



# Modeling Daily Reference Evapotranspiration from Climate Variables: Assessment of Bagging and Boosting Regression Approaches

Jayashree T R<sup>1</sup> · NV Subba Reddy<sup>2</sup> · U Dinesh Acharya<sup>1</sup>

Received: 20 July 2022 / Accepted: 25 November 2022 / Published online: 27 January 2023  
© The Author(s) 2023

## Abstract

The increasing frequency of droughts and floods due to climate change has severely affected water resources across the globe in recent years. An optimal design for the scheduling and management of irrigation is thus urgently needed to adapt agricultural activities to the changing climate. The accurate estimation of reference crop evapotranspiration (ET<sub>0</sub>), a vital hydrological component of the water balance and crop water need, is a tiresome task if all the relevant climatic variables are unavailable. This study investigates the potential of four ensemble techniques for estimating precise values of the daily ET<sub>0</sub> at representative stations in 10 agro-climatic zones in the state of Karnataka, India, from 1979 to 2014. The performance of these models was evaluated by using several combinations of climatic variables as inputs by using tenfold cross-validation. The outcomes indicated that predictions of ET<sub>0</sub> by all four ensemble models based on all climatic variables were the most accurate in comparison with other input combinations. The random forest regressor was found to deliver the best performance among the four models on all measures considered (Nash–Sutcliffe efficiency, 1.0, root-mean-squared error, 0.016 mm/day, and mean absolute error, 0.011 mm/day). However, it incurred the highest computational cost, whereas the computational cost of the bagging model for linear regression was the lowest. The extreme gradient-boosting model delivered the most stable performance with a modified training dataset. The work here shows that these models can be recommended for daily ET<sub>0</sub> estimation based on the users' interests.

**Keywords** Climate variables · Ensemble learning · Modeling · Penman–Monteith (P-M) method · Reference crop evapotranspiration

---

✉ NV Subba Reddy  
nvs.reddy@manipal.edu

Jayashree T R  
jayastr@gmail.com

U Dinesh Acharya  
dinesh.acharya@manipal.edu

<sup>1</sup> Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, India

<sup>2</sup> Department of Information Technology Manipal Institute of Technology Bengaluru, Manipal Academy of Higher Education, Manipal, India

## 1 Introduction

Agriculture activities majorly depend on adequate irrigation practices for optimal crop growth and production. The dearth and uneven rainfall patterns across diverse climates demand the water conservation necessary for irrigation. Evapotranspiration plays the most crucial role in estimating and scheduling the available water resources as well as assessing the impact of climate change on the agriculture sector (Wang et al. 2012). The accurate computation of the reference evapotranspiration (ET<sub>0</sub>) is crucial for managing irrigation-related activities. Though there exist several empirical methods, the Penman–Monteith (P-M) method by the Food and Agriculture Organization (FAO) is the benchmark method that can be applied to diverse climatic conditions across the world. However, it requires a large number of meteorological variables, including data on the atmospheric temperature, solar radiation, wind speed, and relative humidity, that may not be available at all meteorological stations (Yamaç and Todorovic 2020). Several other popular empirical equations to calculate ET<sub>0</sub> require few climatic variables, and have been used in regions for which scant data are available (Priestley and Taylor 1972; Hamid 2011; Pandey et al. 2014; Almorox et al. 2015). However, these models exhibited wide variations in performances across diverse climatic regimes (Chen et al. 2020). This has led to the development of models that can estimate the precise value of ET<sub>0</sub> irrespective of the amount of available climatic data. Remarkable developments in Artificial Intelligence (AI) in recent decades have enabled researchers to handle large datasets and non-linear relationships between various climatic variables to predict ET<sub>0</sub> with significant accuracy.

AI-powered solutions for agriculture have created a revolution in recent years that has enabled improvements in the volume of production and quality of the yield by automating weeding- and irrigation-related practices (Talaviya et al. 2020). Several researchers have proposed AI-based methods in recent years for modeling ET<sub>0</sub>, which is a prime component in irrigation scheduling.

The techniques such as Extreme Learning Machine (ELM), Generalized Regression Neural Network (GRNN), Artificial Neural Network (ANN), and Linear Regression models demonstrated their effectiveness in the evaluation of ET<sub>0</sub> using only temperature data (Feng et al. 2017; Kim et al. 2022) under different time scales in a variety of climatic zones. The modeling of ET<sub>0</sub> by using input combinations composed of various climatic variables, such as the temperature, relative humidity, solar radiation, and wind speed, has been explored by using the ANN (Gocić and Amiri 2021; Elbeltagi et al. 2022a, b), support vector machine (SVM) and tree-based models (Fan et al. 2018), adaptive neuro-fuzzy inference system (ANFIS)–firefly algorithm (Tao et al. 2018), hybrid ELM with optimization (Zhu et al. 2020), and fuzzy logic-based hierarchical fuzzy system (HFS)–particle swarm optimization (PSO) algorithm (Roy et al. 2021). Although these heuristic methods perform better than the empirical models, they require several climatic variables to estimate ET<sub>0</sub> (Wu et al. 2019; Muhammad Adnan et al. 2020).

The water cycle is influenced by rising temperatures that increase the rate of evaporation in the atmosphere and ultimately result in climate change. Mitigation strategies need to be implemented to combat the changing climatic conditions which have affected water availability and distribution in the form of precipitation over the land. Managing the available water resources and accurately predicting ET<sub>0</sub> has thus become critical environmental issues. Although individual machine learning models can accurately predict ET<sub>0</sub>, they fail to handle the errors incurred due to overfitting and uncertainties in the parameters of training, especially in hydrological systems involving complex non-linear variables. Ensemble

machine learning models provide better predictions than individual models by reducing bias and variance and have thus been gaining popularity as robust tools for modeling hydrological systems (Sharafati et al. 2020).

The decision tree (Fan et al. 2018) as well as hybrid metaheuristic machine learning algorithms, such as additive regression (AR), AR bagging, AR random subspace (AR-RSS), and AR-M5 Prime (M5P), predicted ET<sub>0</sub> with high precision (Elbeltagi et al. 2022a, b). Agrawal et al. (2022) comparatively assessed tree-based and boosting-based ensemble methods in terms of modeling ET<sub>0</sub>. However, Boosted Ensemble Neural Networks reported lower minimum predictive errors than the Bagged and individual Neural Networks models (Manikumari et al. 2017). It is useful to assess the performance of various Boosting Ensemble techniques on the classification/regression problems of predicting ET<sub>0</sub> under different time spans (Fan et al. 2019; Ponraj and Vigneswaran 2019; Wu et al. 2020). Han et al. (2019) coupled a bat algorithm with the Extreme Gradient-Boosting (XGB) method (the BAXGB model) to estimate the monthly ET<sub>0</sub> in arid and semi-arid regions of China and found that its results were more accurate than those of the multivariate adaptive regression splines (MARS) and the Gaussian process regression (GPR) models.

This study seeks to answer the following questions: How do ensemble regression models perform on data over 10 meteorological stations in locations that exhibit significant heterogeneity in climatic conditions? Which combinations of meteorological variables have the greatest impact on estimating ET<sub>0</sub>? Which models can accurately predict ET<sub>0</sub> while maintaining stability and incurring a low computational cost?

Karnataka, a state in southwest India, is characterized by heterogeneous conditions among 10 agro-climatic zones (Energy and Wetlands Research @ CES, IISc, Bangalore, India. 2005). Climatic variables play a significant role in estimating ET<sub>0</sub>, which in turn is used to inform farming practices in different climatic zones. The prediction of values of ET<sub>0</sub> for these climatic zones by using ensemble machine learning techniques is useful as they play a crucial role in irrigation planning and water resource management. Few studies to date have assessed the predictive capabilities of bagging-based models of linear regression for predicting ET<sub>0</sub>. Tree-based models can accurately predict ET<sub>0</sub> because the climatic variables have non-linear relationships among themselves. This study evaluates the performance of bagging-based models of linear regression in comparison with that of tree-based boosting models of regression in terms of predicting ET<sub>0</sub>.

## 2 Study Area and Data Collection

Karnataka is a state in southwestern India with significant variations in climate owing to its geographic and physiographic conditions. The climate ranges from arid to semi-arid in the plateau, sub-humid to humid tropical in the Western Ghats, and humid tropical monsoon in the coastal plains (Bangalore Climate Change Initiative–Karnataka, B. C. C. 2011). A trend of decreasing rainfall has been observed in prominent parts of the state due to climate change and has led to an increase in the number of drought-prone regions. The changing patterns and distribution of rainfall can have adverse impacts on natural water resources in these regions. Managing water resources for irrigation purposes is thus a major concern as agriculture is the means of livelihood for a majority of people in the state. Hence, ten climatic stations representing the ten agro-climatic zones comprising arid, semi-arid, humid, sub-humid, tropical, and humid tropical monsoon conditions are taken under study. Daily climate data on maximum and minimum temperature(°C), mean relative humidity (%), wind

speed(m/sec), and solar radiation ( $\text{MJ}/\text{m}^2/\text{day}$ ), for periods of 35 years (1979–2014) were collected for the ten climate stations of Karnataka from “Global Weather Data for SWAT (<https://globalweather.tamu.edu/>)”. The details of the 10 climate stations of Karnataka are shown in Figs. 1 and 2 and the salient features of the same are discussed by Ramachandra et al. (2022).

The detailed procedure for computing the values of  $\text{ET}_0$  is given in the FAO–56 document (Allen et al. 1998). The daily  $\text{ET}_0$  values at these ten stations were calculated by using the FAO P–M method in Python (Ramachandra et al. 2022). The P–M equation is expressed as

$$\text{ET}_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

where,  $\text{ET}_0$  is reference evapotranspiration ( $\text{mm day}^{-1}$ ),  $R_n$  is the net solar radiation at the crop surface ( $\text{MJ m}^{-2} \text{day}^{-1}$ ),  $G$  is the soil heat flux density ( $\text{MJ m}^{-2} \text{day}^{-1}$ ),  $(e_s - e_a)$  is the vapor pressure deficit of the air (kPa),  $\gamma$  is the psychrometric constant ( $\text{kPa } ^\circ\text{C}^{-1}$ ),  $\Delta$  is the slope of the vapor pressure curve ( $\text{kPa } ^\circ\text{C}^{-1}$ ),  $\lambda$  is the latent heat of vaporization ( $\text{MJ kg}^{-1}$ ),  $T$  is the air temperature at 2 m height ( $^\circ\text{C}$ ),  $u_2$  is the wind speed at 2 m height ( $\text{ms}^{-1}$ ),  $e_s$  is the saturation vapor pressure (kPa),  $e_a$  is the actual vapor pressure (kPa).

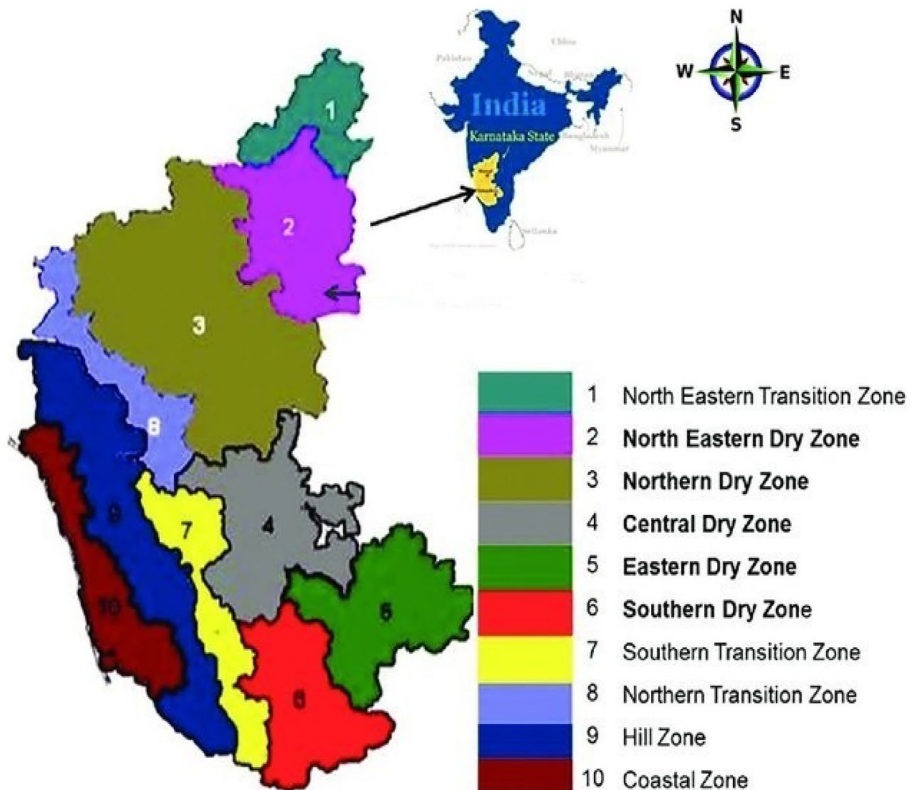


Fig. 1 Map of Karnataka showing its agro-climatic zones

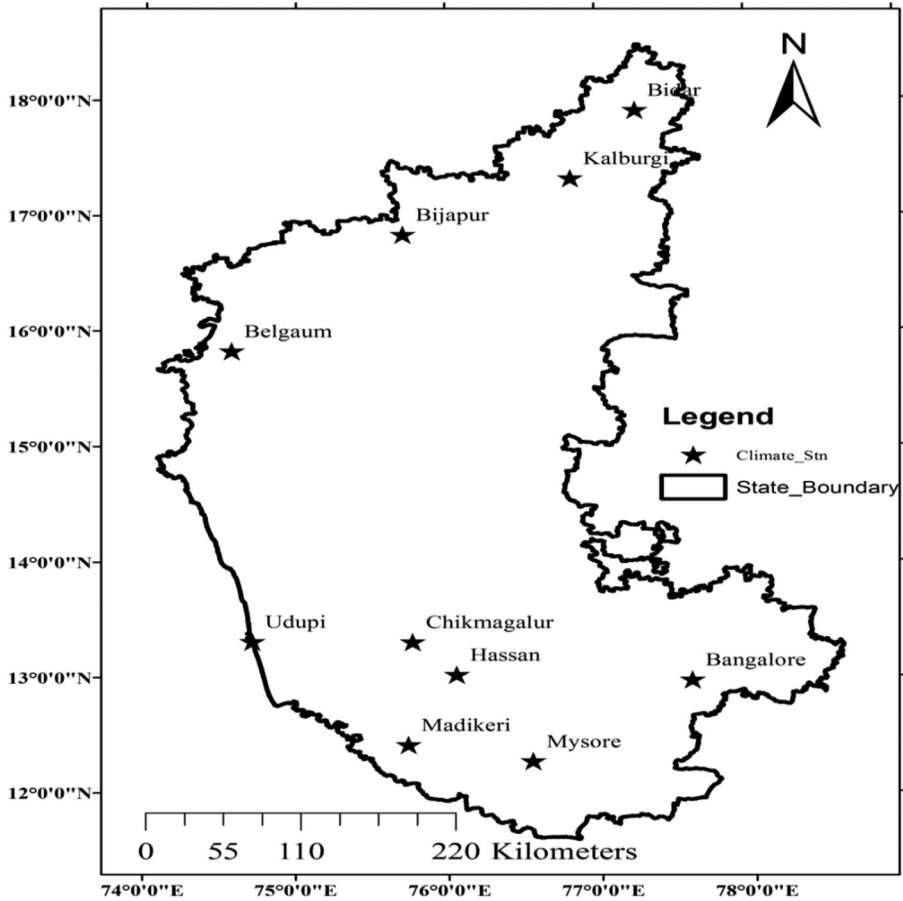


Fig. 2 Geographical locations of the ten stations taken for study

### 3 Methodology

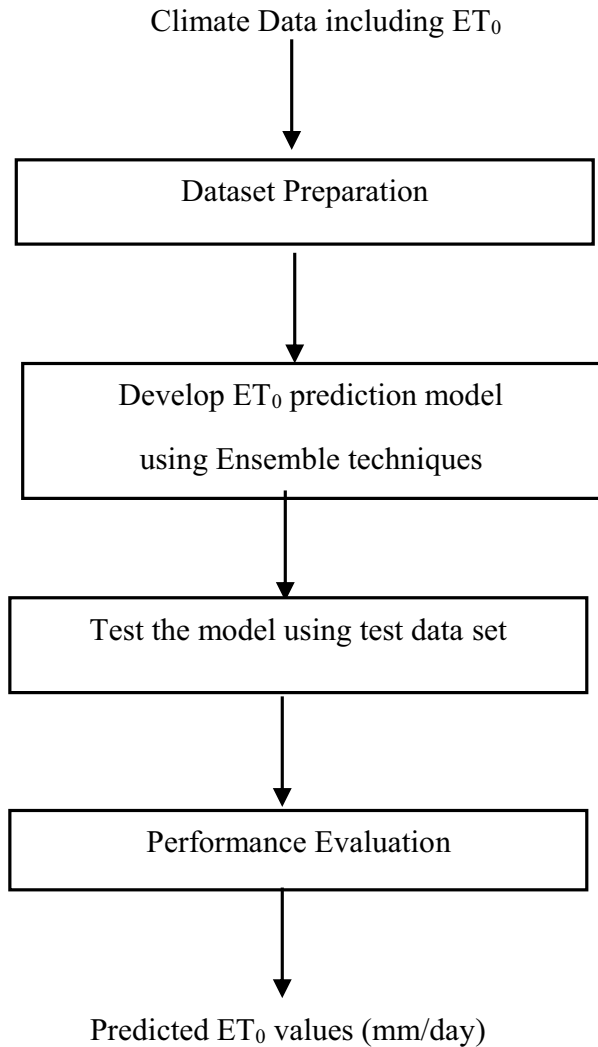
#### 3.1 Modelling $ET_0$ Using Machine Learning Techniques

The climate dataset needed for modeling  $ET_0$  comprised data on the maximum and minimum temperatures ( $^{\circ}C$ ), mean relative humidity (%), wind speed (m/s), solar radiation ( $MJ/m^2/day$ ), longitude ( $^{\circ}$ ), latitude ( $^{\circ}$ ), elevation (m), and  $ET_0$  values (mm/day) computed at the 10 stations. The present study aims to develop models to predict  $ET_0$  at ten stations on a daily scale based on climatic variables by using machine learning techniques. The outline of the overall methodology involves the following steps (Fig. 3).

Step 1: Data preparation: The numerous meteorological parameters needed for the computation of  $ET_0$  by the FAO P-M method were selected here. The irrelevant columns holding date and year attributes were not accounted for in the calculation and hence have been dropped out from the CSV file. Any missing values in the file were filled with null values.

Step 2: Design and development of a machine learning model for the prediction of  $ET_0$ : A regression model has to be developed for the calculation of  $ET_0$ . The data is trained on

**Fig. 3** Flow diagram of modeling of  $ET_0$  using climate variables



the filtered attributes from step 1. The overall dataset is made to split on an 80:20 ratio (80% for training and 20% for testing) to train a regression model.

Step 3: Testing the model: After the construction of the training model, it is tested against the remaining 20% of the dataset, and evaluation is done.

Step 4: Performance Evaluation: The values of statistical measures such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Nash–Sutcliffe coefficient (NSE) were calculated to check the model performance against test data.

### 3.2 Ensemble Techniques for Predicting $ET_0$

Combining the results of numerous models into a single prediction is the goal of ensemble learning, a meta-approach to machine learning. Bagging, stacking, and boosting are

the three basic classes of ensemble learning methods; it is crucial to have a thorough grasp of each approach and to take them into account in any predictive modeling project. Multiple predictions made by the individual members are combined to achieve a better prediction either by majority voting or by averaging the predictions for any classification/regression problem.

### 3.2.1 Ensembles of Regression Models

An ensemble of the regression model is created using different subsets of the original training set with a single machine-learning algorithm. In the present work, the ensembles of the two most known regression models viz., linear regression and decision tree models are developed which can attain low variance and low bias. As part of the optimization procedure, the models are trained with carefully chosen hyperparameter values and make predictions about the desired feature. Then the results of the prediction error analysis are provided to the optimizer for further consideration.

### 3.2.2 Bagging Linear Regressor

Bagging (or bootstrap aggregation) is a sampling method in which many samples are drawn periodically with replacement according to a uniform probability distribution, and then a model is fit to the data (Breiman 1996). It uses the aggregate of predictions or the results of a majority vote to arrive at a more accurate prediction. When applied to models with an overfitting problem, this method improves their performance (high variance models). Bagging may minimize uncertainty in model prediction without sacrificing accuracy.

### 3.2.3 Random Forest Regressor

A random forest fits a number of decision trees on various sub-samples of the dataset to enhance the predictive accuracy and reduce overfitting by averaging over the predictions (Breiman 2001; Geurts et al. 2006). Increasing the number of trees always have an impact on regression problems.

The random forest algorithm functions as follows:

- (i) Pick N random records from the dataset.
- (ii) Build individual decision trees for each record.
- (iii) Specify the number of trees required in the algorithm, and repeat steps i and ii.
- (iv) In case of a regression problem, use each tree in the forest to predict a value for Y (output) for a new record. The final value can be calculated by taking the average of all the predictions by all the nodes in the tree.

### 3.2.4 Boosting Technique

Boosting is an ensemble technique where the most accurate and strong models are built by combining several weak models. Each model is built by correcting the errors present in previous models until the complete training set is predicted correctly. The weights are redistributed after each training stage, and the outcome depends on the weighted average of their estimates.

**Light Gradient Boosting Machine Technique** Light Gradient Boosting Machine (LGBM) is a tree-based boosting technique that offers many advantages such as faster training speed with more accurate results, reduced memory consumption, large-scale data handling, and support for parallel, distributed, and GPU learning (Ke et al. 2017). It develops the model in a tree leaf-wise pattern (node-wise), thus differs from the XGBOOST which does row-wise tree construction.

**Extreme Gradient Boosting** Extreme gradient boosting (XGBOOST) is an improved open-source version of the gradient boosting algorithm proposed by (Chen et al. 2015). XGBOOST facilitates the development of a robust, adaptable, and portable model. For making accurate predictions, XGBOOST is superior to other algorithms and ML frameworks. The improved performance and precise accuracy are the main reasons behind this. To address the flaws in previous approaches, it unifies numerous models into a single one. However, training time is very high for large data sets when compared to LGBM and gradient boosting with categorical features support (CatBoost). In the present work, the XGBOOST technique served as a regression model to predict daily  $ET_0$  from the climate data set. Non-linear interactions between the various climate variables can be handled effectively by making use of this algorithm.

### 3.3 Input Combinations and Validation of the Models

The following parameters have been used for training and testing the model: air temperature (min, max), wind speed, solar radiation, and relative humidity. The 8 input meteorological combinations for the 4 models are:

1. Maximum temperature ( $T_{max}$ ), minimum temperature ( $T_{min}$ )
2.  $T_{max}$ ,  $T_{min}$ , Wind speed (WS)
3.  $T_{max}$ ,  $T_{min}$ , Relative humidity (RH)
4.  $T_{max}$ ,  $T_{min}$ , Solar Radiation (SR)
5.  $T_{max}$ ,  $T_{min}$ , RH, WS
6.  $T_{max}$ ,  $T_{min}$ , RH, SR
7.  $T_{max}$ ,  $T_{min}$ , WS, SR
8.  $T_{max}$ ,  $T_{min}$ , WS, RH, SR

Cross-validation is a resampling method to evaluate machine learning models for a given new set of data (test data). In this study, a ten-fold cross-validation approach was applied to train and validate the ensemble models. The climate dataset containing data on Karnataka from 1979 to 2014 was divided into ten equal folds. Nine folds were used to train the models, while the 10th was used to test them. The entire procedure was repeated 10 times and the average values of the results yielded more reliable results. Random combinations of the hyperparameters were considered in each iteration to find the optimized solution of each model.

### 3.4 Performance Metrics for Prediction

The predictive performance of the ensemble models was compared by using the MAE, RMSE, and NSE. The RMSE assesses the accuracy of the prediction. It is the



root-mean-squared deviation in the differences between the predicted and the actual values and is defined as

$$RMSE = \sqrt{(1/n) \sum_{i=1}^n (ET_{calculated} - ET_{predicted})^2} \tag{2}$$

MAE is the sum of absolute differences between the actual value and predicted values of the number of observations. It is defined as

$$MAE = \frac{1}{n} \sum_{i=1}^n |ET_{calculated} - ET_{predicted}| \tag{3}$$

NSE, the coefficient of efficiency, is a commonly used statistical parameter in hydrology. It indicates the relative assessment of model performance in dimensionless measures. NSE=1 refers to a perfect match between the model and the observed data. A negative NSE value indicates unacceptable model performance. The NSE is given by (Nash and Sutcliffe 1970)

$$NSE = 1 - \frac{\sum_{i=1}^N (ET_{calculated} - ET_{predicted})^2}{\sum_{i=1}^N (ET_{calculated} - meanET_{predicted})^2}, \quad -\infty \leq NSE \leq 1 \tag{4}$$

## 4 Results

### 4.1 Effects of Input Variables on Model Performance Across 10 Stations

Table 1 presents the results of the linear regression bagging (Bagging LR), Random Forest-based Regressor Bagging (Bagging RF), Extreme Gradient Boosting (XGBOOST), and Light Gradient Boosting (LGBM) models for eight input combinations of climatic variables for Bangalore station. The best statistical values of all models in the training and testing phases are marked in bold. Similarly, the performance metrics of the above four models for the eight input combinations at the remaining nine stations are tabulated and given in the supplement file.

The above outcomes reveal that the models with complete climatic data as inputs exhibited better predictive accuracy than those that used incomplete datasets at all 10 stations. The models based on the four input combinations– (Tmax, Tmin, SR), (Tmax, Tmin, RH, SR), (Tmax, Tmin, WS, SR), and (Tmax, Tmin, WS, RH, SR)–delivered better daily estimations of ET0 compared with the other combinations of climatic inputs. Similar results were obtained by using the same input combinations during the testing phase at all 10 stations, as indicated in Tables S1–S9. The results are in accordance with the findings of previous studies (dos Santos Farias et al. 2020; Mehdizadeh et al. 2021; Elbeltagi et al. 2022a, b), which revealed that the use of all climatic variables in the models led to higher prediction accuracies.

The prediction-related performance obtained by the input and output variables in the model was visualized by using a scatter plot. It is the most widely used graph to show the relationship between the predicted and the observed outputs (Wu et al. 2019; Roy et al. 2021; Sattari et al. 2021). The Python implementation of the scatter plots, representing the actual values of ET0 and those predicted by all four models by using the complete dataset of Bangalore station is depicted in Fig. 4. The ET0 values predicted by the LGBM model were close to those that were estimated by using the FAO P–M method in the testing phase,

**Table 1** Performance of Bagging and Boosting based models at Bangalore station

Models	Training Phase			Testing Phase		
	RMSE (mm/day)	MAE (mm/day)	NSE	RMSE (mm/day)	MAE (mm/day)	NSE
<b>T<sub>max</sub>, T<sub>min</sub></b>						
Bagging LR1	0.599	0.468	0.823	1.073	0.869	0.43
Bagging RF1	<b>0.210</b>	<b>0.152</b>	<b>0.978</b>	0.734	0.583	0.734
XGBOOST1	0.499	0.378	0.877	0.526	0.397	0.864
LGBM1	0.471	0.345	0.891	<b>0.523</b>	<b>0.379</b>	<b>0.865</b>
<b>T<sub>max</sub>, T<sub>min</sub>, WS</b>						
Bagging LR2	0.589	0.463	0.829	1.025	0.830	0.48
Bagging RF2	<b>0.190</b>	<b>0.135</b>	<b>0.982</b>	0.989	0.795	0.517
XGBOOST 2	0.463	0.349	0.894	0.489	0.365	0.882
LGBM2	0.421	0.306	0.913	<b>0.485</b>	<b>0.349</b>	<b>0.884</b>
<b>T<sub>max</sub>, T<sub>min</sub>, RH</b>						
Bagging LR3	0.594	0.464	0.826	0.793	0.656	0.689
Bagging RF3	<b>0.189</b>	<b>0.135</b>	<b>0.982</b>	0.813	0.655	0.674
XGBOOST 3	0.463	0.349	0.894	0.494	0.372	0.879
LGBM3	0.422	0.307	0.912	<b>0.482</b>	<b>0.347</b>	<b>0.885</b>
<b>T<sub>max</sub>, T<sub>min</sub>, SR</b>						
Bagging LR4	0.112	0.087	0.994	0.426	0.337	0.91
Bagging RF4	<b>0.023</b>	<b>0.017</b>	<b>1.0</b>	0.259	0.193	0.967
XGBOOST4	0.131	0.109	0.991	0.134	0.110	0.991
LGBM4	0.049	0.036	0.999	<b>0.058</b>	<b>0.043</b>	<b>0.998</b>
<b>T<sub>max</sub>, T<sub>min</sub>, RH, WS</b>						
Bagging LR5	0.573	0.453	0.838	0.792	0.659	0.69
Bagging RF5	<b>0.173</b>	<b>0.123</b>	<b>0.985</b>	0.697	0.547	0.76
XGBOOST5	0.432	0.327	0.908	0.462	0.349	0.894
LGBM5	0.374	0.274	0.931	<b>0.444</b>	<b>0.322</b>	<b>0.903</b>
<b>T<sub>max</sub>, T<sub>min</sub>, RH, SR</b>						
Bagging LR6	0.112	0.086	0.994	0.418	0.328	0.914
Bagging RF6	<b>0.023</b>	<b>0.016</b>	<b>1.0</b>	0.348	0.256	0.94
XGBOOST6	0.131	0.109	0.992	0.135	0.111	0.991
LGBM6	0.047	0.035	0.999	<b>0.056</b>	<b>0.041</b>	<b>0.998</b>
<b>T<sub>max</sub>, T<sub>min</sub>, WS, SR</b>						
Bagging LR7	0.111	0.086	0.994	0.401	0.312	0.92
Bagging RF7	<b>0.023</b>	<b>0.017</b>	<b>1.0</b>	0.504	0.375	0.874
XGBOOST7	0.131	0.109	0.992	0.135	0.111	0.991
LGBM7	0.047	0.035	0.999	<b>0.056</b>	<b>0.042</b>	<b>0.998</b>
<b>T<sub>max</sub>, T<sub>min</sub>, WS, RH, SR</b>						
Bagging LR8	0.111	0.086	0.994	0.222	0.162	0.976
Bagging RF8	<b>0.023</b>	<b>0.017</b>	<b>1.0</b>	0.161	0.115	0.987
XGBOOST8	0.131	0.109	0.992	0.135	0.111	0.991
LGBM8	0.046	0.034	0.999	<b>0.055</b>	<b>0.041</b>	<b>0.998</b>

**Table 1** (continued)

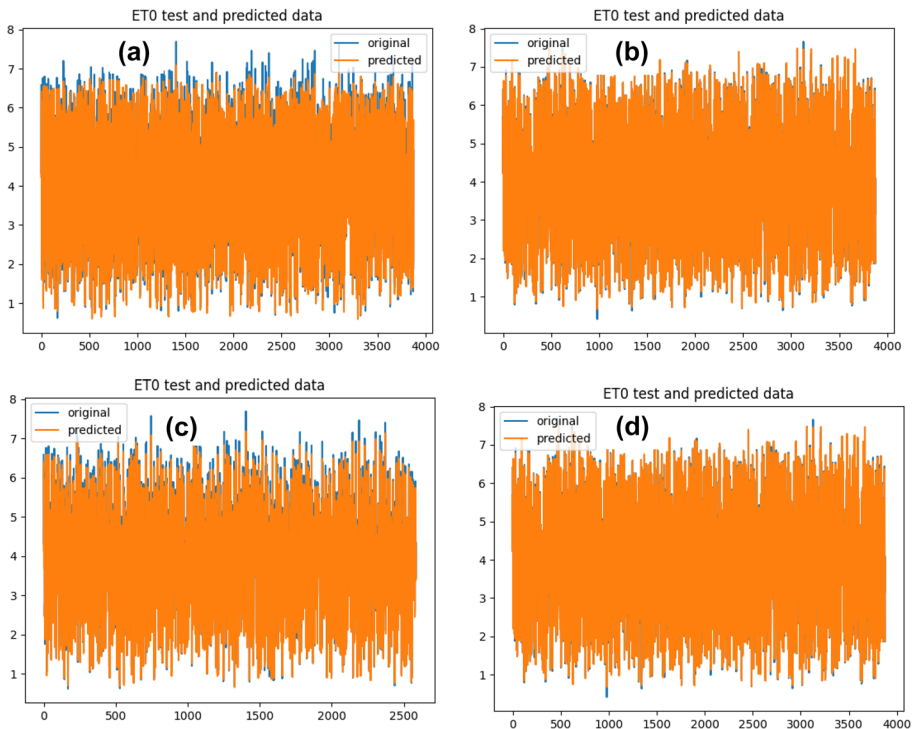
*Bagging LR* linear regression bagging, *Bagging RF* Random Forest- based Regressor Bagging, *XGBOOST* Extreme Gradient Boosting, *LGBM* Light Gradient Boosting Model  
Best statistical values of RMSE (mm/day), MAE (mm/day) and NSE of both training and testing phases

as observed in the above-mentioned figures. The predictive accuracies of the Bagging RF and XGBOOST models were reasonably high. The Bagging LR model generated more scattered ET<sub>0</sub> estimates and delivered the worst performance. Similarly, the scatter plots for the remaining nine stations revealed that the LGBM model was found to be the best performing and the Bagging LR model was found to be the least model. The scatterplots for the nine stations are provided in the supplement file.

## 5 Discussion

### 5.1 Comparison of Model Stability Under Different Input Combinations Across 10 Stations

The stability of a model is an important criterion that indicates how far the predictions have deviated from the measurements when the training dataset is modified. Stability can be determined by the percentage of increase in the values of statistical parameters in the testing phase with respect to those in the training phase. The average training and testing

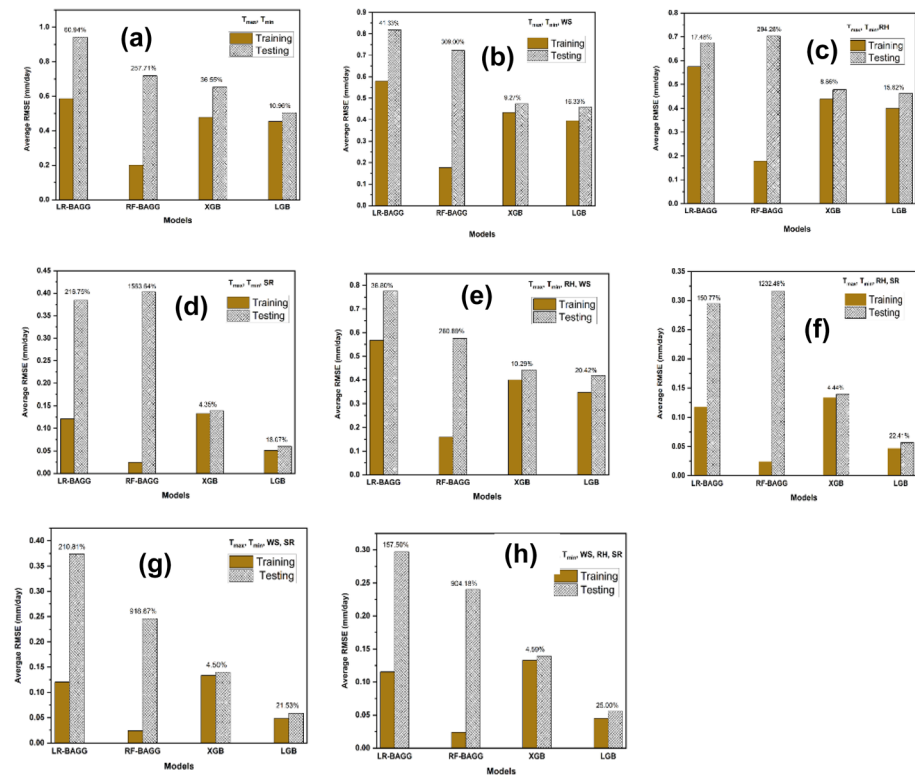


**Fig. 4** Actual and predicted values of ET<sub>0</sub> of Bangalore using **a** Linear Regression Bagging, **b** Random Forest Regressor Bagging, **c** Extreme Gradient Boosting, **d** Light Gradient Boosting models

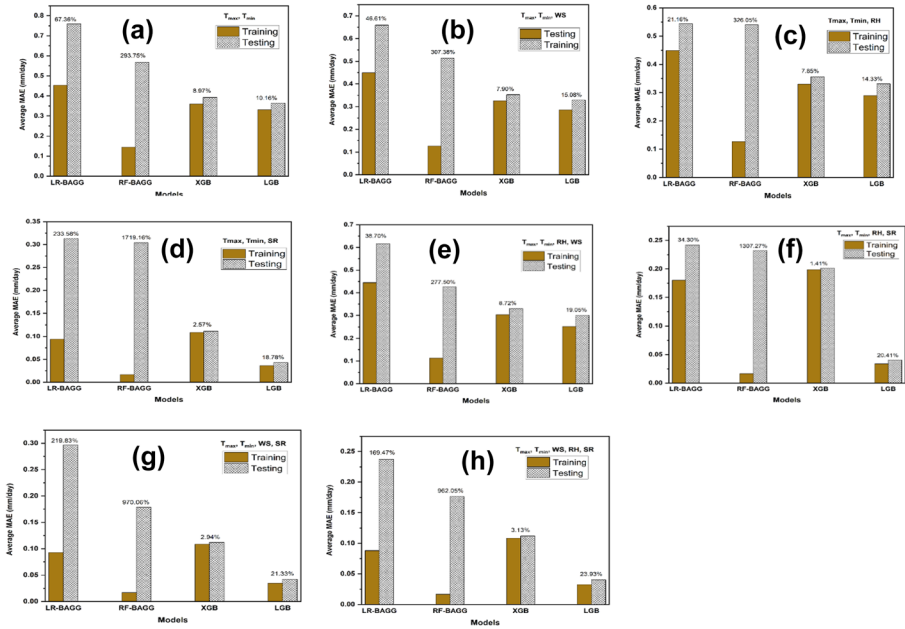
RMSEs for the 10 weather stations and the percentage of increases in the testing RMSE over the training RMSE (average for all weather stations) for the eight input combinations are presented in Fig. 5.

Figures 5a–h indicate that the XGBOOST model was the most stable, followed by the LGBM model. They provided the smallest percentage of increase in the MAE, ranging from 4.35% to 36.55% and 10.96% to 25%, respectively, for most input combinations in comparison with the RF regressor. Bagging RF was the most unstable model and failed to produce accurate values of  $ET_0$  when new input data (test data) were provided. The results agreed well with those obtained by Antonopoulos and Antonopoulos (2017), and Huang et al. (2019), highlighting that Bagging RF was the most unstable of the models considered. It also recorded a notably large percentage of increase in the MAE at all stations under the eight input combinations, as depicted in Figs. 6a–h. This implies that the Bagging RF model failed to produce precise predictions of  $ET_0$  on the test dataset across the 10 stations.

It is remarkable to note that Bagging RF was the most unstable, followed by the Bagging LR model, for the input combination (Tmax, Tmin, SR), and contributed the highest percentage of increase in the MAE. LGBM showed a reasonably high percentage of increase, ranging from 10.16% to 23.93%, for the input combinations. XGBOOST was the most stable model with the smallest percentage of gain, in the range of 1.41% to 8.97%.



**Fig. 5** a–h Percentage increase in testing RMSE over training RMSE values averaged over 4 models under 8 input combinations



**Fig. 6** a–h Percentage increase in testing MAE over training MAE values averaged over 4 models under 8 input combinations

### 5.2 Comparison of Overall Rank of Models Based on Three Metrics

In this study, four ensemble models were developed by using eight input combinations of the climate variables collected from ten climate stations to predict the  $ET_0$  on a daily scale. A ranking system was established to evaluate the overall performance of a total of 32 models (4×8), and the model with the smallest error value or the highest NSE was ranked number 1 (i.e., ranking score = 1) (Elagib and Mansell 2000). The average value of each statistical indicator over the 10 stations was taken to rank the models. The ranking procedure is summarized in the Tables 2, 3 and 4.

### 5.3 Comparison of Computational Costs of Ensemble Models

A comparison of computational costs of the four ensemble models (i.e., the time needed for computation, in seconds) for predicting  $ET_0$  across the 10 stations was done. The average computing time was dependent on the volume of the training data. Figure 7 shows the average computational costs of the four models over 10 stations under the eight input combinations. The results reveal that the average time taken by the Bagging LR model was much shorter than those of the other algorithms for all input combinations across all stations. The Bagging RF model incurred the highest computational cost, which can be attributed to its large number of decision trees.

**Table 2** Ranking scores averaged over 10 stations for RMSE

Models	Rank
Bagging RF6	1
Bagging RF8	2
Bagging RF7	3
Bagging RF4	4
LGBM8	5
LGBM6	6
LGBM7	7
LGBM4	8
Bagging LR8	9
Bagging LR6	10
Bagging LR7	11
Bagging LR4	12
XGBOOST6	13
XGBOOST8	13
XGBOOST7	14
XGBOOST4	15
Bagging RF5	16
Bagging RF2	17
Bagging RF3	18
Bagging RF1	19
LGBM5	20
LGBM2	21
LGBM3	22
XGBOOST5	23
XGBOOST2	24
XGBOOST3	25
LGBM1	26
XGBOOST1	27
Bagging LR5	28
Bagging LR3	29
Bagging LR2	30
Bagging LR1	31

*Bagging LR* linear regression bagging, *Bagging RF* Random Forest-based Regressor Bagging, *XGBOOST* Extreme Gradient Boosting, *LGBM* Light Gradient Boosting Model

Huang et al. (2019) also reported that the average time taken by the Bagging RF model was much higher than the corresponding boosting-based ensemble technique in predicting the daily  $ET_0$ . By contrast, the Bagging RF model took less time than the SVM, ELM, and Gradient-Boosting Decision Tree (GBDT). This contradicts the results of this work (Fan et al. 2018). However, in general, an increase in the number of climatic variables in the models contributed to rising computational costs.

**Table 3** Ranking scores averaged over 10 stations for MAE

<b>Models</b>	<b>Rank</b>
Bagging RF6	1
Bagging RF8	2
Bagging RF4	3
Bagging RF7	3
LGBM8	4
LGBM6	5
LGBM7	6
LGBM4	7
Bagging LR8	8
Bagging LR7	9
Bagging LR4	10
XGBOOST8	11
XGBOOST7	12
XGBOOST4	13
Bagging RF5	14
Bagging RF2	15
Bagging RF3	16
Bagging RF1	17
Bagging LR6	18
XGBOOST6	19
LGBM5	20
LGBM2	21
LGBM3	22
XGBOOST5	23
XGBOOST2	24
XGBOOST3	25
LGBM1	26
XGBOOST1	27
Bagging LR5	28
Bagging LR3	29
Bagging LR2	30
Bagging LR1	31

*Bagging LR* linear regression bagging, *Bagging RF* Random Forest-based Regressor Bagging, XGBOOST- Extreme Gradient Boosting, LGBM- Light Gradient Boosting Model

**Table 4** Ranking scores averaged over 10 stations for NSE

<b>Models</b>	<b>Rank</b>
Bagging RF4	1
Bagging RF6	1
Bagging RF7	1
Bagging RF8	1
LGBM4	2
LGBM6	2
LGBM7	2
LGBM8	2
Bagging LR8	3
Bagging LR6	4
Bagging LR4	5
Bagging LR7	5
XGBOOST6	6
XGBOOST7	6
XGBOOST8	6
XGBOOST4	7
Bagging RF5	8
Bagging RF3	9
Bagging RF2	9
Bagging RF1	10
LGBM5	11
LGBM3	12
LGBM2	13
XGBOOST5	14
XGBOOST2	15
XGBOOST3	16
LGBM1	17
XGBOOST1	18
Bagging LR5	19
Bagging LR3	20
Bagging LR2	21
Bagging LR1	22

*Bagging LR* linear regression bagging, *Bagging RF* Random Forest-based Regressor Bagging, *XGBOOST* Extreme Gradient Boosting, *LGBM* Light Gradient Boosting Model



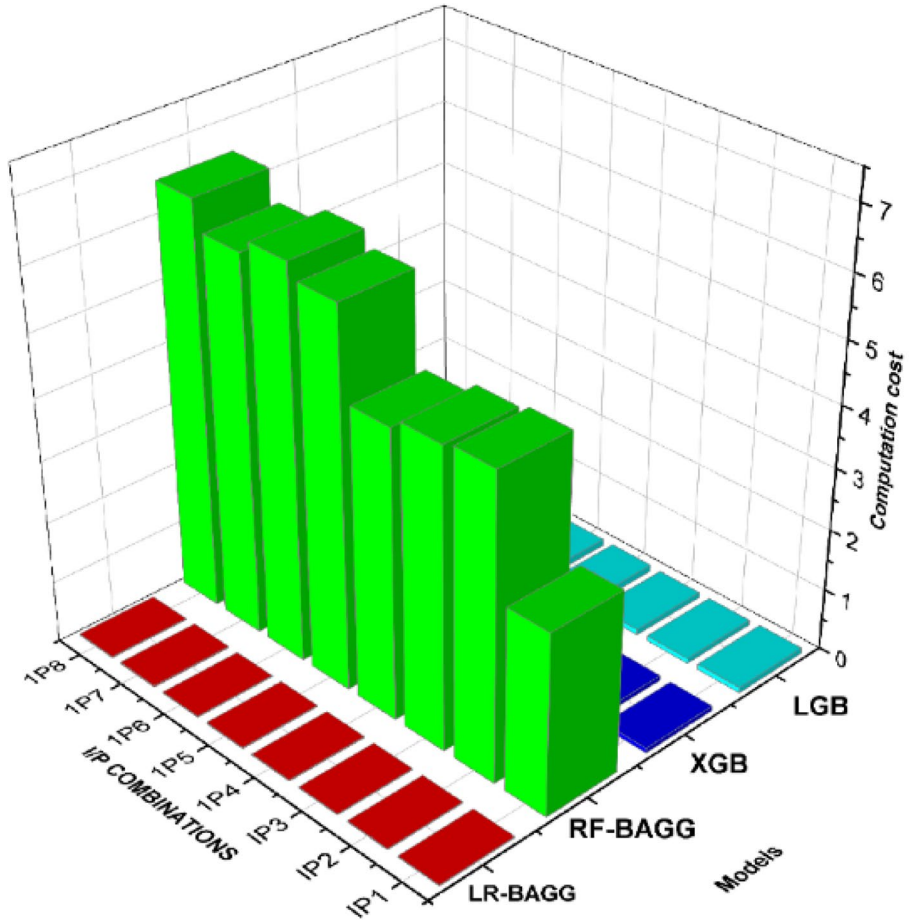


Fig. 7 Comparison of computational costs of four machine learning models under 8 input combinations

### 6 Conclusions

The present study has assessed the performances of four ensemble techniques- Linear Regressor Bagging, Random Forest Regressor Bagging, Extreme Gradient Boosting, and Light Gradient Boosting models in predicting values of the daily  $ET_0$  at 10 agro-climatic stations in Karnataka over 35 years (1979–2014). All four models yielded accurate values of the daily  $ET_0$  with the complete climate dataset (Tmax, Tmin, WS, RH, SR). However, the models that used four sets of input combinations, viz., Tmax, Tmin, and SR; Tmax, Tmin, RH, and WS; Tmax, Tmin, RH, and SR; and Tmax, Tmin, WS, and SR, yielded satisfactory estimates of  $ET_0$ . Of the four models, the Bagging RF and LGBM models had the most accurate predictions. Even though the Bagging RF model was ranked the highest and delivered the best performance based on statistical evaluations,

it was the most expensive model in terms of computation cost while Bagging LR was the least expensive. The XGBOOST model was the most stable of the four models on modified training datasets. Each of the models considered here had unique advantages in terms of predictive accuracy, stability, and computational cost. Thus, the results obtained here may provide useful information for researchers in adopting models for the irrigation-related management of agricultural activities.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11269-022-03399-4>.

**Authors Contributions** Conceptualization: Jayashree T R. Data Collection and formal analysis: Jayashree T R. Methodology: Jayashree T R. Programming and Validation: Jayashree T R. Visualization: Jayashree T R. Supervision: Subba Reddy N V. Writing- Original Draft: Jayashree T R. Writing – review and editing: Dinesh Acharya U. All authors read and approved the final manuscript.

**Funding** Open access funding provided by Manipal Academy of Higher Education, Manipal.

**Data Availability** The climate datasets used in this study are freely available in CFSR Global Weather Data for SWAT (<https://swat.tamu.edu/data/cfsr>).

## Declarations

**Competing Interests** The authors declare that they have no known competing financial interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Agrawal Y, Kumar M, Ananthkrishnan S, Kumarapuram G (2022) Evapotranspiration modeling using different tree based ensemble machine learning algorithm. *Water Resour Manag* 36(3):1025–1042. <https://doi.org/10.1007/s11269-022-03067-7>
- Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao Rome* 300(9):D05109
- Almorox J, Quej VH, Martí P (2015) Global performance ranking of temperature-based approaches for evapotranspiration estimation considering Köppen climate classes. *J Hydrol* 528:514–522. <https://doi.org/10.1016/j.jhydrol.2015.06.057>
- Antonopoulos VZ, Antonopoulos AV (2017) Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. *Comput Electron Agric* 132:86–96. <https://doi.org/10.1016/j.compag.2016.11.011>
- Breiman L (1996) Bagging predictors. *Mach Learn* 24(2):123–140. <https://doi.org/10.1007/BF00058655>
- Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32. <https://doi.org/10.1023/A:1010933404324>
- Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H, Chen K (2015) Xgboost: extreme gradient boosting. *R package version 0.4–2*, 1(4), 1–4
- Chen Z, Zhu Z, Jiang H, Sun S (2020) Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. *J Hydrol* 591:125286. <https://doi.org/10.1016/j.jhydrol.2020.125286>
- dos Santos Farias DB, Althoff D, Rodrigues LN, Filgueiras R (2020) Performance evaluation of numerical and machine learning methods in estimating reference evapotranspiration in a Brazilian agricultural frontier. *Theoret Appl Climatol* 142(3):1481–1492. <https://doi.org/10.1007/s00704-020-03380-4>

- Elagib NA, Mansell MG (2000) New approaches for estimating global solar radiation across Sudan. *Energy Convers Manag* 41(5):419–434. [https://doi.org/10.1016/S0196-8904\(99\)00123-5](https://doi.org/10.1016/S0196-8904(99)00123-5)
- Elbeltagi A, Raza A, Hu Y, Al-Ansari N, Kushwaha NL, Srivastava A, Zubair M (2022a) Data intelligence and hybrid metaheuristic algorithms-based estimation of reference evapotranspiration. *Appl Water Sci* 12(7):1–18. <https://doi.org/10.1007/s13201-022-01667-7>
- Elbeltagi A, Nagy A, Mohammed S, Pande CB, Kumar M, Bhat SA, Juhász C (2022b) Combination of limited meteorological data for predicting reference crop evapotranspiration using artificial neural network method. *Agronomy* 12(2):516. <https://doi.org/10.3390/agronomy12020516>
- Fan J, Yue W, Wu L, Zhang F, Cai H, Wang X, Xiang Y (2018) Evaluation of SVM, ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in different climates of China. *Agric for Meteorol* 263:225–241. <https://doi.org/10.1016/j.agrformet.2018.08.019>
- Fan J, Ma X, Wu L, Zhang F, Yu X, Zeng W (2019) Light Gradient Boosting Machine: An efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data. *Agric Water Manag* 225:105758. <https://doi.org/10.1016/j.agwat.2019.105758>
- Feng Y, Peng Y, Cui N, Gong D, Zhang K (2017) Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. *Comput Electron Agric* 136:71–78. <https://doi.org/10.1016/j.compag.2017.01.027>
- Geurts P, Ernst D, Wehenkel L (2006) Extremely randomized trees. *Mach Learn* 63(1):3–42. <https://doi.org/10.1007/s10994-006-6226-1>
- Gocić M, Amiri MA (2021) Reference evapotranspiration prediction using neural networks and optimum time lags. *Water Resour Manag* 35(6):1913–1926. <https://doi.org/10.1007/s11269-021-02820-8>
- Hamid RF (2011) Evaluation of Blaney-Criddle equation for estimating evapotranspiration in south of Iran. *Afr J Agric Res* 6(13):3103–3109. <https://doi.org/10.5897/AJAR11.421>
- Han Y, Wu J, Zhai B, Pan Y, Huang G, Wu L, Zeng W (2019) Coupling a bat algorithm with xgboost to estimate reference evapotranspiration in the arid and semiarid regions of china. *Adv Meteorol*. <https://doi.org/10.1155/2019/9575782>
- Huang G, Wu L, Ma X, Zhang W, Fan J, Yu X, Zhou H (2019) Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions. *J Hydrol* 574:1029–1041. <https://doi.org/10.1016/j.jhydrol.2019.04.085>
- Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Liu TY (2017) Lightgbm: A highly efficient gradient boosting decision tree. *Adv Neural Inf Process Syst* 30:3146–3154
- Kim SJ, Bae SJ, Jang MW (2022) Linear regression machine learning algorithms for estimating reference evapotranspiration using limited climate data. *Sustainability* 14(18):11674. <https://doi.org/10.3390/su141811674>
- Manikumari N, Murugappan A, Vinodhini G (2017) Time series forecasting of daily reference evapotranspiration by neural network ensemble learning for irrigation system. In *IOP Conference Series: Earth and Environmental Science* (Vol. 80, No. 1, p. 012069). IOP Publishing. <https://doi.org/10.1088/1755-1315/80/1/012069>
- Mehdizadeh S, Mohammadi B, Pham QB, Duan Z (2021) Development of boosted machine learning models for estimating daily reference evapotranspiration and comparison with empirical approaches. *Water* 13(24):3489. <https://doi.org/10.3390/w13243489>
- Muhammad Adnan R, Chen Z, Yuan X, Kisi O, El-Shafie A, Kuriqi A, Ikram M (2020) Reference evapotranspiration modeling using new heuristic methods. *Entropy* 22(5):547. <https://doi.org/10.3390/e22050547>
- Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I—A discussion of principles. *J Hydrol* 10(3):282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Pandey V, Pandey PK, Mahanta AP (2014) Calibration and performance verification of Hargreaves Samani equation in a humid region. *Irrig Drain* 63(5):659–667. <https://doi.org/10.1002/ird.1874>
- Ponraj AS, Vigneswaran T (2019) Daily evapotranspiration prediction using gradient boost regression model for irrigation planning. *J Supercomput* 1–13. <https://doi.org/10.1007/s11227-019-02965-9>
- Priestley CHB, TAYLOR, R. J. (1972) On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon Weather Rev* 100(2):81–92. [https://doi.org/10.1175/1520-0493\(1972\)100%3c0081:OTAOSH%3e2.3.CO;2](https://doi.org/10.1175/1520-0493(1972)100%3c0081:OTAOSH%3e2.3.CO;2)
- Ramachandra JT, Veerappa SRN, Udipi DA (2022) Assessment of spatiotemporal variability and trend analysis of reference crop evapotranspiration for the southern region of Peninsular India. *Environ Sci Pollut Res* 1–18. <https://doi.org/10.1007/s11356-021-15958-0>
- Roy DK, Saha KK, Kamruzzaman M, Biswas SK, Hossain MA (2021) Hierarchical fuzzy systems integrated with particle swarm optimization for daily reference evapotranspiration prediction: A novel approach. *Water Resour Manag* 35(15):5383–5407. <https://doi.org/10.1007/s11269-021-03009-9>

- Sattari MT, Apaydin H, Band SS, Mosavi A, Prasad R (2021) Comparative analysis of kernel-based versus ANN and deep learning methods in monthly reference evapotranspiration estimation. *Hydrol Earth Syst Sci* 25(2):603–618. <https://doi.org/10.5194/hess-25-603-2021>
- Sharafati A, Asadollah HS, S. B., Motta, D., & Yaseen, Z. M. (2020) Application of newly developed ensemble machine learning models for daily suspended sediment load prediction and related uncertainty analysis. *Hydrol Sci J* 65(12):2022–2042. <https://doi.org/10.1080/02626667.2020.1786571>
- Talaviya T, Shah D, Patel N, Yagnik H, Shah M (2020) Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides. *Artif Intell Agric* 4:58–73. <https://doi.org/10.1016/j.aiia.2020.04.002>
- Tao H, Diop L, Bodian A, Djaman K, Ndiaye PM, Yaseen ZM (2018) Reference evapotranspiration prediction using hybridized fuzzy model with firefly algorithm: Regional case study in Burkina Faso. *Agric Water Manag* 208:140–151. <https://doi.org/10.1016/j.agwat.2018.06.018>
- Wang W, Shao Q, Peng S, Xing W, Yang T, Luo Y, Xu J (2012) Reference evapotranspiration change and the causes across the Yellow River Basin during 1957–2008 and their spatial and seasonal differences. *Water Resour Res* 48(5). <https://doi.org/10.1029/2011WR010724>
- Wu L, Peng Y, Fan J, Wang Y (2019) Machine learning models for the estimation of monthly mean daily reference evapotranspiration based on cross-station and synthetic data. *Hydrol Res* 50(6):1730–1750. <https://doi.org/10.2166/nh.2019.060>
- Wu T, Zhang W, Jiao X, Guo W, Hamoud YA (2020) Comparison of five Boosting-based models for estimating daily reference evapotranspiration with limited meteorological variables. *PLoS ONE* 15(6):e0235324. <https://doi.org/10.1371/journal.pone.0235324>
- Yamaç SS, Todorovic M (2020) Estimation of daily potato crop evapotranspiration using three different machine learning algorithms and four scenarios of available meteorological data. *Agric Water Manag* 228:105875. <https://doi.org/10.1016/j.agwat.2019.105875>
- Zhu B, Feng Y, Gong D, Jiang S, Zhao L, Cui N (2020) Hybrid particle swarm optimization with extreme learning machine for daily reference evapotranspiration prediction from limited climatic data. *Comput Electron Agric* 173:105430. <https://doi.org/10.1016/j.compag.2020.105430>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.