  | 4.5                | 2.7  | 19.4 | 16.7  |
|                     |           |          |               | $U$ (m s⁻¹)                 | 1.2  | 1.1                | 0.0  | 5.3  | 5.3   |
|                     |           |          |               | $\text{ET}_o$ (mm day⁻¹)    | 1.9  | 1.31               | 0.05 | 6.85 | 6.80  |
where \( ET_0 \) is the reference evapotranspiration (mm day\(^{-1}\)), \( R_n \) is the net radiation at the crop surface (MJ m\(^{-2}\) day\(^{-1}\)), \( G \) is the soil heat flux density (MJ m\(^{-2}\) day\(^{-1}\)), \( T_{\text{mean}} \) is the mean daily air temperature at a 2-m height (°C), \( u_2 \) is the wind speed at a 2-m height (m s\(^{-1}\)), \( e_s \) is the saturation vapour pressure (kPa), \( e_a \) is the actual vapour pressure (kPa), \( e_s - e_a \) is the saturation vapour pressure deficit (kPa), \( \Delta \) is the slope of the vapour pressure curve (kPa °C\(^{-1}\)), and \( \gamma \) is the psychrometric constant (kPa °C\(^{-1}\)). Several kinds of methods are suggested by Allen et al. (2005) to evaluate the parameters of Equation 1 from the observed climatic data. According to Allen et al. (1998), for daily estimates of the evapotranspiration rate of reference crop, the term \( G \) can be neglected, that is, we set \( G = 0 \). The net radiation at the surface \( (R_n) \) is given as the difference between incoming net short wave radiation and outgoing net long wave radiation. The computation of all data required for the calculation of the net radiation at the surface followed the procedure given in Chapter 3 of the FAO paper 56 (Allen et al., 1998).

2.2.2 | Artificial neural network

The ANN is distinguished as a set of similar structural components called neurons resembling the humans’ neural system. In general, it is a biologically inspired network of artificial neurons put together to accomplish particular roles. An ANN has three fundamental elements including classes of connection weights, a threshold, and an activation function. In other words, the most popular employed ANN is the three-layer feed-forward back propagation (BP) training, which generally has, different layers such as input layer (where the data are offered to the network), hidden layer (with the number of neurons where the data are processed) and output layer (where the results are created), and some nodes in each layer. The structure of ANN depends on the nature of the problem and is distinguished by user. There are many ways to train the neural network, and the error BP algorithm was used in this research.

2.2.3 | Adaptive neuro-fuzzy inference system

A combination of ANNs and fuzzy inference systems (FISs), which was introduced by Jang (1993), creates the ANFIS system. This intelligence system uses a non-linear process from input to output, and it is a multi-layer form connected with a number of nodes that has three sections (Lski & Czogala, 1999). In the first part, a rule base including fuzzy rules is selected. The second part contains the database. In this case, a learning algorithm with the compatible membership functions (MFs) is used to create a set of fuzzy if-then laws from certain input–output subsets. For this purpose, the ANFIS receives primary FIS and modifies it with a propagation algorithm. Finally, in the last section, a logical inference system is used to remove it from the rules and input data to reach an acceptable output. The most famous of them, which are used in fuzzy systems, are Mamdani, Sugeno, and Takagi models. These systems recognize patterns and help to the adaptation of natural world. Figure 2 illustrates an ANFIS structure with two inputs \((x_1, x_2)\) and
one output ($F$) in which a circle represents a stable node, and a square shows an adaptive node.

A description of the ANFIS layers is available in Łęski and Czogala (1999).

2.2.4 | Support vector machine

A SVM theory is based on the classification technique, which is applied a linear model to divide the main data into a feature space through a non-linear mapping function (Vapnik, 1998). This technique maximizes the boundary between the two opposing classes and minimizes the risk of a structure. SVM can also be used in regression relationships. The regression method used in SVM is different from the usual regression methods (Mustafa et al., 2018). In general, in order to reduce the number of errors in the training phase, this model uses non-linear machines to find the most suitable hyperplane. In the next step, after constructing a Lagrange function and optimizing the target function, a non-linear regression function is obtained. The functions used for optimization are presented in the scientific resources (Tezel & Buyukyildiz, 2016). The non-linear part of the SVM is the recognition of Kernel functions. The simplest form of this function is the linear kernel (2005). Other types of kernels include the Polynomial, Gaussian radial basis function (RBF), and sigmoid. If the points could not be separated by a kernel, it would have to do so until it was properly segmented with other kernels. In this study, the kernel function of the RBF type was selected.

2.2.5 | M5 tree model

The tree model is a data-driven technique based on machine learning, presented by Quinlan (1992). The algorithm for learning this model is known as the M5. Decision trees are used to predict outputs from the input data. In these trees, each node (leaf) represents a class or internal decision that performs certain specified tests. The advantage of this model is that it is a white-box method in which the dependence between the variables is determined by using linear and non-linear functions. The M5 tree is the generalized regression tree and uses the linear regression functions in the leaves. In order to implement the model tree generation, three stages of splitting, pruning, and smoothing should be considered. Initially, an inductive algorithm is used to construct the decision tree. For this purpose, a splitting criterion is used. This criterion reduces the internal variation in the model. The criterion is based on the standard deviation (SD) behaviour of each class's values (Pal & Deswal, 2009). When some samples are left very small, the splitting process stops. As a matter of fact, in this stage, the M5 model calculates the error values based on the minimum error rate at each node. In the final step, the accuracy of the estimation of the tree model is improved using the smoothing of the sharp nodes. Further information about the M5 tree model can be found in the study by Quinlan (1992).

2.3 | Models’ performance evaluation

Various techniques are recommended to assess the performance of models in the calibration and validation phases, which is generally considered as goodness of fit (Shcherbakov et al., 2013). Moreover, the performances of the data-driven methods were compared with the PMF-56 method. Five performance evaluation criteria were used in this study. The coefficient of determination ($R^2$) has been widely applied for model evaluation. It is used to describe how differences in a simulated variable can be explained by a difference in the measured variable. The range of this coefficient is between 0 and 1. If $R^2 = 0$, There is no linear relationship between observed and calculated values. If $R^2 = 1$, a very significant regression relationship exists. The root mean squared error (RMSE) evaluates the residual between observed (by PMF-56) and forecasted (using data-driven methods) ET0. The model and observational data are better suited if this statistic is smaller. The mean absolute percentage error (MAPE) has been widely used for relative error comparison in the estimation in respect of the actual rate of the variable. Consequently, The MAPE is an unbiased statistic for quantifying the predictive ability of the model. Smaller values of this index imply higher model efficiency (Lefebvre & Bensalma, 2015). The original index of agreement (IA) is a standardized measure of the degree of model estimation error and varies between...
Combinations of input variables considered in the applied methods

| No. of model | Input structure | Output |
|--------------|-----------------|--------|
| 1            | \(T_{\text{mean}}\), RH\(_{\text{mean}}\), U | \(E\_T\_o\) |
| 2            | \(T_{\text{mean}}\), \(R_u\), U | \(E\_T\_o\) |
| 3            | \(T_{\text{mean}}\), RH\(_{\text{mean}}\), \(R_u\) | \(E\_T\_o\) |
| 4            | \(T_{\text{mean}}\), RH\(_{\text{mean}}\), \(R_u\), U | \(E\_T\_o\) |

(ANN, ANFIS, SVM, and M5) at the three stations. The different combinations of input variables (as models No. 1 to 4) are presented in Table 2.

In this study, for the ANN method, after studying and testing various structures, the best one was obtained as 4-10-10-1. The most common network transfer function, which is known as the tan-sigmoid function, was selected as the network activation function. Moreover, for the calculation of the output, the feed-forward with learning algorithm back propagation error (BP) type with two hidden layers was used. In this algorithm, the output computation was done from the input to the output in the forward direction. The BP algorithm using the 10 nodes during the training phase gave the least mean square error and was used for network training.

Similarly, an optimal architecture was chosen after analysing different structures in the ANFIS method. In this case, the MF parameters were iteratively learned using a hybrid method. In accordance with a number of tests performed, the Gaussian type with two MFs was selected. A Sugeno’s type FIS was selected for functioning, and the forward pass of hybrid learning algorithms was applied to identify the parameters. The resulting parameters are determined by the least squares estimate.

For the SVM method, the RBF kernel was used. The general performance of this method depends on an appropriate setting of the regulation parameters of kernel function. Therefore, two parameters including the penalty parameter \((C = 10)\) and the regulation parameter \((\epsilon = 0.1)\) were obtained for the study stations.

3 | RESULTS AND DISCUSSION

Before presenting the main results of this paper, it should be noted that the sensitivity analysis of the estimation of daily \(E\_T\_o\) in the study area was previously performed by Bakhtiari and Baghizadeh (2012). Sensitivity analysis indicated that \(R_u\), RH\(_{\text{mean}}\), and \(T_{\text{mean}}\) were effective parameters in the study area. However, the sensitivity of the variable U was less than other variables and was less important.

The results of the four different data-driven methods for estimating daily \(E\_T\_o\) are presented in Tables 3–5 in
| Method | Model | Training | Validation | Min. (mm day⁻¹) | Max. (mm day⁻¹) | Range (mm day⁻¹) |
|--------|-------|----------|------------|----------------|----------------|-----------------|
|        |       | RMSE (mm day⁻¹) | MAPE | IA | NSE | RMSE (mm day⁻¹) | MAPE | IA | NSE |
|        |       | R²          |          |    |    | R²          |          |    |    |
| Observed |       | 0.12        | 8.81     | 8.69 | 0.012 | 18.61     | 9.018 | 8.69 | 0.012 |
| ANN     | 4     | 0.978       | 0.481    |     |    | 0.975       | 0.414  | 0.932 | 0.49 |
|         | 3     | 0.970       | 0.518    |     |    | 0.965       | 0.501  | 0.695 | 0.32 |
|         | 2     | 0.969       | 0.504    |     |    | 0.951       | 0.586  | 0.934 | 0.46 |
| ANFIS   | 4     | 0.980       | 0.382    |     |    | **12.284**  | 0.989  | 0.950 | 0.079 |
|         | 3     | 0.972       | 0.477    |     |    | 0.967       | 0.482  | 0.739 | 0.46 |
|         | 2     | 0.961       | 0.628    |     |    | 0.951       | 0.570  | 0.904 | 0.01 |
| SVM     | 4     | **0.982**   | **0.312**|     |    | 0.977       | 0.403  | 0.949 | 0.28 |
|         | 3     | 0.978       | 0.472    |     |    | 0.967       | 0.482  | 0.701 | 0.32 |
|         | 2     | 0.962       | 0.624    |     |    | 0.948       | 0.598  | 0.932 | 0.09 |
| M5      | 4     | 0.968       | 0.512    |     |    | 0.972       | 0.445  | 0.941 | 0.30 |
|         | 3     | 0.961       | 0.594    |     |    | 0.952       | 0.577  | 0.705 | 0.30 |
|         | 2     | 0.939       | 0.792    |     |    | 0.932       | 0.671  | 0.902 | 0.29 |

**Note:** Min. is minimum ET₀, and Max. is maximum ET₀. The values of the performance's indices for the cases of the best inputs and the best model were shown in bold.

**Abbreviations:** ANFIS, adaptive neuro-fuzzy inference system; ANN, artificial neural network; IA, index of agreement; MAPE, mean absolute percentage error; NSE, Nash–Sutcliffe efficiency; SVM, support vector machine.
TABLE 4  Summary of statistical performance indices of different models during training and validation periods for Babolsar synoptic station.

| Method | Model | Training | Validation |
|--------|-------|----------|------------|
|        |       | $R^2$ | RMSE (mm day$^{-1}$) | MAPE | IA | NSE | $R^2$ | RMSE (mm day$^{-1}$) | MAPE | IA | NSE | Min. (mm day$^{-1}$) | Max. (mm day$^{-1}$) | Range (mm day$^{-1}$) |
|        |       |      |            |        |    |    |      |            |        |    |    |           |            |               |
| Observed |       | 0.985 | 0.317 | 11.842 | 0.988 | 0.971 | 0.978 | 0.227 | 10.636 | 0.994 | 0.970 | 0.16 | 6.02 | 5.92 |
| ANN     | 4     | 0.972 | 0.385 | 13.508 | 0.982 | 0.928 | 0.963 | 0.276 | 12.446 | 0.990 | 0.923 | 0.16 | 5.57 | 5.41 |
|         | 3     | 0.885 | 0.592 | 20.118 | 0.968 | 0.965 | 0.878 | 0.512 | 22.608 | 0.964 | 0.972 | 0.30 | 5.06 | 4.76 |
| ANFIS   | 4     | 0.987 | 0.371 | 10.624 | 0.996 | 0.976 | 0.982 | 0.202 | 9.826  | 0.995 | 0.977 | 0.05 | 6.10 | 6.05 |
|         | 3     | 0.963 | 0.394 | 12.862 | 0.990 | 0.911 | 0.968 | 0.257 | 11.628 | 0.992 | 0.914 | 0.42 | 6.06 | 5.64 |
|         | 2     | 0.889 | 0.615 | 20.602 | 0.961 | 0.960 | 0.888 | 0.502 | 18.838 | 0.966 | 0.967 | 0.02 | 6.24 | 6.22 |
| SVM     | 4     | 0.985 | 0.281 | 11.520 | 0.997 | 0.979 | 0.982 | 0.215 | 10.501 | 0.995 | 0.977 | 0.37 | 5.63 | 5.26 |
|         | 3     | 0.962 | 0.340 | 11.128 | 0.993 | 0.925 | 0.967 | 0.267 | 12.122 | 0.991 | 0.923 | 0.43 | 5.76 | 5.33 |
|         | 2     | 0.865 | 0.612 | 20.182 | 0.962 | 0.964 | 0.884 | 0.519 | 21.749 | 0.952 | 0.966 | 0.01 | 5.12 | 5.11 |
| M5      | 4     | 0.970 | 0.392 | 12.181 | 0.983 | 0.970 | 0.971 | 0.250 | 11.972 | 0.992 | 0.970 | 0.21 | 5.98 | 5.77 |
|         | 3     | 0.947 | 0.412 | 14.512 | 0.979 | 0.903 | 0.954 | 0.309 | 13.782 | 0.988 | 0.899 | 0.32 | 5.74 | 5.42 |
|         | 2     | 0.872 | 0.594 | 20.950 | 0.954 | 0.951 | 0.834 | 0.602 | 21.749 | 0.952 | 0.957 | 0.20 | 5.26 | 5.06 |

Note: Min. is minimum $ET_o$ and Max. is maximum $ET_o$. The values of the performance’s indices for the cases of the best inputs and the best model were shown in bold. 
Abbreviations: ANFIS, adaptive neuro-fuzzy inference system; ANN, artificial neural network; IA, index of agreement; MAPE, mean absolute percentage error; NSE, Nash–Sutcliffe efficiency; SVM, support vector machine.
**TABLE 5** Summary of statistical performance indices of different models during training and validation periods for Rasht synoptic station.

| Method | Model | $R^2$ | RMSE (mm day$^{-1}$) | MAPE | IA  | NSE | $R^2$ | RMSE (mm day$^{-1}$) | MAPE | IA  | NSE | Min. (mm day$^{-1}$) | Max. (mm day$^{-1}$) | Range (mm day$^{-1}$) |
|--------|-------|-------|----------------------|------|-----|-----|-------|----------------------|------|-----|-----|---------------------|---------------------|----------------------|
| Observed |       |       |                      |      |     |     |       |                      |      |     |     | 0.05                | 6.85                | 6.80                 |
| ANN    | 4     | 0.964 | 0.412                | 16.620 | 0.981 | 0.932 | 0.962 | 0.328                | 15.264 | 0.985 | 0.952 | 0.01                | 4.50                | 4.49                 |
|        | 3     | 0.951 | 0.470                | 17.812 | 0.976 | 0.885 | 0.959 | 0.388                | 16.571 | 0.979 | 0.921 | 0.37                | 8.04                | 7.67                 |
|        | 2     | 0.894 | 0.652                | 21.512 | 0.948 | 0.908 | 0.926 | 0.521                | 23.251 | 0.959 | 0.941 | 0.30                | 5.06                | 4.76                 |
| ANFIS  | 4     | 0.988 | **0.386**            | 14.914 | **0.988** | 0.793 | **0.980** | **0.274**            | **15.102** | **0.989** | **0.967** | **0.05**            | 6.13                | **6.08**             |
|        | 3     | 0.947 | 0.526                | 15.938 | 0.983 | 0.885 | 0.962 | 0.375                | 17.506 | 0.981 | 0.924 | 0.01                | 5.68                | 5.67                 |
|        | 2     | 0.886 | 0.682                | 24.930 | 0.947 | 0.900 | 0.928 | 0.513                | 23.478 | 0.961 | 0.936 | 0.01                | **6.25**            | **6.24**             |
| SVM    | 4     | 0.988 | 0.388                | **14.879** | 0.981 | **0.936** | 0.977 | 0.297                | 16.977 | 0.984 | 0.960 | 0.12                | 5.88                | 5.76                 |
|        | 3     | 0.965 | 0.407                | 17.250 | 0.972 | 0.872 | 0.958 | 0.417                | 17.818 | 0.977 | 0.907 | 0.01                | 5.74                | 5.73                 |
|        | 2     | 0.925 | 0.558                | 23.914 | 0.951 | 0.886 | 0.909 | 0.587                | 26.258 | 0.942 | 0.926 | 0.01                | 5.12                | 5.11                 |
| M5     | 4     | 0.968 | 0.472                | 17.827 | 0.980 | 0.935 | 0.942 | 0.322                | 16.741 | 0.986 | 0.962 | 0.13                | 5.69                | 5.56                 |
|        | 3     | 0.942 | 0.604                | 19.015 | 0.972 | 0.859 | 0.955 | 0.410                | 18.945 | 0.977 | 0.903 | 0.17                | 5.65                | 5.48                 |
|        | 2     | 0.872 | 0.696                | 28.382 | 0.926 | 0.933 | 0.902 | 0.596                | 27.869 | 0.948 | 0.884 | 0.01                | 5.26                | 5.25                 |

*Note*: Min. is minimum $E_{T_o}$ and Max. is maximum $E_{T_o}$. The values of the performance’s indices for the cases of the best inputs and the best model were shown in bold.

*Abbreviations*: ANFIS, adaptive neuro-fuzzy inference system; ANN, artificial neural network; IA, index of agreement; MAPE, mean absolute percentage error; NSE, Nash–Sutcliffe efficiency; SVM, support vector machine.
two training and validation phases. At the Gorgan synoptic station, the minimum and maximum of \( ET_o \) obtained from the PMF-56 method were 0.21 and 8.81 mm day\(^{-1}\), respectively. The minimum and maximum values at the

**FIGURE 3** Observed (PMF-56) and estimated daily \( ET_o \) of the study synoptic stations in the validation phase.
Bablosar Station were 0.10 and 6.02 mm day$^{-1}$, respectively, and the values related to the Rasht station were 0.05 and 6.85 mm day$^{-1}$, respectively. Since models No. 2–4 showed better performance at the three stations, the figures presented in Tables 3–5 were based on the results of these models. The values of the performance’s indices for the cases of the best inputs and the best model were shown in bold.

These tables present the statistical values of both training and validation phase, concerning various performance criteria. It can be observed that all the different data-driven methods have good performance during both training and validation phases.

For Gorgan synoptic station, during the training phase, the SVM model-4 obtained the best $R^2$, RMSE, IA, and NSE statistics of 0.982, 0.312, 0.990, and 0.956 respectively, while the ANFIS model obtained the best MAPE of 12.284. As can be seen from Table 3, during the validation, the ANFIS model-4 performs the best on the whole, followed by the ANFIS model-3 (with three input factors including $T_{\text{mean}}$, $R_{\text{H}}$, and $U$) and ANFIS model-2 (with three input factors including $T_{\text{mean}}$, $R_{\text{v}}$, and $U$) models. ANFIS model-4 obtained the best values of RMSE = 0.381, MAPE = 13.313, and NSE = 0.946 (very good) during the validation phase. M5 model-3 is able to obtain the maximum $ET_o$ among all models created in this station.

Similarly, for Babolsar synoptic station, in the training phase, the SVM model-4 obtained the best RMSE, IA, and NSE statistics of 0.281, 0.997, and 0.979, respectively, while ANFIS model-4 got the best $R^2$ and MAPE statistics of 0.987 and 10.624, respectively. During the validation phase, the results indicated that both SVM and ANFIS models-4 obtained the best $R^2$, IA, and NSE statistics of 0.982 and 0.995, and 0.97 (very good), respectively, whereas ANFIS model-4 obtained the best RMSE, and MAPE statistics of 0.202 and 9.826, respectively. For Rasht synoptic station, in the training phase, SVM model-4 obtained the best $R^2$, MAPE, and NSE of 0.998, 14.879, and 0.936, respectively, while ANFIS model-4 obtained the best RMSE and IA statistics of 0.386 and 0.988, respectively. According to the figures in Table 3, we can conclude that ANFIS model-4 outperforms all other applied models in the validation phase.

Moreover, in the validation phase, as identified in Tables 3–5, the results of the ANN, ANFIS, and SVM method estimation were able to recognize a good, near estimate, as compared with those with M5 tree method; however, it can be achieved that the ANFIS method obtained the best RMSE between the observed and estimated $ET_o$ in all the study stations. In addition, it can be concluded that the ANFIS method obtained the best minimum absolute error between the observed and modelled...
The permanent need to increase agricultural yield, together with the several more drought occurrences, requires a more careful evaluation of near-term irrigation demands and, accordingly, a more accurate estimation of evapotranspiration. Generally, climatic data in many regions is limited to a reduced number of variables measured; therefore, the current or future evapotranspiration estimation is restricted. Thus, in practical applications, data-driven methods may be a powerful tool for the estimation of precise evapotranspiration when a time series of a few years is accessible. The first question is what approach should be appropriately chosen from a variety of machine learning methods in practical applications. This is very important in usage, mainly in developing countries, where some weather data may be removed because of technical reasons.

In this study, four soft computing techniques were carried out based upon the data-driven methodology to identify the most relevant climatic variables for estimation of daily reference evapotranspiration (ET₀). The high number of climatic variables for estimating ET₀ has many disadvantages. At first, data collection may be costly in some areas. Secondly, the various input climatic variables may reduce the capability of generalizability of a model. It is therefore helpful to apply methods, which allow simplifying the number of input variables. Therefore, the performances of several data-driven methods were investigated for modelling daily ET₀ time series in order to select the most important climate factors on ET₀ estimation in the

minimum and maximum ET₀ in Babolsar and Rasht synoptic stations, and the SVM and ANFIS methods obtained the best minimum absolute error between the observed and modelled minimum and maximum ET₀, respectively, in Gorgan weather station. From the other viewpoint, it can be achieved that the range of ET₀ (difference between the maximum and minimum values = 8.67 mm day⁻¹) calculated from the ANN model-4 was near those of the observed values in Gorgan station (Table 3). Similarly, the nearest range of ET₀ to the observed values were 6.05 (in ANFIS model-4) and 6.24 (in ANFIS model-2) mm day⁻¹ for Babolsar and Rasht stations, respectively (Tables 4 and 5).

According to the values in Tables 3–5, we can conclude that ANFIS model-4 had the best performance in the entire study region compared with the used models. The comparison between the observed and estimated ET₀ in the validation phase shows that the ANFIS method performs superior to the other data-driven methods in estimating ET₀ in all the study stations. The best scores were obtained using ANFIS model-4 in the validation phase for estimating Babolsar’s ET₀. The results indicate that the ANN model underestimated the daily ET₀ of PMF-56 in the Rasht station, especially in values greater than 3 mm day⁻¹ (Figure 3).

Overall, the results of the data-driven methods indicated that during the training phase, the ANFIS or SVM method is suitable to get the best output, and during the validation phase, the ANFIS method can obtain the best output in terms of different performance statistics. Besides, these statistics illustrated that the degree of goodness or defect of estimation accuracy varies according to the different performance criteria during both the training and validation phases. ANFIS method with different input factors (model no. 1–4) can obtain a better output in terms of different evaluation criteria not only during the training but also during the validation phase. The statistics measures indicated that the results of the SVM method during the validation phase are lower-ranking than the results of the training phase. ANN is in the middle level in both the training and validation phases compared with the other three methods, and is able to obtain the better RMSE and MAPE in validation phase in all the stations, especially ANN model-3 can obtain the maximum ET₀ among all models developed in Rasht synoptic station.

Figure 4a–d shows the results of four method’s regression error on the basis of RMSE during the validation phase for each station, which illustrates the effect of each meteorological variables on ET₀ estimation. It is clear that model-4 with input structure of T_mean, RH_mean, Rₛ, and U has the lowest RMSE for all the study stations; therefore, they are the most relevant meteorological variables to predict daily ET₀. Clearly, there is no significant difference between RMSE values for the first and even the second (with input variables including T_mean, RH_mean, Rₛ) most significant parameters. Moreover, it is found that U is a less relevant input variable since it has the highest RMSE. It is found that for all the stations, RH_mean is the second unimportant parameter. This emphasizes the importance of selecting the appropriate input variables for predicting ET₀ at the study stations. It should be considered that a model with more simplicity is likely to be noted in terms of required input variables; therefore, considering the most relevant two input variables would be acceptable for many researchers in terms of optimum number of inputs to prepare a balance between the simplicity and high accuracy. In this case, the model two inputs of T_mean and Rₛ can be used in the study areas. It should be mentioned again that smaller rates of RMSE illustrate further exactness of the predicted ET₀, and in a perfect case, they are zero. Thus, according to the results, it is found that the ANFIS method provides outperforming results to predict daily ET₀ than the ANN, SVM, and M5 tree methods.

4 CONCLUSIONS

The permanent need to increase agricultural yield, together with the several more drought occurrences, requires a more careful evaluation of near-term irrigation demands and, accordingly, a more accurate estimation of evapotranspiration. Generally, climatic data in many regions is limited to a reduced number of variables measured; therefore, the current or future evapotranspiration estimation is restricted. Thus, in practical applications, data-driven methods may be a powerful tool for the estimation of precise evapotranspiration when a time series of a few years is accessible. The first question is what approach should be appropriately chosen from a variety of machine learning methods in practical applications. This is very important in usage, mainly in developing countries, where some weather data may be removed because of technical reasons.

In this study, four soft computing techniques were carried out based upon the data-driven methodology to identify the most relevant climatic variables for estimation of daily reference evapotranspiration (ET₀). The high number of climatic variables for estimating ET₀ has many disadvantages. At first, data collection may be costly in some areas. Secondly, the various input climatic variables may reduce the capability of generalizability of a model. It is therefore helpful to apply methods, which allow simplifying the number of input variables. Therefore, the performances of several data-driven methods were investigated for modelling daily ET₀ time series in order to select the most important climate factors on ET₀ estimation in the
Mediterranean and humid climates. The data-driven methods used in this study were the ANN, ANFIS, SVM, and M5 tree. The conventional standardized PMF-56 method, known for its accuracy, is also used as a benchmarking yardstick for comparison goals. The daily ET₀ data from actual weather observed data in the tree synoptic stations in the southern Caspian Sea were used to investigate various methods. Prior to data-driven modelling, the data set was divided into two subsets of training and validation. Four different combinations of climatic variables were delineated as models-1 to 4. Four standard performance evaluation criteria were used to assess the performances of each model. Based on the results obtained, it is noticeable that the most influential single variable for estimating daily ET₀ in Babolsar (with the humid climate) and Rasht (with the per-humid climate) is solar radiation, while for the Gorgan station (with the Mediterranean climate), the daily mean air temperature is the most effective factor for ET₀ estimation.

Furthermore, if the number of climatic variables was increased, the combination of the daily mean air temperature and solar radiation would indicate the optimal structure of the variables at all three stations. Finally, the combination of four variables, daily mean air temperature, mean relative humidity, solar radiation, and wind speed are the optimal combination for ET₀ estimation. The study indicated that the ANFIS method is able to compute better estimating accuracy in terms of different evaluation statistics during both the training and validation phases. ANFIS model-4 performs better than the others in all the study stations. The results of the SVM method during the validation phase were inferior to the results during the training phase. All-inclusive, the results indicated that the data-driven methods are strong tools for modelling daily evapotranspiration time series, and this may give valuable reference for researchers who apply data-driven methods for modelling ET₀. The results of this study can provide a solution for accurate ET₀ estimation in the absence of complete climatological data. In further study, the artificial intelligence methods may be applied in estimating hourly evapotranspiration and comparing it with other empirical equations in different climates.

**AUTHOR CONTRIBUTIONS**

BB suggested the topic and was the main supervisor, and helped in writing and editing the manuscript. AM-D acquired and analyzed the data, performed the computations, and prepared the initial draft. KQ supervised data driven methods analysis and edited the final manuscript. All authors provided critical feedback and helped shape the research, analysis and manuscript.

**ORCID**

Bahram Bakhtiari 🌟 https://orcid.org/0000-0001-6555-4328

**REFERENCES**

Allen, R.G. (1996) Assessing integrity of weather data for reference evapotranspiration estimation. *Journal of Irrigation and Drainage Engineering*, 122, 97–106.

Allen, R.G., Clemmens, A.J., Burt, C.M., Solomon, K. & O’Halloran, T. (2005) Prediction accuracy for project wide evapotranspiration using crop coefficients and reference evapotranspiration. *Journal of Irrigation and Drainage Engineering*, 131, 24–36. [https://doi.org/10.1061/ASCE0733-94372005131:124](https://doi.org/10.1061/ASCE0733-94372005131:124)

Allen, R.G., Pereira, L.S., Raes, D., & Smith, M. (1998) Crop evapotranspiration guidelines for computing crop water requirements, FAO irrigation and drainage paper No. 56. Rome: Food and Agriculture Organization of the United Nations FAO.

Alquraish, M.M., Abuhasel, K.A., Alqahtani, A.S. & Khadr, M. (2021) A comparative analysis of hidden Markov model, hybrid support vector machines, and hybrid artificial neural fuzzy inference system in reservoir inflow forecasting (case study: the King Fahd Dam, Saudi Arabia). *Water*, 13, 1236.

ASCE Task Committee. (2000a) Artificial neural networks in hydrology—I: preliminary concepts. *Journal of Hydrologic Engineering, ASCE*, 5, 115–123.

ASCE Task Committee. (2000b) Artificial neural networks in hydrology—II: hydrological applications. *Journal of Hydrologic Engineering, ASCE*, 5, 124–137.

Bakhtiari, B. & Baghizadeh, A. (2012) Daily penman-Monteith sensitivity analysis in many subclasses climates based on extended DeMartonne classification. *International conference of agricultural engineering, July, 8–12, Valencia, Spain.*

Celestin, S., Qi, F., Li, R., Yu, T. & Cheng, W. (2020) Evaluation of 32 simple equations against the Penman–Monteith method to estimate the reference evapotranspiration in the Hesi corridor, Northwest China. *Water*, 12, 2772.

Chauhan, S. & Shrivastava, R.K. (2009) Performance evaluation of reference evapotranspiration estimation using climate-based methods and artificial neural networks. *Water Resources Management*, 23, 825–837.

Doorenbos, J. & Pruitt, W.O. (1977) Guidelines for prediction of crop water requirements. F.A.O. Irrig. and Drain. Paper No. 24. 2nd ed. Rome, Italy: FAO, pp. 156.

Eslamian, S.S., Abedi-Koupai, J., Amiri, M.J. & Gohari, S.A. (2009) Estimation of daily reference evapotranspiration using support vector machines and, artificial neural networks in greenhouse. *Research Journal of Environmental Sciences*, 3, 439–447.

Ghadiri, H., Dordipour, E., Bybordi, M. & Malakouti, M. (2006) Potential use of Caspian Sea water for supplementary irrigation in northern Iran. *Agricultural Water Management*, 79, 209–224. Islamic Republic of Iran Meteorological Organization. (2021) Meteorological data from synoptic station during years 1987–2016. Available at: [http://irimo.ir/eng/wd/720-Products-Services.html](http://irimo.ir/eng/wd/720-Products-Services.html).

Jang, J.S.R. (1993) ANFIS: adaptive-network-based fuzzy inference systems. *IEEE Transactions on Systems, Man and Cybernetics*, 23, 665–685.

Jiang, W., Niu, Z., Wang, L., Yao, R., Gui, X., Xiang, F. et al. (2022) Impacts of drought and climatic factors on vegetation dynamics.
in the yellow Rivar Basin and Yangtze river basin, China. Remote Sensing, 14, 930.

Keshtegar, B., Kisi, O., Arab, H.G. & Zounemat-Kermani, M. (2017) Subset modeling basis ANFIS for prediction of the reference evapotranspiration. Water Resources Management, 323, 1101–1116.

Kisi, O. (2008) The potential of different ANN techniques in evapotranspiration modelling. Hydrological Processes, 22, 2449–2460.

Kisi, O., Keshtegar, M., Zounemat-Kermani, M., Heddam, S. & Trung, N.T. (2021) Modeling reference evapotranspiration using a novel regression-based method: radial basis M5 model tree. Theoretical and Applied Climatology, 145, 659–659.

Kumar, M., Raghuvanshi, N.S., Singh, R., Wallender, W.W. & Pruitt, W.O. (2002) Estimating evapotranspiration using artificial neural network. Journal of Irrigation and Drainage Engineering, 128(4), 224–233.

Ladmani, I., Houichi, L., Djemili, L., Heddam, S. & Belouz, K. (2014) Estimation of daily reference evapotranspiration ET0 in the north of Algeria using adaptive neuro-fuzzy inference system ANFIS and multiple linear regression MLR models: a comparative study. Arabian Journal for Science and Engineering, 39, 5959–5969.

Lefebvre, M. & Bensalma, F. (2015) Modelling and fore-casting river flows by means of filtered Poisson processes. Applied Mathematical Modelling, 39, 230–243.

L ski, J. & Czogala, E. (1999) A new artificial neural network based fuzzy inference system with moving consequents in if-then rules and selected applications. Fuzzy Sets and Systems, 108, 289–297.

Lin, J.Y., Cheng, C.T. & Chau, K.W. (2006) Using support vector machines for long-term discharge prediction. Hydrological Sciences Journal, 51, 599–612.

Lopez-Urrea, R., de Santa, M., Olalla, F., Fabeiro, C. & Moratalla, A. (2006) An evaluation of two hourly reference evapotranspiration equations for semiarid conditions. Agricultural Water Management, 86, 277–282.

Ma, N., Zhang, Y.S., Xu, C.Y. & Szilagyi, J. (2015) Modeling actual evapotranspiration with routine meteorological variables in the data-scarce region of the Tibetan plateau: comparisons and implications. Journal of Geophysical Research, 120, 1638–1657.

Majhi, B. & Naidu, D. (2021) Differential evolution based radial basis function neural network model for reference evapotranspiration estimation. SN Applied Sciences, 3, 56.

Murat, C. (2011) Evapotranspiration estimation by two different neuro-fuzzy inference systems. Journal of Hydrology, 398, 292–302.

Mustafa, A., Rienow, A., Saadi, I., Cools, M. & Teller, J. (2018) Comparing support vector machines with logistic regression for calibrating cellular automata land use change models. European Journal of Remote Sensing, 51, 391–401.

Narvekar, M. & Fargose, P. (2015) Daily weather forecasting using artificial neural network. International Journal of Computers and Applications, 11, 9–13.

Nash, J. & Sutcliffe, J. (1970) River flow forecasting through conceptual models part I–A discussion of principles. Journal of Hydrology, 10, 282–290.

Pal, M. & Deswal, S. (2009) M5 model tree-based modelling of reference evapotranspiration. Hydrological Processes, 23, 1437–1443.

Pulido-Calvo, I. & Gutiérrez-Estrada, J.C. (2009) Improved irrigation water demand forecasting using a soft-computing hybrid model. Biosystems Engineering, 102, 202–218.

Quinlan, J.R. (1992) Learning with continuous classes. Proceedings of Australian Joint Conference on Artificial Intelligence. Singapore: World Scientific Press, 92, pp. 343–348.

Rahimi, J., Ebrahimpour, M. & Khalili, A. (2013) Spatial changes of extended De Martonne climatic zones affected by climate change in Iran. Theoretical and Applied Climatology, 112, 409–418.

Rahimikhhoob, A. (2014) Comparison between M5 model tree and neural networks for estimating reference evapotranspiration in an arid environment. Water Resources Research, 28, 657–669.

Raza, A., Shoaib, M., Faiz, M.A., Baig, F., Khan, M.M., Ullah, M.K. et al. (2020) Comparative assessment of reference evapotranspiration estimation using conventional method and machine learning algorithms in four climatic regions. Pure and Applied Geophysics, 177, 4479–4508.

Roy, D.K., Lal, A., Sarker, K.K., Saha, K.K. & Datta, B. (2021) Optimization algorithms as training approaches for prediction of reference evapotranspiration using adaptive neuro fuzzy inference system. Agricultural Water Management, 225, 107003.

Sekertekin, A., Bilgili, M., Arsalan, N., Yildirim, A., Celebi, K. & Ozbek, A. (2021) Short-term air temperature prediction by adaptive neuro-fuzzy inference system (ANFIS) and long short-term memory (LSTM) network. Meteorology and Atmospheric Physics, 133, 943–959.

Shcherbakov, M.V., Brebels, A., Shcherbakova, N.L., Tyukov, A.P., Janovsky, T.A. & Kamaev, V.A.A. (2013) Survey of forecast error measures. World Applied Sciences Journal, 24, 171–176.

Shiri, J., Sadraddini, A.A., Nazemi, A.H., Kisi, O., Martí, P., Fard, A.F. et al. (2013) Evaluation of different data management scenarios for estimating daily reference evapotranspiration. Hydrology Research, 44, 1058–1070.

Smith, M., Allen, R.G., Monteith, J.L., Perrier, A., Pereira, L. & Segeren, A. (1992) Expert consultation on revision of FAO methodologies for crop water requirements. Rome: Land and Water Development Division, Food and Agriculture Organization of the United Nations.

Solomatine, D.P. & Xue, Y. (2004) M5 model trees and neural networks: application to flood forecasting in the upper reach of the Huai river in China. Journal of Hydrologic Engineering, 9, 491–501.

Tezel, G. & Buyukyildiz, M. (2016) Monthly evaporation forecasting using artificial neural networks and support vector machines. Theoretical and Applied Climatology, 124, 69–80.

Valipour, M. (2017) Analysis of potential evapotranspiration using limited weather data. Applied Water Science, 7, 187–197.

Vapnik, V. (1998) Statistical learning theory. New York: Wiley.

Walls, S., Binns, A.D., Levison, J. & MacRitchie, S. (2020) Prediction of actual evapotranspiration by artificial neural network models using data from a Bowen ratio energy balance station. Neural Computing and Applications, 32, 14001–14018.

Wen, C.G. & Lee, C.S. (1998) A neural network approach to multi objective optimization for water quality management in a river basin. Water Resources Research, 34, 427–436.

Willmott, C.J. (1982) Some comments on the evaluation of model performance. Bulletin of the American Meteorological Society, 63, 1309–1313.
Wu, C.L., Chau, K.W. & Li, Y.S. (2008) River stage prediction based on a distributed support vector regression. *Journal of Hydrology*, 358, 96–111.

Xu, L.H., Shi, Z.J., Wang, Y.H., Zhang, S.L., Chu, X.Z., Yu, P.T. et al. (2015) Spatio-temporal variation and driving forces of reference evapotranspiration in Jing River basin, Northwest China. *Hydrological Processes*, 29, 4846–4862.

Yang, C.C., Prasher, S.O. & Lacroix, R. (1996) Applications of artificial neural networks to simulate water-table depths under sub-irrigation. *Canadian Water Resources Journal*, 21, 27–44.

**How to cite this article:** Bakhtiari, B., Mohebbi-Dehaghani, A., & Qaderi, K. (2022). Comparative analysis of data-driven methods for daily reference evapotranspiration estimation of Southern Caspian Sea. *Meteorological Applications*, 29(4), e2091. [https://doi.org/10.1002/met.2091](https://doi.org/10.1002/met.2091)