

## Research Article

# Application of Artificial Intelligence Models for Evapotranspiration Prediction along the Southern Coast of Turkey

Mohammed Majeed Hameed <sup>1</sup>, Mohamed Khalid AlOmar <sup>1</sup>, Siti Fatin Mohd Razali,<sup>2</sup>  
Mohammed Abd Kareem Khalaf,<sup>3</sup> Wajdi Jaber Baniya,<sup>4</sup> Ahmad Sharafati,<sup>5</sup>  
and Mohammed Abdulhakim AlSaadi <sup>6</sup>

<sup>1</sup>Department of Civil Engineering, Al-Maaref University College, Ramadi, Iraq

<sup>2</sup>Department of Civil Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

<sup>3</sup>Iraqi Federal Board of Supreme Audit, Baghdad, Iraq

<sup>4</sup>Department of Civil Engineering, Cairo University, Cairo, Egypt

<sup>5</sup>Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

<sup>6</sup>National Chair of Materials Science and Metallurgy, University of Nizwa, Nizwa, Oman

Correspondence should be addressed to Mohamed Khalid AlOmar; mohd.alomar@yahoo.com

Received 18 August 2020; Revised 1 October 2020; Accepted 6 August 2021; Published 24 August 2021

Academic Editor: Honglei Xu

Copyright © 2021 Mohammed Majeed Hameed et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Reference evapotranspiration ( $ET_o$ ) is one of the most significant factors in the hydrological cycle since it has a great influence on water resource planning and management, agriculture and irrigation management, and other processes in the hydrological sector. In this study, an efficient and local predictive model was established to forecast the monthly mean  $ET_o$  over Turkey based on the data collected from 35 locations. For this purpose, twenty input combinations including hydrological and geographical parameters were introduced to three different approaches called multiple linear regression (MLR), random forest (RF), and extreme learning machine (ELM). Moreover, in this study, large investigation was done, involving the establishment of 60 models and their assessment using ten statistical measures. The outcome of this study revealed that the ELM approach achieved high accurate estimation in accordance with the Penman–Monteith formula as compared to other models such as MLR and RF. Moreover, among the 10 statistical measures, the uncertainty at 95% ( $U_{95}$ ) indicator showed an excellent ability to select the best and most efficient forecast model. The superiority of ELM in the prediction of mean monthly  $ET_o$  over MLR and RF approaches is illustrated in the reduction of the  $U_{95}$  parameter to 49.02% and 34.07% for RF and MLR models, respectively. Furthermore, it is possible to develop a local predictive model with the help of computer to estimate the  $ET_o$  using the simplest and cheapest meteorological and geographical variables with acceptable accuracy.

## 1. Introduction

**1.1. Background.** Global warming has become a great concern of researchers and world leaders. It is well known that the Earth surface temperature is increasing significantly during the last decades [1, 2]. Water storage, hydrological cycles, and, consequently, water availability are directly affected by global warming [3–5]. Thus, one of the most essential indicators of climate change is the referenced

evapotranspiration ( $ET_o$ ), which is considered as the most complicated element in the hydrological cycle [6–8].

$ET_o$  mainly occurs due to two complicated processes. The first is when water evaporates from the surface of the soil, lakes, rivers, etc., and this process is called physical evaporation. The second process is the transpiration phenomenon in crops and plants, which is called biological transpiration [9, 10]. Undeniably, the evaporation process requires energy to convert water from the liquid phase to the

vapor phase. Therefore, the main parameters that affect the  $ET_o$  process is the sun radiation, wind speed and direction, air temperature, and humidity [9–11]. In conclusion, the  $ET_o$  also represents the link between surface energy and carbon cycle [12, 13]. Based on the stated literature, a precise measurement and prediction of  $ET_o$  is essential for quantifying surface energy and water reserves worldwide [14–16]. Thus, providing accurate models for weather and climate change diagnosis is crucial [14, 17–21].

*1.2. Literature Review.* Due to the significant effect of  $ET_o$  on climate change, Earth temperature, crops and plants, water management, and runoff quantity, many researchers have studied the  $ET_o$  prediction over the last decades [9, 22–24]. The Penman–Monteith ( $P - M$ ) is the most widely used model, and it is considered as a physical  $ET_o$  model as it is an approximate linearized solution governing energy balance, thermodynamic state, vertical heat, and water-vapor diffusion [9, 23]. However,  $P - M$  requires many meteorological data to be applied, which can be considered as a drawback for this equation [25, 26]. In any case, there are many models applied to estimate the  $ET_o$  around the world. Examples of such are the constant heat method by including heat pulse [27, 28] and the Shuttleworth–Wallace S-W method to estimate the transpiration from plants [29–31]. It is worth mentioning that the number of empirical equations for modeling the evaporation has exceeded 100 due to the importance of  $ET_o$  measurements and the variety of meteorological data around the world. Therefore, it is impossible to decisively compare these models [32–34].

Recently, the development of artificial intelligence has received significant attention from communities in the hydrological and environmental sciences, including water treatment [35, 36], hydrology [37–41], water reservoir optimization [42–44], remote sensing applications [45, 46], etc. Consequently, due to the highly nonlinear characteristics associated with the  $ET_o$  data, AI technology presented a suitable modeling approach to solve many issues with the empirical equations that has been used before [47]. Kumar et al. utilized the artificial neural network (ANN) in 2002 for predicting  $ET_o$ , where different ANN architectures were implemented for evaporation simulation. The radial neural network yielded the best results for evaporation simulation, and it calculates the number of layers and neurons based on a trial and error process. [48]. Many researchers have followed his footprint in predicting  $ET_o$  [49–52]. In addition, an adaptive network-based fuzzy inference system (ANFIS) has been used to predict  $ET_o$  [53, 54]. The ANFIS and ANN techniques and empirical equations were used in the evapotranspiration field, and it was found that the ANFIS and ANN methods were much better than the empirical equations. [55, 56]. However, it is well established that ANN models easily get stuck in a local minimum, and, therefore, recent studies have employed new models adopting other AI techniques for  $ET_o$  modeling [47]. Many approaches have been utilized for this purpose, including support vector machine SVM [53, 57, 58]. The SVM is well known to have a

basic form, but one of the drawbacks of SVM is the unknown parameter [59]. Another approach in the field of simulation  $ET_o$  data is genetic programming (GP). This approach consists of measurement programming capable of obtaining input data and producing a nonlinear interaction between data to determine the outputs [55, 60–62].

In order to enhance the ability of AI models, many algorithms were associated with different AI methodologies, including the use of wavelet transform regression model [63], wavelet coupled with ANN [64, 65], and wavelet enhanced extreme learning machine [66]. Others utilize random forest (RF) algorithm to enhance the AI techniques. Due to its success over a variety of datasets, high precision estimation, a small range of user-defined parameters, the ability to estimate relative value of the variables, and its ability to preclude overfitting, the RF approach has become extremely popular in recent years [26, 34, 67, 68]. Recently, due to its higher efficiency and much quicker calculation speed, a newly proposed machine learning technology called the extreme learning machine (ELM) has confirmed it to be a promising  $ET_o$  estimation tool [69]. First, Abdullah et al. (2015) used ELM to forecast  $ET_o$  at three Iraqi stations and concluded that the ELM model is highly efficient and computerized at high generalization speeds [70, 71]. Ever since, the ELM for  $ET_o$  predictions has been used by many studies in different climate environments [72–74]. To the best of the authors' knowledge, all models presented in the literature were established to simulate the evapotranspiration using a single model for each location or case study. Furthermore, some researchers employed modern techniques and used different case studies, but they could not produce a general model that could take into account more than one case study. In this study, an effort is being made to include a robust modeling methodology using a variety of locations along the southern coast of Turkey to create a comprehensive general model to forecast  $ET_o$ . The other objective of the study is to predict mean monthly  $ET_o$  from limited data, which can be easily available.

*1.3. Motivation of the Study.* Due to the significance of  $ET_o$ , there are numerous studies that have been conducted to estimate it, such as AI techniques and empirical models. In general, these approaches achieved a satisfactory success. However, creating one robust model for conducting local  $ET_o$  throughout a specific country based on data collected from different sites is still considered as a challenging issue that needs to be addressed. Moreover, recognizing the statistical measures that are effectively used to assess the feasibility of a certain model are also very significant in the selection of the best predictive model. Thus, in this study, a broad investigation is performed using three different approaches, multiple linear regression (MLR), random forest (RF), and extreme learning machine (ELM), based on twenty combinations of meteorological and geographical indicators constituting 60 predictive models. Moreover, there were 10 efficient statistical measures employed to assess the accuracy of the performance for each model separately in accordance

with the Penman–Monteith equation. Although the formula of Penman–Monteith is well-known in the prediction of  $ET_o$ , it does pose some issues regarding the measurement of some factors, which may not be relevant to every site such as solar radiation and sensible heat flux into the soil. Therefore, a robust local model is established in this study, which can efficiently predict the local mean monthly  $ET_o$  over the southern coast of Turkey using conveniently and inexpensively measurable parameters. Furthermore, the evaluation of the outcomes of the models will be further assessed and validated against the actual  $ET_o$  values, which are calculated by the Penman–Monteith equation.

## 2. Case Study and Data Collection

In this study, the data collected from 35 meteorological stations in Turkey are used. The geographical location of these stations covers large area located between a latitude from  $36^\circ$  to  $38^\circ$ . Figure 1 shows the location of each meteorological station separately. It can be seen that the majority of these stations are located in the south of Turkey on the coasts of the Mediterranean Sea. The data collected from the general directorate of Turkish state meteorological includes several long-term monthly meteorological data such as temperature, humidity, wind speed, and reference evapotranspiration. Furthermore, the dataset comprises of the long-term mean monthly variables covering the period from 1975 to 2010. The highest temperature is recorded at the Mut station at  $46.7^\circ\text{C}$ , while the lowest temperature is recorded at Goksun station at  $-33.5^\circ\text{C}$ . Figure 2 showed the long-term monthly mean ( $ET_o$ ), overall in Turkey. It can be observed that the highest value of ( $ET_o$ ) occurred in July followed by June, August, and May, consecutively.

It is worth mentioning that Turkey primarily has a complex climate due to its location and topography. The Mediterranean climate in southern Turkey is predominant with warm and dry summers and wet and moderate to cold winters. Continental weather is predominant in central Turkey with warm and dry summers and cold winters. The oceanic atmosphere in northern Turkey is seen through warm and rainy summers and cold and wet winters. This study ensures that the stations chosen to measure  $ET_o$  are spread nearly uniformly across southern Turkey in order to determine spatial differences in mean monthly  $ET_o$  values and their time characteristics.

Due to the lack of  $ET_o$  data, the FAO embraced the PM equation as a standard methodology for calculating the reference  $ET_o$ . The FAO56-PM can be used on hourly or daily scales to supply the data needed for machine learning approaches. The equation is suggested for hourly time measures as expressed in the following equation [7, 75]:

$$ET_o = \frac{0.408\Delta(R_N - G) + \gamma(900/(T + 273))U_2(e_a - e_d)}{\Delta + \gamma(1 + 0.34U_2)}, \quad (1)$$

where  $ET_o$  is the reference evapotranspiration in (mm/day),  $\Delta$  represents the slope of the saturation vapor pressure function at air temperature  $T$  in ( $\text{kPa} \cdot \text{C}^{-1}$ ),  $R_N$  is the net

solar radiation in ( $\text{MJ} \cdot \text{m}^{-2} \cdot \text{day}^{-1}$ ),  $G$  is the soil heat flux density in ( $\text{MJ} \cdot \text{m}^{-2} \cdot \text{day}^{-1}$ ),  $\gamma$  is the psychrometric constant in ( $\text{kPa} \cdot \text{C}^{-1}$ ),  $T$  is the mean air temperature in ( $^\circ\text{C}$ ),  $U_2$  is the average 24-hour wind speed at 2m height above the ground surface in ( $\text{m} \cdot \text{s}^{-1}$ ),  $e_a$  is the saturation vapor pressure in (kPa), and  $e_d$  is the actual vapor pressure in (kPa). Finally, it is important to mention that the dataset is subjected to the normalization process (between 0 and 1), for all input variables and their appropriate targets. This process is very important in order to boost the predictive model performance. Later, the data are simulated using three modeling approaches, namely, RF, ELM, and MLR.

## 3. Methodology

**3.1. Random Forest.** Random Forests (RF) is an algorithm that manages regression issues of high dimensions. This method is tree-based, where all trees have random variables in the selection, and the forest is established from several trees of regression and is clustered together [34, 76]. The tree is chosen as a random subset of variables that will be used to determine the result of the prediction. Whereas, in the random forest learning process, two significant parameters are identified: the first one is the number of trees ( $n_{\text{tree}}$ ) and the second one is the number of variables in each division ( $m_{\text{try}}$ ). After fitting single tree into the ensemble (bagging procedure), the final decision is made by averaging the output. The bias between the bagged trees is equal to that of the single tree, while the variance is reduced as the correlation between the trees is reduced [77].

For regression-based RF formation, in order to get the tree predictor  $h(X, \Theta)$  numerical value, the process starts by growing trees on the bias of a random vector ( $\Theta$ ). For any given numerical estimator, the mean squared generalization error can be expressed as follows [78]:

$$E_{(X,Y)} = (Y - h(X))^2. \quad (2)$$

The RF predictor is generated by taking an average over  $j$  of a single tree. Here, in this respect, the following theorems line up:

**Theorem 1.** *By expanding the number of trees in the forest, the error will then be expressed as follows:*

$$E_{(X,Y)} \left( \text{Yav}_j h(X, \Theta_j) \right)^2 \Theta_{X,Y} (Y - E_\Theta h(X, \Theta))^2. \quad (3)$$

This equation's right hand demonstrates the generalization error of the forest. Similarly, the average tree generalization error can be found from the following equation:

$$PE * (\text{tree}) = E_\Theta E_{X,Y} (Y - h(X, \Theta))^2. \quad (4)$$

**Theorem 2.** *If we assume  $E_Y = E_X h(X, \Theta)$  for every  $\Theta$ , then*

$$PE * (\text{forest}) \leq \bar{p} \cdot PE(\text{tree}), \quad (5)$$

where  $\bar{p}$  represent the weight correlation [76].

**3.2. Extreme Learning Machine (ELM).** The ELM, an advanced learning algorithm, is a machine learning modeling approach assembled with single-layer feedforward neural network (SLFN) [79]. ELM's primary strength is that the weight of the input parameters is determined arbitrarily, while the output parameter weights are measured analytically using the Moore–Penrose approach (Huang et al., 2006). The SLFN function combines the hidden node additives and activation function, which can be represented mathematically as follows:

$$f(x) = \sum_{i=1}^L \beta_i G(a_i, b_i, x), \quad (6)$$

where  $f(x)$  represents the ELM model output function;  $x$  represents the input variable,  $a_i$  and  $b_i$  represent the learning parameters hidden nodes, and  $L$  stands for the number of hidden nodes. The  $\beta_i$  governs the connecting weight to the  $i$ -hidden node between the output nodes. The output vector  $G(a_i, b_i, x)$  implies the output of the hidden node. The sigmoid activation solves the additional hidden node as follows:

$$G(a_i, b_i, x) = g(a_i \cdot x + b_i), \quad (7)$$

where  $a_i$  and  $b_i$  represent weight and basis values for each  $i$ -node in its hidden layer. In the algorithm of ELM, the weight of the input layer and bias values are randomly generated. Figure 3 showed the basic structure of the ELM.

For the input and output variables, the arbitrary distinct sample  $N$  is denoted as  $(x_i, y_i)$ . Based on the above findings, equation (6) can be rewritten as follows:

$$H\beta = T, \quad (8)$$

where the Hussain matrix ( $H$ ) is represented as

$$H = \begin{bmatrix} G(a_1, b_1 + x_1) & \cdots & g_L(a_L, b_L, x_1) \\ \vdots & \ddots & \vdots \\ G(a_N, b_N + x_N) & \cdots & g_L(a_L, b_L, x_N) \end{bmatrix}_{N \times L}. \quad (9)$$

In addition,

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m}, \quad (10)$$

$$T = \begin{bmatrix} Y_1^T \\ \vdots \\ Y_L^T \end{bmatrix}_{N \times m}.$$

$G$  represents the activation function and in this current study, log-sigmoid transfer function is used, which is expressed as

$$G(y) = \frac{1}{1 + \exp(-y)}, \quad (11)$$

where  $y = ax + b$  as indicated in equation (6).

The ELM model with independent  $L$  sets, which can result in zero learning errors, can be trained in the form of SLFN [80]. In addition, even with fewer hidden neurons ( $L$ ) than the number of separate samples ( $N$ ), the ELM will specify random parameters for the hidden nodes. Ultimately, the output weights can be calculated by means of pseudoinverse of  $H$ , which makes a limited error range  $\varepsilon > 0$ . The hidden node parameters (weights and bias) are given random values during the training phase [81].

**3.3. Multiple Linear Regression.** In the form of the local climatic parameters of minimum and maximum temperature, relative humidity, wind speed and sun hours, the  $ET_o$  data can be formulated using multiple linear regression (MLR) models. Herein, the  $ET_o$  is calculated by MLR. The MLR can be employed to describe the relationship between dependent and independent variables as a multivariate statistical tool described by the following equation:

$$Y = \alpha_o + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_j X_j, \quad (12)$$

where the response variables are represented by  $Y$ , which is also the predicted mean monthly  $ET_o$ . The independent variables and the predictors are represented by  $X_1, X_2, X_n$ , and the coefficients of regression are represented by  $\alpha_o, \alpha_1, \alpha_2, \dots, \alpha_j$ , which can be acquired by the following equations [82]:

$$e = \sum_{i=1}^n (Y - y_i)^2,$$

$$\frac{\partial e}{\partial \alpha_o} = 0,$$

$$\frac{\partial e}{\partial \alpha_1} = 0,$$

$$\frac{\partial e}{\partial \alpha_2} = 0,$$

$$\frac{\partial e}{\partial \alpha_j} = 0,$$
(13)

where the error of the estimated and real values of  $ET_o$  is represented by  $e$  and  $y_i$ , respectively.

**3.4. Model Evaluation.** The selection process of the best predictive model is of great importance to achieve high accurate predictions. Therefore, in this study, ten statistical parameters have been used to assess the performances of each predictive approach [83, 84]. The quantitative parameters are as follows:

- (i) Mean absolute error (MAE). It can be expressed by assuming the absolute errors divided by the

number of total observations. This indicator is widely used in water resources and hydrological sectors to assess the predictive models because it provides significant information on how closely the simulated data points match with the actual ones. The mathematical expression of MAE is shown in the following equations [85, 86]:

$$e = ET_{oi,m} - ET_{oi,p}, \quad (14)$$

where  $e$  is the forecasted error,

$$\text{MAE (mm/month)} = \sum_{i=1}^n \frac{|e_i|}{n}. \quad (15)$$

- (ii) Root mean square error (RMSE). It is a statistical parameter often used to compare the forecasting errors of several models. The lower RMSE value usually points out to better predictions. The RMSE can be derived using the following equation [87, 88]:

$$\text{RMSE (mm/month)} = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2}. \quad (16)$$

- (iii) Mean absolute relative error (MARE). This indicator can be expressed as an absolute relative computer error (difference between actual and forecasted points). When the MARE parameter is represented as a percentage, it is called the mean absolute relative error (MARE), which is expressed by the following equation:

$$\text{MARE} = \frac{1}{n} \sum_{i=1}^n \frac{e_i}{ET_{oi,m}}. \quad (17)$$

- (iv) Root mean square relative error (RMSRE) is expressed as

$$\text{RMSRE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left( \frac{e_i}{ET_{oi,m}} \right)^2}. \quad (18)$$

- (v) Relative root mean square error (RRMSE) can be calculated by dividing the RMSE criteria over the mean of actual data points. This parameter is very vital in assessing the accuracy of a model. In accordance to [89], the model is considered excellent if  $\text{RRMSE} > 10\%$ , good if  $10\% < \text{RRMSE} < 20\%$ , fair if  $20\% < \text{RRMSE} < 30\%$ , and last, the model can be considered poor if  $\text{RRMSE} > 30\%$ . The mathematical expression of the RRMSE is depicted as follows:

$$\text{RRMSE (\%)} = \sqrt{\frac{\text{RMSE}}{1/n \sum_{i=1}^n ET_{oi,m}}}. \quad (19)$$

- (vi) Mean bias error (MBE). It discloses the tendency of a model and explains whether it overestimates the data or underestimates them and is expressed by the following equation:

$$\text{MBE (mm/month)} = \frac{1}{n} \sum_{i=1}^n e_i. \quad (20)$$

- (vii) Correlation of coefficient (CC). The CC is a significant factor that can be utilized to efficiently discover the robustness of the relationship between predicted and simulated data points. The mathematical expression of CC is shown in the following equation [82, 90]:

$$\text{CC} = \frac{\sum_{i=1}^n (ET_{oi,m} - \overline{ET_{oi,m}})(ET_{oi,p} - \overline{ET_{oi,p}})}{\sum_{i=1}^n (ET_{oi,m} - \overline{ET_{oi,m}})^2 \sum_{i=1}^n (ET_{oi,p} - \overline{ET_{oi,p}})^2}. \quad (21)$$

- (viii) Maximum absolute relative error (erMAX) is expressed as

$$\text{erMAX} \left( \left| \frac{e_i}{ET_{oi,m}} \right| \right). \quad (22)$$

- (ix)  $t$ -statistic ( $t$ -stat). This statistical test is beneficial for validating and testing the broadband models [91]. As  $t$ -stat indicator approaches zero, the desired model would be achieved and it is expressed by the following equation:

$$t\text{-stat} = \sqrt{\frac{(n-1)\text{MBE}^2}{\text{MBE}^2 - \text{RMSE}^2}}, \quad (23)$$

where  $n$  is the total number of observations, and  $ET_{oi,m}$   $ET_{oi,p}$  are the  $i_{th}$  actual and estimated mean monthly reference of evapotranspiration, respectively.

- (x) Uncertainty at 95% ( $U_{95}$ ). This quantitative criterion is very efficient in terms of selecting an efficient predictive model among several models. The  $U_{95}$  indicator provides very useful information on the deviation of a certain model [91]. The  $U_{95}$  can be calculated by the following equation:

$$U_{95} = 1.96(\text{SD}^2 + \text{RMSE}^2)^{1/2}, \quad (24)$$

where SD represents the standard deviation of the difference between true and simulated data points. The value 1.96 is a coverage factor corresponding to 95% confidence level. Last, for visual evaluation, boxplots and scatter plots are presented in the Results and Discussion sections. The coefficient of determination ( $R^2$ ) is calculated based on the following equation:

$$R^2 = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (ET_{oi,p} - \overline{ET_{oi,p}})^2}. \quad (25)$$

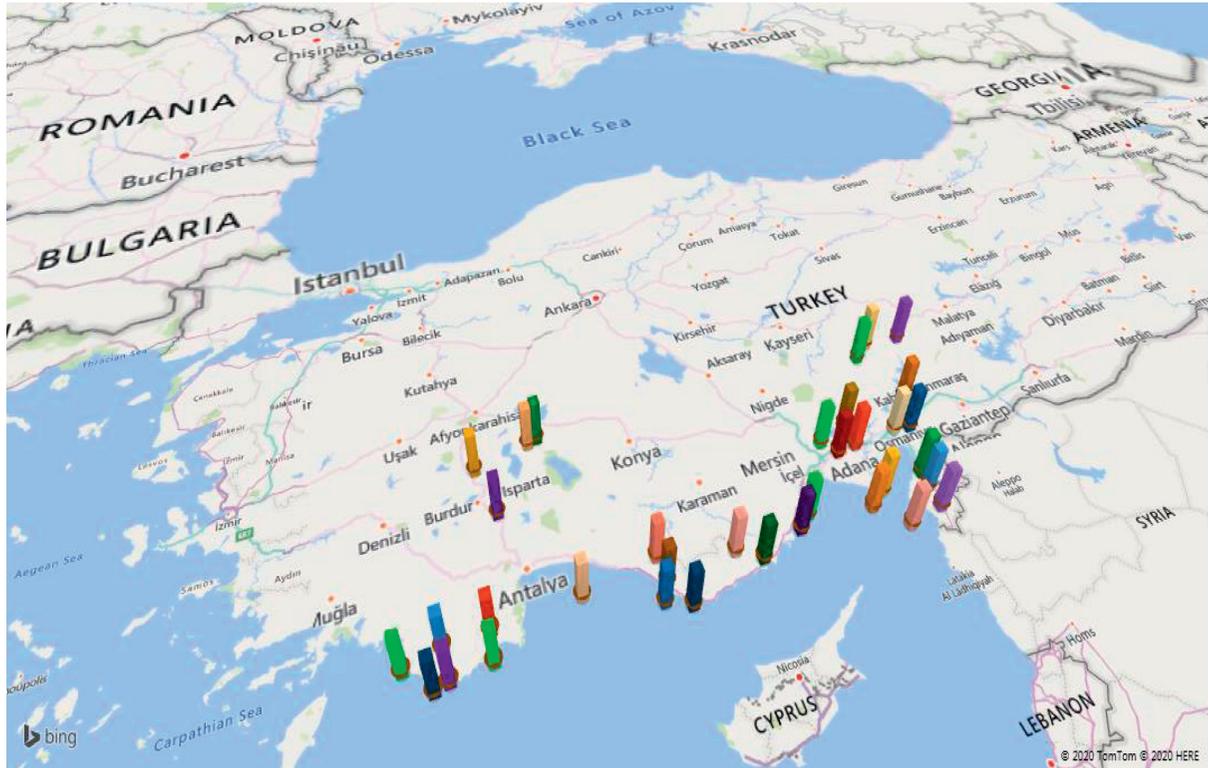


FIGURE 1: The distribution of meteorological stations in Turkey.

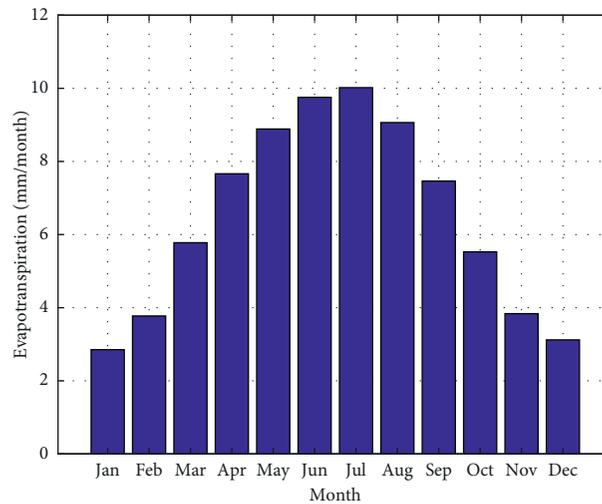


FIGURE 2: Distribution of the mean monthly ( $ET_0$ ) over Turkey.

**3.5. Model Development.** To accurately predict the monthly mean  $ET_0$ , three different models were assigned called MLR, RF, and ELM approaches. In accordance with the nature of dataset collected from different locations and sites, and in addition to having different characters, it is a significant and difficult task to create a reliable predictive model. Therefore, the current work is carried out by establishing 20 different input groups for each predictive approach as shown in Table 1. The dataset is randomly divided into phases called training set and testing set. The

training set is used for calibration process and model construction, while the testing set is used for examining the performance accuracy of each candidate model. Table 2 shows the statistical description of the variables used in this study. In addition, Table 3 gives information about the Pearson correlation coefficient between all variables and  $ET_0$  used in the suggested case study. As shown in the table, the maximum temperature variable has the highest correlation with  $ET_0$  for both training and testing datasets. The table shows that the longitude variable has the lowest correlation

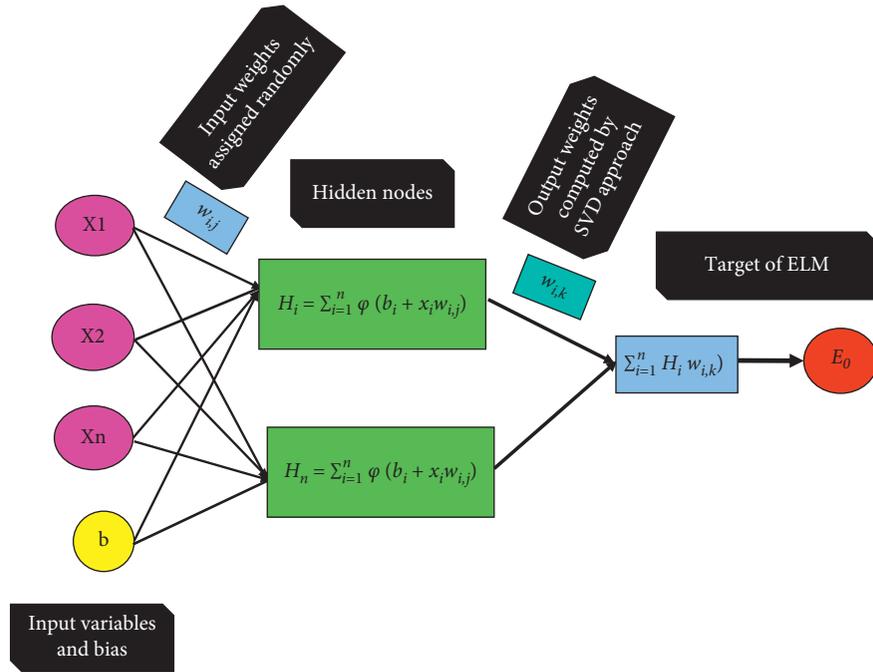


FIGURE 3: Fundamental structure of ELM.

TABLE 1: Statistical description of the predictors and evapotranspiration that used in this current study.

Variables		Unit	Max	Min	Mean	Std	Skewness	Kurtosis
Latitude	Training	—	38.21	36.05	36.8	0.69	0.89	2.44
	Testing		38.21	36.05	36.87	0.75	0.73	2.05
	Overall		38.21	36.05	36.82	0.7	0.85	2.34
Longitude	Training	—	37.11	29.07	33.37	2.57	-0.24	1.56
	Testing		37.11	29.07	33.8	2.58	-0.47	1.68
	Overall		37.11	29.07	33.48	2.58	-0.29	1.58
Elevation	Training	m	1552	3	369.35	489.37	0.97	2.33
	Testing		1552	3	426.25	528.16	0.77	1.94
	Overall		1552	3	383.57	499.32	0.92	2.22
$T_{\max}$	Training	$^{\circ}\text{C}$	46.7	12.5	32.65	7.54	-0.44	2.32
	Testing		45.4	14.4	31.7	8.01	-0.17	1.84
	Overall		46.7	12.5	32.41	7.67	-0.37	2.17
$T_{\min}$	Training	$^{\circ}\text{C}$	20.6	-33.5	1.46	10.27	-0.68	3.65
	Testing		17	-29.4	-0.11	11.03	-0.67	3.15
	Overall		20.6	-33.5	1.07	10.47	-0.68	3.52
$T_{\text{mean}}$	Training	$^{\circ}\text{C}$	30.2	-3.8	16.78	7.81	-0.31	2.35
	Testing		28.9	-0.5	15.73	8.25	-0.16	2.01
	Overall		30.2	-3.8	16.52	7.93	-0.27	2.25
$W_s$	Training	m/sec	5.6	0.7	1.97	0.76	1.46	6.2
	Testing		5.1	0.7	2.15	0.93	1.26	4.35
	Overall		5.6	0.7	2.02	0.81	1.44	5.68
$\text{RH}_{\max} \%$	Training	—	88.4	53.8	73.33	7.62	-0.39	2.52
	Testing		90.1	55.8	75.17	7.19	-0.44	2.8
	Overall		90.1	53.8	73.79	7.55	-0.41	2.59
$\text{RH}_{\min} \%$	Training	—	38	0	7.76	5.67	1.43	6.47
	Testing		26	1	7.81	5.44	0.77	3.16
	Overall		38	0	7.77	5.61	1.28	5.75
$\text{ET}_O$	Training	mm/month	15.98	1.5	6.47	2.79	0.27	2.59
	Testing		14.8	1.45	6.5	2.98	0.26	2.34
	Overall		15.98	1.45	6.48	2.84	0.27	2.52

TABLE 2: Different input combinations.

Combinations	Input parameters
C1	$E_O = f(\text{Latitude, Longitude, Elevation, } T_{\max}, T_{\min}, T_{\text{mean}}, \text{Ws, RH}_{\max}, \text{RH}_{\min})$
C2	$E_O = f(\text{Latitude, } T_{\max}, T_{\min}, T_{\text{mean}}, \text{Ws, RH}_{\max}, \text{RH}_{\min})$
C3	$E_O = f(\text{Longitude, } T_{\max}, T_{\min}, T_{\text{mean}}, \text{Ws, RH}_{\max}, \text{RH}_{\min})$
C4	$E_O = f(\text{Elevation, } T_{\max}, T_{\min}, T_{\text{mean}}, \text{Ws, RH}_{\max}, \text{RH}_{\min})$
C5	$E_O = f(\text{Latitude, } T_{\max}, \text{Ws, RH}_{\max})$
C6	$E_O = f(\text{Longitude, } T_{\max}, \text{Ws, RH}_{\max})$
C7	$E_O = f(\text{Elevation, } T_{\max}, \text{Ws, RH}_{\max})$
C8	$E_O = f(\text{Latitude, } T_{\min}, \text{Ws, RH}_{\min})$
C9	$E_O = f(\text{Longitude, } T_{\min}, \text{Ws, RH}_{\min})$
C10	$E_O = f(\text{Elevation, } T_{\min}, \text{Ws, RH}_{\min})$
C11	$E_O = f(\text{Latitude, } T_{\text{mean}}, \text{Ws})$
C12	$E_O = f(\text{Longitude, } T_{\text{mean}}, \text{Ws})$
C13	$E_O = f(\text{Elevation, } T_{\text{mean}}, \text{Ws})$
C14	$E_O = f(T_{\max}, \text{Ws, RH}_{\max})$
C15	$E_O = f(T_{\min}, \text{Ws, RH}_{\max})$
C16	$E_O = f(T_{\text{mean}}, \text{Ws, RH}_{\max})$
C7	$E_O = f(T_{\text{mean}}, \text{Ws, RH}_{\min})$
C18	$E_O = f(T_{\max}, \text{Ws, Latitude})$
C19	$E_O = f(T_{\max}, \text{Ws})$
C20	$E_O = f(T_{\max})$

coefficient with an  $ET_o$  of 0.09, while the maximum temperature is significantly correlated with  $ET_o$  with an  $R$  of 0.841.

The process of selecting a more accurate model is relatively tough with ten statistical matrices and three different approaches including several input combinations. Thus, the assessment process is carried out based on two stages. First, during the training set, from each approach, the best three different models (in total nine models) are being selected. In the second step, it is crucial to monitor the performance of those models, which are selected throughout the training phase, during the testing phase and to select the three most efficient models for each approach. Finally, this process would provide much information about each adopted approach, apart from providing a clear and realistic impression of the performance of each predictive model separately.

Furthermore, a robust and effective performance measure is used to assess the capability of each model. Among these statistical metrics, the uncertainty at 95% ( $U_{95}$ ) has been used to assess the performance accuracy for each model and subsequently recognize the best predictive model. Figure 4 describes the prediction process of  $ET_o$  along the southern coast of Turkey. Last, it is imperative to emphasize that all input variables and their corresponding targets are normalized between 0 and 1. This process is very important to enhance the effectiveness of the predictive models [92, 93]. All models are developed using MATLAB 2017a.

#### 4. Results and Discussion

This section of the study is dedicated to illustrating the forecast results obtained for mean monthly  $ET_o$  over Turkey via three different predictive models, namely, MLR, RF, and ELM. Twenty scenarios have been presented including different input variables and are introduced to the mentioned models. 60 predictive models are assessed and

validated against the Penman–Monteith equation using ten efficient statistical indicators and graphical presentations. A qualified model is one that meets the requirements of most of the mentioned statistical parameters. The dataset collected from 35 stations is divided randomly into two sets: the training phase (75%) is used to calibrate the models, and the rest of the data is used for validation purposes.

In terms of quantitative assessment, Table 4 exhibits further information about the performances of the proposed 60 different models based on different input variables through the training phase. The three predictive modeling approaches have achieved different accuracies in accordance with statistical measures. Moreover, it is difficult to rank the models in accordance with the achieved accuracies, but the RF approach relatively showed the best accuracy predictions. However, for conducting a fair comparative analysis for each adopted approach, from each modeling technique, the best three models with different input combinations have been selected for performing further and fair comparisons.

Thus, among the 60 predictive models, only the best nine models are selected for optimally carrying out the efficient quantitative analysis. Reducing the number of models has many advantages. For instance, it ensures conducting a powerful and excellent compilation and, thereby, optimally choosing the best statistical matrices. Table 5 shows the performance accuracy of three different approaches based on several input variables.

In general, remarkably, it can be noted from Table 5 that the most frequent combination is C1, C4, and C7, consecutively. This means that the component includes all useful parameters that have effective impact on mean monthly  $ET_o$ . At a glance, it can be said that the RF models provided more accurate estimations of mean monthly  $ET_o$  than MLR and ELM approaches. Here, it is essential to mention that the uncertainty at 95% ( $U_{95}$ ) indicator is the most efficient factor, which plays a major role in the evaluation process of

TABLE 3: Crosscorrelation matrix of the variables used in this study.

	Latitude	Longitude	Elevation	$T_{max}$	$T_{min}$	$T_{mean}$	WS	$RH_{max}$	$RH_{min}$	$ET_O$
Training										
Latitude	1.000									
Longitude	0.206	1.000								
Elevation	0.673	-0.112	1.000							
$T_{max}$	-0.194	0.095	-0.301	1.000						
$T_{min}$	-0.468	-0.021	-0.522	0.852	1.000					
$T_{mean}$	-0.319	0.036	-0.393	0.943	0.954	1.000				
WS	-0.114	0.004	0.024	0.027	0.125	0.089	1.000			
$RH_{max}$	0.115	0.157	-0.012	-0.279	-0.341	-0.339	-0.024	1.000		
$RH_{min}$	0.000	-0.046	0.084	-0.143	-0.058	-0.068	0.118	0.079	1.000	
$ET_O$	-0.101	0.052	-0.071	0.815	0.696	0.778	0.459	-0.208	-0.065	1.000
Testing										
Latitude	1.000									
Longitude	0.181	1.000								
Elevation	0.717	-0.068	1.000							
$T_{max}$	-0.168	0.166	-0.374	1.000						
$T_{min}$	-0.450	0.064	-0.625	0.877	1.000					
$T_{mean}$	-0.283	0.118	-0.493	0.958	0.951	1.000				
WS	-0.354	0.013	-0.169	0.024	0.176	0.113	1.000			
$RH_{max}$	0.194	0.217	0.083	-0.175	-0.268	-0.237	-0.005	1.000		
$RH_{min}$	0.157	-0.042	0.164	-0.279	-0.210	-0.233	0.093	0.204	1.000	
$ET_O$	-0.166	0.091	-0.206	0.841	0.738	0.820	0.419	-0.144	-0.208	1.000

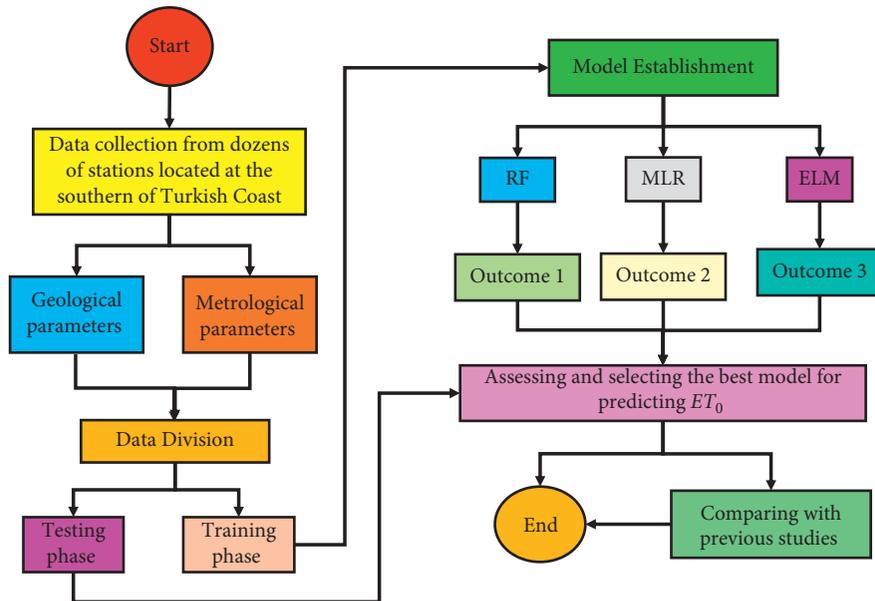


FIGURE 4: Prediction process of the monthly mean evapotranspiration using several AI approaches.

the best model. Furthermore, the RF-C7 produced the lowest value (1.441) of  $U_{95}$  as compared with other predictive models. Besides, the model mentioned also presented high performance based on the rest of the statistical measures. Nevertheless, the accuracy of MLR was unsatisfactory and recorded a high uncertainty with  $U_{95}$  of 6.077. This is due to the fact that the nonlinear relationship between predictors and their corresponding targets was not considered. Finally, the performances of ELM modeling approaches were

satisfactory according to the  $U_{95}$  indicator. With respect to the ELM approach, the best model was ELM-C1, which recorded 4.126 of  $U_{95}$ .

Essentially, it should be taken into consideration that promising estimations were obtained during the training phase, which is primarily employed to effectively calibrate the models based on the known input variables and targets. However, the testing step is vital in assessing the performance of a model since it examines the model's accuracy

TABLE 4: Performance skill description of each predictive model according to different input combinations: training set.

Model	MAE (mm/month)	RMSE (mm/month)	MARE	RMSRE	RRMSE (%)	CC	erMAX	MBE (mm/month)	t_stat	U <sub>95</sub> (%)
MLR-C1	0.801	0.938	0.158	0.22	14.490	0.942	1.692	6.47	17.909	6.077
RF-C1	0.561	0.688	0.102	0.134	10.628	0.972	1.573	6.468	17.821	1.758
ELM-C1	0.722	0.851	0.141	0.193	13.154	0.952	1.776	6.47	17.875	4.126
MLR-C2	0.839	1.011	0.164	0.232	15.630	0.932	1.838	6.47	17.941	8.227
RF-C2	0.534	0.671	0.095	0.124	10.373	0.973	1.529	6.469	17.816	1.596
ELM-C2	0.772	0.93	0.144	0.202	14.380	0.943	1.621	6.47	17.906	5.893
MLR-C3	0.84	1.02	0.164	0.24	15.766	0.931	1.815	6.47	17.944	8.516
RF-C3	0.541	0.69	0.097	0.127	10.663	0.971	1.5	6.471	17.822	1.781
ELM-C3	0.822	0.99	0.159	0.222	15.303	0.935	1.726	6.47	17.931	7.557
MLR-C4	0.801	0.938	0.158	0.219	14.503	0.942	1.698	6.47	17.909	6.1
RF-C4	0.525	0.66	0.095	0.126	10.197	0.974	1.531	6.481	17.813	1.49
ELM-C4	0.731	0.86	0.135	0.18	13.293	0.951	1.591	6.47	17.879	4.305
MLR-C5	0.842	1.014	0.165	0.235	15.678	0.931	1.829	6.47	17.942	8.325
RF-C5	0.565	0.705	0.103	0.134	10.890	0.971	1.689	6.484	17.826	1.937
ELM-C5	0.761	0.923	0.136	0.175	14.263	0.944	1.532	6.47	17.903	5.704
MLR-C6	0.876	1.053	0.173	0.252	16.270	0.926	1.799	6.47	17.959	9.657
RF-C6	0.556	0.697	0.101	0.13	10.779	0.971	1.489	6.465	17.824	1.861
ELM-C6	0.816	0.968	0.156	0.213	14.956	0.938	1.762	6.47	17.922	6.896
MLR-C7	0.8	0.942	0.158	0.22	14.559	0.941	1.705	6.47	17.911	6.194
RF-C7	0.524	0.654	0.095	0.124	10.112	0.975	1.531	6.472	17.811	1.441
ELM-C7	0.788	0.923	0.156	0.214	14.270	0.944	1.635	6.47	17.903	5.716
MLR-C8	1.248	1.508	0.254	0.352	23.310	0.841	2.396	6.47	18.222	40.689
RF-C8	0.831	1.006	0.159	0.209	15.542	0.938	1.64	6.464	17.938	8.043
ELM-C8	1.088	1.346	0.205	0.278	20.809	0.876	2.113	6.47	18.117	25.846
MLR-C9	1.436	1.687	0.298	0.412	26.069	0.796	2.352	6.47	18.355	63.653
RF-C9	0.843	1.02	0.166	0.225	15.769	0.936	1.89	6.47	17.945	8.52
ELM-C9	1.133	1.382	0.216	0.294	21.357	0.868	2.105	6.47	18.138	28.67
MLR-C10	1.216	1.461	0.245	0.337	22.586	0.852	2.25	6.47	18.19	35.863
RF-C10	0.801	0.975	0.157	0.212	15.076	0.943	1.898	6.482	17.924	7.118
ELM-C10	1.086	1.341	0.205	0.272	20.724	0.877	1.95	6.47	18.113	25.431
MLR-C11	1.071	1.26	0.2	0.252	19.473	0.892	1.852	6.47	18.066	19.818
RF-C11	0.836	1.013	0.159	0.21	15.659	0.943	1.829	6.463	17.942	8.284
ELM-C11	1.004	1.206	0.177	0.219	18.645	0.901	1.688	6.47	18.036	16.653
MLR-C12	1.169	1.367	0.223	0.29	21.130	0.871	1.942	6.47	18.129	27.486
RF-C12	0.837	1.005	0.159	0.21	15.539	0.94	1.793	6.461	17.938	8.033
ELM-C12	0.976	1.189	0.176	0.232	18.380	0.904	1.912	6.47	18.027	15.735
MLR-C13	1.013	1.208	0.18	0.225	18.676	0.901	1.807	6.47	18.037	16.768
RF-C13	0.783	0.934	0.151	0.202	14.427	0.952	1.849	6.478	17.907	5.973
ELM-C13	0.898	1.073	0.161	0.203	16.591	0.923	1.742	6.47	17.969	10.441
MLR-C14	0.875	1.057	0.172	0.249	16.330	0.925	1.822	6.47	17.961	9.801
RF-C14	0.669	0.837	0.122	0.16	12.940	0.961	1.726	6.449	17.871	3.862
ELM-C14	0.75	0.921	0.131	0.165	14.235	0.944	1.457	6.47	17.902	5.66
MLR-C15	1.451	1.706	0.297	0.405	26.362	0.791	2.271	6.47	18.37	66.586
RF-C15	0.976	1.189	0.192	0.258	18.383	0.913	1.986	6.467	18.028	15.74
ELM-C15	1.232	1.502	0.236	0.324	23.222	0.842	2.301	6.47	18.218	40.075
MLR-C16	1.162	1.36	0.22	0.287	21.011	0.873	1.907	6.47	18.125	26.868
RF-C16	0.845	1.045	0.161	0.217	16.150	0.935	1.824	6.469	17.956	9.381
ELM-C16	0.999	1.201	0.181	0.23	18.563	0.902	1.823	6.47	18.033	16.366
MLR-C17	1.16	1.358	0.226	0.3	20.988	0.873	1.972	6.47	18.124	26.749
RF-C17	0.867	1.052	0.165	0.218	16.254	0.936	1.8	6.463	17.959	9.622
ELM-C17	1.038	1.247	0.187	0.239	19.278	0.894	1.902	6.47	18.059	19.034
MLR-C18	1.071	1.26	0.2	0.252	19.473	0.892	1.852	6.47	18.066	19.818
RF-C18	0.824	0.989	0.156	0.205	15.281	0.942	1.741	6.471	17.931	7.516
ELM-C18	0.943	1.141	0.167	0.21	17.637	0.912	1.73	6.47	18.002	13.341
MLR-C19	1.169	1.369	0.223	0.291	21.155	0.871	1.953	6.47	18.13	27.616
RF-C19	0.827	1.022	0.149	0.197	15.794	0.933	1.795	6.468	17.945	8.577
ELM-C19	1.05	1.274	0.186	0.236	19.684	0.889	1.866	6.47	18.074	20.69
MLR-C20	1.394	1.751	0.273	0.38	27.065	0.778	2.651	6.47	18.407	73.96
RF-C20	1.092	1.425	0.199	0.279	22.028	0.86	2.297	6.445	18.17	32.439
ELM-C20	1.329	1.709	0.25	0.347	26.416	0.79	2.409	6.47	18.373	67.125

TABLE 5: Selection of the best three models for each predictive approach through the calibration process.

Model	MAE (mm/month)	RMSE (mm/month)	MARE	RMSRE	RRMSE (%)	CC	erMAX	MBE (mm/month)	t <sub>-stat</sub>	U <sub>95</sub> (%)	Rank
MLR-C1	0.801	0.938	0.158	0.22	14.490	0.942	1.692	6.47	17.909	6.077	1
MLR-C4	0.801	0.938	0.158	0.219	14.503	0.942	1.698	6.47	17.909	6.1	2
MLR-C7	0.8	0.942	0.158	0.22	14.559	0.941	1.705	6.47	17.911	6.194	3
RF-C7	0.524	0.654	0.095	0.124	10.112	0.975	1.531	6.472	17.811	1.441	1
RF-C4	0.525	0.66	0.095	0.126	10.197	0.974	1.531	6.481	17.813	1.49	2
RF-C2	0.534	0.671	0.095	0.124	10.373	0.973	1.529	6.469	17.816	1.596	3
ELM-C1	0.722	0.851	0.141	0.193	13.154	0.952	1.776	6.47	17.875	4.126	1
ELM-C4	0.731	0.86	0.135	0.18	13.293	0.951	1.591	6.47	17.879	4.305	2
ELM-C14	0.75	0.921	0.131	0.165	14.235	0.944	1.457	6.47	17.902	5.66	3

based on unseen target values. This advantage does not exist in the training set. Therefore, the reliable model should have a stable and balanced performance in both training and testing phases.

After demonstrating the performance of the models during the calibration (training) phase, it is very important to see the accuracy of the adopted models during the testing phase. Table 6 demonstrated the performance skill of each predictive model using different input parameters.

It is also important to carefully follow-up the performance of nine efficient models that were chosen in the training set (MLR-C1, MLR-C4, MLR-C7, RF-C7, RF-C4, RF-C2, ELM-C1, ELM-C4, and ELM-C14). Besides the mentioned models, there were additional three models (RF-C5, RF-C6, and ELM-C2) that have effectively provided satisfactory estimations. The heat-map diagram as presented in Figure 5 provided significant information about the best modeling performance based on ten statistical parameters.

Although RF models generated acceptable precision during the training set, it exhibited the worst accuracies as compared to ELM and MLR techniques in the testing set. Moreover, these models produced high uncertainty, and the values of  $U_{95}$  are 19.59, 27.1, 22.74, 22.35, and 26.72% for RF-C7, RF-C4, RF-C5, RF-C6, and RF-C2, respectively. Additionally, the other statistical parameters such as RMSE and RRMSE also gave further information about the weaknesses of RF models. It is undeniable that the RF approach suffers from overfitting issue. On the other hand, the MLR models showed much better performance capacity than RF models. Finally, the ELM models achieved high precision in the prediction of mean monthly  $ET_o$  in accordance with the P-M equation. Moreover, the ELM-C1 is considered the best predictive model and recorded the highest CC (0.957), lowest RMSE (1.155), MAE (0.946 mm/month),  $t - \text{stat}$ (10.37), RRMSE (16.54%), and  $U_{95}$ (9.989%), respectively. Moreover, among ten statistical parameters, the most efficient parameters, which can easily recognize the best predictive model are  $U_{95}$ , RMSE, and RRMSE. The supremacy of ELM approaches was evaluated in accordance with its ability of reducing the most significant statistical measures ( $U_{95}$ , RMSE, and RRMSE) throughout the testing phase. The results obtained as shown in Figure 6 illustrated

the superiority of ELM-C1 over other predictive models in reducing the value of the mentioned three statistical metrics. The effectiveness of the ELM-C1 model over MLR-C1 powerfully appeared in reducing the RMSE and  $U_{95}$  parameters to 10.05% and 34.07%, respectively. Moreover, the prediction accuracy of ELM-C1 accomplished better outcome when it is compared with the RF-C7 model, where there was a significant improvement in the reduction of RMSE and  $U_{95}$  parameters, which reached 16.36% and 49.02%, respectively.

The box plot diagram, shown in Figure 7, presented the best candidate models, which were employed to predict the average monthly  $ET_o$ . It can be remarkably noted that the RF models could not perform well as compared to the other approaches such as MLR and ELM. However, the MLR models had the modest performance and poor accuracy as compared to ELM approaches. It can be said that the best performance approach in prediction average monthly  $ET_o$  is ELM, followed by MLR and RF techniques, consecutively. The ELM-C1 achieved the best estimation accuracy where the median and interquarter range (IQR) were found to be very close to the actual median and IQR.

Line graph and scatterplot of predicted and actual average monthly  $ET_o$  are provided in Figures 8(a) and 8(b) during the testing phase. The scatter plot effectively presents useful visualization information on the diversion between observed and predicted values, and the coefficient of determination ( $R^2$ ) showed the determination between them. Based on Figures 8(a) and 8(b), the proposed ELM-C1 has better prediction performance than other comparative models in terms of providing a higher value of  $R^2$  (0.915).

Considering the best ELM models, it can be noted that this approach often requires a relatively higher number of input parameters in comparison with RF and MLR approaches. The logical explanation of this phenomenon is that the data size includes 35 locations of different characteristics. Besides, the geographical factors are very important when developing local and robust models based on a dataset collected from several locations.

For further assessment, it is vital to examine the ability of the suggested model (ELM-C1) against several predictive models that were proposed in the previous studies to estimate

TABLE 6: Performance skill description of each predictive model according to different input combinations: testing set.

Model	MAE (mm/month)	RMSE (mm/month)	MARE	RMSRE	RRMSE (%)	CC	erMAX	MBE (mm/month)	t_stat	U <sub>95</sub> (%)
MLR-C1	1.078	1.284	0.172	0.214	18.392	0.945	1.449	6.254	10.42	15.151
RF-C1	1.281	1.561	0.187	0.213	22.359	0.938	1.556	5.992	10.563	29.955
ELM-C1	<b>0.946</b>	<b>1.155</b>	0.149	0.191	<b>16.544</b>	<b>0.956</b>	1.634	6.333	<b>10.372</b>	<b>9.989</b>
MLR-C2	1.117	1.358	0.173	0.216	19.458	0.936	1.545	6.242	10.448	19.537
RF-C2	1.229	1.513	0.178	0.206	21.673	0.938	1.503	6.032	10.535	26.723
ELM-C2	0.962	1.216	<b>0.146</b>	0.18	17.417	0.95	1.422	6.322	10.392	12.574
MLR-C3	1.167	1.407	0.181	0.22	20.159	0.936	1.501	6.158	10.475	21.306
RF-C3	1.253	1.541	0.182	0.209	22.080	0.935	1.492	6.018	10.55	28.899
ELM-C3	1.111	1.357	0.168	0.203	19.444	0.946	1.456	6.131	10.458	17.33
MLR-C4	1.074	1.282	0.172	0.214	18.371	0.945	1.456	6.266	10.419	15.248
RF-C4	1.256	1.523	0.184	0.21	21.820	0.939	1.516	6.013	10.542	27.104
ELM-C4	0.999	1.248	0.147	0.177	17.884	0.955	1.467	6.228	10.409	12.848
MLR-C5	1.118	1.348	0.178	0.221	19.312	0.934	1.56	6.283	10.441	19.589
RF-C5	1.144	1.423	0.168	0.198	20.391	0.946	1.526	6.168	10.481	22.741
ELM-C5	1.052	1.269	0.163	0.19	18.186	0.952	1.447	6.197	10.419	13.441
MLR-C6	1.199	1.429	0.195	0.246	20.475	0.926	1.551	6.232	10.477	24.555
RF-C6	1.134	1.422	0.164	0.195	20.371	0.947	1.677	6.155	10.482	22.346
ELM-C6	1.066	1.282	0.174	0.218	18.367	0.95	1.496	6.197	10.423	14.126
MLR-C7	1.071	1.281	0.168	0.204	18.348	0.945	1.461	6.274	10.417	15.282
RF-C7	1.103	1.381	0.16	0.194	19.785	0.951	1.649	6.165	10.464	19.593
ELM-C7	1.047	1.253	0.164	0.199	17.947	<b>0.956</b>	1.445	6.185	10.414	12.401
MLR-C8	1.524	1.803	0.276	0.366	25.837	0.851	1.871	6.348	10.636	73.725
RF-C8	1.389	1.684	0.217	0.272	24.132	0.901	1.784	6.255	10.589	52.417
ELM-C8	1.316	1.633	0.218	0.273	23.396	0.891	1.554	6.256	10.564	45.735
MLR-C9	1.751	2.069	0.315	0.419	29.644	0.796	2.267	6.296	10.798	129.564
RF-C9	1.468	1.781	0.229	0.279	25.512	0.886	1.727	6.207	10.646	65.221
ELM-C9	1.368	1.731	0.212	0.28	24.804	0.872	2.188	6.302	10.606	60.569
MLR-C10	1.462	1.746	0.264	0.367	25.015	0.867	1.811	6.283	10.616	62.258
RF-C10	1.414	1.698	0.221	0.274	24.321	0.906	2.015	6.241	10.598	53.811
ELM-C10	1.237	1.56	0.198	0.249	22.356	0.903	1.593	6.296	10.526	38.316
MLR-C11	1.342	1.66	0.213	0.269	23.785	0.901	1.66	6.083	10.6	43.783
RF-C11	1.467	1.836	0.209	0.251	26.298	0.911	1.767	5.958	10.719	64.059
ELM-C11	1.246	1.583	0.179	0.213	22.677	0.929	1.349	5.976	10.576	31.619
MLR-C12	1.505	1.801	0.248	0.307	25.806	0.876	1.788	6.057	10.681	62.81
RF-C12	1.408	1.763	0.2	0.234	25.254	0.917	1.524	5.976	10.673	53.541
ELM-C12	1.302	1.605	0.201	0.26	22.997	0.927	1.476	5.947	10.591	32.973
MLR-C13	1.264	1.589	0.196	0.246	22.765	0.918	1.577	6.044	10.57	34.417
RF-C13	1.423	1.783	0.202	0.239	25.547	0.919	1.546	5.944	10.69	55.172
ELM-C13	1.277	1.527	0.19	0.227	21.880	0.949	1.302	5.878	10.561	23.495
MLR-C14	1.197	1.429	0.194	0.244	20.470	0.925	1.569	6.248	10.476	24.857
RF-C14	1.247	1.55	0.184	0.22	22.210	0.94	1.622	6.146	10.539	33.389
ELM-C14	1.015	1.259	0.148	0.176	18.038	0.952	1.397	6.209	10.414	13.109
MLR-C15	1.734	2.079	0.308	0.409	29.792	0.789	2.079	6.339	10.795	134.22
RF-C15	1.599	1.924	0.266	0.336	27.565	0.858	2.048	6.264	10.716	93.894
ELM-C15	1.419	1.76	0.229	0.298	25.220	0.858	1.979	6.382	10.61	67.441
MLR-C16	1.487	1.805	0.244	0.305	25.854	0.878	1.776	6.034	10.687	62.369
RF-C16	1.518	1.888	0.227	0.28	27.053	0.891	2.048	5.994	10.745	74.975
ELM-C16	1.351	1.633	0.207	0.251	23.394	0.923	1.52	5.932	10.608	35.411
MLR-C17	1.534	1.825	0.262	0.338	26.149	0.879	1.663	5.985	10.708	63.591
RF-C17	1.539	1.896	0.229	0.274	27.166	0.901	1.858	5.962	10.756	74.846
ELM-C17	1.363	1.672	0.205	0.241	23.954	0.916	1.44	5.936	10.628	40.003
MLR-C18	1.342	1.66	0.213	0.269	23.785	0.901	1.66	6.083	10.6	43.783
RF-C18	1.407	1.775	0.203	0.245	25.426	0.915	1.656	5.957	10.683	54.491
ELM-C18	1.254	1.577	0.185	0.227	22.590	0.929	1.535	5.975	10.573	31.011
MLR-C19	1.505	1.805	0.248	0.309	25.867	0.877	1.776	6.047	10.685	63.046
RF-C19	1.369	1.736	0.195	0.23	24.875	0.906	1.49	5.99	10.655	50.356
ELM-C19	1.265	1.583	0.186	0.221	22.674	0.923	1.45	6.029	10.569	33.28
MLR-C20	1.877	2.329	0.295	0.361	33.366	0.82	2.197	5.567	11.228	154.758
RF-C20	1.859	2.398	0.264	0.316	34.363	0.813	2.089	5.499	11.333	171.226
ELM-C20	1.8	2.275	0.261	0.317	32.599	0.832	2.109	5.613	11.156	142.188

Bold numbers refer to the best results.

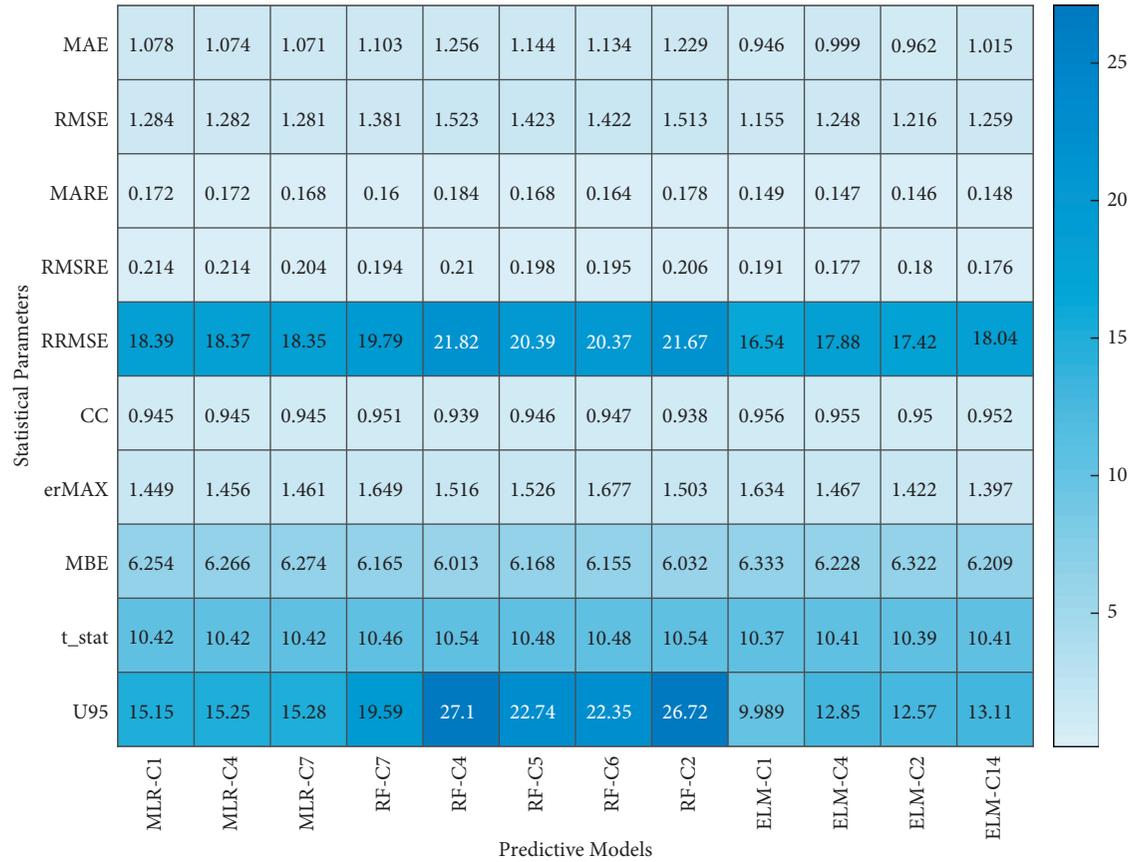


FIGURE 5: Heat-map diagram illustrating the best three models for each predictive approach through the calibration process.

$ET_O$  over different locations around the world. In this respect, the findings obtained during the test step by using the ELM-C1 model are validated against several predictive models carried out in the literature by some researchers. Khoob [94] conducted a study using ANN to predict  $ET_O$  in Safiabad station, which is located in Southern Iran. The outcome of this study revealed that the suggested ANN model performed very well when compared with the actual values of  $ET_O$ . The most significant observation is that the accuracy of the model was very good with a high  $R^2$  of 0.9135. Moreover, another study was carried out using the ANN technique to predict  $ET_O$  in the Reynolds Creek Experimental Watershed in Southwestern Idaho, USA [95]. The obtained results showed that the ANN model managed to effectively estimate  $ET_O$  with high accuracy of prediction and fewer forecasted error reported ( $R^2 = 0.82$ ). Additionally, another study was conducted in Southeast Asia using different AI models; Wu and Fan [96] introduced a comprehensive study to accurately predict  $ET_O$  over 14 stations located in China. The researchers employed eight different techniques such as ANFIS, MLP, GRNN, KNEA, SVR, XGBoost, M5Tree, and Mars. These models were developed based on the collected data, which includes precipitation and temperature variables. The outcomes of the study illustrated that the machine learning techniques could provide satisfactory prediction of  $ET_O$  by using only one variable (temperature). In addition, the estimated  $ET_O$  values were close to the actual ones, and the reported average  $R^2$  indicator

was equal to 0.829. On top of that, Izadifar and Elshorbagy [97] employed three data-driven models called MLR, GP, and ANN to estimate  $ET_O$  over Northern Alberta, Canada. Statistical analyses of the study revealed that the most influencing variables on  $ET_O$  were net radiation and temperature. Besides, the ANN approached trained Bayesian regularization algorithm presented more accurate results than other comparable models, which recorded the highest accuracy of prediction ( $R^2 = 0.82$ ).

Significant attempts were made to increasingly constitute powerful models to accurately estimate  $ET_O$  based on few parameters, which can easily be measured such as temperature. Zhu et al. [69] managed to develop a hybrid model called PSO-ELM to predict  $ET_O$  over five meteorological stations located at the Northwest of China. The proposed model produced more accurate estimations than other AI models and employed empirical equations. The proposed model (PSO-ELM) were built based on temperature only, in general, and produced fewer forecasted errors and higher correlation with the actual one (mean  $R^2$  of 0.8905).

Overall, the published studies, as illustrated above, have been carried out to estimate  $ET_O$  over several case studies. Even though researchers developed powerful approaches that attained high accuracy, the approaches mentioned could not be applied to constitute a general model instead of simulation of each case study separately. However, in this

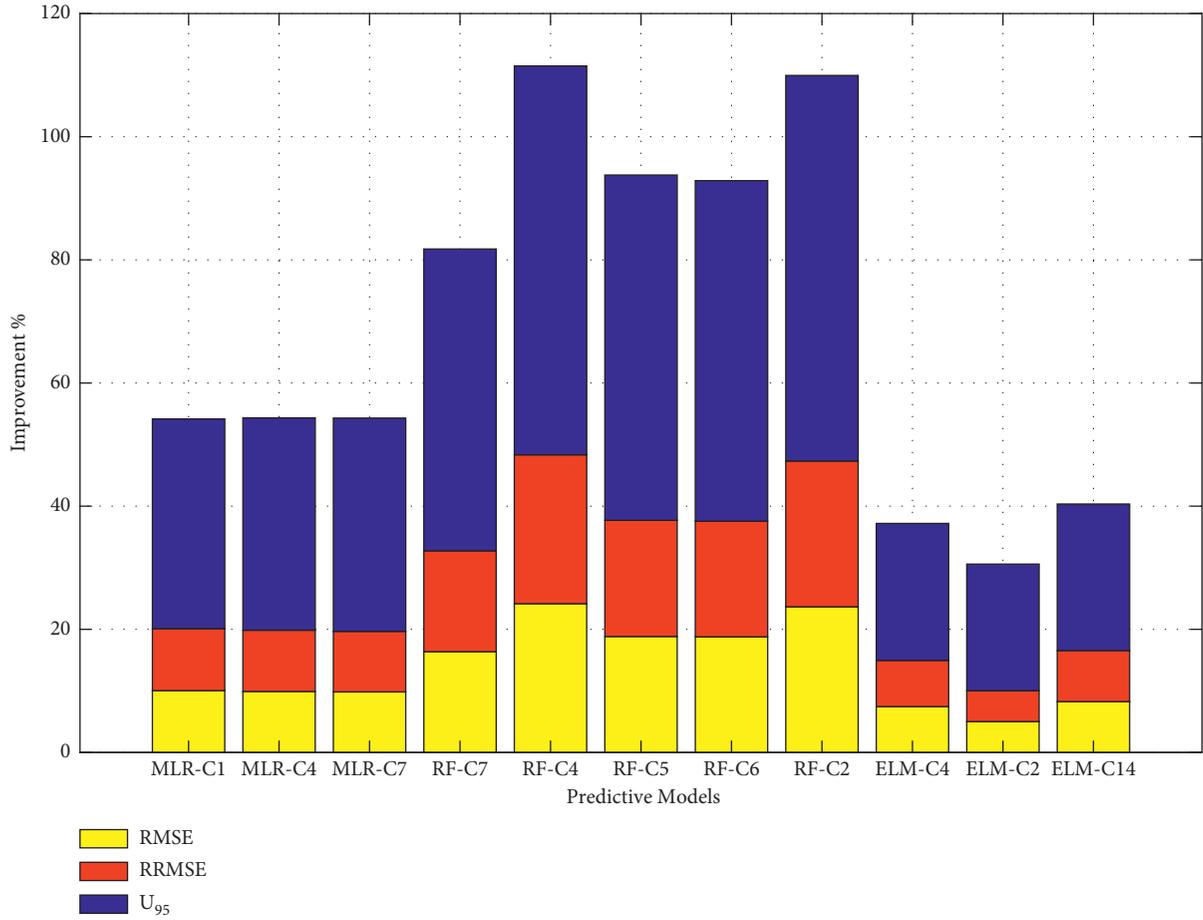


FIGURE 6: The enhancement percentage of the ELM-C1 predictive model over other models based on RMASE, RRMSE, and  $U_{59}$ .

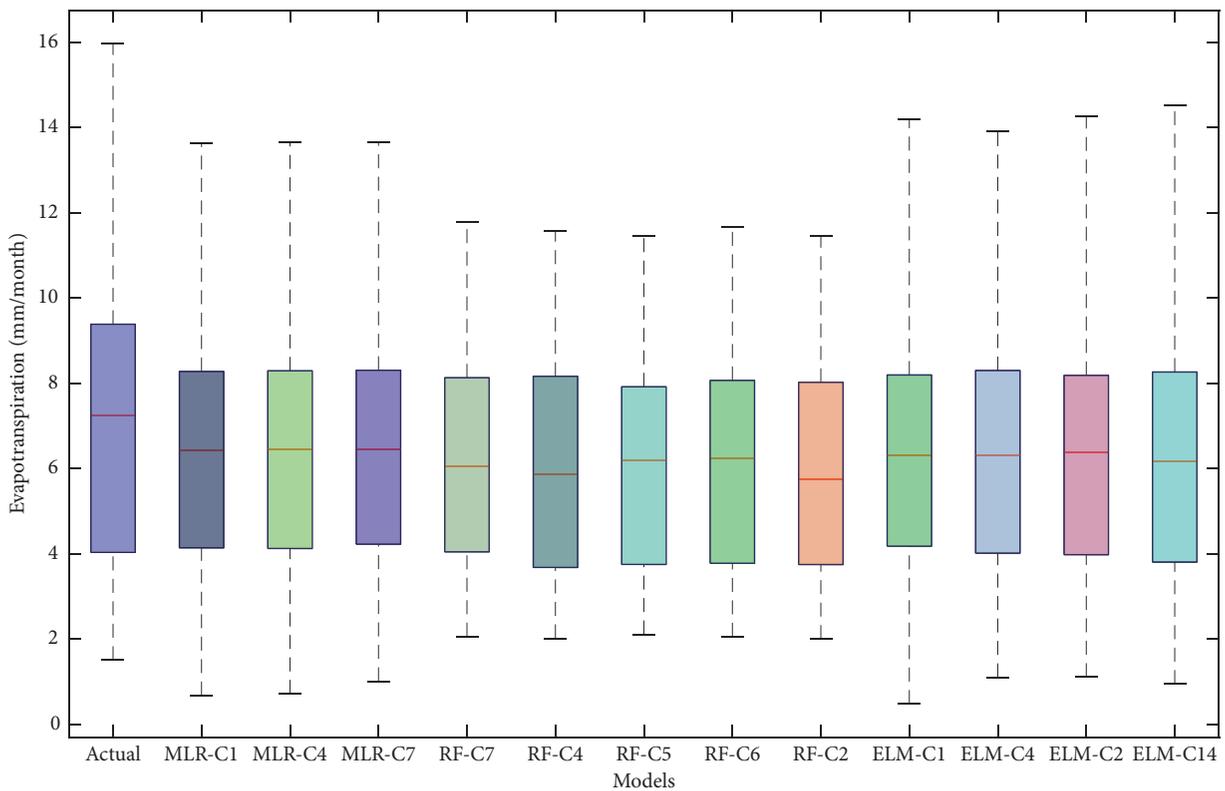
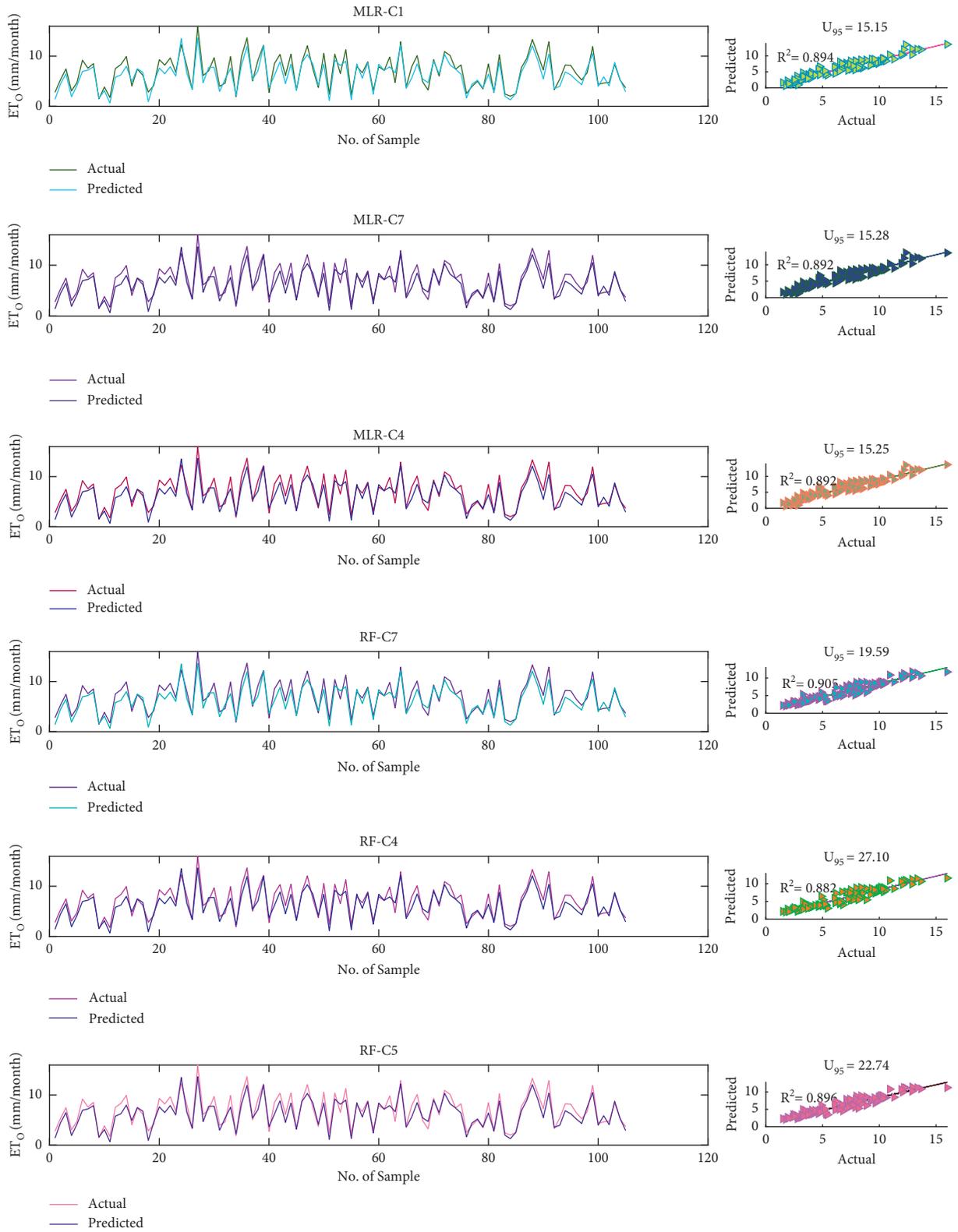
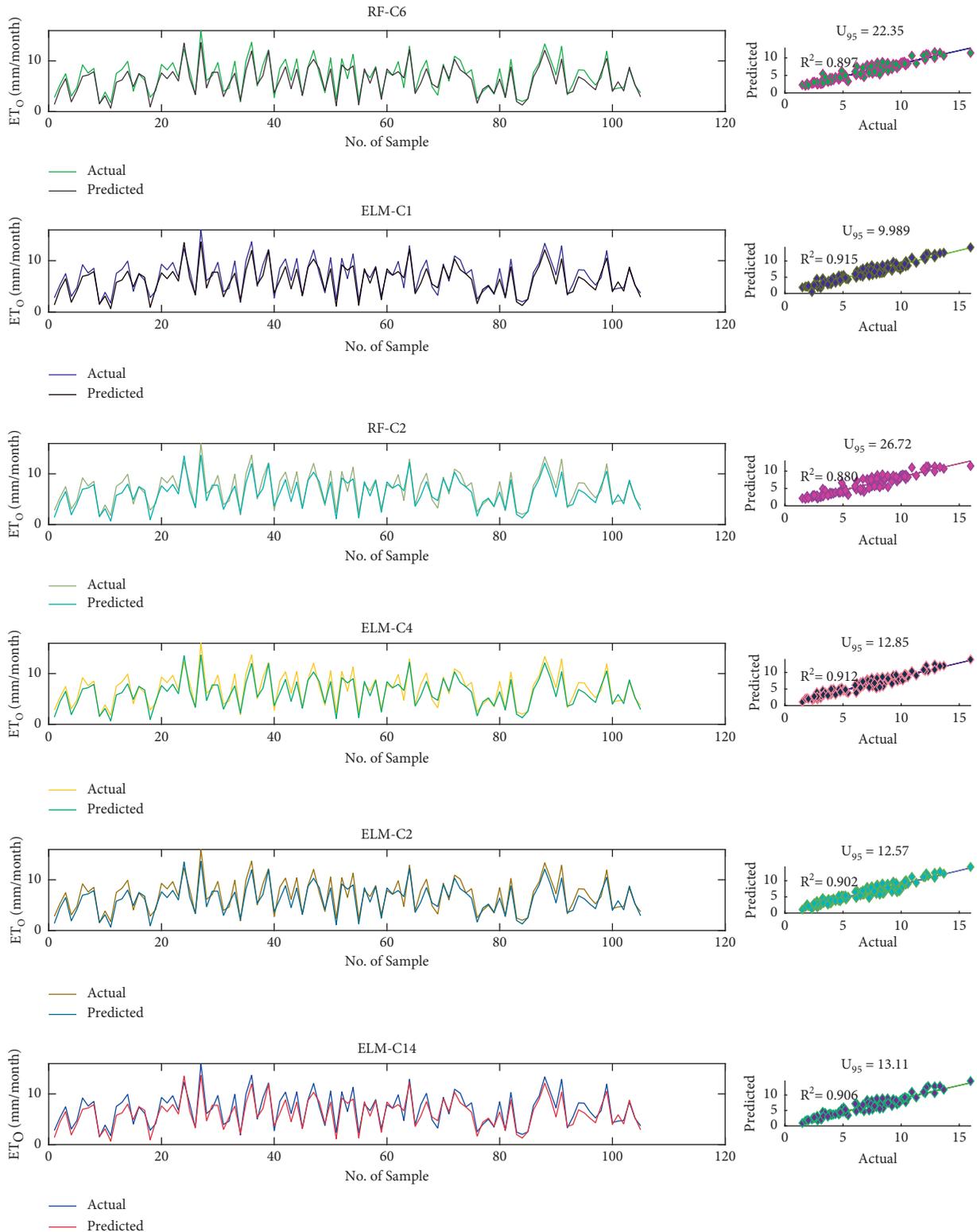


FIGURE 7: Boxplot diagram presented the superiority of ELM over other comparable models during the testing set.



(a)

FIGURE 8: Continued.



(b)

FIGURE 8: (a): Observed vs. predicted values of mean monthly  $ET_0$  using different models: testing set. (b) Observed vs. predicted values of mean monthly  $ET_0$  using different models: testing set.

current study, a single model (ELM-C1) successfully managed to simulate the  $ET_O$  based on the data collected from dozens of stations. In addition to the simulation process of 35 stations in one single model, the other interesting feature is that the predictability of the proposed model was recorded high ( $R^2 > 0.91$ ).

## 5. Conclusion

Evapotranspiration is considered as one of the most significant factors in the hydrological cycle. Although the well-known Penman–Monteith equation for computing the evapotranspiration exists, there are some difficulties in accurately calculating some of its parameters such as solar radiation and sensible heat flux into the soil. Therefore, in this study, three different approaches were employed, namely, ELM, RF, and MLR, based on geographical and meteorological parameters for the prediction of the mean monthly evapotranspiration over southern Mediterranean coast of Turkey. Besides, twenty different input combinations assigned according to data collected from 35 meteorological stations were established. Ten efficient statistical parameters have been used to assess the 60 predictive models. The outcomes of this study revealed that the performance of the ELM approach outperformed the MLR and RF models. In addition, the efficiency of the proposed ELM produced less computed errors during the testing set. It is worth mentioning that the  $U_{95}$  measure played a vital role in the selection of the best model accuracy among the established 60 models in this study. The other essential observation that can be drawn is that the RF approach provided the perfect accuracy during the training phase, and when it comes to examining its performances during the testing phase, the predictions' accuracies were very poor and disappointing. Finally, this study suggested the use of ELM in building local models based on different stations and weather conditions. In addition, a study that covers the whole area of the Mediterranean Sea via one robust model is needed. For further investigation, feature selection approaches can possibly be integrated prior to the predictive learning process to extract the essential input variables for the prediction matrix [98, 99].

## Abbreviations

$ET_o$ :	Reference evapotranspiration
$T_{max}$ :	Maximum monthly temperature
$T_{min}$ :	Minimum monthly temperature
$T_{mean}$ :	Average monthly temperature
Std:	Standard deviation
$W_s$ :	Wind speed
$RH_{max}$ :	% Maximum relative humidity
$RH_{min}$ :	% Minimum relative humidity
MAE:	Mean absolute error
$e$ :	Estimated error between actual and predicted $ET_O$
RMSE:	Root mean square error
MARE:	Mean absolute relative error
RMSRE:	Root mean square relative error
RRMSE:	Relative root mean square error

MBE:	Mean bias error
CC:	Correlation of coefficient
$R^2$ :	Correlation of determination
erMAX:	Maximum absolute relative error
$t$ – stat:	$t$ -statistic
$U_{95}$ :	Uncertainty at 95%
ELM:	Extreme learning machine
RF:	Random Forests
MLR:	Multiple-linear regression
P – M:	Penman–Monteith.

## Data Availability

All the data are available upon request.

## Conflicts of Interest

The authors have no conflicts of interest.

## Acknowledgments

The authors would like to thank AlMaaref University College (AUC) for funding this research.

## References

- [1] U. Maček, N. Bezak, and M. Šraj, “Reference evapotranspiration changes in Slovenia, Europe,” *Agricultural and Forest Meteorology*, vol. 260–261, pp. 183–192, 2018.
- [2] H. Sanikhani, R. C. Deo, P. Samui et al., “Survey of different data-intelligent modeling strategies for forecasting air temperature using geographic information as model predictors,” *Computers and Electronics in Agriculture*, vol. 152, pp. 242–260, 2018.
- [3] J. Hall, B. Arheimer, M. Borga et al., “Understanding flood regime changes in Europe: a state-of-the-art assessment,” *Hydrology and Earth System Sciences*, vol. 18, no. 7, pp. 2735–2772, 2014.
- [4] J. de Leeuw, J. Methven, and M. Blackburn, “Variability and trends in England and Wales precipitation,” *International Journal of Climatology*, vol. 36, no. 8, pp. 2823–2836, 2016.
- [5] S. M. Awadh, H. Al-Mimar, and Z. M. Yaseen, “Groundwater availability and water demand sustainability over the upper mega aquifers of Arabian Peninsula and West Region of Iraq,” *Environment, Development and Sustainability*, vol. 23, pp. 1–21, 2021.
- [6] Z. Yang, Q. Liu, and B. Cui, “Spatial distribution and temporal variation of reference evapotranspiration during 1961–2006 in the Yellow River Basin, China,” *Hydrological Sciences Journal*, vol. 56, no. 6, pp. 1015–1026, 2011.
- [7] K. Xiang, Y. Li, R. Horton, and H. Feng, “Similarity and difference of potential evapotranspiration and reference crop evapotranspiration - a review,” *Agricultural Water Management*, vol. 232, Article ID 106043, 2020.
- [8] A. Malik, A. Kumar, S. Kim et al., “Modeling monthly pan evaporation process over the Indian central Himalayas: application of multiple learning artificial intelligence model,” *Engineering Applications of Computational Fluid Mechanics*, vol. 14, no. 1, pp. 323–338, 2020.
- [9] F. Granata, “Evapotranspiration evaluation models based on machine learning algorithms-A comparative study,” *Agricultural Water Management*, vol. 217, pp. 303–315, 2019.

- [10] M. E. Jensen and R. G. Allen, *Evaporation, Evapotranspiration, and Irrigation Water Requirements*, American Society of Civil Engineers, Reston, VA, USA, 2016.
- [11] M. Valipour, "Analysis of potential evapotranspiration using limited weather data," *Applied Water Science*, vol. 7, no. 1, pp. 187–197, 2017.
- [12] S. Jasechko, Z. D. Sharp, J. J. Gibson, S. J. Birks, Y. Yi, and P. J. Fawcett, "Terrestrial water fluxes dominated by transpiration," *Nature*, vol. 496, no. 7445, pp. 347–350, 2013.
- [13] M. Jung, M. Reichstein, P. Ciais et al., "Recent decline in the global land evapotranspiration trend due to limited moisture supply," *Nature*, vol. 467, no. 7318, pp. 951–954, 2010.
- [14] K. Zhang, J. S. Kimball, and S. W. Running, "A review of remote sensing based actual evapotranspiration estimation," *WIREs Water*, vol. 3, no. 6, pp. 834–853, 2016.
- [15] G. R. Diak, J. R. Mecikalski, M. C. Anderson et al., "Estimating land surface energy budgets from space: review and current efforts at the University of Wisconsin-madison and USDA-ARS," *Bulletin of the American Meteorological Society*, vol. 85, no. 1, pp. 65–78, 2004.
- [16] K. Khosravi, P. Daggupati, M. T. Alami et al., "Meteorological data mining and hybrid data-intelligence models for reference evaporation simulation: a case study in Iraq," *Computers and Electronics in Agriculture*, vol. 167, p. 105041, 2019.
- [17] R. A. Pielke, fnm Sr, R. Avissar et al., "Interactions between the atmosphere and terrestrial ecosystems: influence on weather and climate," *Global Change Biology*, vol. 4, no. 5, pp. 461–475, 1998.
- [18] C. Schumacher, R. A. Houze Jr., and I. Kraucunas, "The tropical dynamical response to latent heating estimates derived from the TRMM precipitation radar," *Journal of the Atmospheric Sciences*, vol. 61, no. 12, pp. 1341–1358, 2004.
- [19] W. J. Shuttleworth, "Putting the "vap" into evaporation," *Hydrology and Earth System Sciences*, vol. 11, no. 1, pp. 210–244, 2007.
- [20] T. Stocker, *Climate Change 2013: The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, UK, 2014.
- [21] S. Naganna, P. Deka, M. Ghorbani, S. Biazar, N. Al-Ansari, and Z. Yaseen, "Dew point temperature estimation: application of artificial intelligence model integrated with nature-inspired optimization algorithms," *Water*, vol. 11, no. 4, p. 742, 2019.
- [22] X. Lian, S. Piao, C. Huntingford et al., "Partitioning global land evapotranspiration using CMIP5 models constrained by observations," *Nature Climate Change*, vol. 8, no. 7, pp. 640–646, 2018.
- [23] P. C. D. Milly and K. A. Dunne, "Potential evapotranspiration and continental drying," *Nature Climate Change*, vol. 6, no. 10, pp. 946–949, 2016.
- [24] W. Jing, Z. M. Yaseen, S. Shahid et al., "Implementation of evolutionary computing models for reference evapotranspiration modeling: short review, assessment and possible future research directions," *Engineering Applications of Computational Fluid Mechanics*, vol. 13, no. 1, pp. 811–823, 2019.
- [25] M. K. Saggi and S. Jain, "Application of fuzzy-genetic and regularization random forest (FG-RRF): estimation of crop evapotranspiration (ET ) for maize and wheat crops," *Agricultural Water Management*, vol. 229, Article ID 105907, 2020.
- [26] S. Karimi, J. Shiri, and P. Marti, "Supplanting missing climatic inputs in classical and random forest models for estimating reference evapotranspiration in humid coastal areas of Iran," *Computers and Electronics in Agriculture*, vol. 176, Article ID 105633, 2020.
- [27] S. Green, B. Clothier, and B. Jardine, "Theory and practical application of heat pulse to measure sap flow," *Agronomy Journal*, vol. 95, no. 6, pp. 1371–1379, 2003.
- [28] H. Chen, J. J. Huang, and E. McBean, "Partitioning of daily evapotranspiration using a modified shuttleworth-wallace model, random forest and support vector regression, for a cabbage farmland," *Agricultural Water Management*, vol. 228, Article ID 105923, 2020.
- [29] S. Kang, B. Gu, T. Du, and J. Zhang, "Crop coefficient and ratio of transpiration to evapotranspiration of winter wheat and maize in a semi-humid region," *Agricultural Water Management*, vol. 59, no. 3, pp. 239–254, 2003.
- [30] W. Zegada-Lizarazu and P. R. Berliner, "Inter-row mulch increase the water use efficiency of furrow-irrigated maize in an arid environment," *Journal of Agronomy and Crop Science*, vol. 197, no. 3, pp. 237–248, 2011.
- [31] N. Raz-Yaseef, D. Yakir, G. Schiller, and S. Cohen, "Dynamics of evapotranspiration partitioning in a semi-arid forest as affected by temporal rainfall patterns," *Agricultural and Forest Meteorology*, vol. 157, pp. 77–85, 2012.
- [32] U. S. Panu and T. Nguyen, "Estimation of mean areal evaporation in Northwestern Ontario," *Canadian Water Resources Journal*, vol. 19, no. 1, pp. 69–82, 1994.
- [33] V. P. Singh and C.-Y. Xu, "Evaluation and generalization of 13 mass-transfer equations for determining free water evaporation," *Hydrological Processes*, vol. 11, no. 3, pp. 311–323, 1997.
- [34] J. Shiri, "Improving the performance of the mass transfer-based reference evapotranspiration estimation approaches through a coupled wavelet-random forest methodology," *Journal of Hydrology*, vol. 561, pp. 737–750, 2018.
- [35] S. S. Fiyadh, M. A. AlSaadi, M. K. AlOmar et al., "The modelling of lead removal from water by deep eutectic solvents functionalized CNTs: artificial neural network (ANN) approach," *Water Science and Technology*, vol. 76, no. 9, pp. 2413–2426, 2017.
- [36] S. S. Fiyadh, M. A. AlSaadi, M. K. AlOmar, S. S. Fayaed, F. S. Mjalli, and A. El-Shafie, "BTPC-based DES-functionalized CNTs for As3+ removal from water: NARX neural network approach," *Journal of Environmental Engineering*, vol. 144, no. 8, Article ID 04018070, 2018.
- [37] Z. M. Yaseen, W. H. M. W. Mohtar, A. M. S. Ameen et al., "Implementation of Univariate paradigm for streamflow simulation using hybrid data-driven model: case study in tropical region," *IEEE Access*, vol. 7, pp. 74471–74481, 2019.
- [38] Z. Yaseen, I. Ebtehaj, S. Kim et al., "Novel hybrid data-intelligence model for forecasting monthly rainfall with uncertainty analysis," *Water*, vol. 11, no. 3, p. 502, 2019.
- [39] H. H. Hasan, S. F. Mohd Razali, N. S. Muhammad, and A. Ahmad, "Research trends of hydrological drought: a systematic review," *Water*, vol. 11, no. 11, p. 2252, 2019.
- [40] M. A. Ghorbani, R. Khatibi, V. Karimi, Z. M. Yaseen, and M. Zounemat-Kermani, "Learning from multiple models using artificial intelligence to improve model prediction accuracies: application to river flows," *Water Resources Management*, vol. 32, no. 13, pp. 4201–4215, 2018.
- [41] S. Maroufpoor, E. Maroufpoor, O. Bozorg-Haddad, J. Shiri, and Z. Mundher Yaseen, "Soil moisture simulation using hybrid artificial intelligent model: hybridization of adaptive neuro fuzzy inference system with grey wolf optimizer algorithm," *Journal of Hydrology*, vol. 575, pp. 544–556, 2019.

- [42] A. Ahmad, S. F. Mohd Razali, A. El-Shafie, and Z. Samba Mohamad, "Optimal timah tasoh reservoir in, perlis: an operation using the Gravitational Search Algorithm (GSA)," *Jurnal Teknologi*, vol. 77, no. 30, 2015.
- [43] A. Ahmad, S. F. M. Razali, Z. S. Mohamed, and A. El-shafie, "The application of artificial bee colony and gravitational search algorithm in reservoir optimization," *Water Resources Management*, vol. 30, no. 7, pp. 2497–2516, 2016.
- [44] M. F. Allawi, O. Jaafar, F. Mohamad Hamzah, S. M. S. Abdullah, and A. El-shafie, "Review on applications of artificial intelligence methods for dam and reservoir-hydro-environment models," *Environmental Science and Pollution Research*, vol. 25, no. 14, pp. 13446–13469, 2018.
- [45] A. M. Muad, S. K. M. Zaki, and S. A. Jasim, "Optimizing hopfield neural network for super-resolution mapping," *Jurnal Kejuruteraan*, vol. 32, no. 1, pp. 91–97, 2020.
- [46] A. Koppa and M. Gebremichael, "A framework for validation of remotely sensed precipitation and evapotranspiration based on the Budyko hypothesis," *Water Resources Research*, vol. 53, no. 10, pp. 8487–8499, 2017.
- [47] Y. Feng, N. Cui, D. Gong, Q. Zhang, and L. Zhao, "Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling," *Agricultural Water Management*, vol. 193, pp. 163–173, 2017.
- [48] M. Kumar, N. S. Raghuvanshi, R. Singh, W. W. Wallender, and W. O. Pruitt, "Estimating evapotranspiration using artificial neural network," *Journal of Irrigation and Drainage Engineering*, vol. 128, no. 4, pp. 224–233, 2002.
- [49] G. Landeras, A. Ortiz-Barredo, and J. J. López, "Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain)," *Agricultural Water Management*, vol. 95, no. 5, pp. 553–565, 2008.
- [50] Ö. Kişi, "Generalized regression neural networks for evapotranspiration modelling," *Hydrological Sciences Journal*, vol. 51, no. 6, pp. 1092–1105, 2006.
- [51] S. Traore, Y.-M. Wang, and T. Kerh, "Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone," *Agricultural Water Management*, vol. 97, no. 5, pp. 707–714, 2010.
- [52] N. Abrishami, A. R. Sepaskhah, and M. H. Shahrokhnia, "Estimating wheat and maize daily evapotranspiration using artificial neural network," *Theoretical and Applied Climatology*, vol. 135, no. 3, pp. 945–958, 2019.
- [53] H. Tabari, O. Kisi, A. Ezani, and P. Hosseinzadeh Talaei, "SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment," *Journal of Hydrology*, vol. 444–445, pp. 78–89, 2012.
- [54] M. F. Allawi and A. El-Shafie, "Utilizing RBF-NN and ANFIS methods for multi-lead ahead prediction model of evaporation from reservoir," *Water Resources Management*, vol. 30, no. 13, pp. 4773–4788, 2016.
- [55] S. Gavili, H. Sanikhani, O. Kisi, and M. H. Mahmoudi, "Evaluation of several soft computing methods in monthly evapotranspiration modelling," *Meteorological Applications*, vol. 25, no. 1, pp. 128–138, 2018.
- [56] A. Moghaddamnia, M. Ghafari Gousheh, J. Piri, S. Amin, and D. Han, "Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques," *Advances in Water Resources*, vol. 32, no. 1, pp. 88–97, 2009.
- [57] Y. Yao, S. Liang, X. Li et al., "Improving global terrestrial evapotranspiration estimation using support vector machine by integrating three process-based algorithms," *Agricultural and Forest Meteorology*, vol. 242, pp. 55–74, 2017.
- [58] Z. Yin, X. Wen, Q. Feng, Z. He, S. Zou, and L. Yang, "Integrating genetic algorithm and support vector machine for modeling daily reference evapotranspiration in a semi-arid mountain area," *Hydrology Research*, vol. 48, no. 5, pp. 1177–1191, 2016.
- [59] M. Ehteram, V. P. Singh, A. Ferdowsi et al., "An improved model based on the support vector machine and cuckoo algorithm for simulating reference evapotranspiration," *PLoS One*, vol. 14, no. 5, Article ID e0217499, 2019.
- [60] K. Parasuraman, A. Elshorbagy, and S. K. Carey, "Modelling the dynamics of the evapotranspiration process using genetic programming," *Hydrological Sciences Journal*, vol. 52, no. 3, pp. 563–578, 2007.
- [61] A. Guven, A. Aytekin, M. I. Yuce, and H. Aksoy, "Genetic programming-based empirical model for daily reference evapotranspiration estimation," *Clean-Soil, Air, Water*, vol. 36, no. 10–11, pp. 905–912, 2008.
- [62] B. Narapusetty, "The role of atmospheric internal variability on the prediction skill of interannual North Pacific sea-surface temperatures," *Theoretical and Applied Climatology*, vol. 133, no. 1, pp. 113–121, 2018.
- [63] Ö. Kişi, "Evapotranspiration modeling using a wavelet regression model," *Irrigation Science*, vol. 29, no. 3, pp. 241–252, 2011.
- [64] Y. Falamarzi, N. Palizdan, Y. F. Huang, and T. S. Lee, "Estimating evapotranspiration from temperature and wind speed data using artificial and wavelet neural networks (WNNs)," *Agricultural Water Management*, vol. 140, pp. 26–36, 2014.
- [65] T. Partal, "Modelling evapotranspiration using discrete wavelet transform and neural networks," *Hydrological Processes*, vol. 23, no. 25, pp. 3545–3555, 2009.
- [66] O. Kisi and M. Alizamir, "Modelling reference evapotranspiration using a new wavelet conjunction heuristic method: wavelet extreme learning machine vs wavelet neural networks," *Agricultural and Forest Meteorology*, vol. 263, pp. 41–48, 2018.
- [67] S. Wang, J. Lian, Y. Peng, B. Hu, and H. Chen, "Generalized reference evapotranspiration models with limited climatic data based on random forest and gene expression programming in Guangxi, China," *Agricultural Water Management*, vol. 221, pp. 220–230, 2019.
- [68] C. Gonzalo-Martín, M. Lillo-Saavedra, A. Garcia-Pedrero, O. Lagos, and E. Menasalvas, "Daily evapotranspiration mapping using regression random forest models," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 12, pp. 5359–5368, 2017.
- [69] B. Zhu, Y. Feng, D. Gong, S. Jiang, L. Zhao, and N. Cui, "Hybrid particle swarm optimization with extreme learning machine for daily reference evapotranspiration prediction from limited climatic data," *Computers and Electronics in Agriculture*, vol. 173, Article ID 105430, 2020.
- [70] S. S. Abdullah, M. A. Malek, N. S. Abdullah, O. Kisi, and K. S. Yap, "Extreme Learning Machines: a new approach for prediction of reference evapotranspiration," *Journal of Hydrology*, vol. 527, pp. 184–195, 2015.
- [71] M. K. AlOmar, M. M. Hameed, N. Al-Ansari, and M. A. AlSaadi, "Data-driven model for the prediction of total dissolved gas: robust artificial intelligence approach," *Advances in Civil Engineering*, vol. 2020, Article ID 6618842, 2020.

- [72] J. Fan, W. Yue, L. Wu et al., "Evaluation of SVM, ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in different climates of China," *Agricultural and Forest Meteorology*, vol. 263, pp. 225–241, 2018.
- [73] M. Gocic, D. Petković, S. Shamshirband, and A. Kamsin, "Comparative analysis of reference evapotranspiration equations modelling by extreme learning machine," *Computers and Electronics in Agriculture*, vol. 127, pp. 56–63, 2016.
- [74] D. Kumar, J. Adamowski, R. Suresh, and B. Ozga-Zielinski, "Estimating evapotranspiration using an extreme learning machine model: case study in north Bihar, India," *Journal of Irrigation and Drainage Engineering*, vol. 142, no. 9, Article ID 04016032, 2016.
- [75] R. G. Allen, L. S. Pereira, D. Raes, and M. Smith, *Crop evapotranspiration—Guidelines for Computing Crop Water Requirements—FAO Irrigation and Drainage Paper 56*, FAO, vol. 300, no. 9, p. D05109, Rome, Italy, 1998.
- [76] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [77] T. Hastie, R. Tibshirani, and J. Friedman, "Random forests," in *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, pp. 587–604, Springer New York, New York, NY, USA, 2009.
- [78] L. Breiman, *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [79] T. Hai, A. Sharafati, A. Mohammed et al., "Global solar radiation estimation and climatic variability analysis using extreme learning machine based predictive model," *IEEE Access*, vol. 8, pp. 12026–12042, 2020.
- [80] N. Nan-Ying Liang, fnm Guang-Bin Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," *IEEE Transactions on Neural Networks*, vol. 17, no. 6, pp. 1411–1423, 2006.
- [81] Z. M. Yaseen, S. M. Awadh, A. Sharafati, and S. Shahid, "Complementary data-intelligence model for river flow simulation," *Journal of Hydrology*, vol. 567, pp. 180–190, 2018.
- [82] M. M. Hameed and M. K. AlOmar, *Prediction of Compressive Strength of High-Performance Concrete: Hybrid Artificial Intelligence Technique*, Springer International Publishing, Cham, Switzerland, 2020.
- [83] A. Ashrafzadeh, M. A. Ghorbani, S. M. Biazar, and Z. M. Yaseen, "Evaporation process modelling over northern Iran: application of an integrative data-intelligence model with the krill herd optimization algorithm," *Hydrological Sciences Journal*, vol. 64, no. 15, pp. 1843–1856, 2019.
- [84] T. M. T. Tiyasha, T. M. Tung, and Z. M. Yaseen, "A survey on river water quality modelling using artificial intelligence models: 2000–2020," *Journal of Hydrology*, vol. 585, p. 124670, 2020.
- [85] G. Zhang, Z. Hasan Ali, M. S. Aldemy et al., "Reinforced concrete deep beam shear strength capacity modelling using an integrative bio-inspired algorithm with an artificial intelligence model," *Engineering with Computers*, 2020.
- [86] K. Onyelowo, G. U. Alaneme, M. Onyia et al., "Comparative modeling of strength properties of hydrated-lime activated rice-husk-ash (HARHA) modified soft soil for pavement construction purposes by artificial neural network (ANN) and fuzzy logic (FL)," *Jurnal Kejuruteraan*, vol. 33, no. 2, pp. 365–384, 2021.
- [87] Z. M. Yaseen, A. El-Shafie, H. A. Afan, M. Hameed, W. H. M. W. Mohtar, and A. Hussain, "RBFNN versus FFNN for daily river flow forecasting at Johor River, Malaysia," *Neural Computing & Applications*, vol. 27, no. 6, pp. 1533–1542, 2016.
- [88] M. M. Hameed, M. K. AlOmar, W. J. Baniya, and M. A. AlSaadi, "Prediction of high-strength concrete: high-order response surface methodology modeling approach," *Engineering with Computers*, 2021.
- [89] M.-F. Li, X.-P. Tang, W. Wu, and H.-B. Liu, "General models for estimating daily global solar radiation for different solar radiation zones in mainland China," *Energy Conversion and Management*, vol. 70, pp. 139–148, 2013.
- [90] M. K. AlOmar, M. M. Hameed, and M. A. AlSaadi, "Multi hours ahead prediction of surface ozone gas concentration: robust artificial intelligence approach," *Atmospheric Pollution Research*, 2020.
- [91] O. Behar, A. Khellaf, and K. Mohammadi, "Comparison of solar radiation models and their validation under Algerian climate—the case of direct irradiance," *Energy Conversion and Management*, vol. 98, pp. 236–251, 2015.
- [92] H. A. Afan, A. El-Shafie, Z. M. Yaseen, M. M. Hameed, W. H. M. Wan Mohtar, and A. Hussain, "ANN based sediment prediction model utilizing different input scenarios," *Water Resources Management*, vol. 29, no. 4, pp. 1231–1245, 2015.
- [93] M. M. Hameed, S. Sharqi, Z. M. Yaseen, and H. A. Afan, "Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia," *Neural Computing & Applications*, vol. 28, no. 1, pp. 893–905, 2017.
- [94] A. Rahimi Khoob, "Artificial neural network estimation of reference evapotranspiration from pan evaporation in a semi-arid environment," *Irrigation Science*, vol. 27, no. 1, pp. 35–39, 2008.
- [95] S. K. Jain, P. C. Nayak, and K. P. Sudheer, "Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation," *Hydrological Processes*, vol. 22, no. 13, pp. 2225–2234, 2008.
- [96] L. Wu and J. Fan, "Comparison of neuron-based, kernel-based, tree-based and curve-based machine learning models for predicting daily reference evapotranspiration," *PLoS One*, vol. 14, no. 5, Article ID e0217520, 2019.
- [97] Z. Izadifar and A. Elshorbagy, "Prediction of hourly actual evapotranspiration using neural networks, genetic programming, and statistical models," *Hydrological Processes*, vol. 24, no. 23, pp. 3413–3425, 2010.
- [98] H. A. Afan, M. F. Allawi, A. El-Shafie et al., "Input attributes optimization using the feasibility of genetic nature inspired algorithm: application of river flow forecasting," *Scientific Reports*, vol. 10, no. 1, p. 4684, 2020.
- [99] S. J. Hadi, S. I. Abba, S. S. Sammen, S. Q. Salih, N. Al-Ansari, and Z. M. Yaseen, "Non-linear input variable selection approach integrated with non-tuned data intelligence model for streamflow pattern simulation," *IEEE Access*, vol. 7, pp. 141533–141548, 2019.