*Computer Modeling in Engineering & Sciences*
DOI:10.32604/cmes.2020.011004
***Article***

Tech Science Press

# Forecasting Multi-Step Ahead Monthly Reference Evapotranspiration Using Hybrid Extreme Gradient Boosting with Grey Wolf Optimization Algorithm

**Xianghui Lu[1], Junliang Fan[2], Lifeng Wu[1,*] and Jianhua Dong[3]**

[1]School of Hydraulic and Ecological Engineering, Nanchang Institute of Technology, Nanchang, 330099, China
[2]Key Laboratory of Agricultural Soil and Water Engineering in Arid and Semiarid Areas of Ministry of Education, Northwest A&F University, Yangling, 712100, China
[3]Faculty of Agriculture and Food, Kunming University of Science and Technology, Kunming, 650500, China
*Corresponding Author: Lifeng Wu. Email: china.sw@163.com
Received: 14 April 2020; Accepted: 24 July 2020

**Abstract:** It is important for regional water resources management to know the agricultural water consumption information several months in advance. Forecasting reference evapotranspiration ($ET_0$) in the next few months is important for irrigation and reservoir management. Studies on forecasting of multiple-month ahead $ET_0$ using machine learning models have not been reported yet. Besides, machine learning models such as the XGBoost model has multiple parameters that need to be tuned, and traditional methods can get stuck in a regional optimal solution and fail to obtain a global optimal solution. This study investigated the performance of the hybrid extreme gradient boosting (XGBoost) model coupled with the Grey Wolf Optimizer (GWO) algorithm for forecasting multi-step ahead $ET_0$ (1–3 months ahead), compared with three conventional machine learning models, i.e., standalone XGBoost, multi-layer perceptron (MLP) and M5 model tree (M5) models in the subtropical zone of China. The results showed that the GWO-XGB model generally performed better than the other three machine learning models in forecasting 1–3 months ahead $ET_0$, followed by the XGB, M5 and MLP models with very small differences among the three models. The GWO-XGB model performed best in autumn, while the MLP model performed slightly better than the other three models in summer. It is thus suggested to apply the MLP model for $ET_0$ forecasting in summer but use the GWO-XGB model in other seasons.

**Keywords:** Reference evapotranspiration; extreme gradient boosting; Grey Wolf Optimizer; multi-layer perceptron; M5 model tree

## 1 Introduction

Reference evapotranspiration ($ET_0$) as well as the soil and plant canopy characteristics are usually the main input parameters for soil water balance models. Accurate estimation of $ET_0$ plays an important role in water resource management and crop irrigation requirement determination, supporting the irrigation scheduling, regional water management decisions, drought early

warning assessments, hydrological models and climate change impact models. The international standard method for calculating $ET_0$ is the Penman–Monteith equation in irrigation and drainage No. 56 published by the Food and Agriculture Organization. This method (hereinafter called FAO 56 PM) is considered to be one of the best methods for estimating daily and monthly $ET_0$ under different climatic conditions [1]. However, the FAO 56 PM equation requires many meteorological variables and high data quality, which includes meteorological data such as air temperature, wind speed, solar radiation and relative humidity [2]. Since weather stations are usually tens to hundreds of kilometers apart, it is not easy to obtain accurate meteorological data in desolate and rural areas, given that meteorological elements vary largely from place to place.

Due to the strong capability of searching for nonlinear relationship unknown to human beings from big data [3], soft computing models have recently started to attract people's interest in hydrological and agricultural fields, e.g., modeling runoff [4], precipitation [5], drought [6], solar radiation [7–9], evaporation [10–12] and crop yield [13], also including $ET_0$ and actual ET prediction/forecasting [14,15]. Especially, modeling $ET_0$ with limited historical meteorological data is the most popular approach. Artificial Neural Network (ANN) has been the most popular soft computing algorithm applied for $ET_0$ simulation. The traditional multilayer perceptron neural network (MLP) has the advantages of simple model structure, fast training speed and strong applicability [16]. The latest deep learning strategies are also based on neural network architecture. Khoob [17] evaluated the ANN model for estimating $ET_0$ in South Iran, and the evaporation and temperature of type A evaporator were used as input parameters. The results showed that the ANN model was much better than the traditional method of multiplying evaporation coefficient by evaporation amount of evaporation. Based on the ANN model, Martí et al. [18] explored how to estimate $ET_0$ in unknown regions when lack of local meteorological data. The relations of $ET_0$ and geographic information were also studied by using geographic information as inputs and ANN model as the tool [19]. Huo et al. [20] used the ANN model for prediction of $ET_0$ in the Shiyanghe Basin of North China. They also found ANN models overestimated $ET_0$ from August to December and underestimated in the other months. Antonopoulos et al. [21] assessed the performances of ANN and Priestley–Taylor, Makkink, Hargreaves and mass transfer models for estimating $ET_0$ in northern Greece. They found the ANN model performed best among the models.

Another type of popular model is the fuzzy system which received extensive attention. Cobaner [22] investigated the performance of two type of neuro-fuzzy inference systems, e.g., grid partition based fuzzy inference system (G-ANFIS) and subtractive clustering based fuzzy inference system (SC-ANFIS) for predicting $ET_0$ in Los Angeles, USA. They confirmed that SC-ANFIS had higher accuracy and less computational time than the other models. Kisi et al. [23] also compared G-ANFIS and SC-ANFIS model at Adana Station, Turkey. They found the G-ANFIS model was less affected by missing data and training data length than the SC-ANFIS model. Shamshirband et al. [24] developed a hybrid model based on the ANFIS model and cuckoo search algorithm (CSA) at twelve meteorological stations in Serbia. They found the ANFIS-CSA model yielded higher accuracy. Keshtegar et al. [25] developed a subset model based on ANFIS for estimation of $ET_0$ in three cities of Turkey. The results showed that the new model was superior to the ANN, single ANFIS and M5 models. The advantages of ANFIS have also been reported in North India [26]. Furthermore, the performances of ANFIS and ANN models can be improved by the wavelet transform technology [27]. A recent study in Burkina Faso using the firefly algorithm (FFA) to optimize ANFIS was also reported [28]. The GEP model is another soft computing method with explicit expression and it has better precision than

traditional empirical models. Karimi et al. [29] evaluated GEP and SVM as well as empirical models for predicting $ET_0$ in a humid region of Korea. They found that GEP was superior to SVM when cross-station application of the developed models was applied. Kiafar et al. [30] compared the GEP models and different types of empirical models for prediction of $ET_0$ in distant humid and arid regions. They found the new model outperformed the other three models, i.e., ANN, M5 and ANFIS models. The GEP model has also been applied in other areas, e.g., island of Iran [31] and Egypt [32], and appropriate results have been achieved.

The kernel-based model is another powerful soft computing method, such as support vector machine (SVM), least squares support vector machines (LSSVM), extreme learning machine (ELM) and kernel-based nonlinear extension of Arps decline model (KNEA). Abdullah et al. [33] firstly introduced ELM as a new method for $ET_0$ prediction. The highest $R^2$ yielded by the ELM model was 0.991 in Iraq and the corresponding value was 0.985 by the ANN model. For the same purpose, Gocic et al. [34] evaluated the performance of ELM for predicting $ET_0$, as well as three empirical models in Nis and Belgrade stations of Serbia. Their results confirmed that ELM was a stable and reliable model. Feng et al. [35] compared three soft computing approaches (ELM, GANN and WNN) for predicting $ET_0$ in Sichuan of China. The result showed that ELM model outperformed the other models under different input combinations. Feng et al. [36] also compared the ELM with GRNN model with only temperature data as input in the same region. They found the ELM model performed better than GRNN in the local application scenario and the opposite result was obtained in the cross-station scenario, both of which were superior to the empirical Hargreaves model. Wen et al. [37] evaluated the performance of SVM in extreme arid regions of China and compared it with ANN and three empirical models. It was concluded that SVM was superior to the other models. Since the parameters of ELM and SVM models are not always easy to obtain optimal values, the evolution algorithms have been used to optimize the parameters, and the results of the optimized models were better than the original models [38,39].

Decision tree-based model is another type of soft computing method, which is supported by mathematical theory. These methods are based on the idea of binary tree, which incorporate the theories of bagging or boosting and has been widely used in many fields. Pal et al. [40] firstly used M5 model for $ET_0$ estimation at Davis station, USA. Rahimikhoob [41] assessed the accuracy of ANN and M5 models for $ET_0$ estimation with temperature data, air humidity and extraterrestrial radiation as inputs. The results showed that ANN had slighter better accuracy than M5 model and they both performed well. On this basis, the accuracy of the former two models using the temperature data from satellite images as inputs was further evaluated, and M5 was proved to be superior to ANN model [42]. In general, the accuracy of mass transfer methods was low. In order to solve this problem, Shiri [43] used the hybrid of wavelet transfer with RF model to improve the accuracy. Wang et al. [44] used data from 24 stations in a karst region of Southwest China to develop generalized model based on RF and GEP models. They found RF model was superior to GEP model under different input combinations. Similar work has also been done in Central Florida, where four soft computing models, i.e., M5, Bagging, RF and SVM were used [45]. Fan et al. [14] evaluated the performances of four tree-based model (M5, GBDT, XGBoost and RF) as well as ELM and SVM in different climatic regions of China. The results showed that the tree-based models had less computing time and suitable accuracy [46].

However, the above studies have mainly focused on $ET_0$ estimation based on the known data in the past, and these methods are difficult to be used for future water resources planning. Many scholars are interested in how to estimate future $ET_0$, so as to plan and allocate water resources in advance to deal with potential water crisis. In recent years, outputs of numerical

weather prediction outputs have been used in estimating $ET_0$ in Australia [47]. Traore et al. [48] compared four kinds of ANN models for forecasting $ET_0$ in large irrigation areas of Texas and MLP achieved the highest accuracy. Similar work has been conducted in six cities of China, where the daily outputs of public weather forecasts were empirically converted into PM models, which yielded RMSE values of 0.65–1.08 mm $d^{-1}$ and MAE values of 0.63–0.84 mm $d^{-1}$ in the future 7 days. Zhao et al. [49] attempted to use the GCMs output to forecast $ET_0$ for the next 1–3 months in Australia, but the accuracy of this method needed to be improved. I drawback of this approach is that it requires a lot of meteorological knowledge, because the output of numerical meteorological forecast usually includes dozens of variables. How to select effective meteorological factors is a complicated and arduous work. On the other hand, using time-series data to forecast possible future results is also an interesting way to solve this problem. Karbasi [50] coupled GPR with wavelet transfer technology to forecast 1–30 days ahead $ET_0$ in Zanjan, Iran. The results showed that new approach was better than the single GPR model. Mehdizadeh [51] compared the accuracies of history weather data-based and lagged $ET_0$ data-based models based on MARS and GEP models in Iran. The found the lagged $ET_0$ data-based models outperformed the former one. Mohammadi et al. [52] developed a couple model based on SVM coupled with the whale optimization algorithm (WOA) for $ET_0$ prediction at three stations of Iran and found that it was better than other models.

Floods and droughts have brought severe challenges to the sustainable economic development in southern China, often causing losses to the personal safety and property of millions of people. In terms of time, flood occurs mainly in spring and summer, while drought occurs mainly in late summer and autumn. For example, in June 2019, rainfall-triggered floods have affected a total of 201.4 million people in the nine cities divided into districts in Jiangxi, with 231,000 people in need of emergency relocation and living assistance. Crops have been affected by 137.2 thousand $ha^2$, with a total area of 15.6 million $ha^2$ of crop died. Meanwhile, thousands of houses collapsed. However, the droughts occurred in autumn and winter, which caused more than 5 million people to be affected. Due to the drought, 845,000 people needed to be saved due to the lack of drinking water. Crops were affected by 472.5 million $ha^2$, and 85.6 million $ha^2$ were lost, resulting in a direct economic loss of 0.7 billion dollars.

As the best knowledge of the authors, studies on forecasting $ET_0$ for the next season have not been reported yet. The XGBoost model has multiple parameters that need to be tuned, and traditional methods can get stuck in a regional optimal solution and fail to obtain a global optimal solution. However, reports on optimizing XGBoost model parameters are also very limited, especially those using gray wolf optimization (GWO) [53], which is a powerful optimization algorithm and has been used to optimize MLP [54], ANFIS [55], SVM [56], ELM [57] and Grey model [58].Therefore, the objectives of this study are: (1) To evaluate whether it is feasible to use historical time-series $ET_0$ data for forecasting $ET_0$ in the next 1–3 months in a humid region of China based on four soft computing models, i.e., ANN, M5, XGB and GWO-XGB; (2) To verify whether the GWO algorithm can improve the accuracy of standalone XGBoost model for forecasting $ET_0$.

## 2 Case Study

In this study, monthly meteorological data were collected from nine weather stations in South China operated by the China Meteorological Administration (CMA) (Fig. 1). This region is characterized by a subtropical monsoon climate [8]. Meteorological data including air temperature, relative humidity, sunshine hours and wind speed were used to calculate $ET_0$ using the FAO No.

56 Penman–Monteith formula. Detailed information on meteorological data at the nine weather stations can be found in Tab. 1. It is not difficult to see from Tab. 1 that the average $ET_0$ at Station 58238 fluctuated greatly during training, testing and validation periods. Especially, the validation period was significantly lower than the other two periods, which may affect the prediction accuracy of the model. In addition, $ET_0$ in South China has undergone drastic changes from 1966 to 2015, mainly showing that $ET_0$ was significantly lower than that in the 1960s and 1970s before and after 1990, and rebounded after 2000 [59], which may also have some impact on $ET_0$ forecasting.

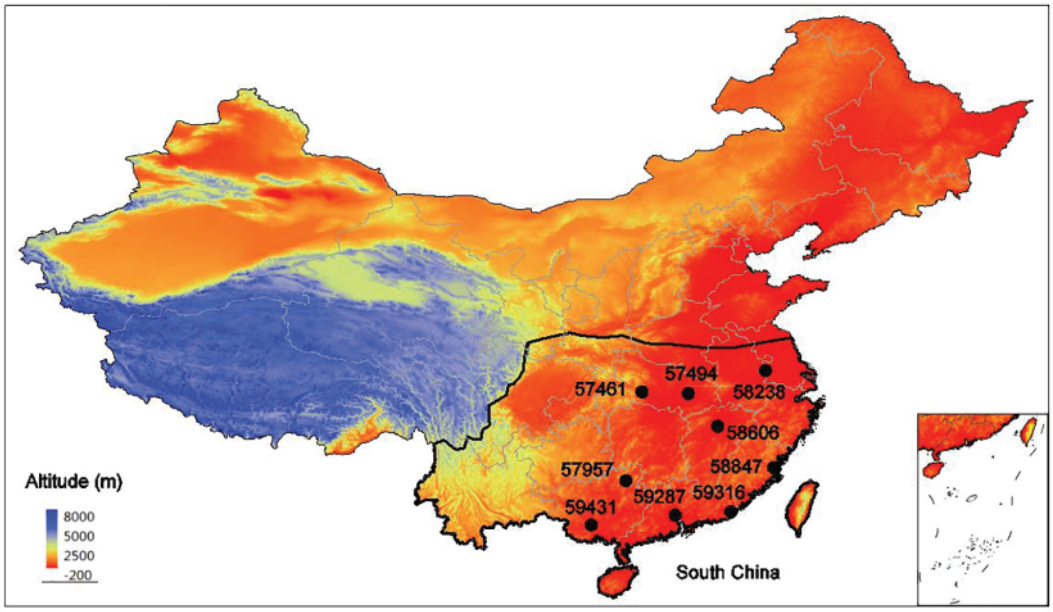


**Figure 1:** Locations of the nine weather stations in South China

**Table 1:** Statistical results of $ET_0$ (mm $d^{-1}$) during various periods at the nine stations selected for this study

| ID | Latitude °N | Langitude °E | Altitude m | $R_s$ MJ $m^{-2}$ $d^{-1}$ | $T_{max}$ °C | $T_{min}$ °C | RH % | $U_2$ $ms^{-1}$ | $ET_0$ mm $d^{-1}$ |
|---|---|---|---|---|---|---|---|---|---|
| 57461 | 30.42 | 111.13 | 134.3 | 10.77 | 21.51 | 13.55 | 75.04 | 0.98 | 2.28 |
| 57494 | 30.37 | 114.06 | 27 | 12.03 | 21.36 | 13.23 | 76.68 | 1.38 | 2.45 |
| 57957 | 25.19 | 110.18 | 166.2 | 11.26 | 23.32 | 16.07 | 74.87 | 1.78 | 2.67 |
| 58238 | 31.93 | 118.49 | 12.5 | 12.47 | 20.49 | 11.88 | 74.93 | 1.86 | 2.50 |
| 58606 | 28.37 | 115.56 | 45.7 | 0.00 | 21.84 | 14.80 | 76.16 | 2.66 | 2.64 |
| 58847 | 26.05 | 119.17 | 85.4 | 12.09 | 24.61 | 17.00 | 75.16 | 1.92 | 2.89 |
| 59287 | 23.09 | 113.2 | 4.2 | 11.60 | 26.53 | 18.99 | 76.72 | 1.32 | 2.65 |
| 59316 | 23.24 | 116.41 | 7.3 | 13.69 | 25.53 | 18.98 | 79.26 | 1.81 | 2.95 |
| 59431 | 22.46 | 108.18 | 73.7 | 12.48 | 26.32 | 18.55 | 79.26 | 1.07 | 2.72 |

## 3 Machine Learning Models for Monthly Reference Evapotranspiration Forecasting

### *3.1 Extreme Gradient Boosting (XGBoost)*

XGBoost, proposed by Chen et al. [60], is a scalable soft computing algorithm for CART type tree boosting. XGBoost model consists of multiple decision trees, each of which pays attention to the residuals of the previous tree and USES the gradient algorithm to find a new decision tree establishment method to reduce the residuals of model training. As seen in Fig. 2, previous predictors are redeveloped to decrease the residuals during each iteration. The algorithm can independently determine the types of loss functions used for model evaluation. To reduce the risk of overfitting, different types of regular terms, e.g., L1 and L2 can be selected. The mean score of each tree is used as the predictive value for classification or regression.

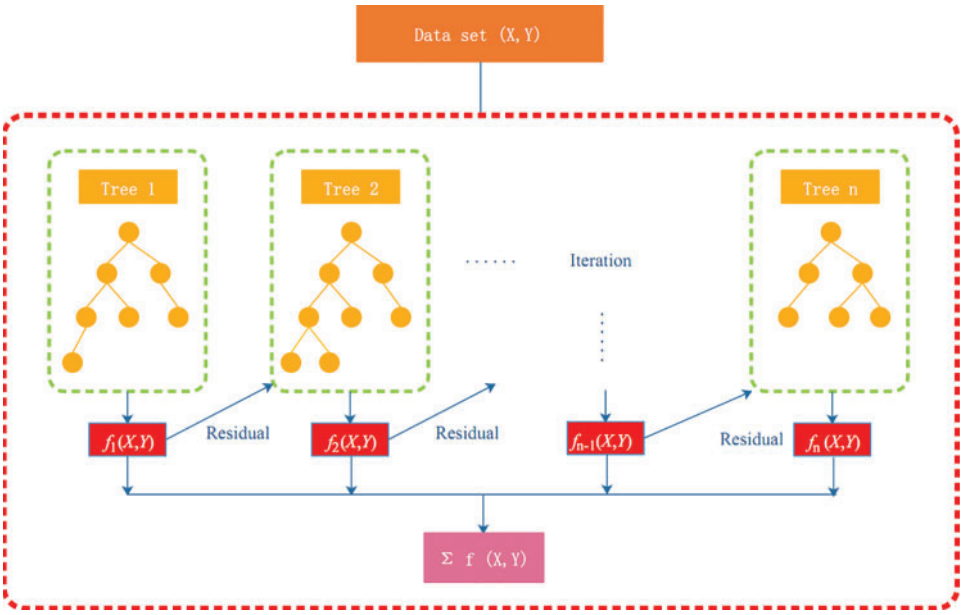


**Figure 2:** The structure of XGB model

For the $m$-th decision tree, its calculation formula can be expressed as:

$$\hat{y}_i = \sum_{i=1}^{m} f_m(x_i), \quad f_m \in W \tag{1}$$

where $m$ is the number of CART trees; $f_m$ is a function in the functional space W, and W is the space of all CART trees.

The objective function of the model at the t-th iteration is written as follow:

$$\Theta(t) = \Phi(t) + \Omega(t) \tag{2}$$

$$\Theta(t) = \sum_{i=1}^{n} \Phi\left(y_i, \hat{y}_i\right) + \sum_{k=1}^{t} \Omega(f_m) \tag{3}$$

where $n$ is the $n$-th prediction, and $\hat{y}_i^{(t)}$ can be given as:

$$\hat{y}_i^{(t)} = \sum_{m=1}^{t} f_m(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i) \tag{4}$$

The regularization term $\Omega(f_k)$ for a CART tree is added by Chen et al. [60] as follows:

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{5}$$

where $\gamma$ is the complexity of each leaf; $T$ is the number of leaves in a tree; $\lambda$ is a parameter to scale the penalty and $w$ is the vector of scores on the leaves. Then, the first-order as well as the second-order Taylor expansions are taken to the loss function in XGBoost. Assuming that the loss function is MSE, then the objective function can be written as:

$$\Theta^{(t)} \approx \sum_{i=1}^{n} \left[ g_i w_{q(x_i)} + \frac{1}{2}\left( h_i w_{q(x_i)}^2 \right) \right] + \gamma T + \frac{1}{2}\sum_{j=1}^{T} w_j^2 \tag{6}$$

where $q(\cdot)$ is a function can assign data points to corresponding leaves; $g_i$ and $h_i$ are the first and second derivative of MSE loss function respectively. In the formula (6), the loss function is determined by the sum of loss values for each data sample. Since each data sample corresponds to only one leaf node, the loss function can also be expressed as the sum of loss values for each leaf node, which is:

$$\Theta^{(t)} \approx \sum_{j=1}^{T} \left[ \left( \sum_{i \in I_j} g_i \right) w_j + \frac{1}{2}\left( \sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T \tag{7}$$

where $I_j$ represents all data samples in the leaf node $j$. Obviously, the objective function is equivalent to finding the minimum value of the quadratic function. In short, the model performance change caused by a node splitting in the CART tree can be evaluated according to the change of objective function value. That is to say, if the model performance of decision tree after node splitting is improved, then it is adopted; otherwise, the splitting will stop.

### *3.2 Gray Wolf Optimization*

The gray wolf optimization (GWO) algorithm is a meta-heuristic optimization algorithm proposed by Mirjalili et al. [53], which reflects the social system and group of gray wolf families in nature body hunting behavior. The gray wolf community has a very strict social hierarchy. In the system, wolves are usually divided into four levels: $\alpha$, $\beta$, $\delta$ and $\omega$, in which $\alpha$ wolf is the first level, mainly responsible for overall decision-making; $\beta$ is the second level, assisting $\alpha$ wolf to make decisions; $\delta$ wolf is the third level, and needs to obey the decisions of $\alpha$ and $\beta$. The lowest ranking wolves $\omega$ in the pack are the d wolves and they have to obey the higher ranking wolves. According to the social hierarchy and hunting process of wolves, the mathematical model can be defined and the optimal solution can be found.

The optimal solution is defined as $\alpha$ and the second and third optimal solution are defined as $\beta$ and $\delta$, respectively. The rest candidate solutions are defined as $\omega$. The distance between the individual and the prey in the Wolf pack is defined as $D$.

$$D = |CX_P(t) - X| \tag{8}$$

$$C = 2r_1 \tag{9}$$

where $D$ is the current number of iterations, $C$ is step length coefficient, $X_P(t)$ is prey location, $X$is the location of a grey Wolf, $r_1$ is a random number ranged in (0,1).

To shorten the distance between themselves and their prey, individuals in a pack are constantly updated according to the following formula:

$$X(t-1) = X_P(t) - A \cdot D \tag{10}$$

$$A = 2ar_2 - a \tag{11}$$

where $A$ is the convergence influence factor, which increases with the number of iterations from 2 to 0 by linear decrease; $r_2$ is a random number ranged in (0,1).

Because $\alpha$, $\beta$ and $\delta$ have a high level in the wolves, so that they can carry more prey location information, lead the pack gradually close to the hunting. They will now have three optimal solutions to save and ignore other solution, and according to the three optimal solutions on the location information to update the wolves, gradually finds the global optimal solution, the process of updating is defined as follows:

$$D_\alpha = |C_1 X_\alpha(t) - X| \tag{12}$$

$$D_\beta = \left|C_2 X_\beta(t) - X\right| \tag{13}$$

$$D_\delta = |C_3 X_\delta(t) - X| \tag{14}$$

Based on the calculated distance $\alpha$, $\beta$ and $\delta$ carry out itself by the following formula which correction of position as:

$$X_1 = X_\alpha - A_1 \cdot D_\alpha \tag{15}$$

$$X_2 = X_\beta - A_2 \cdot D_\beta \tag{16}$$

$$X_3 = X_\delta - A_3 \cdot D_\delta \tag{17}$$

The remaining individuals in the pack will then be based on the joint decision of $\alpha$, $\beta$ and $\delta$. The next step to move the position shown in:

$$X(t-1) = \frac{X_1 + X_2 + X_3}{3} \tag{18}$$

To sum up, Grey Wolf Optimizer algorithm continuously updates the location search solution space during the optimization process, and finally finds the optimal solution (Fig. 3).

### *3.3 Multi-Layer Perceptron (MLP)*

Neural Networks architectures are the general term for a series of machine learning methods based on the activation and transmission of information by simulated neurons. Most neural networks have an input layer, certain hidden layers and an output layer. Among them, the most popular is the single-hidden-layer neural networks, and one layer to another layer using activation

function and summation function connection. In this study, we used a multilayer perceptron (MPL) neural networks model, and the activation function used the sigmoid function [61].

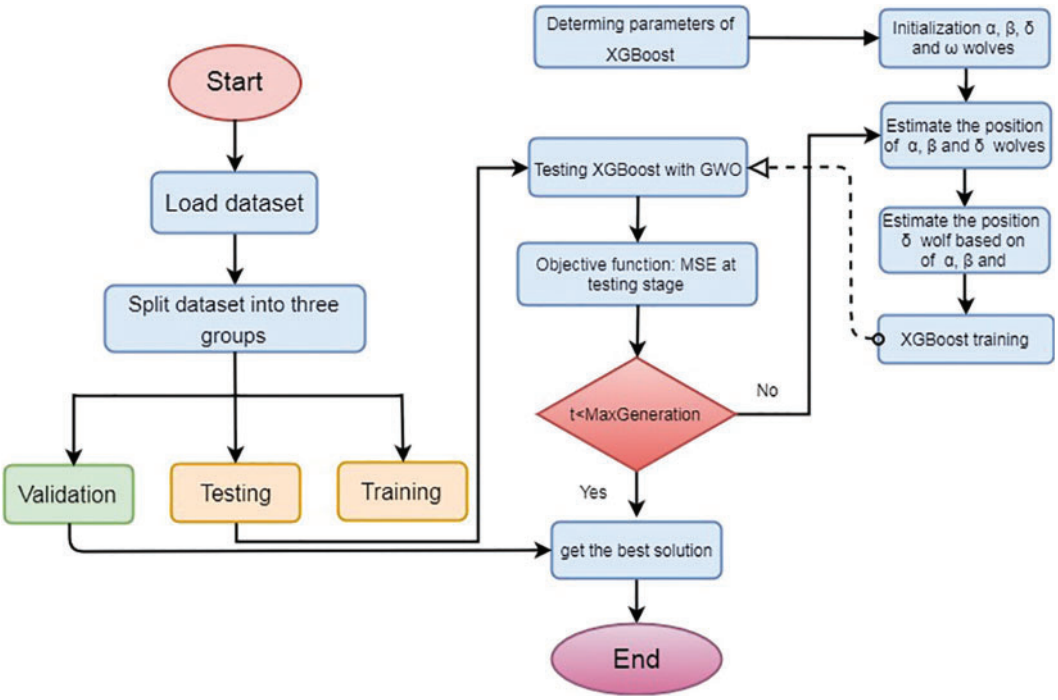


**Figure 3:** The flow chart of the hybrid GWO-XGB model

### *3.4 M5 Model Tree (M5)*

M5 model tree is a kind of decision tree that adopts linear regression function at leaf node. This technique is very successful in predicting continuous values. It can be implemented by adopting a standard method of transforming a classification problem into a functional optimization problem. The model tree represents a piecewise linear function. Like a typical regression equation, it predicts the value of a variable (called a class) by a set of independent variables called attributes. The training data in table form can be directly used to construct the decision tree [62]. For a given data set, a typical linear regression algorithm can only give a single regression equation, but the model tree divides the sample space into rectangular areas with parallel edges, and determines a corresponding regression model for each partition.

The structure of the model tree is generated recursively, starting with the entire training sample set. At each level of the model tree, the most discriminative attribute is selected as the root node of the subtree, and the samples arriving at this node are divided into several subsets according to the value of node attributes. Model tree algorithm is a global model combined with a series of piecewise linear models. It differs from linear regression in that the input space is divided automatically by the algorithm. It has the advantages of high efficiency, good robustness, can be effective learning, can handle the input attributes up to several hundred dimensions.

### 3.5 Model Setup and Parameter Optimization

In this study, the data were first divided into three groups, the first (1966–1995) for training the models, the second (1996–2005) for testing the models, and the third (2006–2015) for validation and forecasting. The input meteorological combination consisted of $T_{\max}$, $T_{\min}$ and $ET_0$ during the previous period. Different from finding extremum of curve, error of $ET_0$ simulation cannot approach 0 in most cases, which will lead to overfitting problem. Normally, with other parameters unchanged, the accuracy of XGBoost model will gradually decrease during the training period as the number of trees increases, while that of testing will first decrease and then increase. Therefore, we chose the minimum MSE error during the test period as the objective function to establish the model, and considered that the corresponding parameter value was the most suitable parameter value at this time. Then, we used the data in the third part (2006–2015) to evaluate the prediction ability of the model. We used the R3.4.2 platform with packages named RSNNS, Rpart, and xgboost for model implementation. Three parameters were optimized, including the number of trees (nrounds), [50, 1500]; the ratio of sub datasets to all data for training the model (subset), [0.5, 1]; and the minimum sum of instance weight needed in a child (min_child_weight), (1, 15). The GWO algorithm had only two parameters, namely the number of population and the number of iterations, which were set as 50 and 500, respectively.

### 3.6 Performance Criteria

The performance of the four soft computing model's efficiency, including accuracy and agreement, was evaluated using statistical criteria, such as RMSE, NSE and MAE. The evaluation criteria of RMSE and MAE are common mean error indicators that indicate how close data points are to a best fit line. The main difference between RMSE and MAE is that the point with large error has more weight than the point with small error, while MAE pays more attention to the average error performance. According to Nash et al. [63], "the NSE is defined as the sum of the absolute squared differences of the observed and estimated data normalized by the variance minus one." As determined by Krause [64], the NSE value ranges 1 to $-\infty$. When NSE is close to 1, it means that the model performs well.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (PMET_0 - ET_{0s})^2}{n}} \tag{19}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n} (ET_{0PM} - ET_{0s})^2}{\sum_{i=1}^{n} \left(ET_{0PM} - \overline{ET_{0PM}}\right)^2} \tag{20}$$

$$MAE = \frac{\sum_{i=1}^{n} |PMET_0 - ET_{0s}|}{n} \tag{21}$$

where $ET_{0s}$, $ET_{0PM}$ are model estimated $ET_0$ and $ET_0$ calculated by the FAO56 PM equation, respectively. $\overline{ET_{0PM}}$ is the average of $ET_0$ calculated by the FAO56 PM equation.

## 4 Results and Discussion

The forecasting ability of the hybrid GWO-XGB model for reference evapotranspiration at different time steps (1, 2 and 3-month) was evaluated. The GWO-XGB model was compared by three models, i.e., M5, MLP and XGB models.

**Table 2:** Statistical indicators of four soft computing models for 1-month ahead reference evapotranspiration forecasting at nine meteorological stations

| Station | Model | Training | | | Testing | | | Validation | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | RMSE mm $d^{-1}$ | NSE | MAE mm $d^{-1}$ | RMSE mm $d^{-1}$ | NSE | MAE mm $d^{-1}$ | RMSE mm $d^{-1}$ | NSE | MAE mm $d^{-1}$ |
| 57461 | M5 | 0.372 | 0.890 | 0.274 | 0.378 | 0.878 | 0.296 | 0.385 | 0.868 | 0.291 |
| | MLP | 0.418 | 0.862 | 0.314 | 0.366 | 0.886 | 0.287 | 0.371 | **0.878** | 0.286 |
| | XGB | **0.349** | **0.904** | **0.258** | 0.371 | 0.883 | 0.295 | 0.371 | **0.878** | **0.285** |
| | GWO-XGB | 0.368 | 0.893 | 0.274 | **0.350** | **0.896** | **0.278** | **0.370** | **0.878** | 0.286 |
| 57494 | M5 | 0.446 | 0.879 | 0.314 | 0.424 | 0.871 | 0.315 | **0.430** | **0.885** | **0.311** |
| | MLP | 0.480 | 0.865 | 0.342 | 0.386 | 0.892 | 0.296 | 0.431 | **0.885** | 0.317 |
| | XGB | **0.380** | **0.915** | **0.271** | 0.419 | 0.874 | 0.299 | 0.437 | 0.882 | 0.319 |
| | GWO-XGB | 0.450 | 0.882 | 0.324 | **0.373** | **0.900** | **0.287** | 0.435 | 0.883 | 0.319 |
| 57957 | M5 | 0.392 | 0.849 | 0.300 | 0.486 | 0.771 | 0.368 | 0.430 | 0.833 | 0.335 |
| | MLP | 0.448 | 0.808 | 0.342 | 0.470 | 0.786 | 0.355 | 0.431 | 0.832 | 0.344 |
| | XGB | **0.357** | **0.878** | **0.276** | 0.433 | 0.819 | 0.321 | 0.405 | 0.852 | 0.323 |
| | GWO-XGB | 0.358 | 0.877 | 0.277 | **0.411** | **0.836** | **0.310** | **0.401** | **0.855** | **0.317** |
| 58238 | M5 | 0.392 | 0.892 | 0.273 | 0.411 | 0.876 | 0.304 | 0.546 | 0.803 | 0.416 |
| | MLP | 0.437 | 0.868 | 0.307 | **0.372** | **0.899** | **0.278** | **0.507** | **0.830** | **0.374** |
| | XGB | **0.350** | **0.916** | **0.247** | 0.441 | 0.857 | 0.324 | 0.566 | 0.788 | 0.430 |
| | GWO-XGB | 0.384 | 0.898 | 0.271 | 0.409 | 0.877 | 0.308 | 0.539 | 0.808 | 0.412 |
| 58606 | M5 | 0.415 | 0.889 | 0.298 | 0.513 | 0.806 | 0.371 | 0.475 | 0.856 | 0.359 |
| | MLP | 0.521 | 0.828 | 0.388 | 0.536 | 0.789 | 0.385 | 0.525 | 0.825 | 0.381 |
| | XGB | **0.376** | **0.911** | **0.284** | 0.476 | 0.833 | 0.348 | 0.438 | 0.878 | 0.340 |
| | GWO-XGB | 0.384 | 0.907 | **0.284** | **0.465** | **0.841** | **0.339** | **0.445** | **0.874** | **0.345** |
| 58847 | M5 | 0.417 | 0.865 | 0.319 | 0.474 | 0.825 | 0.356 | 0.473 | 0.832 | 0.368 |
| | MLP | 0.540 | 0.773 | 0.405 | 0.523 | 0.786 | 0.377 | 0.542 | 0.779 | 0.410 |
| | XGB | **0.382** | **0.887** | **0.299** | 0.516 | 0.792 | 0.386 | 0.472 | 0.832 | 0.365 |
| | GWO-XGB | 0.389 | 0.883 | 0.303 | **0.467** | **0.830** | **0.347** | **0.441** | **0.854** | **0.349** |
| 59287 | M5 | 0.361 | 0.787 | 0.284 | 0.395 | 0.724 | 0.298 | 0.439 | 0.724 | 0.349 |
| | MLP | 0.453 | 0.673 | 0.373 | 0.407 | 0.708 | 0.320 | 0.429 | 0.736 | 0.341 |
| | XGB | **0.330** | **0.827** | **0.265** | 0.386 | 0.738 | 0.289 | 0.435 | 0.729 | 0.349 |
| | GWO-XGB | 0.348 | 0.807 | 0.282 | **0.360** | **0.771** | **0.280** | **0.423** | **0.743** | **0.338** |
| 59316 | M5 | 0.348 | 0.840 | 0.266 | 0.489 | 0.716 | 0.387 | 0.496 | 0.744 | 0.386 |
| | MLP | 0.413 | 0.776 | 0.316 | **0.427** | **0.782** | **0.314** | 0.453 | 0.786 | 0.349 |
| | XGB | **0.316** | **0.869** | **0.246** | 0.463 | 0.745 | 0.345 | 0.442 | 0.797 | 0.332 |
| | GWO-XGB | 0.330 | 0.857 | 0.260 | 0.437 | 0.772 | 0.329 | **0.422** | **0.814** | **0.321** |
| 59431 | M5 | 0.354 | 0.882 | 0.272 | 0.364 | 0.852 | 0.281 | 0.398 | 0.849 | 0.304 |
| | MLP | 0.392 | 0.855 | 0.313 | 0.375 | 0.843 | 0.296 | **0.359** | **0.877** | **0.277** |
| | XGB | **0.328** | **0.899** | **0.252** | 0.344 | 0.868 | 0.268 | 0.374 | 0.866 | 0.292 |
| | GWO-XGB | 0.346 | 0.887 | 0.260 | **0.339** | **0.872** | **0.262** | 0.388 | 0.856 | 0.310 |

### *4.1 1-Month Ahead Forecasting*

Tab. 2 shows the statistical indicators of the GWO-XGB hybrid model as well as the three single soft computing models (M5, MLP and XGB) for forecasting 1-month ahead reference evapotranspiration. The results during training and testing periods were compared to observe whether the model was over-fitted or under-fitted. The validation period was used to evaluate

the prediction ability of the model for independent datasets. Overall, RMSE, NSE and MAE values of the four soft computing models ranged 0.316–0.540 mm $d^{-1}$, 0.673–0.916, 0.246–0.405 mm $d^{-1}$ in the training stage, and corresponding values ranged 0.339–0.536 mm $d^{-1}$, 0.708–0.900, 0.262–0.387 mm $d^{-1}$ in the testing stage, respectively. It can be concluded that there were no significant over-fitting or under-fitting. In addition, for the forecasting of 1-month ahead $ET_0$, GWO-XGB model provided the best result with RMSE = 0.350 mm $d^{-1}$, NSE = 0.896 and MAE = 0.278 mm $d^{-1}$ in the testing stage at Station 57461. However, MLP, XGB and GWO-XGB models had similar performances with average RMSE = 0.370 mm $d^{-1}$, NSE = 0.878 and MAE = 0.286 mm $d^{-1}$. On the other hand, the M5 model was slightly worse than the other three models with RMSE = 0.385 mm $d^{-1}$, NSE = 0.868 and MAE = 0.291 mm $d^{-1}$. At Station 57494, the GWO-XGB model also attained the best result among the four models in the testing stage (RMSE = 0.373 mm $d^{-1}$, NSE = 0.900 and MAE = 0.287 mm $d^{-1}$), while the M5 model was slightly better than the other models in the validation stage with RMSE = 0.430 mm $d^{-1}$, NSE = 0.885 and MAE = 0.311 mm $d^{-1}$. At Stations 57957, 58606, 58847 and 59287, GWO-XGB model was superior to the other models, with RMSE ranging 0.360–0.467 and 0.401–0.445 mm $d^{-1}$, NSE ranging 0.771–0.840 and 0.743–0.874, MAE ranging 0.280–0.347 and 0.317–0.349 mm $d^{-1}$ in the testing and validation stages, respectively. In addition, the GWO-XGB model yielded the best results at Station 59316 in the validation stage with RMSE = 0.422 mm $d^{-1}$, NSE = 0.814 and MAE = 0.321 mm $d^{-1}$. However, the MLP model had the best results at Stations 58238 and 59431.

The comparison of forecasted $ET_0$ value and PM $ET_0$ values by different soft computing models in each month and the error fluctuation of each month are presented in Figs. 4–7. It can be seen that all the four models well described the fluctuation of $ET_0$ in different seasons. The errors of each model mainly came from the underestimation of peak values. It should be noted that this underestimation was common in all four models, with an error of up to 1.5 mm $d^{-1}$ in extreme years, and may increase water stress in use. Overall, it can be concluded that the GWO-XGB model performed better than the other models, followed by the MLP model.

### *4.2 2-Month Ahead Forecasting*

Tab. 3 shows the statistical indicators of the GWO-XGB hybrid model as well as three single soft computing models (M5, MLP and XGB) for forecasting 2-month ahead reference evapotranspiration. As shown in Tab. 3, similar to the forecasting of 1-month ahead $ET_0$, there were no obvious over-fitting and under-fitting problems in the four soft computing models for forecasting 2-month ahead $ET_0$. At Station 58238, MLP model yielded the best accuracy with RMSE = 0.383 mm $d^{-1}$, NSE = 0.983 and 0.284 mm $d^{-1}$ in the testing stage, and the corresponding values in the validation stage were 0.532 mm $d^{-1}$, 0.813 and 0.398 mm $d^{-1}$, respectively. Besides, the M5 model showed the same RMSE and NSE as well as slightly worse MAE than the MLP model in the validation stage. At Stations 57494 and 58606, XGB model was superior to the other models with RMSE = 0.443 and 0.432 mm $d^{-1}$, 0.878 and 0.881, 0.315 and 0.333 mm $d^{-1}$ in the validation stage. At the other six stations, the GWO-XGB model was more accurate in both the testing and validation stages with RMSE ranging 0.327–0.459 mm $d^{-1}$ and 0.355–0.454 mm $d^{-1}$, NSE ranging 0.778–0.901 and 0.733–0.888, MAE ranging 0.254–0.345 mm $d^{-1}$ and 0.276–0.360 mm $d^{-1}$. It can be concluded that the GWO-XGB model performed best, followed by the XGB model.

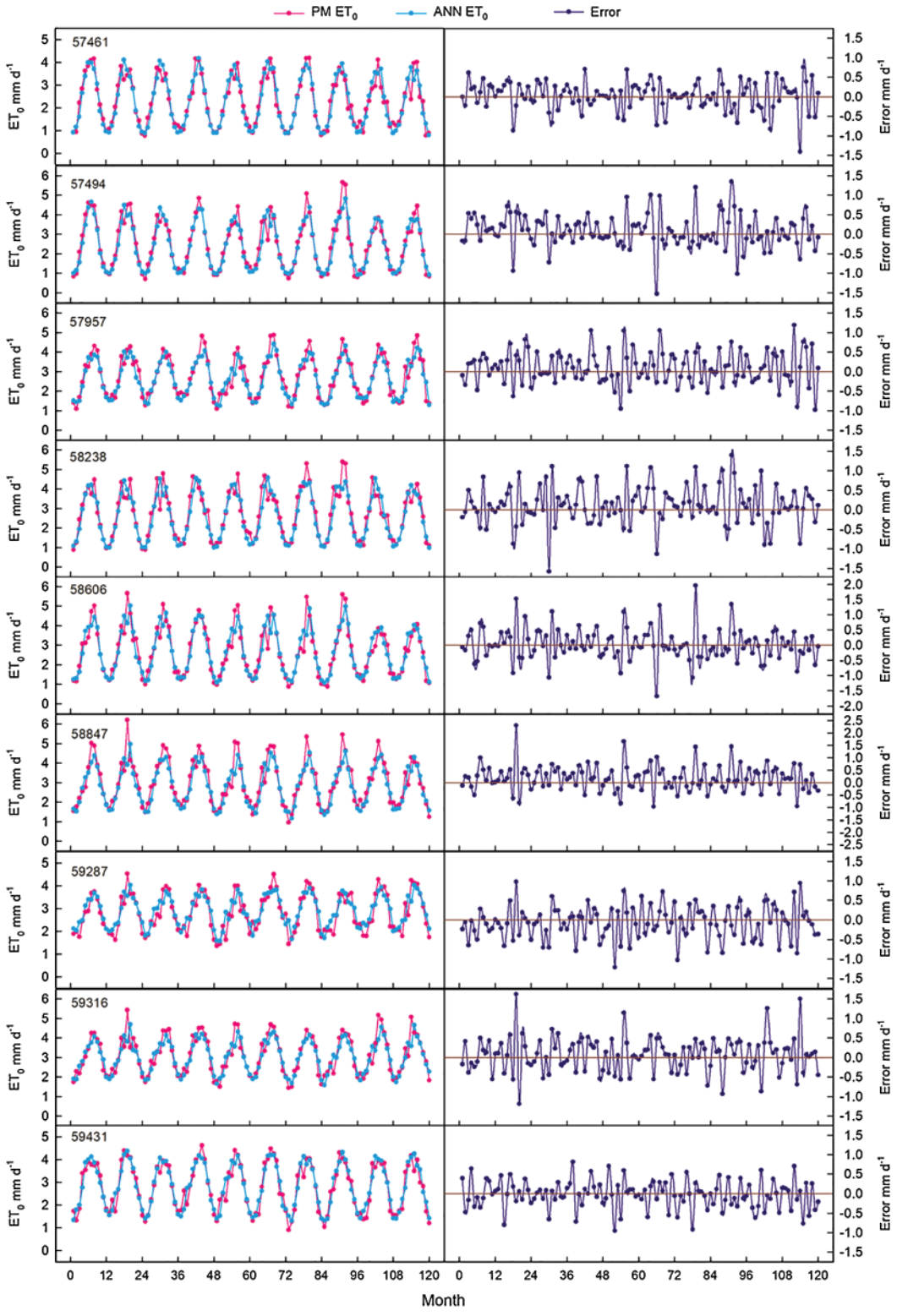


**Figure 4:** Comparison of forecasted $ET_0$ value and the PM $ET_0$ values by ANN model in each month and the error fluctuation of each month

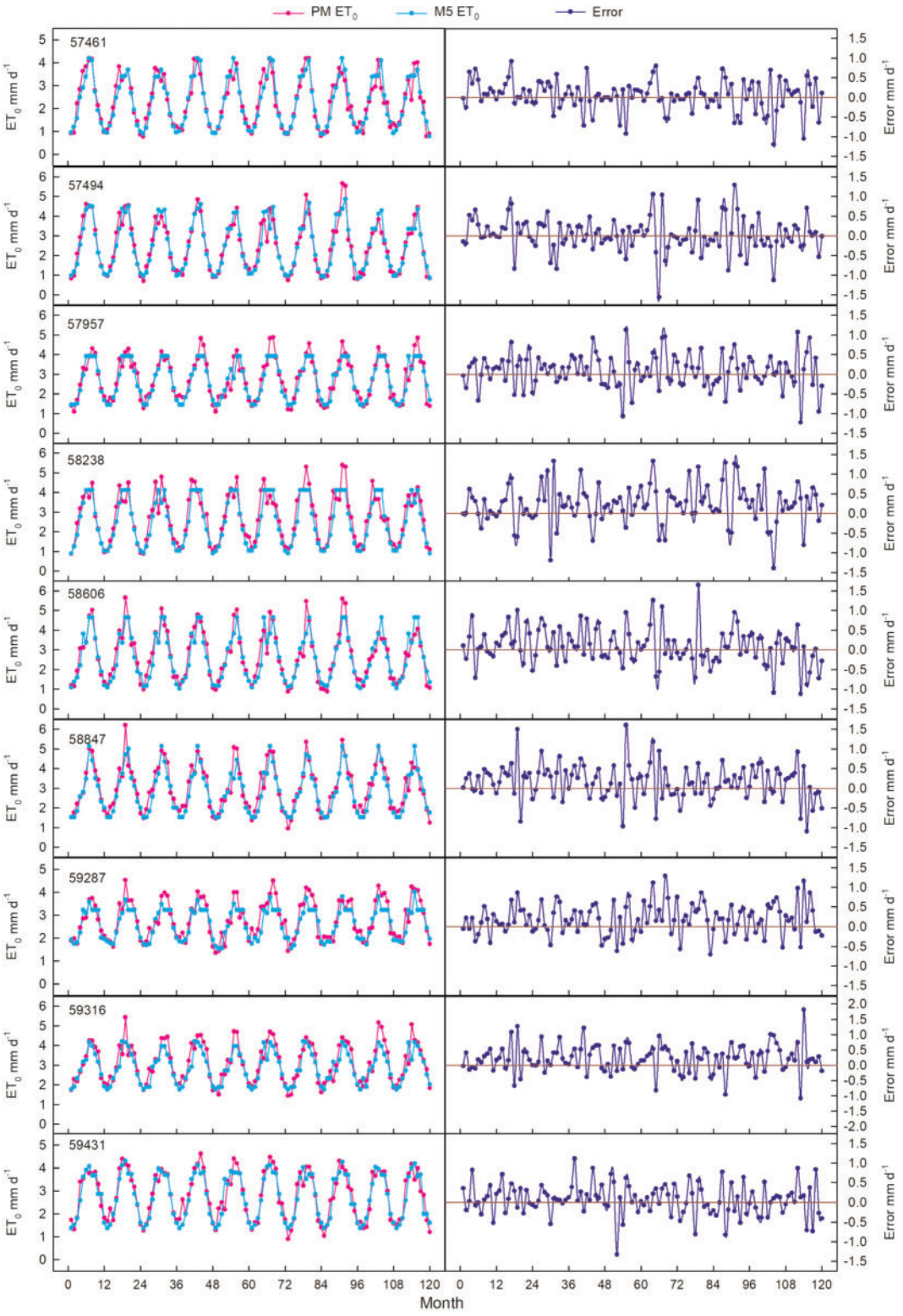


**Figure 5:** Comparison of forecasted $ET_0$ value and the PM $ET_0$ values by M5 model in each month and the error fluctuation of each month

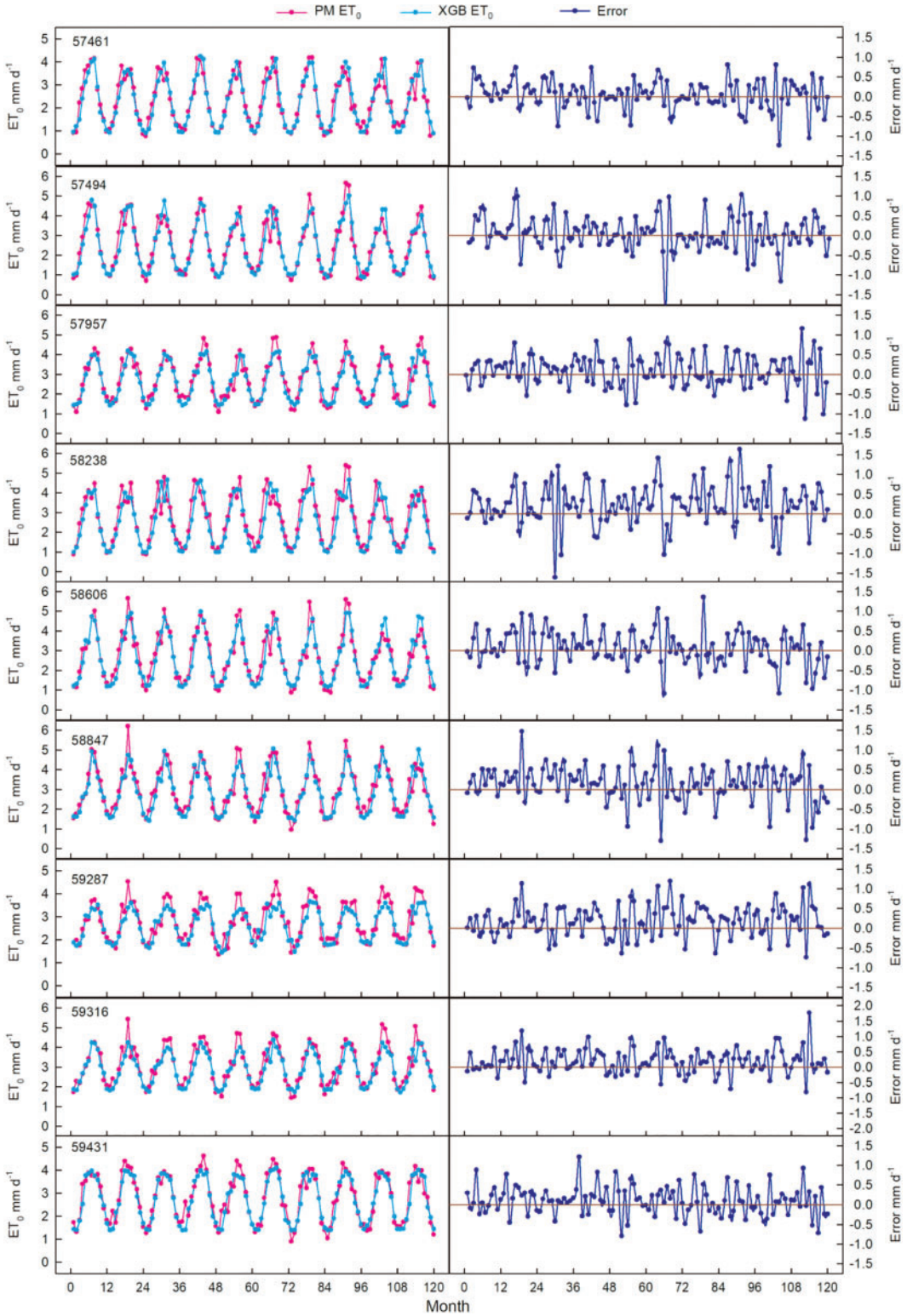


**Figure 6:** Comparison of forecasted $ET_0$ value and the PM $ET_0$ values by XGB model in each month and the error fluctuation of each month

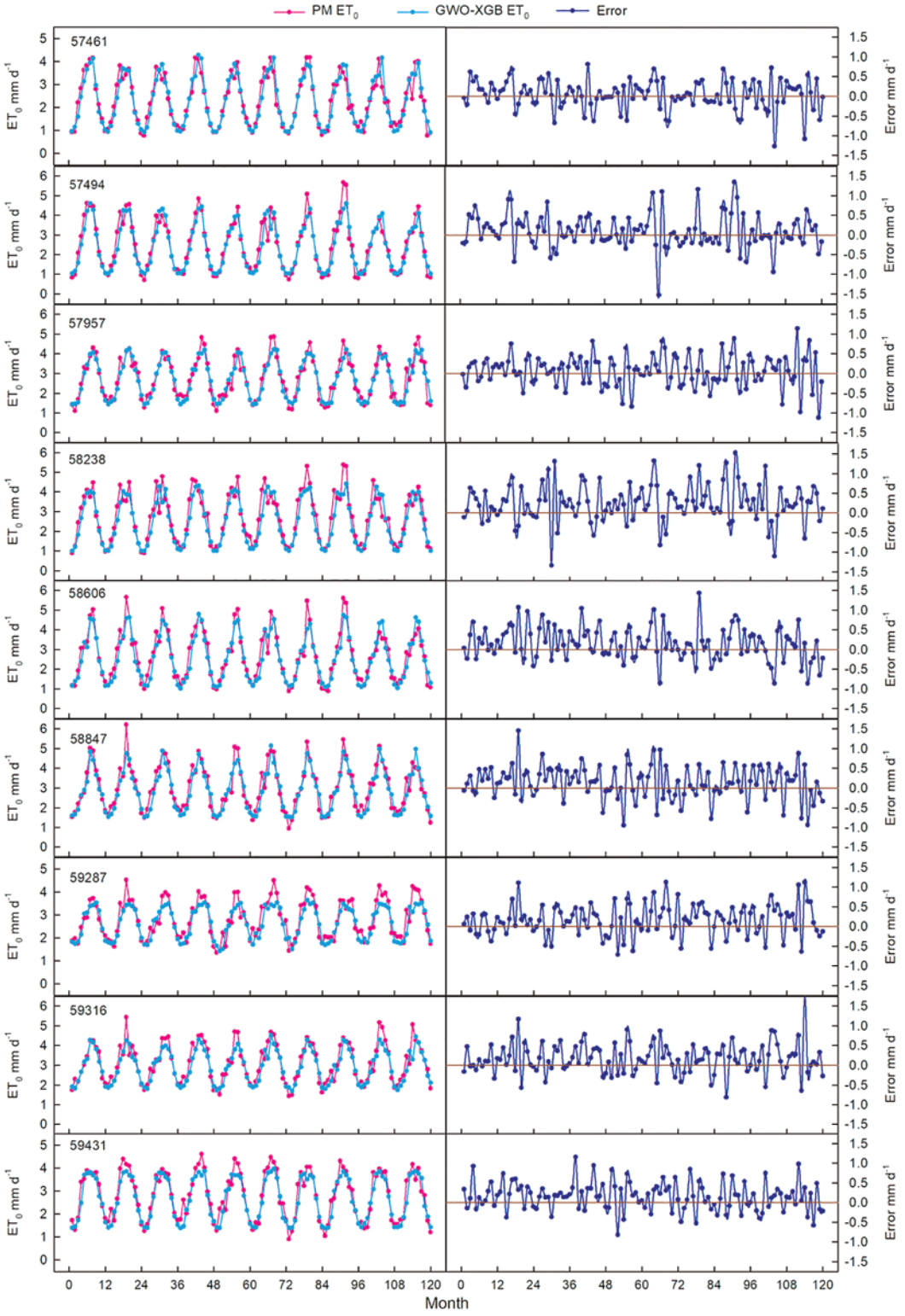


**Figure 7:** Comparison of forecasted $ET_0$ value and the PM $ET_0$ values by GWO-XGB model in each month and the error fluctuation of each month

**Table 3:** Statistical indicators of four soft computing models for 2-month ahead reference evapotranspiration forecasting at nine meteorological stations

| Station | Model | Training | | | Testing | | | Validation | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | RMSE mm $d^{-1}$ | NSE | MAE mm $d^{-1}$ | RMSE mm $d^{-1}$ | NSE | MAE mm $d^{-1}$ | RMSE mm $d^{-1}$ | NSE | MAE mm $d^{-1}$ |
| 57461 | M5 | 0.392 | 0.877 | 0.287 | 0.371 | 0.883 | 0.287 | 0.365 | 0.882 | 0.271 |
| | MLP | 0.429 | 0.855 | 0.314 | 0.369 | 0.884 | 0.281 | 0.384 | 0.869 | 0.291 |
| | XGB | **0.361** | **0.897** | **0.265** | 0.374 | 0.881 | 0.281 | 0.362 | 0.883 | 0.277 |
| | GWO-XGB | 0.383 | 0.884 | 0.283 | **0.341** | **0.901** | **0.264** | **0.355** | **0.888** | **0.276** |
| 57494 | M5 | 0.469 | 0.867 | 0.325 | 0.407 | 0.881 | 0.298 | 0.423 | 0.889 | 0.307 |
| | MLP | 0.510 | 0.848 | 0.367 | 0.386 | 0.893 | 0.288 | 0.481 | 0.857 | 0.352 |
| | XGB | **0.415** | **0.899** | **0.299** | 0.374 | 0.899 | 0.278 | **0.443** | **0.878** | **0.315** |
| | GWO-XGB | 0.447 | 0.883 | 0.319 | **0.367** | **0.903** | **0.276** | 0.448 | 0.876 | 0.319 |
| 57957 | M5 | 0.390 | 0.851 | 0.298 | 0.459 | 0.796 | 0.350 | 0.404 | 0.853 | 0.318 |
| | MLP | 0.469 | 0.790 | 0.361 | 0.489 | 0.769 | 0.358 | 0.472 | 0.799 | 0.388 |
| | XGB | 0.364 | 0.873 | 0.278 | 0.461 | 0.795 | 0.349 | 0.411 | 0.847 | 0.318 |
| | GWO-XGB | **0.358** | **0.877** | **0.274** | **0.447** | **0.807** | **0.345** | **0.406** | **0.851** | **0.319** |
| 58238 | M5 | 0.412 | 0.881 | 0.288 | 0.404 | 0.880 | 0.304 | **0.532** | **0.813** | 0.414 |
| | MLP | 0.447 | 0.862 | 0.315 | **0.383** | **0.893** | **0.284** | **0.532** | **0.813** | **0.398** |
| | XGB | **0.366** | **0.908** | **0.258** | 0.406 | 0.879 | 0.300 | 0.563 | 0.790 | 0.424 |
| | GWO-XGB | 0.395 | 0.892 | 0.279 | 0.398 | 0.884 | 0.302 | 0.543 | 0.805 | 0.416 |
| 58606 | M5 | 0.409 | 0.892 | 0.303 | 0.520 | 0.801 | 0.375 | 0.445 | 0.874 | 0.340 |
| | MLP | 0.534 | 0.820 | 0.399 | 0.515 | 0.805 | 0.365 | 0.514 | 0.831 | 0.383 |
| | XGB | **0.394** | **0.902** | **0.288** | 0.484 | 0.827 | 0.344 | **0.432** | **0.881** | **0.333** |
| | GWO-XGB | 0.429 | 0.884 | 0.309 | **0.459** | **0.844** | **0.329** | 0.452 | 0.870 | 0.351 |
| 58847 | M5 | 0.483 | 0.819 | 0.349 | 0.532 | 0.779 | 0.386 | 0.525 | 0.792 | 0.411 |
| | MLP | 0.601 | 0.720 | 0.452 | 0.572 | 0.745 | 0.422 | 0.584 | 0.743 | 0.447 |
| | XGB | **0.399** | **0.877** | **0.305** | 0.504 | 0.802 | 0.378 | 0.493 | 0.817 | 0.389 |
| | GWO-XGB | 0.421 | 0.862 | 0.323 | **0.459** | **0.836** | **0.342** | **0.445** | **0.851** | **0.360** |
| 59287 | M5 | 0.346 | 0.804 | 0.268 | 0.390 | 0.732 | 0.308 | 0.439 | 0.724 | 0.366 |
| | MLP | 0.448 | 0.681 | 0.359 | 0.365 | 0.765 | 0.285 | 0.450 | 0.710 | 0.382 |
| | XGB | **0.326** | **0.831** | **0.252** | 0.366 | 0.763 | 0.286 | 0.459 | 0.698 | 0.378 |
| | GWO-XGB | 0.353 | 0.801 | 0.280 | **0.327** | **0.811** | **0.254** | **0.432** | **0.733** | **0.358** |
| 59316 | M5 | 0.373 | 0.816 | 0.284 | 0.482 | 0.724 | 0.368 | 0.490 | 0.750 | 0.372 |
| | MLP | 0.437 | 0.749 | 0.334 | 0.454 | 0.754 | 0.348 | 0.488 | 0.752 | 0.393 |
| | XGB | **0.338** | **0.850** | **0.259** | 0.475 | 0.731 | 0.364 | 0.472 | 0.768 | 0.363 |
| | GWO-XGB | 0.361 | 0.829 | 0.279 | **0.432** | **0.778** | **0.326** | **0.454** | **0.785** | **0.355** |
| 59431 | M5 | 0.392 | 0.877 | 0.287 | 0.371 | 0.883 | 0.287 | 0.365 | 0.882 | 0.271 |
| | MLP | 0.429 | 0.855 | 0.314 | 0.369 | 0.884 | 0.281 | 0.384 | 0.869 | 0.291 |
| | XGB | **0.361** | **0.897** | **0.265** | 0.374 | 0.881 | 0.281 | 0.362 | 0.883 | 0.277 |
| | GWO-XGB | 0.383 | 0.884 | 0.283 | **0.341** | **0.901** | **0.264** | **0.355** | **0.888** | **0.276** |

### *4.3 3-Month Ahead Forecasting*

Tab. 4 shows the statistical indicators of the GWO-XGB hybrid model as well as three single soft computing models (M5, MLP and XGB) for forecasting 3-month ahead $ET_0$. Similar to the forecasting of 1, 2-month ahead $ET_0$, there were no obvious over-fitting and under-fitting problems in the four soft computing models for forecasting 3-month ahead. Based on the

tabulated results, the GWO-XGB model performed best at Station 57461 in the testing stage, while the MLP model had the highest accuracy in terms of statistical indicators in the validation stage. Similar results occurred at Station 57957, where the M5 model performed slightly better than the GWO-XGB model based on the MAE value in the validation stage. At the other stations, the GWO-XGB model was generally superior to the other three models with RMSE ranging 0.388–0.540 mm $d^{-1}$, NSE ranging 0.734–0.872, 0.303–0.423 mm $d^{-1}$ at the validation stage. Overall, the GWO model performed best, while the other three models were close to each other.

**Table 4:** Statistical indicators of four soft computing models for 3-month ahead reference evapotranspiration forecasting at nine meteorological stations

| Station | Model | Training | | | Testing | | | Validation | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | RMSE mm $d^{-1}$ | NSE | MAE mm $d^{-1}$ | RMSE mm $d^{-1}$ | NSE | MAE mm $d^{-1}$ | RMSE mm $d^{-1}$ | NSE | MAE mm $d^{-1}$ |
| 57461 | M5 | 0.356 | 0.881 | 0.270 | 0.381 | 0.838 | 0.280 | 0.458 | 0.800 | 0.337 |
| | MLP | 0.398 | 0.851 | 0.320 | 0.385 | 0.835 | 0.288 | **0.380** | **0.862** | **0.307** |
| | XGB | **0.333** | **0.895** | **0.257** | 0.361 | 0.855 | 0.277 | 0.407 | 0.841 | 0.315 |
| | GWO-XGB | 0.374 | 0.868 | 0.280 | **0.337** | **0.874** | **0.259** | 0.398 | 0.848 | 0.322 |
| 57494 | M5 | 0.389 | 0.879 | 0.277 | 0.393 | 0.869 | 0.303 | 0.361 | 0.884 | 0.279 |
| | MLP | 0.429 | 0.855 | 0.309 | 0.361 | 0.889 | 0.284 | 0.363 | 0.883 | 0.279 |
| | XGB | **0.360** | **0.898** | **0.264** | 0.384 | 0.874 | 0.297 | 0.355 | 0.888 | 0.270 |
| | GWO-XGB | 0.388 | 0.881 | 0.284 | **0.358** | **0.891** | **0.278** | **0.351** | **0.891** | **0.270** |
| 57957 | M5 | 0.477 | 0.862 | 0.326 | 0.420 | 0.873 | 0.312 | **0.451** | **0.874** | **0.324** |
| | MLP | 0.510 | 0.848 | 0.361 | 0.417 | 0.875 | 0.330 | 0.453 | 0.873 | 0.337 |
| | XGB | **0.434** | **0.890** | **0.301** | 0.421 | 0.873 | 0.312 | 0.446 | 0.877 | 0.326 |
| | GWO-XGB | 0.509 | 0.848 | 0.354 | **0.365** | **0.904** | **0.274** | **0.451** | **0.874** | 0.331 |
| 58238 | M5 | 0.420 | 0.827 | 0.315 | 0.440 | 0.813 | 0.342 | 0.426 | 0.836 | 0.339 |
| | MLP | 0.542 | 0.719 | 0.409 | 0.550 | 0.707 | 0.415 | 0.502 | 0.773 | 0.409 |
| | XGB | 0.387 | 0.856 | 0.291 | 0.444 | 0.810 | 0.336 | 0.401 | 0.855 | 0.310 |
| | GWO-XGB | **0.384** | **0.859** | **0.290** | **0.432** | **0.820** | **0.329** | **0.388** | **0.865** | **0.303** |
| 58606 | M5 | 0.406 | 0.884 | 0.280 | 0.423 | 0.869 | 0.316 | 0.550 | 0.800 | 0.417 |
| | MLP | 0.440 | 0.866 | 0.311 | 0.403 | 0.881 | 0.297 | 0.545 | 0.804 | 0.414 |
| | XGB | 0.362 | **0.910** | 0.259 | 0.421 | 0.870 | 0.302 | 0.546 | 0.803 | 0.423 |
| | GWO-XGB | **0.361** | **0.910** | **0.257** | **0.402** | **0.881** | **0.293** | **0.540** | **0.807** | **0.423** |
| 58847 | M5 | 0.439 | 0.876 | 0.321 | 0.515 | 0.805 | 0.366 | 0.532 | 0.820 | 0.388 |
| | MLP | 0.523 | 0.827 | 0.394 | 0.539 | 0.786 | 0.395 | 0.495 | 0.844 | 0.376 |
| | XGB | **0.389** | **0.905** | **0.290** | 0.506 | 0.811 | 0.352 | **0.452** | **0.870** | **0.346** |
| | GWO-XGB | 0.421 | 0.888 | 0.311 | **0.475** | **0.834** | **0.339** | 0.471 | 0.859 | 0.365 |
| 59287 | M5 | 0.424 | 0.860 | 0.321 | 0.545 | 0.768 | 0.413 | 0.443 | 0.852 | 0.354 |
| | MLP | 0.528 | 0.784 | 0.404 | 0.500 | 0.805 | 0.384 | 0.494 | 0.816 | 0.376 |
| | XGB | **0.392** | **0.881** | **0.299** | 0.502 | 0.803 | 0.381 | 0.421 | 0.866 | 0.335 |
| | GWO-XGB | 0.416 | 0.866 | 0.318 | **0.477** | **0.823** | **0.356** | **0.412** | **0.872** | **0.321** |
| 59316 | M5 | 0.355 | 0.794 | 0.280 | 0.425 | 0.681 | 0.326 | 0.471 | 0.682 | 0.372 |
| | MLP | 0.556 | 0.508 | 0.433 | 0.455 | 0.635 | 0.369 | 0.520 | 0.613 | 0.422 |
| | XGB | **0.348** | **0.808** | **0.273** | 0.384 | 0.740 | 0.304 | **0.426** | 0.740 | 0.344 |
| | GWO-XGB | 0.378 | 0.773 | 0.297 | **0.375** | **0.751** | **0.300** | 0.431 | **0.734** | **0.341** |
| 59431 | M5 | 0.373 | 0.816 | 0.281 | 0.531 | 0.664 | 0.424 | 0.452 | 0.787 | 0.348 |
| | MLP | 0.451 | 0.733 | 0.357 | **0.435** | **0.775** | **0.335** | 0.468 | 0.772 | 0.370 |
| | XGB | **0.343** | **0.846** | **0.260** | 0.511 | 0.689 | 0.397 | 0.454 | 0.785 | 0.356 |
| | GWO-XGB | 0.362 | 0.828 | 0.280 | 0.461 | 0.747 | 0.359 | **0.438** | **0.800** | **0.354** |

### *4.4 Model Evaluation in Each Month*

Plants have different water requirements in different seasons, especially in summer and autumn. Therefore, taking RMSE value as an example, we evaluated the performances of various models for forecasting $ET_0$ in summer and autumn at the nine stations (Fig. 8). As shown in Fig. 8, the data fluctuated in different models and at various forecasting time steps of monthly $ET_0$. I MLP model performed best in summer, especially for forecasting 1-month ahead $ET_0$, where the median values and lower-quartile values of RMSE were significantly lower than those of the other models. In addition, the median values of RMSE for forecasting 2- and 3-month ahead $ET_0$ were also lower than those of the other models. However, the MLP model performed worst in summer; particularly, the upper quartile line of RMSE exceeded 0.7 mm $d^{-1}$, followed by the M5 model. The GWO-XGB model attained the lowest RMSE for forecasting 2-month ahead $ET_0$ in autumn. In other cases, the GWO-XGB model showed similar results to the XGB model. The inspiration from the above was that different models could be considered for different seasons, namely, MLP model in summer and GWO-XGB model in autumn. Huang et al. [65] evaluated the CatBoost model, random forest and support vector machines for prediction of $ET_0$. In their study, meteorological data from 2001 to 2015 were collected from 12 stations in humid regions of China and eight input combinations were investigated to find the suitable model with limited data. The RMSE was 0.35 mm $d^{-1}$ with maximum temperature, minimum temperature, global solar radiation as inputs and 0.52 mm $d^{-1}$ with maximum temperature, minimum temperature, wind speed and relative humid as inputs. It is worth mentioning that historical meteorological data are used as inputs, and the errors will further increase if numerical weather forecasting data are used as inputs. This indicates that the error of this study is within the appropriate and applicable range.

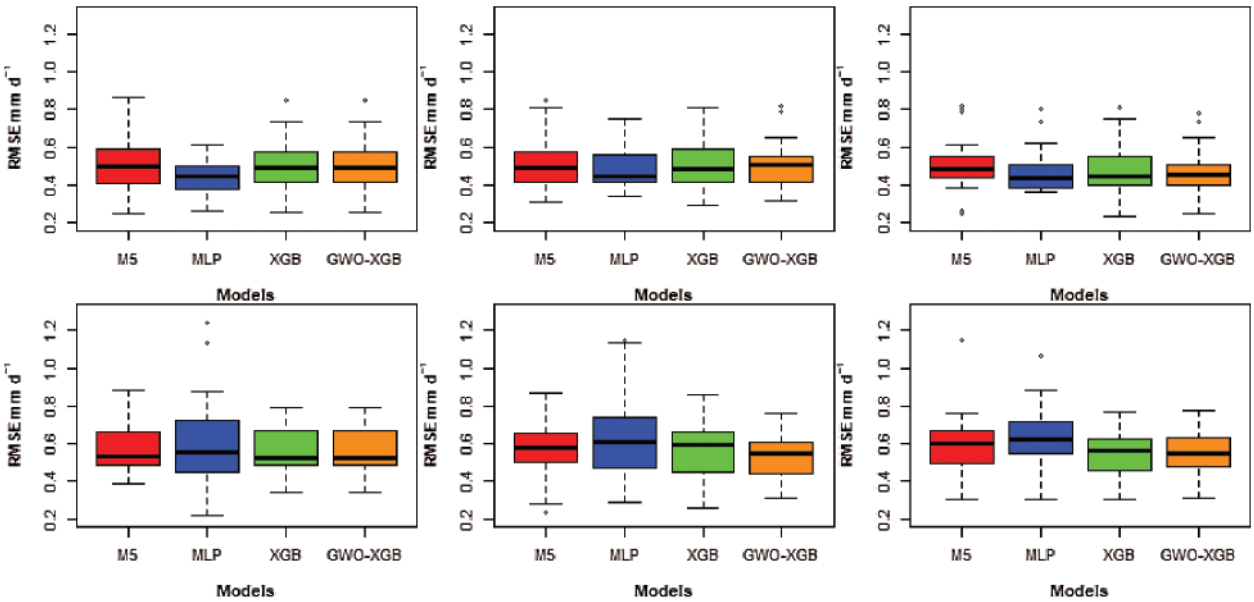


**Figure 8:** Accuracy performance of different models in the seasons of summer and autumn at the nine stations

## 5 Discussion

Accurate forecasting of $ET_0$ in the future is of great significance for hydrological simulation, water resource management and agricultural water management. A higher accuracy can be obtained by using future weather forecasts; however, this method usually increases the error significantly after 7d [47,66,67]. Karbasi [50] developed a hybrid model of the Gaussian Process Regression (GPR) and Wavelet-GPR to forecast multi-step ahead daily (1–30 days ahead) reference evapotranspiration in Iran. I obtained RMSE values by the models ranged from 0.068 mm/d to 0.816 mm/d from 1 to 30 days. Zhao et al. [49] proposed a post-processing method for 1–3 month $ET_0$ forecasting based on GCM outputs, which has obvious advantages over the original method. The RMSE dropped from 0.83–0.97 mm/d (GCM) to 0.34–0.76 mm/d with the new approach at Aero Station in Townsville, Australia. In this study, we proposed a method to predict 1–3 month $ET_0$ ahead based on historical $ET_0$ in a humid region of China, and the method also obtained good accuracy.

Boosting model is a new method based on decision tree, which uses boosting thought to integrate decision tree. Compared with the earlier random forest model based on bagging method, XGBoost model had a great advantage in running speed and a slight improvement in accuracy and controlling overfitting [14,68–71].

The machine learning model needs parameters based on the dataset. A lot of research showed that heuristic algorithms can improve the accuracy and stability of machine learning models [7,72,73]. Similar results have been obtained in this study, where GWO improved the stability of XGBoost models and slightly improve the accuracy of the GWO model. The disadvantage is that it takes more time to adjust parameters. If the standalone XGB model uses grid search to determine the most appropriate model parameters, only a few hundred parameters are needed. However, the GWO-XGB model needs 25,000 parameters in this study, which takes hundreds of times more than the standalone XGB model. The next research goal is to use new techniques to speed up the efficiency of parameter tuning, such as parallel algorithms, to save the time required for parameter tuning.

Meteorological factors such as temperature and solar radiation change dramatically in four seasons of the year, resulting in huge differences in $ET_0$ among the seasons. For models like MLP, the algorithm itself has only one model, which may be more prone to the problem of unbalanced performance in different seasons. Similar results also occurred in Sichuan, China as found by Feng et al. [36]. Models such as XGBoost method, which have multiple sub-models, can build more specific models for different seasons, and combine all of them together so that the differences among the seasons are more balanced.

## 6 Conclusions

The forecasting of several months ahead reference evapotranspiration is helpful in water resources management and allocation for irrigated areas. This study investigated and compared the performance of the XGBoost model hybridized with the Grey Wolf Optimization algorithm, along with three traditional models, e.g., single XGBoost, ANN and M5 models for forecasting 1–3 month ahead $ET_0$. The meteorological data obtained from nine stations in different subtropical zones were used as inputs for training, testing and validating the above models. The results showed that the newly developed GWO-XGBoost was a reliable and stable approach for $ET_0$ forecasting. To forecast future 1–3 month ahead, the GWO-XGBoost model had the smallest error in the validation stage, with RMSE = 0.431 mm $d^{-1}$, NSE = 0.840 and MAE = 0.335 mm $d^{-1}$, followed by the XGBoost model with RMSE = 0.438 mm $d^{-1}$, NSE = 0.842 and

MAE = 0.338 mm $d^{-1}$, respectively. The MLP model was the worse model in this study, with RMSE = 0.465 mm $d^{-1}$, NSE = 0.811 and MAE = 0.359 mm $d^{-1}$. However, the error of the MLP model was the smallest in summer among the four models.

**Acknowledgement:** Thanks to the National Meteorological Information Center of China Meteorological Administration for offering the meteorological data.

**Funding Statement:** This study was jointly supported by the National Natural Science Foundation of China (Nos. 51879196, 51790533, 51709143), Jiangxi Natural Science Foundation of China (No. 20181BAB206045).

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

1. Allen, R., Pereira, L., Raes, D., Smith, M., Allen, R. G. et al. (1998). Crop evapotranspiration: Guidelines for computing crop water requirements. *FAO irrigation and drainage paper 56*. Vol. 300, pp. 1–50. Rome: Fao.
2. Júnior, L., Ventura, T., Gomes, R., Nogueira, J., Lobo, F. et al. (2020). Comparative assessment of modelled and empirical reference evapotranspiration methods for a brazilian savanna. *Agricultural Water Management, 232,* 1–13. DOI 10.1016/j.agwat.2020.106040.
3. Xiao, Q., Li, C., Tang, Y., Li, L., Li, L. (2019). A knowledge-driven method of adaptively optimizing process parameters for energy efficient turning. *Energy, 166,* 142–156. DOI 10.1016/j.energy.2018.09.191.
4. Khan, M. M., Shamseldin, A. Y., Khan, Z. M., Khan, S., Sultan, T. et al. (2018). A comparative study of various hybrid wavelet feedforward neural network models for runoff forecasting. *Water Resources Management, 32(1),* 83–103. DOI 10.1007/s11269-017-1796-1.
5. Choubin, B., Zehtabian, G., Azareh, A., Rafiei-Sardooi, E., Sajedi-Hosseini, F. et al. (2018). Precipitation forecasting using classification and regression trees (CART) model: A comparative study of different approaches. *Environmental Earth Sciences, 77(8),* 314–334. DOI 10.1007/s12665-018-7498-z.
6. Ali, M., Deo, R. C., Downs, N. J., Maraseni, T. (2018). An ensemble-ANFIS based uncertainty assessment model for forecasting multi-scalar standardized precipitation index. *Atmospheric Research, 207,* 155–180. DOI 10.1016/j.atmosres.2018.02.024.
7. Fan, J. L., Wu, L. F., Ma, X., Zhou, H., Zhang, F. C. (2020). Hybrid support vector machines with heuristic algorithms for prediction of daily diffuse solar radiation in air-polluted regions. *Renewable Energy, 145,* 2034–2045. DOI 10.1016/j.renene.2019.07.104.
8. Fan, J. L., Wu, L. F., Zhang, F. C., Cai, H. J., Ma, X. et al. (2019). Evaluation and development of empirical models for estimating daily and monthly mean daily diffuse horizontal solar radiation for different climatic regions of China. *Renewable and Sustainable Energy Reviews, 105,* 168–186. DOI 10.1016/j.rser.2019.01.040.
9. Fan, J. L., Wu, L. F., Zhang, F. C., Cai, H. J., Zeng, W. Z. et al. (2019). Empirical and machine learning models for predicting daily global solar radiation from sunshine duration: A review and case study in China. *Renewable and Sustainable Energy Reviews, 100,* 186–212. DOI 10.1016/j.rser.2018.10.018.
10. Lu, X. H., Ju, Y., Wu, L. F., Fan, J. L., Zhang, F. C. et al. (2018). Daily pan evaporation modeling from local and cross-station data using three tree-based machine learning models. *Journal of Hydrology, 566,* 668–684. DOI 10.1016/j.jhydrol.2018.09.055.
11. Valipour, M., Sefidkouhi, M., Raeini-Sarjaz, M., Guzman, S. (2019). A hybrid data-driven machine learning technique for evapotranspiration modeling in various climates. *Atmosphere, 10(6),* 311–326. DOI 10.3390/atmos10060311.

12. Zhao, W., Gentine, P., Reichstein, M., Zhang, Y., Zhou, S. et al. (2020). Physics-constrained machine learning of evapotranspiration. *Geophysical Research Letters, 1–28(24),* 14496–14507. DOI 10.1029/2019GL085291.
13. Kouadio, L., Deo, R. C., Byrareddy, V., Adamowski, J. F., Mushtaq, S. et al. (2018). Artificial intelligence approach for the prediction of Robusta coffee yield using soil fertility properties. *Computers and Electronics in Agriculture, 155,* 324–338. DOI 10.1016/j.compag.2018.10.014.
14. Fan, J. L., Yue, W., Wu, L. F., Zhang, F. C., Cai, H. J. et al. (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. *Agricultural and Forest Meteorology, 263,* 225–241. DOI 10.1016/j.agrformet.2018.08.019.
15. Granata, F., Gargano, R., Marinis, G. (2020). Artificial intelligence based approaches to evaluate actual evapotranspiration in wetlands. *Science of the Total Environment, 703,* 1–15. DOI 10.1016/j.scitotenv.2019.135653.
16. Wu, L. F., Zhou, H., Ma, X., Fan, J. L., Zhang, F. C. (2019). Daily reference evapotranspiration prediction based on hybridized extreme learning machine model with bio-inspired optimization algorithms: Application in contrasting climates of China. *Journal of Hydrology, 577,* 1–13. DOI 10.1016/j.jhydrol.2019.123960.
17. Khoob, A. R. (2008). Comparative study of Hargreaves's and artificial neural network's methodologies in estimating reference evapotranspiration in a semiarid environment. *Irrigation Science, 26(3),* 253–259. DOI 10.1007/s00271-007-0090-z.
18. Martí, P., González-Altozano, P., Gasque, M. (2011). Reference evapotranspiration estimation without local climatic data. *Irrigation Science, 29(6),* 479–495. DOI 10.1007/s00271-010-0243-3.
19. Martí, P., Zarzo, M. (2012). Multivariate statistical monitoring of $ET_0$: A new approach for estimation in nearby locations using geographical inputs. *Agricultural and Forest Meteorology, 152,* 125–134. DOI 10.1016/j.agrformet.2011.08.008.
20. Huo, Z., Feng, S., Kang, S., Dai, X. (2012). Artificial neural network models for reference evapotranspiration in an arid area of Northwest China. *Journal of Arid Environments, 82,* 81–90. DOI 10.1016/j.jaridenv.2012.01.016.
21. Antonopoulos, V. Z., Antonopoulos, A. V. (2017). Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. *Computers and Electronics in Agriculture, 132,* 86–96. DOI 10.1016/j.compag.2016.11.011.
22. Cobaner, M. (2011). Evapotranspiration estimation by two different neuro-fuzzy inference systems. *Journal of Hydrology, 398(3),* 292–302. DOI 10.1016/j.jhydrol.2010.12.030.
23. Kisi, O., Zounemat-Kermani, M. (2014). Comparison of two different adaptive neuro-fuzzy inference systems in modelling daily reference evapotranspiration. *Water Resources Management, 28(9),* 2655–2675. DOI 10.1007/s11269-014-0632-0.
24. Shamshirband, S., Amirmojahedi, M., Gocic, M., Akib, S., Petkovic, D. et al. (2016). Estimation of reference evapotranspiration using neural networks and Cuckoo Search Algorithm. *Journal of Irrigation and Drainage Engineering, 142(2),* 1–11. DOI 10.1061/(ASCE)IR.1943-4774.0000949.
25. Keshtegar, B., Kisi, O., Arab, H. G., Zounemat-Kermani, M. (2018). Subset modeling basis ANFIS for prediction of the reference evapotranspiration. *Water Resources Management, 32(3),* 1101–1116. DOI 10.1007/s11269-017-1857-5.
26. Ladlani, I., Houichi, L., Djemili, L., Heddam, S., Belouz, K. (2014). Estimation of daily reference evapotranspiration ($ET_0$) in the North of Algeria using adaptive neuro-fuzzy inference system (ANFIS) and multiple linear regression (MLR) models: A comparative study. *Arabian Journal for Science & Engineering, 39(8),* 5959–5969. DOI 10.1007/s13369-014-1151-2.
27. Patil, A. P., Deka, P. C. (2015). Performance evaluation of hybrid wavelet-ANN and wavelet-ANFIS models for estimating evapotranspiration in arid regions of India. *Neural Computing and Applications, 28(2),* 275–285. DOI 10.1007/s00521-015-2055-0.

28. Tao, H., Diop, L., Bodian, A., Djaman, K., Ndiaye, P. M. et al. (2018). Reference evapotranspiration prediction using hybridized fuzzy model with firefly algorithm: Regional case study in Burkina Faso. *Agricultural Water Management, 208,* 140–151. DOI 10.1016/j.agwat.2018.06.018.
29. Karimi, S., Kisi, O., Kim, S., Nazemi, A. H., Shiri, J. (2016). Modelling daily reference evapotranspiration in humid locations of South Korea using local and cross-station data management scenarios. *International Journal of Climatology, 37(77),* 3238–3246. DOI 10.1002/joc.4911.
30. Kiafar, H., Babazadeh, H., Marti, P., Kisi, O., Shiri, J. (2016). Evaluating the generalizability of GEP models for estimating reference evapotranspiration in distant humid and arid locations. *Theoretical and Applied Climatology, 130(1–2),* 377–389. DOI 10.1007/s00704-016-1888-5.
31. Shiri, J. (2019). Modeling reference evapotranspiration in island environments: Assessing the practical implications. *Journal of Hydrology, 570,* 265–280. DOI 10.1016/j.jhydrol.2018.12.068.
32. Mattar, M. A. (2018). Using gene expression programming in monthly reference evapotranspiration modeling: A case study in Egypt. *Agricultural Water Management, 198,* 28–38. DOI 10.1016/j.agwat.2017.12.017.
33. Abdullah, S. S., Malek, M. A., Abdullah, N. S., Kisi, O., Yap, K. S. (2015). Extreme learning machines: A new approach for prediction of reference evapotranspiration. *Journal of Hydrology, 527,* 184–195. DOI 10.1016/j.jhydrol.2015.04.073.
34. Gocic, M., Petković, D., Shamshirband, S., Kamsin, A. (2016). Comparative analysis of reference evapotranspiration equations modelling by extreme learning machine. *Computers and Electronics in Agriculture, 127,* 56–63. DOI 10.1016/j.compag.2016.05.017.
35. Feng, Y., Cui, N. B., Zhao, L., Hu, X. T., Gong, D. Z. (2016). Comparison of ELM, GANN, WNN and empirical models for estimating reference evapotranspiration in humid region of Southwest China. *Journal of Hydrology, 536,* 376–383. DOI 10.1016/j.jhydrol.2016.02.053.
36. Feng, Y., Peng, Y., Cui, N. B., Gong, D. Z., Zhang, K. D. (2017). Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. *Computers and Electronics in Agriculture, 136,* 71–78. DOI 10.1016/j.compag.2017.01.027.
37. Wen, X., Si, J., He, Z., Wu, J., Shao, H. et al. (2015). Support-vector-machine-based models for modeling daily reference evapotranspiration with limited climatic data in extreme arid regions. *Water Resources Management, 29(9),* 3195–3209. DOI 10.1007/s11269-015-0990-2.
38. Dong, J. H., Wu, L. F., Liu, X. G., Li, Z. J., Gao, Y. L. et al. (2020). Estimation of daily dew point temperature by using bat algorithm optimization based extreme learning machine. *Applied Thermal Engineering, 165,* 1–15. DOI 10.1016/j.applthermaleng.2019.114569.
39. Liu, T., Ding, Y., Xin, C., Zhu, Y., Zhang, X. (2017). Extreme learning machine based on particle swarm optimization for estimation of reference evapotranspiration. *36th Chinese Control Conference*, pp. 4567–4572. Dalian. DOI 10.23919/ChiCC.2017.8028076.
40. Pal, M., Deswal, S. (2009). M5 model tree based modelling of reference evapotranspiration. *Hydrological Processes, 23(10),* 1437–1443. DOI 10.1002/hyp.7266.
41. Rahimikhoob, A. (2014). Comparison between M5 model tree and neural networks for estimating reference evapotranspiration in an arid environment. *Water Resources Management, 28(3),* 657–669. DOI 10.1007/s11269-013-0506-x.
42. Rahimikhoob, A. (2016). Comparison of M5 model tree and artificial neural network's methodologies in modelling daily reference evapotranspiration from NOAA satellite images. *Water Resources Management, 30(9),* 3063–3075. DOI 10.1007/s11269-016-1331-9.
43. Shiri, J. (2018). Improving the performance of the mass transfer-based reference evapotranspiration estimation approaches through a coupled wavelet-random forest methodology. *Journal of Hydrology, 561,* 737–750. DOI 10.1016/j.jhydrol.2018.04.042.
44. Wang, S., Lian, J., Peng, Y., Hu, B., Chen, H. (2019). Generalized reference evapotranspiration models with limited climatic data based on random forest and gene expression programming in Guangxi. *China Agricultural Water Management, 221,* 220–230. DOI 10.1016/j.agwat.2019.03.027.

45. Granata, F. (2019). Evapotranspiration evaluation models based on machine learning algorithms—A comparative study. *Agricultural Water Management, 217,* 303–315. DOI 10.1016/j.agwat.2019.03.015.
46. Dong, J. H., Wu, L. F., Liu, X. G., Fan, C., Leng, M. H. et al. (2020). Simulation of daily diffuse solar radiation based on three machine learning models. *Computer Modeling in Engineering & Sciences, 123(1),* 49–73. DOI 10.32604/cmes.2020.09014.
47. Perera, K. C., Western, A. W., Nawarathna, B., George, B. (2014). Forecasting daily reference evapotranspiration for Australia using numerical weather prediction outputs. *Agricultural and Forest Meteorology, 194,* 50–63. DOI 10.1016/j.agrformet.2014.03.014.
48. Traore, S., Luo, Y., Fipps, G. (2016). Deployment of artificial neural network for short-term forecasting of evapotranspiration using public weather forecast restricted messages. *Agricultural Water Management, 163,* 363–379. DOI 10.1016/j.agwat.2015.10.009.
49. Zhao, T., Wang, Q. J., Schepen, A., Griffiths, M. (2019). Ensemble forecasting of monthly and seasonal reference crop evapotranspiration based on global climate model outputs. *Agricultural and Forest Meteorology, 264,* 114–124. DOI 10.1016/j.agrformet.2018.10.001.
50. Karbasi, M. (2018). Forecasting of multi-step ahead reference evapotranspiration using wavelet—Gaussian process regression model. *Water Resources Management, 32(3),* 1035–1052. DOI 10.1007/s11269-017-1853-9.
51. Mehdizadeh, S. (2018). Estimation of daily reference evapotranspiration ($ET_0$) using artificial intelligence methods: Offering a new approach for lagged ETo data-based modeling. *Journal of Hydrology, 559,* 794–812. DOI 10.1016/j.jhydrol.2018.02.060.
52. Mohammadi, B., Mehdizadeh, S. (2020). Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm. *Agricultural Water Management, 1–13,* 106145. DOI 10.1016/j.agwat.2020.106145.
53. Mirjalili, S., Mirjalili, S. M., Lewis, A. (2014). Grey Wolf Optimizer. *Advances in Engineering Software, 69,* 46–61. DOI 10.1016/j.advengsoft.2013.12.007.
54. Faris, H., Mirjalili, S., Aljarah, I. (2019). Automatic selection of hidden neurons and weights in neural networks using Grey Wolf Optimizer based on a hybrid encoding scheme. *International Journal of Machine Learning and Cybernetics, 10(10),* 2901–2920. DOI 10.1007/s13042-018-00913-2.
55. Maroufpoor, S., Maroufpoor, E., Bozorg-Haddad, O., Shiri, J., Mundher Yaseen, Z. (2019). Soil moisture simulation using hybrid artificial intelligent model: Hybridization of adaptive neuro fuzzy inference system with Grey Wolf Optimizer algorithm. *Journal of Hydrology, 575,* 544–556. DOI 10.1016/j.jhydrol.2019.05.045.
56. Deng, S., Wang, X., Zhu, Y., Lv, F., Wang, J. (2019). Hybrid Grey Wolf Optimization algorithm-based support vector machine for groutability prediction of fractured rock mass. *Journal of Computing in Civil Engineering, 33(2),* 4018061–4018065. DOI 10.1061/(ASCE)CP.1943-5487.0000814.
57. Zhou, J., Huo, X., Xu, X., Li, Y. (2019). Forecasting the carbon price using extreme-point symmetric mode decomposition and extreme learning machine optimized by the Grey Wolf Optimizer algorithm. *Energies, 12(5),* 950–972. DOI 10.3390/en12050950.
58. Ma, X., Mei, X., Wu, W., Wu, X., Zeng, B. (2019). A novel fractional time delayed grey model with Grey Wolf Optimizer and its applications in forecasting the natural gas and coal consumption in Chongqing China. *Energy, 178,* 487–507. DOI 10.1016/j.energy.2019.04.096.
59. Fan, J. L., Wu, L. F., Zhang, F. C., Xiang, Y. Z., Zheng, J. (2016). Climate change effects on reference crop evapotranspiration across different climatic zones of China during 1956–2015. *Journal of Hydrology, 542,* 923–937. DOI 10.1016/j.jhydrol.2016.09.060.
60. Chen, T., Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *ACM Sigkdd International Conference on Knowledge Discovery & Data Mining, 11(34),* 785–794. DOI 10.1145/2939672.2939785.
61. Behrang, M. A., Assareh, E., Ghanbarzadeh, A., Noghrehabadi, A. R. (2010). The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. *Solar Energy, 84(8),* 1468–1480. DOI 10.1016/j.solener.2010.05.009.

62. Wang, Y., Witten, I. H. (1997). Induction of model trees for predicting continuous lasses. *Proceedings of the Poster Papers of the European Conference on Machine Learning*, Prague: University of Economics, Faculty of Informatics and Statistics.
63. Nash, J. E., Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology, 10(3),* 282–290. DOI 10.1016/0022-1694(70)90255-6.
64. Despotovic, M., Nedic, V., Despotovic, D., Cvetanovic, S. (2015). Review and statistical analysis of different global solar radiation sunshine models. *Renewable and Sustainable Energy Reviews, 52,* 1869–1880. DOI 10.1016/j.rser.2015.08.035.
65. Huang, G. M., Wu, L. F., Ma, X., Zhang, W., Fan, J. L. et al. (2019). Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions. *Journal of Hydrology, 574,* 1029–1041. DOI 10.1016/j.jhydrol.2019.04.085.
66. Cai, J., Liu, Y., Lei, T., Pereira, L. S. (2007). Estimating reference evapotranspiration with the FAO Penman–Monteith equation using daily weather forecast messages. *Agricultural & Forest Meteorology, 145(1–2),* 22–35. DOI 10.1016/j.agrformet.2007.04.012.
67. Luo, Y., Chang, X., Peng, S., Khan, S., Wang, W. et al. (2014). Short-term forecasting of daily reference evapotranspiration using the Hargreaves–Samani model and temperature forecasts. *Agricultural Water Management, 136,* 42–51. DOI 10.1016/j.agwat.2014.01.006.
68. Shimoda, A., Ichikawa, D., Oyama, H. (2018). Using machine-learning approaches to predict non-participation in a nationwide general health check-up scheme. *Computer Methods and Programs in Biomedicine, 163,* 39–46. DOI 10.1016/j.cmpb.2018.05.032.
69. Joharestani, M., Cao, C., Ni, X., Bashir, B., Talebiesfandarani, S. (2019). $PM_{2.5}$ prediction based on random forest, XGBoost, and deep learning using multisource remote sensing data. *Atmosphere (Basel), 10(7),* 12–18. DOI 10.3390/atmos10070373.
70. Fan, J. L., Wang, X. K., Zhang, F. C., Ma, X., Wu, L. F. (2020). Predicting daily diffuse horizontal solar radiation in various climatic regions of China using support vector machine and tree-based soft computing models with local and extrinsic climatic data. *Journal of Cleaner Production, 248,* 1–24. DOI 10.1016/j.jclepro.2019.119264.
71. Zheng, H., Wu, Y. (2019). A XGboost model with weather similarity analysis and feature engineering for short-term wind power forecasting. *Applied Sciences, 9(15),* 1–12. DOI 10.3390/app9153019.
72. Deo, R. C., Ghorbani, M. A., Samadianfard, S., Maraseni, T., Bilgili, M. et al. (2018). Multi-layer perceptron hybrid model integrated with the Firefly Optimizer Algorithm for windspeed prediction of target site using a limited set of neighboring reference station data. *Renewable Energy, 116,* 309–323. DOI 10.1016/j.renene.2017.09.078.
73. Ghorbani, M. A., Deo, R. C., Yaseen, Z. M., Kashani, M. H., Mohammadi, B. (2017). Pan evaporation prediction using a hybrid multilayer perceptron-firefly algorithm (MLP-FFA) model: Case study in North Iran. *Theoretical & Applied Climatology, 133(3–4),* 1119–1131. DOI 10.1007/s00704-017-2244-0.