Pol. J. Environ. Stud. Vol. 32, No. 2 (2023), 1009-1023
DOI: 10.15244/pjoes/156927 ONLINE PUBLICATION DATE: 2023-01-16

Original Research

Evaluation of Tree-Based Machine Learning and Deep Learning Techniques in Temperature-Based Potential Evapotranspiration Prediction

Funda Akar[1], Okan Mert Katipoğlu[2]*, Sefa Nur Yeşilyurt[2], Mehmet Bilge Han Taş[1]

[1]Erzincan Binali Yıldırım University, Department of Computer Engineering, Erzincan, Turkey
[2]Erzincan Binali Yıldırım University, Department of Civil Engineering, Erzincan, Turkey

Received: 3 October 2022
Accepted: 27 November 2022

Abstract

In this study, Extreme Gradient Boosting (XGBoost), Gradient Boosting Machine (GBM), Random Forest (RF), Bagged Trees (BT), and Custom Deep Learning methods were used to estimate the potential evapotranspiration (PET) values at Diyarbakir airport station in the Tigris basin. In establishing the models, the average temperature, maximum temperature, minimum temperature, maximum wind speed, relative humidity, average wind speed, and total precipitation values in the monthly time period were chosen as inputs, and PET values were used as output. The data set is divided into 70% training and 30% testing. 10-fold cross-validation to avoid overfitting problems. Training and test data were randomly selected. The prediction performances of the models were evaluated according to the statistical criteria of determination coefficient (R^2), root mean square error (RMSE), mean absolute error (MAE), and rank analysis. The best PET estimates were obtained using the inputs of mean, min, maximum temperature, relative humidity, total precipitation, average, and maximum wind speed. It was also concluded that XGBoost was the highest performance. When the R^2 values were examined, it was seen that the Deep Learning model had higher performance. But for RMSE and MAE, XGBoost did better. As a result of the rank analysis, it was seen that XGBoost got a higher score.

Keywords: potential evapotranspiration, Thornthwaite equation, machine learning, deep learning, meteorological data

*e-mail: okatipoglu@erzincan.edu.tr

Introduction

Evapotranspiration refers to the plant's water consumption and evaporation, which is the combination of evaporation and transpiration. Accurate estimation of evapotranspiration (ET) values is vital in terms of the water budget, calculation of water needs, and management of water resources and irrigation systems [1]. ET values can be estimated through meteorological and climatological parameters, empirical equations, and artificial intelligence methods. However, ET values are a complex and non-linear process as they depend on various meteorological variables such as wind speed, air temperature, relative humidity, and solar radiation [2].

Machine learning methods have been widely employed for estimating ET. Yassin et al. [3] evaluated the performance of gene expression programming (GEP) and ANN methods to predict daily ET values in the Kingdom of Saudi Arabia. According to statistical indicators, it has been determined that ANN produces more accurate results than GEP. As a result of the study, it was revealed that SVM-RBF and MARS methods were mostly more successful than SVM-Poly and GEP. Afzaal et al. [4] estimated the ET value in the North Cape, Summerside, Harrington, and Saint Peters regions with recurrent neural networks (RNNs) and long short-term memory (LSTM). As a result, it has been determined that LSTM can successfully predict the ET value in almost all regions. Chen et al. [5] developed deep learning (DL), a temporal convolution neural network (TCNN), and LSTM to predict daily PET. In addition, the performances of three DL models were evaluated against SVM, random forest (RF), and empirical equations. It was obtained that the TCNN and LSTM models performed significantly better than the temperature-based experimental models. Chia et al. [6] was applied the adaptive neuro-fuzzy inference system (ANFIS) model to estimate evapotranspiration values in Labuan, East Malaysia. It has been concluded that the ANFIS model can obtain accurate predictions with only one year of training data, and there is no need to make long-term predictions. Zhou et al. [7] evaluated the performances of DeepFM, three gradient boosting models (CatBoost, LightGBM, and XGBoost), three tree-based models (GBDT, RF, and ET), and SVM-RBF models to predict daily ET values in China. As a result, it has been determined that the CatBoost and LightGBM models are relatively more successful in ET estimation and the sunshine duration is an effective parameter. Agrawal et al. [8] RF, DT, AdaBoost, GBM, and XGBoost potential were evaluated in daily ET prediction. XGBoost was the most superior among the ensembled machine learning models used. When the studies in the literature are examined, it has been determined that the studies on comparing the prediction performances of XGBoost and GBM, RF and Deep Learning methods in PET prediction are lacking. In addition, the current study is the first for PET estimation using various meteorological variables in the province of Diyarbakir in terms of the region's agricultural productivity. Therefore, this study contributes to the literature comparing machine learning and deep learning methods in PET estimation in semi-arid climate regions.

Measuring and forecasting weather variables, such as rain, temperature, and humidity, are essential for present and future planning. With the development of deep learning methods in recent years, very good predictions can be made with such datasets [9]. It has a wide range of usage in prediction processes such as soil moisture, signals, student performance, house price, and meteorological variables [10-13]. The deep learning method used in this study can also be performed with other prediction operations. For this reason, the technique used has been generalized as much as possible.

This study aimed to evaluate the performance of machine learning such as XGBoost, GBM, RF, Bagged Trees, and Custom Deep Learning methods to estimate PET values of the Diyarbakır Airport meteorological observation station located in the Tigris Basin. For this purpose, the average temperature (T_{mean}), maximum temperature (T_{max}), minimum temperature (T_{min}), maximum wind speed (WS_{max}), Relative humidity (RH_{mean}), Average wind speed (WS_{mean}), Total precipitation (P_s) values were calculated in the monthly time period as input and Thornthwaite-based PET values as output. In addition, the effect of various input combinations on PET estimation was evaluated. Finally, which meteorological variables were most effective in PET estimation were assessed according to different statistical indicators.

Material and Methods

Study Area and Data

This study used monthly meteorological data of the Diyarbakir airport, numbered 17280, located in the Tigris Basin, in the 1964-2017 time period. The data used were obtained from the Turkish Meteorology General Directorate. The digital elevation model expresses the geographical location of the Diyarbakır airport station used in the study (Fig. 1).

Diyarbakir airport meteorology station, chosen as the study area, is important in energy and agricultural production because it is located in the Southeastern Anatolia Project region. Due to the effect of the fronts coming from the Mediterranean, the most precipitation in Diyarbakır occurs in the winter. However, in the summer season, very high temperatures are generally seen due to the effect of the Basra low pressure [14]. The statistics of the station are presented in Fig. 2.

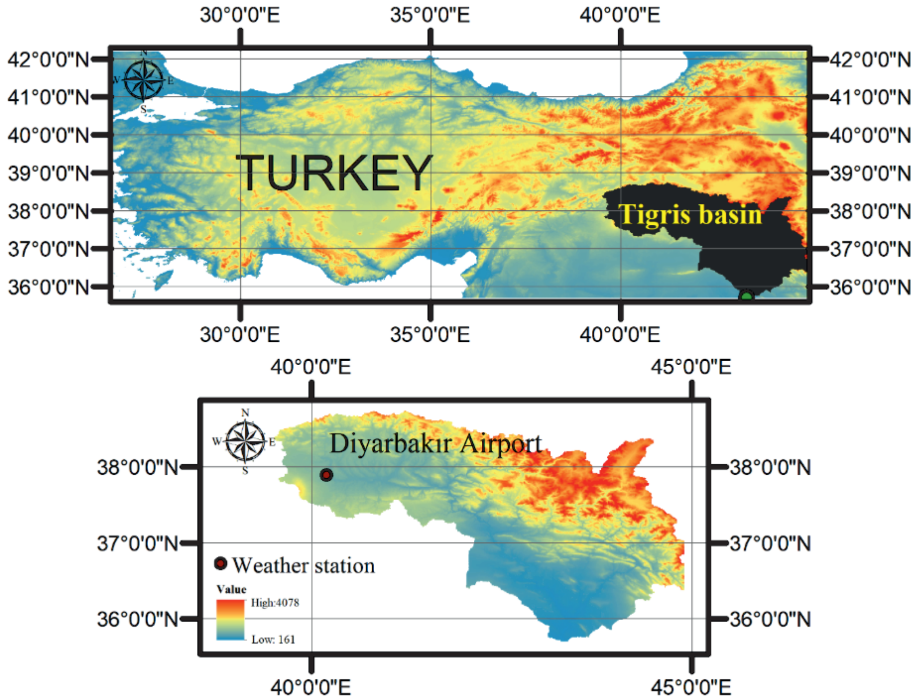

Fig. 1. Geographical locations of the Diyarbakır Airport weather station.

Calculation of PET with the Thornthwaite Method

This study calculated PET values with the Thornthwaite method due to limited data. Equation (1) proposed by [15] was used to calculate PET. This equation relates the monthly average temperature with the PET values determined by the water balance for the valleys considered to be adequately fed by precipitation and groundwater.

$$PET = 16\, d\, \left({}^{10T}/_{I}\right)^{a} \tag{1}$$

Here T is the average temperature of the month (ºC), I is the annual thermal index, that is, the sum of the monthly thermal index (*i*) and is calculated by the equation [$i = (T/5)^{1.514}$]. The *d* values are *a* correction factors based on latitude and month. The *a* values are calculated using Equation (2).

$$a = 0.49 + 0.0179\, I - 0.0000771\, I^2 + 0.000000675\, I^3 \tag{2}$$

Drought Index Calculator (DrinC) software was used to calculate PET values. In this study, various meteorological variables were input to the artificial intelligence models, and prediction models were developed by presenting the PET values calculated with the Thornthwaite Method as output. There is no definitive approach to selecting the training and testing ratio. The training rate is usually chosen between 70-90%. In this study, 70% of the data was selected for training in establishing the artificial intelligence model. The cross-validation technique has been developed for more effective use of machine learning with higher accuracy or performance. For this reason, 10-fold cross-validation technique was used in the study.

Gradient Boosting Machine (GBM)

GBM is a supervised learning method put forward by [16]. Thanks to this method, it can be used with various classes or in solving regression problems. Besides being a supervised machine learning technique, it can be used in many fields. This method basically includes three components. A loss function for optimization, a weak learner for estimating, and an additive model for optimizing the loss function [17]. GBR creates a powerful learner to perform the loss function appropriate to the problem. For iteration, it is applied to a group of weak learner iterations [18, 19]. In the algorithm, ntrees, interaction depth, and shrinkage parameters are optimized. In the first step, by defining the loss function, it examines how well the model coefficients fit the data as given in Equation (3).

$$L(y_i, F(x)) \tag{3}$$

Here y_i the observed value and $F(x)$ is the predicted value. As a second step, by determining the constant variable in Equation (4), all observations are determined and it is aimed to find the situation where the loss function value is minimum.

$$F_0(x) = argmin \sum_{i=1}^{n} L(y_i, \gamma) \tag{4}$$

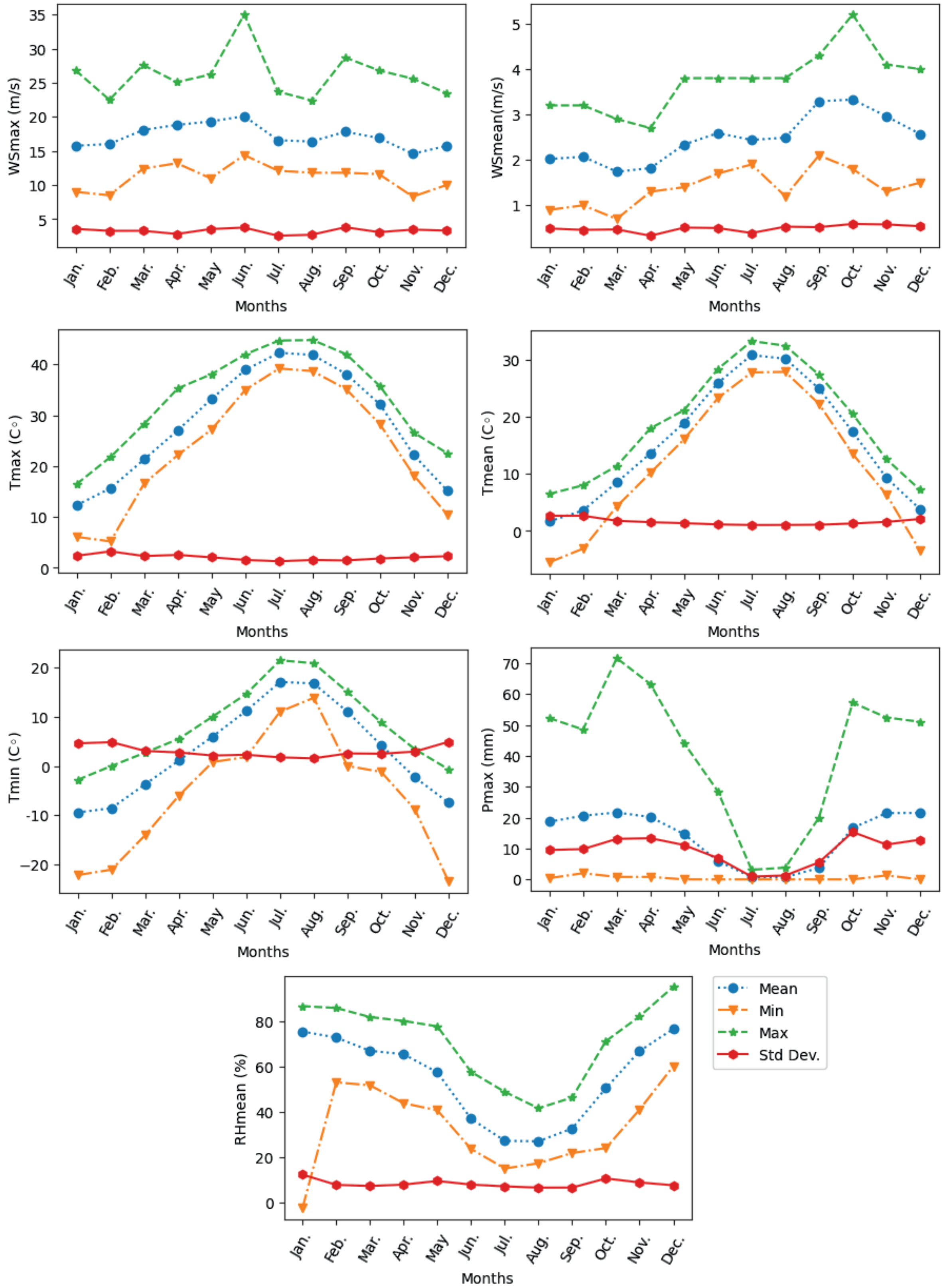

Fig. 2. Statistics of monthly meteorological data of Diyarbakir Airport station.

Modeled and other machine-learning methods were modeled with the help of R-Studio (4.7-1.1) software. Numerous packages were used for GBM. The basic libraries are "gmb", "caret (Classification And REgression Training)" and libraries containing various statistical calculations. In addition, the method has advanced hyperparameter optimization in itself.

Extreme Gradient Boosting (XGBoost)

XGBoost is built on the concept of "boosting," which involves merging all of a group of "weak" learners' predictions to create a "strong" learner through an adaptive training procedure. XGBoost is a quick and high-performance decision tree-based algorithm. Transactions are completed rapidly because this approach does concurrent and distributed computation while generating decision trees [20, 21]. Equation is given below.

$$L^t = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega\,(f_t) \tag{5}$$

The algorithm first detects the depth of the tree with max_depth. If the tree is too deep in the down direction, it continues its operations by pruning backward. Optimizing the max_depth, eta, gamma, min_child_weight, nround, and subsample hyper-parameters in the XGBoost algorithm reduces complexity [22]. Numerous packages were used for XGBoost. The basic libraries are "XGBoost", "gbm", "caret" libraries. The method has advanced hyperparameter optimization in itself.

Random Forest (RF)

RF is a set of tree estimators, each of which is built from the values of a randomly sampled random vector. This method begins at the tree's root node and works its way up, considering all the information. Following that, each predictive variable is calculated. This approach also uses cross-validation pruning to reduce the likelihood of the tree overfitting the data [23]. Many packages were used in Random Forest in R-Studio. But mainly used package "randomForest" which can be used for classification and regression.

Bagged Trees (BT)

BT is a meta-algorithm that improves the efficiency and accuracy of a single decision tree by constructing numerous models and combining them into a single and coherent aggregate estimator [24, 25]. It is also an essential algorithm in terms of solving the overfitting problem. The installation stages of the BA; the random resampling of the training dataset, designing and training an individual model for each subset, and creating a coherent aggregate estimator [26]. The hyperparameters are optimized in the machine learning methods and the methods' main features (Fig. 3). These methods were selected by trial-and-error method-cross validation method. "ipred" library was used as the basis for Bagged Tree. R-Studio software was used for all machine learning algorithms and codes were created in accordance with the data. Code details can be shared upon readers' requests.

Deep Learning Method (DLM)

Deep learning was employed in the study and numerous other machine learning technologies. Only Dense layers are utilized as a technique, not Convolution layers. High Epoch values might be employed in this approach, and better results could be produced. Low-level characteristics are transferred to thick layers in the dense layer to get higher-level features [27].

Large-scale and real-world data were supposed to make the task more essential while the data was being gathered. The chosen stations and their values are true data as a result. Many elements, including the type of data, its organization, and the presence of missing data, are crucial when constructing the deep learning algorithm. For this reason, it is essential to analyze the data collection and prepare optimum accurately. First, the dataset should be filled in with the proper techniques for missing data points. The dataset used for the study contained no missing data. Hence

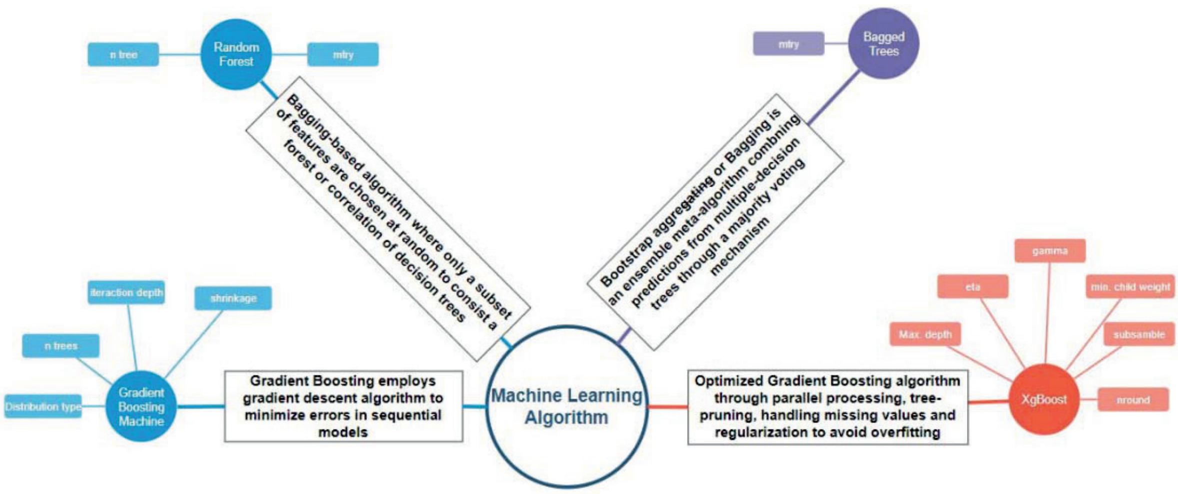

Fig. 3. Machine learning algorithms.

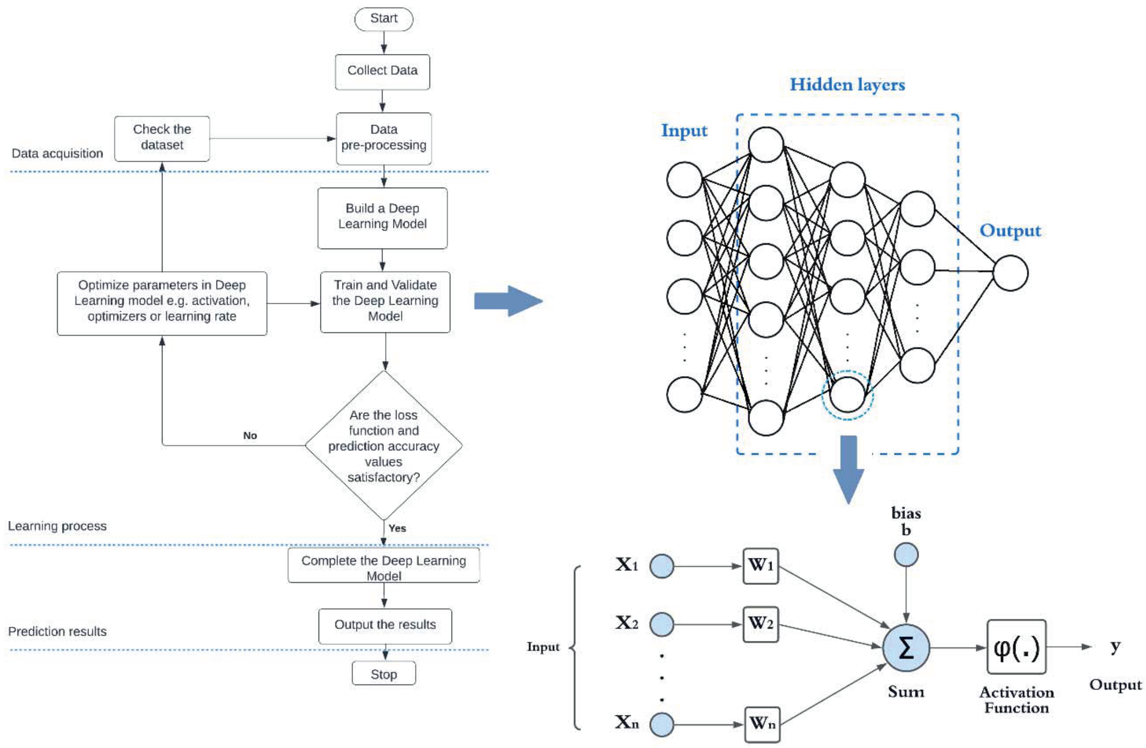

Fig. 4. Flow chart of deep learning model [28].

no filling method was applied. A deep learning method is then developed using the preprocessed dataset. The best appropriate model is chosen. The data set is split into test and training sections, and modifications are made to the model's parameters until the training produces satisfactory results. If the adjustments are insufficient, another attempt is made. If additional preprocessing is required, it is applied after the data set is returned if necessary to provide a suitable result. If the outcomes are adequate and appropriate for solving real-world issues, a deep learning model is considered to have been built. The activation functions and optimizers in the model are tested in various combinations to produce these results. Hidden layers and dataset inputs make up the deep learning model. A weighting process is carried out according to each level's parameters. Artificial intelligence selects the optimum weighted path, and the result is used. Predictions are made and outputs appropriate for performance metrics are obtained after all procedures have been completed. The flowchart used is given in Fig. 4.

Three major Dense layers were employed and used to provide predictions with more consistent and lower loss values, as illustrated in Figure 5. Because the model is assigned seven classes because the initial input layer has data in seven columns. The first layer examines the data's input values for qualities to use.

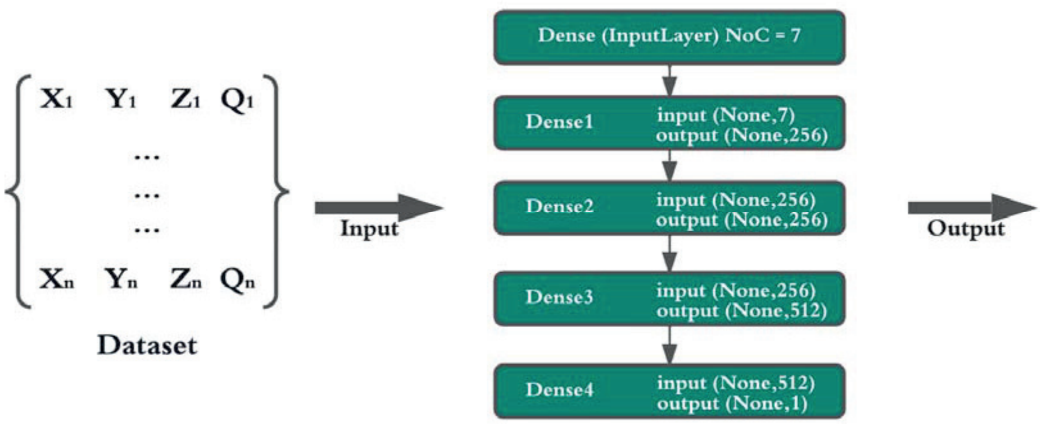

Fig. 5. Proposed deep learning model [28].

For the deep learning model to be made, collecting data from selected stations is necessary first. It is required to test the collected data's suitability and if parts of the data need to be corrected, it is necessary to correct them. Then it is essential to develop a deep learning model suitable for the data. When the collected data is fitted to the model, it is necessary to try again by changing the parameters until satisfactory results are obtained. If the results can make predictions as desired, the operation and setup of the model is correct. Then, the model learning ends, and the prediction process is made.

Many optimizers, activation, and loss measures were attempted throughout the training. The applied methods and their parameters are given in Table 2. In libraries, default values are used for parameters that are not specified. The model combinations used to establish the deep learning model are presented in Table 1. Changes to the optimizer and loss functions were used to build the models shown here. Many parameters have been tested, and the best findings have been presented. The optimizer settings are based on "mean absolute error" [29], which displays the most consistent values as the Loss value. On the other hand, the mean absolute percentage error [30] loss function returns the greatest number. ReLU, the most suitable for the inputs, is preferred as the activation function. "Adam" is mainly used as the optimizer [31]. Adadelta, Adagrad, Adamax, and Nadam were the other optimizers utilized [31-34]. Tensorflow and the keras library in it were used to run the deep learning algorithm. The graphics card used in the study is Nvidia 1660ti 6GB and the processor is Intel i7 9750.

Performance Evaluation

When using machine learning methods, it is challenging to select each parameter manually because there are many parameters. Therefore, using Grid Search is essential to find the best parameters. Therefore, cross-validation is used with grid search, and the k-fold value is set to 10. The performance evaluation of the applied methods was made according to various statistical criteria. These criteria are determination coefficient (R^2), root mean square error (RMSE), and mean absolute error (MAE). Error-values relative to 0 and R^2 values relative to 1 indicate the accuracy of estimation [35]. Equation (6-8) is used to calculate the stated statistical criteria.

$$R^2 = \frac{\sum_{i=1}^{n}(c_i - c_{avg})^2 - \sum_{i=1}^{n}(c_i - p_i)^2}{\sum_{i=1}^{n}(c_i - c_{avg})^2} \tag{6}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(c_i - p_i)^2} \tag{7}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} \mid c_i - p_i \mid \tag{8}$$

where c_i: calculated values of ET, p_i: the predicted values of ET, c_{avg}: the average of the c_i values, p_{avg}: the average of the p_i values, $c_i - p_i$: the value of the errors, and n: number of data.

Rank analysis is a scoring method applied to determine the best performing model among the models by considering all evaluation metrics. It aims to determine the models' score while evaluating the best model's performance. The method is performed by assigning a rank to the models according to their closeness to the best value for each data set, summing the scores for all data sets, and comparing them. R_i denotes the rank in the selected model of each data set and the total rank if n is represented as the number of models [36];

$$Model\ Total\ Rank = \sum_{i=1}^{n} R_i \tag{9}$$

Results and Discussion

Within the scope of the study, XGBoost, GBM, RF, BT, and proposed Deep Learning Model methods were used to estimate Diyarbakır Airport station PET values. The correlation status of the data was examined to achieve the best result. The correlation matrix of the data is given in Fig. 6. Furthermore, combinations were created according to the input with the highest correlation with the PET values in the correlation matrix. The input combinations and model numbers created are given in Table 1.

When Table 1 and Fig. 6 were examined, it was seen that the input with the highest correlation with PET was T_{mean}. It was observed that there was a 0.96 correlation between the T_{mean} and PET. This value shows an excellent relationship between the two values. It was observed that the lowest correlation with PET was in the WS_{max} value.

The model hyper-parameters were determined according to the best results obtained from the experiments. The hyper-parameters used are given in Table 2.

When the results were examined, it was seen that all model results provided sufficient success for estimation. The best values; are 6.5434 RMSE, 0.9844 R^2, and 5.2812 MAE. The results obtained by the rank analysis according to the input combinations are given in Table 3. Looking at the input combinations, the best input combination was seen in the M7. On the other hand, the data combined with the worst results were seen in M2. This situation proves that the artificial intelligence models (M6, M7) obtained by using all meteorological parameters in the study area to make a practical PET estimation give more successful results than the predictions made with limited input variables (M1, M2). In addition, when the M3 and M4 models are compared, it is seen that the M3 model, installed with only temperature components,

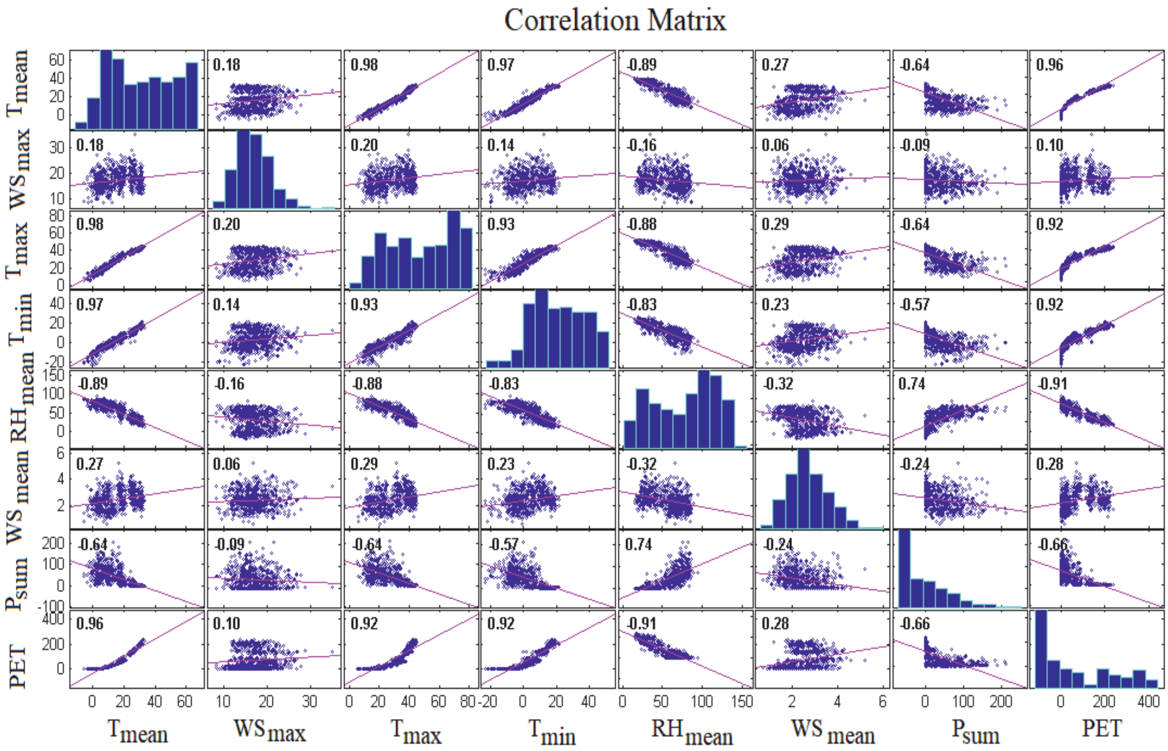

Fig. 6. Correlation matrix.

Table 1. Input Combinations.

Model	Input	Output
M1	T_{mean}	PET
M2	T_{mean}, T_{max}	PET
M3	T_{mean}, T_{min}, T_{max}	PET
M4	T_{mean}, T_{min}, T_{max}, RH_{mean}	PET
M5	T_{mean}, T_{min}, T_{max}, RH_{mean}, P_{sum}	PET
M6	T_{mean}, T_{min}, T_{max}, RH_{mean}, P_{sum}, WS_{mean}	PET
M7	T_{mean}, T_{min}, T_{max}, RH_{mean}, P_{sum}, WS_{mean}, WS_{max}	PET

is superior to the M4 model, which is established with temperature and relative humidity. This result shows that the model's performance decreases when the relative humidity values are used together with the temperature values.

The average temperature value was not used in the study. However, this study shows that the relationship between average temperature and PET is stronger. Table 2 defines the acronyms "me_ab_er," "me_sq_er," and "me_ab_pe_er" as mean absolute error, mean squared error, and mean absolute percentage error, respectively. While creating the deep learning method was determined by considering the smooth curves in the Loss Function as given in Fig. 7. Parameters were fixed with more reasonable curves and other parameters that would affect this was revised. In addition, the original and predicted values of the models given in Fig. 7 and the loss functions are also provided. It is essential to show the change of the loss values parameters and progress within the framework of the best loss value. In this way, the activation functions and optimizers were selected.

When evaluated according to the methods used, it was seen that the best model was XGBoost. GBM, seen as the XGBoost model's ancestor, is in second place. Finally, although the deep learning model is in third place, it has been seen that it is not as successful as XGBoost and GBM. At the same time, the best RMSE and RANK values were seen in the XGBoost model (Table 3).

When RMSE, MAE and R results are analyzed with the help of (Fig. 8), it is seen that Deep Learning R^2 values are better than other models but lag behind XGBoost and GBM algorithms in RMSE and MAE metrics. This shows that the Deep Learning model exhibits high performance, but the error values are larger. In this case, the rank analysis clearly shows the result when all models are evaluated together.

The change in the performance statistics of the established artificial intelligence models is shown in Fig. 8. Among these statistics, the model with the lowest error rate and the highest R^2 value was evaluated as the most successful. Accordingly, it has been determined that XGBoost models are slightly more successful than other models.

This study aims to evaluate the effect of various meteorological variables in the estimation of PET values and to reveal which artificial intelligence technique makes more accurate predictions in semi-arid regions.

Table 2. The best parameters.

XgBoost							
	M1	M2	M3	M4	M5	M6	M7
Booster	Gblinear	Gblinear	Gblinear	Gblinear	Gblinear	Gblinear	Gblinear
Max depth	2	3	2	2	3	3	3
Eta	0.025	0.01	0.05	0.025	0.025	0.025	0.025
Gamma	0.9	0.5	1	0	0	0	0
Min child weight	3	1	1	3	1	3	1
Subsample	0.5	0.5	0.75	1	1	1	1
Nround	350	650	100	250	200	200	200
Other	Default	Default	Default	Default	Default	Default	Default
Gbm							
	M1	M2	M3	M4	M5	M6	M7
Distirbution	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian	Gaussian
N trees	5000	5000	5000	5000	5000	5000	5000
Iteraction depth	1	3	1	1	1	1	3
Shrinkage	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Other	Default	Default	Default	Default	Default	Default	Default
Rf							
	M1	M2	M3	M4	M5	M6	M7
Mtry	500-1000						
Ntree	100-1000						
Other	Default	Default	Default	Default	Default	Default	Default
Bagged trees							
	M1	M2	M3	M4	M5	M6	M7
Mtry	500-1000						
Deep learning (custom model)							
	M1	M2	M3	M4	M5	M6	M7
Activation	Relu						
Batch size	1						
Epoch	10000						
Optimizer	Adam	Adam	Adam	Adadelta	Adagrad	Adamax	Nadam
Loss	Me_ab_er	Me_sq_er	Me_ab_pe_er	Me_ab_err	Me_ab_er	Me_ab_er	Me_ab_er

PET estimation results were evaluated according to various statistical criteria. The most successful artificial intelligence technique was determined as XGBoost. The results confirmed the power of XGBoost stated in the [20] study and the [37] study. Furthermore, XGBoost gave much better results than other algorithms due to its high predictive power, avoidance of over-learning, fastness, and system optimization. In the study by [38], RF and XGBoost models were also used in PET estimation. In this study, the highest performance was achieved in the XGBoost algorithm. In the study of [8], XGBoost, and Random Forest algorithms were used, and it was concluded that the data with the highest correlation with PET values was the maximum temperature. Wang et al. [39] used various algorithms and meteorological variables to estimate reference evapotranspiration values. As a result, the RF-based reference evapotranspiration prediction models gave slightly more successful prediction results than the gene-expression programming model. In addition, the order of importance of meteorological variables in the estimation of reference evapotranspiration: sunshine

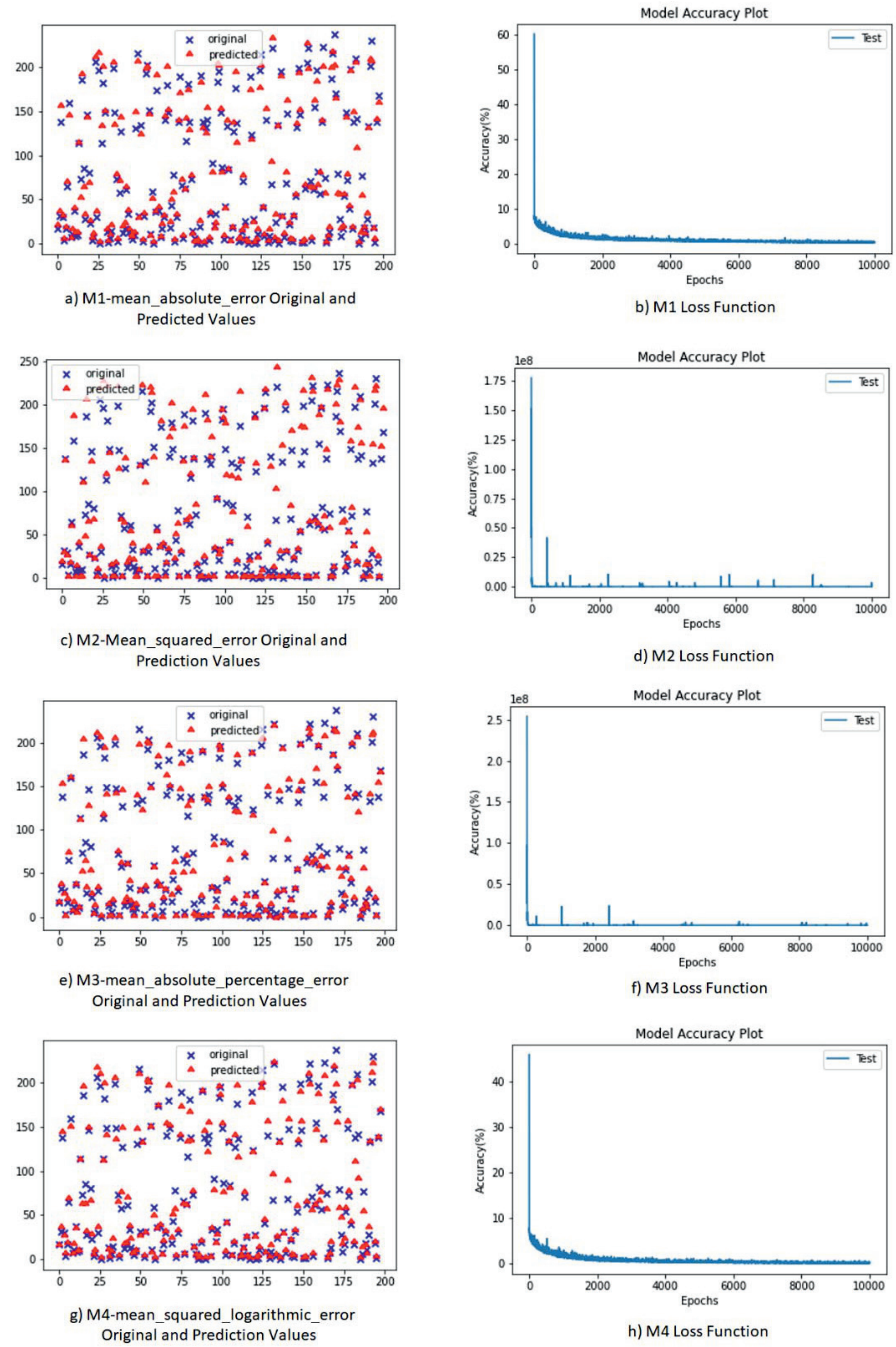

Fig. 7. Original value, prediction value and loss functions.

Table 3. Rank analysis by input combinations.

NO	MODEL	MODEL TYPE	MODEL					
			RMSE	RMSE Rank	R^2	R^2 Rank	MAE	MAE Rank
1	M1	XGBOOST	9.1833	3	0.9841	6	6.4161	3
2	M2		12.0148	1	0.9727	5	7.7594	1
3	M3		9.2090	2	**0.9844**	7	6.4522	2
4	M4		7.4873	4	0.8615	1	6.0140	4
5	M5		6.7060	5	0.8868	2	5.3938	5
6	M6		**6.5434**	7	0.8933	4	**5.3026**	7
7	M7		6.6363	6	0.8900	3	5.3529	6
8	M1	GBM	9.3175	2	0.9837	6	6.4573	3
9	M2		12.02762	1	0.9729	5	8.0426	1
10	M3		9.1978	3	**0.9843**	7	6.6698	2
11	M4		7.3980	4	0.8651	1	5.8522	4
12	M5		**6,7324**	7	0,8858	2	5,3779	6
13	M6		6,7452	6	0.8880	4	**5.3714**	7
14	M7		6.7901	5	0.8878	3	5.4031	5
15	M1	RF	10.0294	2	0.9809	6	6.4592	2
16	M2		12.7501	1	0.9696	5	8.4116	1
17	M3		9.7624	3	**0.9822**	7	6.6109	3
18	M4		7.5712	4	0.8570	1	6.0059	4
19	M5		7.0536	5	0.8758	2	5.6653	5
20	M6		6.9726	6	0.8811	3	5.5690	6
21	M7		**6.8473**	7	0.8863	4	**5.2812**	7
22	M1	BT	10.0188	2	**0.9809**	7	6.4404	3
23	M2		13.3724	1	0.9663	5	8.4116	1
24	M3		9.8349	3	0.9817	6	6.4557	2
25	M4		7.6840	4	0.8538	1	6.1249	4
26	M5		6.9397	5	0.8801	2	5.5029	6
27	M6		6.8947	6	0.8840	3	5.5137	5
28	M7		**6.8545**	7	0.8859	4	**5.4667**	7
29	M1	DL	10.0754	5	0.9811	5	6.6291	5
30	M2		13.0063	1	0.9685	1	8.9850	1
31	M3		**9.1884**	7	**0.9843**	7	**6.3861**	7
32	M4		10.0365	6	0.9812	6	6.4213	6
33	M5		10.8662	3	0.97802	3	7.4091	3
34	M6		12.0255	2	0.9730	2	7.51526	2
35	M7		10.1332	4	0.9809	4	6.7611	4
TOTAL RANK/ MODEL			M1	M2	M3	M4	M5	M6	**M7**
			60	31	68	54	61	70	**76**

Table 4. Rank analysis according to algorithms.

NO	MODEL	MODEL TYPE	RMSE	RMSE Rank	R^2	R^2 Rank	MAE	MAE Rank	TOTAL Rank
1	M1	XGBOOST	**9.1833**	5	**0.9841**	5	**6.4161**	5	**91**
2	M2		**12.0148**	5	0.9727	4	**7.7594**	5	
3	M3		9.2090	3	**0.9844**	5	6.4522	4	
4	M4		7.4873	4	0.8615	3	6.0140	3	
5	M5		**6.7060**	5	0.8868	4	5.3938	4	
6	M6		**6.5434**	5	0.8933	4	**5.3026**	5	
7	M7		**6.6363**	5	0.8900	4	5.3529	4	
8	M1	GBM	9.3175	4	0.9837	4	6.4573	3	79
9	M2		12.02762	4	**0.9729**	5	8.0426	4	
10	M3		9.1978	4	0.9843	3	6.6698	1	
11	M4		**7.3980**	5	0.8651	4	**5.8522**	5	
12	M5		6.7324	4	0.8858	3	**5.3779**	5	
13	M6		6.7452	4	0.8880	3	5.3714	4	
14	M7		6,7901	4	0,8878	3	5,4031	3	
15	M1	RF	10,0294	2	0,9809	3	6,4592	2	49
16	M2		12,7501	3	0,9696	3	8,4116	3	
17	M3		9,7624	2	0,9822	2	6,6109	1	
18	M4		7,5712	3	0.8570	2	6.0059	4	
19	M5		7.0536	2	0.8758	1	5.6653	2	
20	M6		6.9726	2	0.8811	1	5.5690	2	
21	M7		6.8473	2	0.8863	2	**5.2812**	5	
22	M1	BT	10.0188	3	0.9809	2	6.4404	4	44
23	M2		13.3724	1	0.9663	1	8.4116	2	
24	M3		9.8349	1	0.9817	1	6.4557	2	
25	M4		7.6840	2	0.8538	1	6.1249	2	
26	M5		6.9397	3	0.8801	2	5.5029	3	
27	M6		6.8947	3	0.8840	2	5.5137	3	
28	M7		6.8545	3	0.8859	1	5.4667	2	
22	M1	DL	10.0754	1	0.9811	1	6.6291	1	50
23	M2		13.0063	2	0.9685	2	8.9850	1	
24	M3		**9.1884**	5	0.9843	4	**6.3861**	5	
25	M4		10.0365	1	**0.9812**	5	6.4213	1	
26	M5		10.8662	1	**0.97802**	5	7.4091	1	
27	M6		12.0255	1	**0.9730**	5	7.51526	1	
28	M7		10.1332	1	**0.9809**	5	6.7611	1	

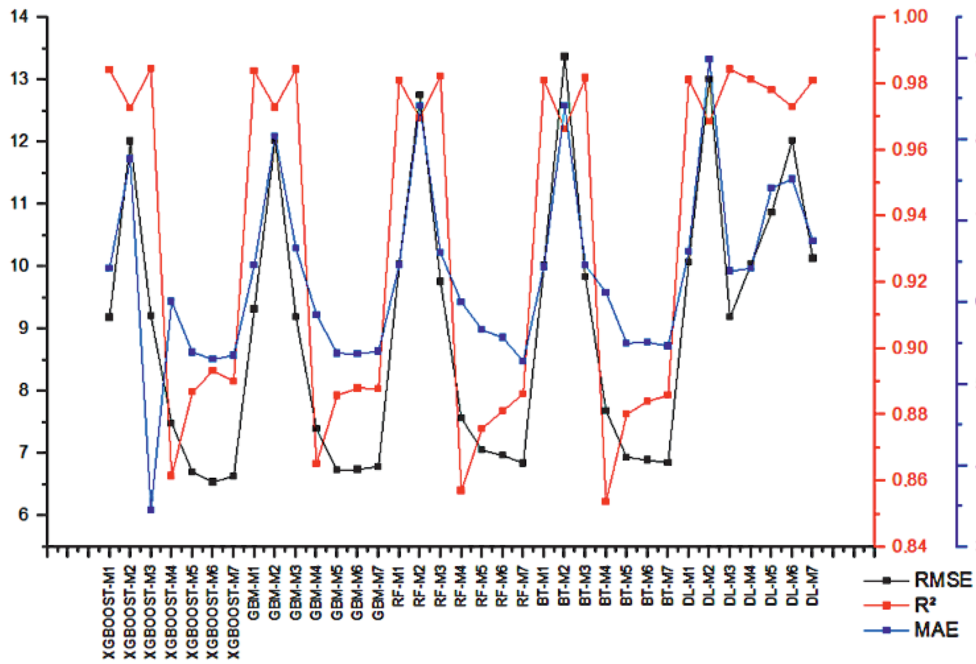

Fig. 8. RMSE, MAE, R^2, values of all models.

duration, air temperature, relative humidity and wind speed were found. In the study of [39], the effect of the variables used in PET estimation and the success of the RF model overlap significantly with the study. Fan et al. [40] evaluated the performance of LightGBM the M5Tree and RF algorithms for predicting daily reference evapotranspiration values in the humid subtropical region of China. As a result, the LightGBM algorithm showed superior results than other tree-based machine learning techniques. The results of the study support the present study. In the Agrawal et al. [8] study, the accuracy of tree-based machine learning techniques was analyzed in estimating daily evapotranspiration values in the meteorological station in Bengaluru, Karnataka, India. For this, RF, GBM and XGBoost, decision tree and Adaptive Boosting (AdaBoost) algorithms were applied. As a result of the study, the XGBoost algorithm showed the highest prediction accuracy. The study's results [8] are largely in line with the study. The outputs of the study are important in terms of meeting the region's needs in terms of agriculture, doing irrigation planning and ensuring the region's economic development.

Conclusion

The study aimed to estimate the monthly PET values in Diyarbakir, which has an important position in agricultural production. For this purpose, XGBoost, GBM, RF, Bagged Trees, CM methods were used. The results obtained are listed below;

- All models with selected input combinations have sufficient accuracy from the PET estimation to have high R^2 (>0.8) values.
- The average temperature is the data with the highest correlation with the PET values. This is likely due to the use of the Thornthwaite Method in PET calculation.
- The model, including all inputs, gave better results for the estimation of PET values. Since T_{mean}, T_{min}, T_{max}, RH_{mean}, P_{sum}, WS_{mean}, WS_{max} inputs for PET estimation offers good results, these inputs can also be applied in different regions by considering the generalization feature of numerical models.
- R^2 is shown in better positions in deep learning tools in the system. This is the case when XGBoost heights are high. Evaluation according to all indexes is seen to be the highest performer.
- When the algorithms are sorted starting from the best performance, it is seen that XGBoost>GBM>DL>RF>BT. This shows that the XGBoost algorithm is superior to other algorithms. Bagged Trees, on the other hand, has the worst performance. This also reveals the difference between bagging and boosting (Bagging is based on the resampling method. boosting is based on residual optimization). It is recommended to use all available meteorological variables in the study area as inputs in the estimation of PET values with artificial intelligence models.

This study estimates the PET values of the Diyarbakır Airport station. These results can be used to analyze and design hydrological events based on the generalization feature of numerical methods. However, more stations or basin-based evaluations will improve the study in future studies. In addition, the use and comparison of other methods for PET estimation may be preferred according to the study area and purpose.

Acknowledgment

The authors thank the General Directorate of Meteorology for providing the meteorological data used in the study.

Conflicts of Interest

The author declares no conflict of interest.

References

1. POUR-ALI BABA A., SHIRI J., KISI O., FARD A.F., KIM S., AMINI R. Estimating daily reference evapotranspiration using available and estimated climatic data by adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN). Hydrology Research, **44** (1), 131, **2013**.
2. MEHDIZADEH S., BEHMANESH J., KHALILI K. Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. Computers and electronics in agriculture **139**, 103, **2017**.
3. YASSIN M.A., ALAZBA A.A., MATTAR M.A. Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. Agricultural Water Management, **163**, 110, **2016**.
4. AFZAAL H., FAROOQUE A.A., ABBAS F., ACHARYA B., ESAU T. Computation of evapotranspiration with artificial intelligence for precision water resource management. Applied Sciences, **10** (5), 1621, **2020**.
5. CHEN Z., ZHU Z., JIANG H., SUN S. Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. Journal of Hydrology, **591**, 125286, **2020**.
6. CHIA M.Y., HUANG Y.F., KOO C.H. Reference evapotranspiration estimation using adaptive neuro-fuzzy inference system with limited meteorological data. In IOP Conference Series: Earth and Environmental Science (Vol. **612**, No. 1, p. 012017). IOP Publishing, **2020**
7. ZHOU Z., ZHAO L., LIN A., QIN W., LU Y., LI J., HE L. Exploring the potential of deep factorization machine and various gradient boosting models in modeling daily reference evapotranspiration in China. Arabian Journal of Geosciences, **13** (24), 1, **2020**.
8. AGRAWAL Y., KUMAR M., ANANTHAKRISHNAN S., KUMARAPURAM G. Evapotranspiration modeling using different tree-based ensembled machine learning algorithms. Water Resources Management, **36** (3), 1025, **2022**.
9. ASWIN S., GEETHA P., VINAYAKUMAR R. Deep learning models for the prediction of rainfall. 2018 International Conference on Communication and Signal Processing (ICCSP), (pp. 0657- 0661), **2018**.
10. CAI Y., ZHENG W., ZHANG X., ZHANGZHONG L., XUE X. Research on soil moisture prediction model based on deep learning PloS one., **14** (4), e0214508, **2019**.
11. AZMAN S.K., ISBEIH Y.J., EL MOURSI M.S., ELBASSIONI K.A. unified online deep learning prediction model for small signal and transient stability. IEEE Transactions on Power Systems, **35** (6), 4585, **2020**.
12. HUSSAIN S., MUHSION Z.F., SALAL Y.K., THEODOROU P., KURTOGLU F., HAZARIKA G.C. Prediction Model on Student Performance based on Internal Assessment using Deep Learning iJET, **14** (8) 4, **2019**.
13. YU L., JIAO C., XIN H., WANG Y., WANG K. Prediction on housing price based on deep learning. International Journal of Computer and Information Engineering, **12** (2), 90, **2018**.
14. DORAN I., KOCA,Y.K., KILIC T. The effects of probable climate changing into the agriculture in Diyarbakır. V. National Geography Symposium **16** (17), 369, **2009**.
15. THORNTHWAITE C.W. An approach toward a rational classification of climate, Geographical review, **38** (1), 55, **1948**.
16. FRIEDMAN J.H. Greedy function approximation: a gradient boosting machine. Annals of statistics 1189, **2001**.
17. HUYNH X.P., PARK S.M., KIM Y.G. Detection of driver drowsiness using 3D deep neural network and semi-supervised gradient boosting machine. In Asian Conference on Computer Vision (pp. 134-145). Springer, Cham. 10.1007 978-3-319-54526-4-10, **2016**.
18. NATEKIN A., KNOLL A. Gradient boosting machines, a tutorial. Frontiers in neurorobotics (pp. 7, 21.), **2013**.
19. PONRAJ A.S., VIGNESWARAN T. Daily evapotranspiration prediction using gradient boost regression model for irrigation planning. The Journal of Supercomputing **76** (8), 5732, **2020**.
20. CHEN T., GUESTRIN C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 785, **2016**.
21. WU L., PENG Y., FAN J., WANG Y. Machine learning models for the estimation of monthly mean daily reference evapotranspiration based on cross-station and synthetic data. Hydrology Research, **50** (6), 1730, **2019**.
22. KUS I., BOZKURT KESER S., YOLACAN E.N. Comparison of Ensemble Learning Methods in Intrusion Detection Systems. Avrupa Bilim ve Teknoloji Dergisi **31**, 725, **2021**.
23. BREIMAN L., FRIEDMAN J., STONE C.J., OLSHEN R.A. Classifcation and regression trees. Wadsworth, Belmont, **1984**.
24. SANCHEZ-MEDINA A.J., GALVAN-SANCHEZ I., FERNANDEZ-MONROY M. Applying artificial intelligence to explore sexual cyberbullying behaviour. Heliyon, **6** (1), e03218, **2020**.
25. SHAHDAD M., SABER B. Drought forecasting using new advanced ensemble-based models of reduced error pruning tree. Acta Geophysica, 1, **2022**.
26. CHEN Z., ZHU Z., JIANG H., SUN S. Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. Journal of Hydrology, **591**, 125286, **2020**.
27. WEI S., WU W., JEON G., AHMAD A., YANG X. Improving resolution of medical images with deep dense convolutional neural network: Improving resolution of medical im- ages with deep dense convolutional neural network. Concurr. Comput. **32** (1), e5084, **2020**.
28. NGUYEN H.N., KIM C.H., KIM J.M. Effective Prediction of Bearing Fault Degradation under Different Crack Sizes Using a Deep Neural Network. Applied Sciences, **8** (11), 2332, **2018**.

29. WILLMOTT C.J., MATSUURA K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate research **30** (1), 79, **2005**.
30. DE MYTTENAERE A., GOLDEN B., LE GRAND B., ROSSI, F. Mean Absolute Percentage Error for regression models. Neurocomputing **192**, 38, **2016**.
31. KINGMA D.P., BA J. Adam: A method for stochastic optimization. (arXiv preprint arXiv:1412.6980.), **2014**.
32. DUCHI J. HAZAN E., SINGER Y. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. Journal of Machine Learning Research **12**, 2121, **2011**.
33. ZEILER M.D. Adadelta: an adaptive learning rate method. arXiv preprint arXiv:1212.5701, **2012**.
34. DOZAT T. Incorporating Nesterov momentum into Adam, **2016**.
35. DAWSON C.W., WILBY R. An artificial neural network approach to rainfall-runoff modelling. Hydrological Sciences Journal 47, **1998**.
36. ZHANG H., ZHOU J., JAHED ARMAGHANI D., TAHIR M.M., PHAM B.T., HUYNH V.V. A combination of feature selection and random forest techniques to solve a problem related to blast-induced ground vibration. Applied Sciences, **10** (3), 869, **2020**.
37. NIELSEN D. Tree boosting with XGBoost-why does XGBoost win" every" machine learning competition (Master's thesis, NTNU), **2016**.
38. FERREIRA L.B., DA CUNHA F.F. New approach to estimate daily reference evapotranspiration based on hourly temperature and relative humidity using machine learning and deep learning. Agricultural Water Management, **234**, 106113, **2020**.
39. WANG S., LIAN J., PENG Y., HU B., CHEN H. Generalized reference evapotranspiration models with limited climatic data based on random forest and gene expression programming in Guangxi, China. Agricultural Water Management, **221**, 220, **2019**.
40. FAN J., MA X., WU L., ZHANG F., YU X., ZENG W. Light Gradient Boosting Machine: An efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data. Agricultural Water Management, **225**, 105758, **2019**.