

## Research Article

# Forecasting Daily and Monthly Reference Evapotranspiration in the Aidoghmoush Basin Using Multilayer Perceptron Coupled with Water Wave Optimization

Fatemeh Sayyahi, Saeed Farzin , and Hojat Karami

Department of Water Engineering and Hydraulic Structures, Faculty of Civil Engineering, Semnan University, Semnan 35131-1911, Iran

Correspondence should be addressed to Saeed Farzin; saeed.farzin@semnan.ac.ir

Received 1 November 2020; Revised 20 January 2021; Accepted 2 August 2021; Published 14 August 2021

Academic Editor: Ning Cai

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

The aim of this study is to evaluate the ability of soft computing models including multilayer perceptron- (MLP-) water wave optimization (MLP-WWO), MLP-particle swarm optimization (MLP-PSO), and MLP-genetic algorithm (MLP-GA), to simulate the daily and monthly reference evapotranspiration (ET) at the Aidoghmoush basin (Iran). Principal component analysis (PCA) was used to find the best input combination including the lagged ETs. According to the results, the ET values with 1, 2, and 3 (days) lags as well as those with 1, 2, and 3 (months) lags were the most effective variables in the formation of the PCs. The total variance proportion of inputs and eigenvalues was used to identify the most important variables. The accuracy of the models was assessed based on multiple statistical indices such as the mean absolute error (MAE), Nash–Sutcliffe efficiency (NSE), and percent bias (PBIAS). The results showed that the performance of hybrid MLP models was better than that of the standalone MLP. The findings confirmed that the MLP-WWO could precisely predict ET.

## 1. Introduction

Soft computing models are widely used to solve various problems in water resource management, such as reservoir operation [1, 2], flood routing [3], irrigation management [4], and drug removal modeling [5]. Reference evapotranspiration (ET) is a key parameter for hydrology [6]. The major significance of ET in estimating the water budget has been well proved. The precise estimation of ET is highly important in managing water resources. In fact, it is essential to predict ET with an acceptable level of precision in managing watersheds [7], which can help decision-makers use accurate predictions to ensure the best allocation of water resources among the stakeholders [8]. For example, models used for predicting ET in relatively humid areas may not be suitable for dry basins where the water shortage is a significant challenge [9].

However, ET is measured only at a very limited number of meteorological stations due to the high cost of the necessary equipment. The FAO-56 Penman–Monteith equation, which is widely used as a reference model for ET computation [10], requires a multitude of hydrological data, which is considered as one of its serious drawbacks. However, many hydrological modelers do not have access to accurate measurement devices to record information regularly [11]. Several empirical methods have been proposed to compute ET based on meteorological variables. The radiation-based and temperature-based ET models are widely used due to the restriction in the availability of meteorological data. A large number of studies have shown that the ET temperature-oriented models cannot achieve a high level of accuracy.

A growing body of research has examined the ability of the soft computing models in estimating ET. These computing models usually use nonlinear inputs to predict ET,

while old data-driven ET predictions mostly use statistical methods such as linear regression and autoregressive integrated moving average models [12] and are restricted by the hypothesis that the inputs are linear. Nonlinear models have received growing attention in the management of water resources using soft computing techniques, such as artificial neural network (ANN) and adaptive neuro-fuzzy interface system (ANFIS).

Shiri et al. [13] compared the performance of gene expression programming (GEP) and FAO-56 Penman–Monteith (PM) in Spain. Their results indicated that the GEP model outperformed other models in most cases. In another study, Tabari et al. [14] applied ANFIS and support vector machine (SVM) to calculate potato crop evapotranspiration and found that the ANFIS had a better performance than the SVM model. In addition, Luo et al. [15] applied ANN and PM models to predict the ET and reported the superior performance of ANN in comparison to the PM model. Furthermore, Patil and Chandra Deka [16] examined the ability of the extreme learning machine (ELM) and ANN model in India and found that the ELM model provided better results.

Patil and Chandra Deka [17] applied ANFIS, wavelet ANFIS, and wavelet ANN to predict ET. Their results showed that the ANFIS-wavelet model outperformed other models. Seifi and Riahi [18] investigated the effectiveness of least square support vector machine (LSSVM), ANFIS, and ANN models in predicting ET using meteorological data. The results indicated that LSSVM model provided a more desirable prediction compared to the other models.

In this study, the MLP model was used to predict monthly and daily ET. Although the SVM models can predict hydrological variables, they have some inherent drawbacks [18]. For example, the SVM models cannot accurately estimate the target variables when the number of features is more than the number of the samples, and the input data have more noise [18]. Therefore, choosing a kernel function is not easy for the modelers. The MLP models are widely used to predict hydrological variables because they can handle multivariate inputs and have multistep forecasts. However, the MLP models have some important drawbacks [19]. The training soft computing models are a real challenge for users. The traditional training algorithms such as the backpropagation algorithm of soft computing models may trap in the local optimum. Optimization algorithms are considered as suitable alternatives for traditional training algorithms due to their advanced operators, which avoid trapping in local optimums. The optimization algorithms are widely used for training soft computing models [19–26]. The genetic algorithm (GA) and particle swarm optimizations (PSO) are powerful optimization algorithms. The positions of particles are considered as the candidate solutions. This study combines the WWO and MLP models to increase the convergence speed and accuracy. In addition, the WWO avoids trapping in local optimums by refraction operator, which increases the population diversity. Therefore, the use of this strategy avoids falling into the local optimums.

Water wave optimization (WWO), as an innovative optimization algorithm, has been recently used in various research fields such as the optimal reactive power dispatch, benchmark functions, and traveling salesman problem [27–29]. Previous research has shown that the WWO could increase the convergence speed and computation accuracy compared to PSO, GA, and other algorithms. WWO has several advantages such as a good balance between exploration and exploitation. Furthermore, it uses different operators such as refraction, propagation, and breaking operators to increase population diversity. Furthermore, the advanced operators of the WWO can reduce premature convergence. The main motivation of this paper is to develop the new hybrid MLP models for predicting ET. In addition, the new hybrid models can be used for predicting target variables in other fields. Although the different studies use different metrological data for predicting ET, this study uses the lagged ET values for predicting ET. Thus, this study is useful when the scholars do not have different climate input data for predicting ET.

In this paper, a new optimization algorithm aiming to combine within the MLP model was developed to predict monthly ET. To this end, meteorological data were collected from Aidoghmoush station in Iran. To the best of our knowledge, scanty research has examined the potential of the combination of MLP model and WWO algorithm to predict daily and monthly ET. Regarding the performance of the combined MLP-WWO, the outputs of the combined MLP-WWO are compared with the standalone MLP model, MLP-genetic algorithm (MLP-GA), and MLP-particle swarm optimization (MLP-PSO).

## 2. Materials and Methods

**2.1. MLP Model.** MLPs include a set of neurons placed in layers. Activation functions are used in each node to transform the weighted inputs to an output characteristic of the mathematical properties of the activation function. The MLP in this study was trained using the back propagation algorithm (BPA) [30]. This network includes input, hidden, and output layers. The input data are received in the first layer. Then, the information is processed in the hidden layer. Finally, the model prediction is produced by the output layer. The applied MLP model is based on the following levels:

- (1) The input-out data were accidentally selected by employing the given training input data. Models with different sizes, which had the highest accuracy, were tested for training and testing levels, among which 70 and 30 of the data were selected for training and testing levels, respectively.
- (2) The outputs were generated for some input patterns after being applied to the transfer function.
- (3) An objective function such as root mean square error (RMSE) was selected.
- (4) The connection weight was updated to obtain the lowest RMSE value.

- (5) For the testing and training levels, each pair consisting of the input-output vectors was continued through levels to the extent in which there was no considerable change in the RMSE in the model (Figure 1).

**2.2. Genetic Algorithm.** GA is an evolutionary algorithm which finds the optimal solutions for a problem based on Darwin's principle via mutation, crossover, and selection operators. To this end, several initial solutions are generated and their corresponding objective function values are computed. The selection operation is used to choose the old population parents. Next, the new individuals are generated by the crossover operator [31]. Finally, the mutation operator is used to maintain the diversity between the current and the next generations. The algorithm ends when the stopping criteria are satisfied (Figure 2).

**2.3. PSO Algorithm.** PSO is widely used for different optimization problems. Similar to other optimization algorithms, this algorithm begins with a randomly initialized population of members and uses the social behaviors of the particles to obtain the best solution by setting the position of each member concerning the best position of the particle in the swarm population. Equations (1) and (2) are used to update the velocity and position of the particles [32]:

$$V_i(t+1) = V_i(t) + c_1 \cdot \text{rand} \cdot (l_{\text{best}_i} - X_i) + c_2 \cdot \text{rand} \cdot (g_{\text{best}_i} - X_i), \quad (1)$$

$$X_i(t+1) = X_i(t) + V_i(t+1), \quad (2)$$

where  $X_i(t+1)$  is the position of the iterating particle ( $t+1$ ),  $V_i(t+1)$  denotes the velocity of the iterating particle ( $t+1$ ),  $c_1$  and  $c_2$  represent the constant value ranges 0–2,  $g_{\text{best}_i}$  shows the global best particle,  $l_{\text{best}_i}$  is local best particle, and  $\text{rand}$  denotes the random number, which is between 0 and 1. Figure 3 shows the optimization process of PSO.

**2.4. WWO Algorithm.** The WWO algorithm aims to enhance the exploration and exploitation of the abilities of the optimization algorithms based on propagation, refraction, and braking operators. The WWO mimics the shallow water wave theory. Every agent in the algorithm has some resemblance to the “water wave” entity with a wavelength (WL) of  $\lambda$  and a wave height of  $h$ . As shown in Figure 4, the objective function value of the water wave is considerably lower, its  $h$  is lower, and WL is shorter than in deep water origins [29].

**2.4.1. Propagation.** It is hypothesized that  $x$  represents the original water wave, and  $x'$  denotes a new one generated by the propagation operator (weights and biases of the MLP) [33].

$$x'(d) = x(d) + \text{rand}(-1, 1) \cdot \lambda L(d), \quad (3)$$

where  $L(d)$  is the length of the search space,  $\text{rand}(-1, 1)$  shows a uniformly distributed random number,  $x(d)$  represents the original water wave, and  $x'(d)$  denotes the new water wave. After propagating, the objective function value of  $x'$  is evaluated. Without losing generality,  $F$  is assumed as a minimization problem with fitness functions of  $f$  and  $F$ . The practical problem can be compared with the shallow water model (Figure 5).

If  $f(x') < f(x)$ ,  $x'$  is used instead of  $x$  in the population. Then, the wave height is adjusted to maxim height ( $h_{\text{max}}$ ). On the contrary, the wave  $x$  is fixed, and the wave height is reduced by one to show the loss of energy. Therefore, the following equation could be used to update the wavelength [33]:

$$\lambda = \lambda \cdot \alpha^{-(f(x) - f_{\text{min}} + \epsilon / f_{\text{max}} - f_{\text{min}} + \epsilon)}, \quad (4)$$

where  $\alpha$  is the wavelength reduction,  $f_{\text{max}}$  shows the maximum objective function value,  $f_{\text{min}}$  represents the minimum objective function value, and  $\epsilon$  denotes the positive integer.

**2.4.2. Refraction.** In the optimization process, the refraction operation can only be conducted on the wave  $x$ , which approaches 0 in its height to prevent search stagnation. The position update is given as follows [33]:

$$x'(d) = N\left(\frac{x_d^* + x_d}{2}, \frac{|x^*(d) - x(d)|}{2}\right), \quad (5)$$

where  $x_d^*$  is the optimal solution and  $N(\sigma, \mu)$  shows a Gaussian random number with the mean ( $\mu$ ) of 0 and standard deviation of 1 ( $\sigma$ ). The  $N$  allows the wave  $x_d$  to learn from the best solution  $x_d^*$ .

**2.4.3. Breaking.** The utility of the breaking operator is to make the population diverse.  $K$  dimensions are randomly chosen, and each dimension is selected for providing each dimension of solitary wave  $x'$  [29]:

$$x'(d) = x(d) + N(0, 1) \cdot \beta L(d), \quad (6)$$

where  $\beta$  is the breaking coefficient. If the objective function value of wave  $x^*$  is much better than the provided solitary waves, the  $x^*$  wave is kept. Figure 6 shows the optimization levels for WWO.

**2.5. Optimization Algorithms for Training MLPs.** It is necessary to consider two technical aspects in order to integrate the optimization algorithms with MLP, namely, the method for encoding the agents/solutions and the procedure for determining the objective function. Although the standalone MLP models have high ability, their training algorithms may have slow convergence or may trap in local optimums [21–34]. Therefore, it is essential to improve the accuracy of the MLP models. These models were trained for 1000 epochs in this study. The learning rate and momentum coefficient were 0.001 and 0.09, respectively.

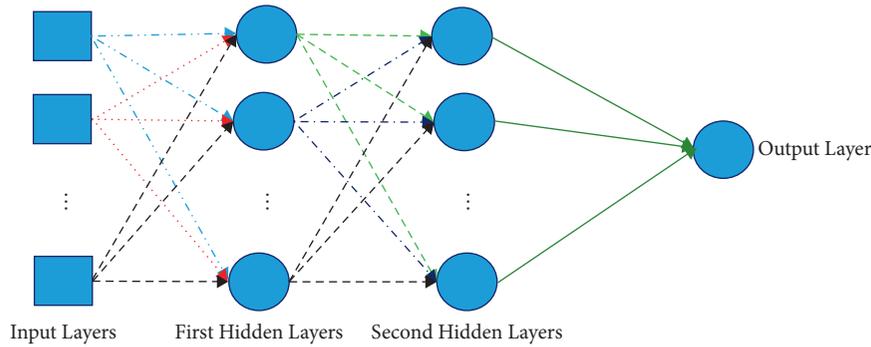


FIGURE 1: The schematic structure of the MLP model.

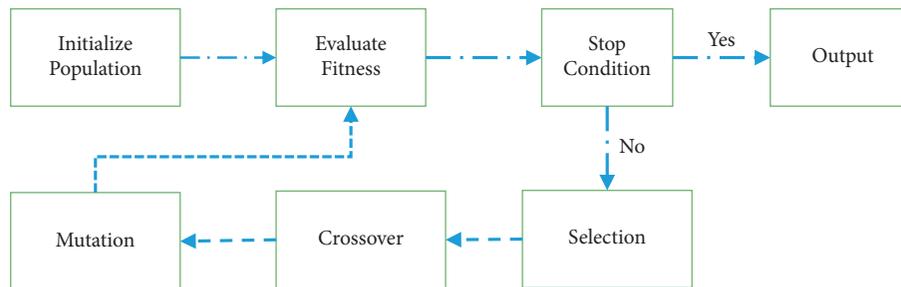


FIGURE 2: GA flowchart.

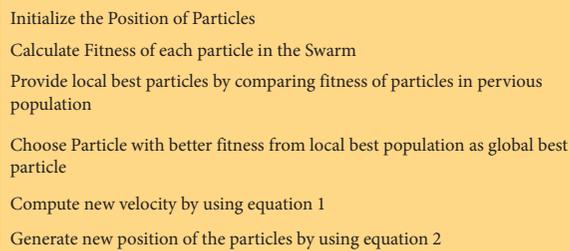


FIGURE 3: The optimization process of PSO.

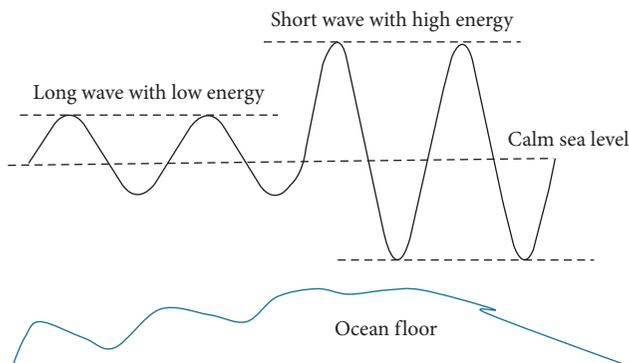


FIGURE 4: The details of WWO.

In evolutionary algorithms-MLP models, every dimensional vector can refer to an agent (e.g., particles (PSO), chromosomes (GA), and 'wave water' objects (WWO)), which may include random numbers in  $[-1, 1]$ . Each agent indicates a candidate MLP (Figure 7). The encoded agents

contain sets of bias values and connection weights. The number of weights and biases determines the length of the vector. To compute the objective function values of the agents, all agents should be transferred to the MLP so that they could be labeled as the connection weights. RMSE is typically the applied objective function in MLP-optimization algorithms. The level of the hybrid model trainer can be explained in the following levels:

- (1) The optimization algorithm-MLP model initiates the random agents
- (2) The agents are mapped to some biases and weight values of a candidate MLP
- (3) The quality of the MLPs is evaluated according to the RMSE
- (4) The optimization algorithm-MLP model constructs the fittest MLP with the minimum RMSE
- (5) The agents are updated
- (6) Steps 2 to 4 are continued until the last iteration (Figure 7)

**2.6. Dataset.** Figure 8 shows the data obtained from the meteorological station located in the Aidoghmosh basin of Iran ( $37^{\circ}16'$  to  $37^{\circ}31'N$ ;  $47^{\circ}33'$  to  $47^{\circ}49'E$ ) during 1987 to 2000.

The average discharge and the average annual rainfall are  $190 \times 10^6 \text{ m}^3$  per year and 340 mm, respectively. The rainy period of the year is from October to May, and August is the driest month. In addition, the peak probability for rainfall is in April, while the rainless period of the year is from June to

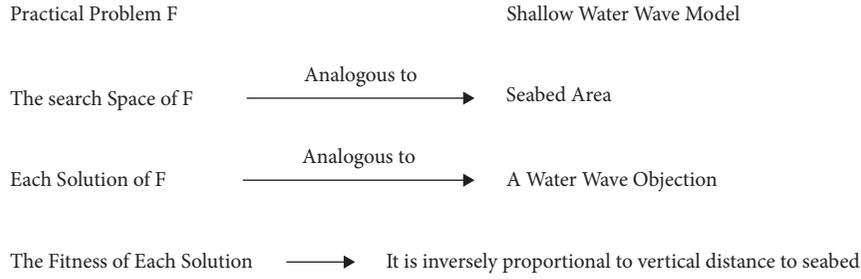


FIGURE 5: The detail of WWO for a practical problem.

- 1- Initialize random parameters and a population  $p$  of NP waves
- 2- While stop criterion is not satisfied do
- 3- For each  $x \in P$  do
- 4- Propagate  $x$  to new  $x'$  based on equation 3
- 5- If  $(x') < f(x^*)$  then
- 6- Break  $x$  based on equation 6
- 7- Update  $x^*$  with  $x'$
- 8- End if
- 9- Replace with  $x'$
- 10- Else
- 11- Decrease  $x.h$  by one
- 12- If  $x.h = 0$  then
- 13- Refract  $x$  to a new  $x'$  based on equation 5
- 14- End if
- 15- End if
- 16- Update the wavelength based on equation 4
- 17- End for
- 18- End while

FIGURE 6: The optimization levels for WWO.

October. The brighter and darker periods are from May to August and October to February, respectively [37].

The MLP, MLP-WWO, MLP-PSO, and MLP-GA were used to forecast the daily and monthly ET. Table 1 shows the tabulation of the statistical characteristics of the input data in this study. The lagged ET values were used to estimate one month or day ahead prediction of ET. The lagged ET values were used because it aimed to evaluate the ability of the models based on the limited input data. In fact, since there was no access to the climate data, the accurate forecasting ET was an important issue.

The data were divided into 70% training and 30% testing. The lagged ET values were used as the input to soft computing models. Different sizes of data were tested, and 70 and 30 achieved the least value of the objective function. Before developing the model, principal component analysis (PCA) was performed on monthly and daily ET values to select the significant lags, that is, the lagged inputs mostly affect daily and monthly ET values. The variables lagged up to 7 days and 7 months were used as the inputs for the soft computing models to predict daily and monthly ET. PCA is a statistical model to transform a given set of  $n$  variables into a new set of PCs which are orthogonal to each other. The PCA was used to choose the best-lagged input variable to predict hydrological variables [37–39]. Previous studies indicated the

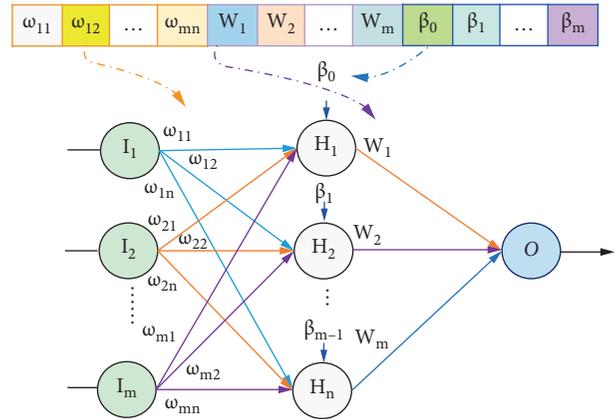


FIGURE 7: Solution structure in MLP-optimization model [35].

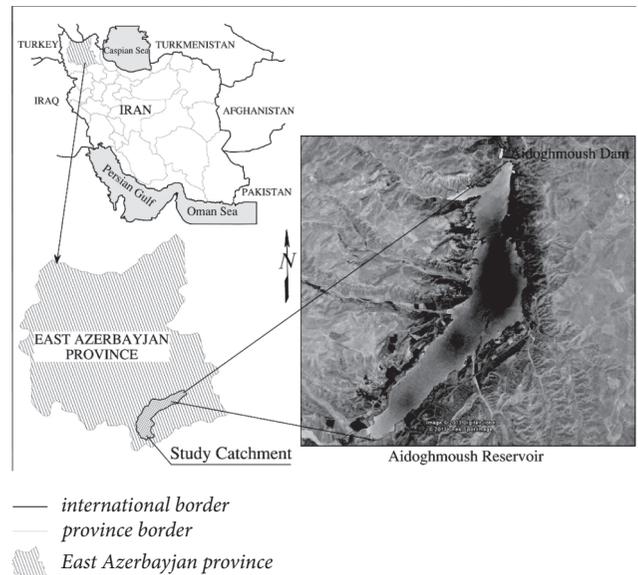


FIGURE 8: Location of the case study [36].

superiority of PCA in comparison to other methods such as gamma test and correlation method [37–39]. The principals with eigenvalues, which are more than one, are considered as significant inputs. In addition, the most effective variables in PCs have a coefficient of  $\geq 0.90$ . Thus, the principals and their variables are selected based on the aforementioned rules [37–39].

TABLE 1: Statistical characteristics for ET: monthly and daily values.

	ET	Maximum (mm)	Minimum (mm)	Coefficient of variation (mm)	Average (mm)
Monthly values	January	18	10	0.23	14
	February	17	12	0.20	15
	March	52	34	0.45	39
	April	89	45	0.56	52
	May	112	90	0.91	104
	June	224	209	1.01	212
	July	232	223	1.04	228
	August	145	138	0.98	142
	September	132	112	1.01	129
	October	98	91	0.65	95
	November	89	82	0.55	85
	December	67	54	0.72	59
Daily values	Value	24	12	0.32	18
	Value	21	14	0.45	19

Preparing the input data, computing the covariance matrix, along with the eigenvalues and vectors, as well as calculating the proportion of total variance for each PC are considered as the main levels of PCA model. The following indexes are used to assess the models.

$$\text{PBIAS} = \frac{\sum_{i=1}^N (Y_i^{\text{obs}} - Y_i^{\text{sim}})}{\sum_{i=1}^N Y_i^{\text{obs}}}, \quad (7)$$

$$\text{NSE} = 1 - \left[ \frac{\sum_{i=1}^n (Y_i^{\text{obs}} - Y_i^{\text{sim}})^2}{\sum_{i=1}^n (Y_i^{\text{obs}} - Y^{\text{mean}})^2} \right], \quad (8)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Y_i^{\text{obs}} - Y_i^{\text{sim}}|, \quad (9)$$

where MAE is the mean absolute error, NSE shows Nash–Sutcliff efficiency,  $Y_i^{\text{obs}}$  indicates the observed data,  $Y^{\text{mean}}$  represents the average of the observed data, and  $Y_i^{\text{sim}}$  denotes the forecasted data.

### 3. Results and Discussion

#### 3.1. Selection of the Inputs

**3.1.1. Daily Scale.** Table 2 shows the outputs of the PCA including the contribution of seven inputs to seven PCs, the described variance of each PC, and the cumulative sum of the described variance. As shown, the PC1 and the first three PCs contributed to 54% and 91% of the total variance, respectively. The results indicated that the ET values with lags 1, 2, and 3 (days) were significant and used as the inputs.

**3.1.2. Monthly Scale.** Based on the results, the PC1 and the first three PCs contributed to 53% and 94% of the total variance, respectively. In addition, the ET values with lags 1, 2, and 3 (month) were significant and used as the inputs to the soft computing models (Table 2).

#### 3.2. Results of the Data Analyses for Soft Computing Models

**3.2.1. Daily Scale.** The random parameters of optimization algorithms are considered as influential coefficients which significantly affect the performance of the algorithms. For example, the random parameters of WWO are the population size, wavelength, wave height, and breaking coefficients. Table 3 shows other important variables. Each MLP-optimization algorithm model was tested by different values of the random parameters. A sensitivity analysis was used to find the optimal value of the parameters. The variation of the objective function was computed versus the variation of interest of parameter. When a value of the interest of parameters varies, the other parameters are fixed. The best values of the parameters minimize the objective function.

The objective function was used to compute the error of performance at the end of each iteration. Each model was run within 2000 iterations to optimize the MLP parameters. Therefore, a series MLP-optimization algorithms was conducted using various population sizes (PS), wave lengths (WL), wave heights (WH), and breaking coefficients (BC) ranging from 100 to 400, 0.1 to 1, 1 to 5, and 0.01 to 0.10, respectively. The obtained results showed that the PS: 200, WL: 0.5, WH: 2, and BC: 0.05 provided the lowest RMSE. Therefore, these values were selected as the optimum parameters in the MLP-WWO model. Regarding a similar process, the optimal parameters of the other algorithms were obtained.

Table 4 demonstrates the outputs related to the training level of the MLP models. The results shown in Tables 4 and 5 are based on the best value of random parameters obtained in Table 3.

Based on Table 4, the hybrid MLP has better and more acceptable results for modeling. In addition, the MAE of the MLP-WWO was 2.1%, 3.2%, and 4.1% lower than those of the MLP-PSO, MLP-GA, and MLP models. The PBIAS of the MLP-WWO was 0.14, while it was 0.35, 0.37, and 0.39 for the MLP-PSO, MLP-GA, and MLP models. The MLP-WWO has the highest NSE and the lowest PBIAS among other

TABLE 2: The daily scale of outputs in the PCA.

	Input	PC (1)	PC (2)	PC (3)	PC (4)	PC (5)	PC (6)	PC (7)
Daily scale	$E(d-1)$	0.94	0.93	0.92	0.62	0.55	0.43	0.42
	$E(d-2)$	0.90	0.92	0.91	0.59	0.54	0.47	0.42
	$E(d-3)$	0.89	0.91	0.90	0.55	0.52	0.40	0.39
	$E(d-4)$	0.65	0.62	0.69	0.49	0.58	0.47	0.42
	$E(d-5)$	0.62	0.45	0.42	0.42	0.45	0.40	0.39
	$E(d-6)$	0.54	0.34	0.31	0.29	0.41	0.36	0.35
	$E(d-7)$	0.43	0.29	0.27	0.25	0.32	0.30	0.29
	Explained variance %	54	22	17	3	3	0.5	0.5
	Eigen value	3.24	1.32	1.02	0.18	0.18	0.03	0.03
	Cumulative variance %	54	76	91	96	99	100	100
Monthly scale	$E(m-1)$	0.95	0.94	0.92	0.76	0.61	0.60	
	$E(m-2)$	0.94	0.92	0.91	0.69	0.55	0.52	
	$E(m-3)$	0.93	0.92	0.90	0.67	0.44	0.41	
	$E(m-4)$	0.85	0.84	0.80	0.65	0.42	0.39	
	$E(m-5)$	0.82	0.81	0.79	0.62	0.40	0.34	
	$E(m-6)$	0.80	0.79	0.77	0.55	0.38	0.36	
	Explained variance (%)	53	27	17	2	0.5	0.5	
	Eigen vale	3.18	1.62	1.02	0.12	0.03	0.03	
	Cumulative variance %	53	80	94	99	99.5	100	

TABLE 3: The optimal value of the random parameters.

WWO, population size: 200, wave length: 0.3, wave height: 0.5, and BC: 0.05
PSO, number of particles: 200, $c_1$ : 2, $c_2$ : 2, and $w$ : 0.4
GA, number of chromosomes: 100, mutation probability: 0.01, and crossover probability: 0.3

TABLE 4: Training level for the statistical results related to soft computing models based on the daily scale.

Index	MAE (mm)	PBIAS	NSE
Training level			
MLP-WWO	3.12	0.14	0.94
MLP-PSO	3.19	0.35	0.91
MLP-GA	3.22	0.37	0.89
MLP	3.25	0.39	0.85
Testing level			
MLP-WWO	5.15	0.17	0.92
MLP-PSO	5.22	0.47	0.90
MLP-GA	5.28	0.49	0.87
MLP	5.32	0.51	0.84

TABLE 5: The training level for the statistical results related to the soft computing models based on the monthly scale.

Index	MAE (mm)	PBIAS	NSE
Training level			
MLP-WWO	0.49	0.19	0.95
MLP-PSO	0.53	0.32	0.92
MLP-GA	0.57	0.33	0.90
MLP	0.59	0.35	0.88
Testing level			
MLP-WWO	0.51	0.21	0.93
MLP-PSO	0.55	0.34	0.90
MLP-GA	0.58	0.37	0.87
MLP	0.61	0.39	0.85

models. The PBIAS of the MLP model is higher than those of the other models.

Table 4 shows the error indexes for soft computing models based on the daily scale at the testing level. As seen in this table, the MAE of the MLP-WWO is 1.3%, 2.5%, and 3.3% lower than those of the MLP-PSO, MLP-GA, and MLP models. Furthermore, the PBIAS of the MLP-GA is higher than those of the other hybrid models. The NSE of the MLP is 0.84, while it was 0.87, 0.90, and 0.92 for MLP-GA, MLP-PSO, and MLP-WWO models.

**3.2.2. Monthly Scale.** Based on Table 5, the hybrid MLP has better and more acceptable testing results for modeling. The PBIAS of the MLP-WWO is lower than those of the other hybrid models. The MAE of the MLP-WWO was 7.2%, 14%, and 17% lower than those of the MLP-PSO, MLP-GA, and MLP models. The highest NSE and lowest PBIAS were obtained for the MLP-WWO model.

Table 5 shows the results related to the testing level of the MLP models. As shown, the MAEs of the WWO are 7.2, 12, and 17% lower than those of the MLP-PSO, MLP-GA, and MLP models, respectively. The NSE of the MLP-PSO is higher than that of the MLP-GA and MLP models.

The results indicated that the optimization algorithms improved the accuracy of the standalone MLP model. In other words, the combination of the hybrid MLP model and preprocessing method such as PCA could be used for practical problems with different input scenarios in water resource management. The modelers may encounter with a large number of input data in estimating different hydrological variables. In addition, the standalone soft computing models may not lead to good results only. Thus, it is essential to use a hybrid framework based on preprocessing methods and optimization algorithms to ensure accurate estimations of the target variables.

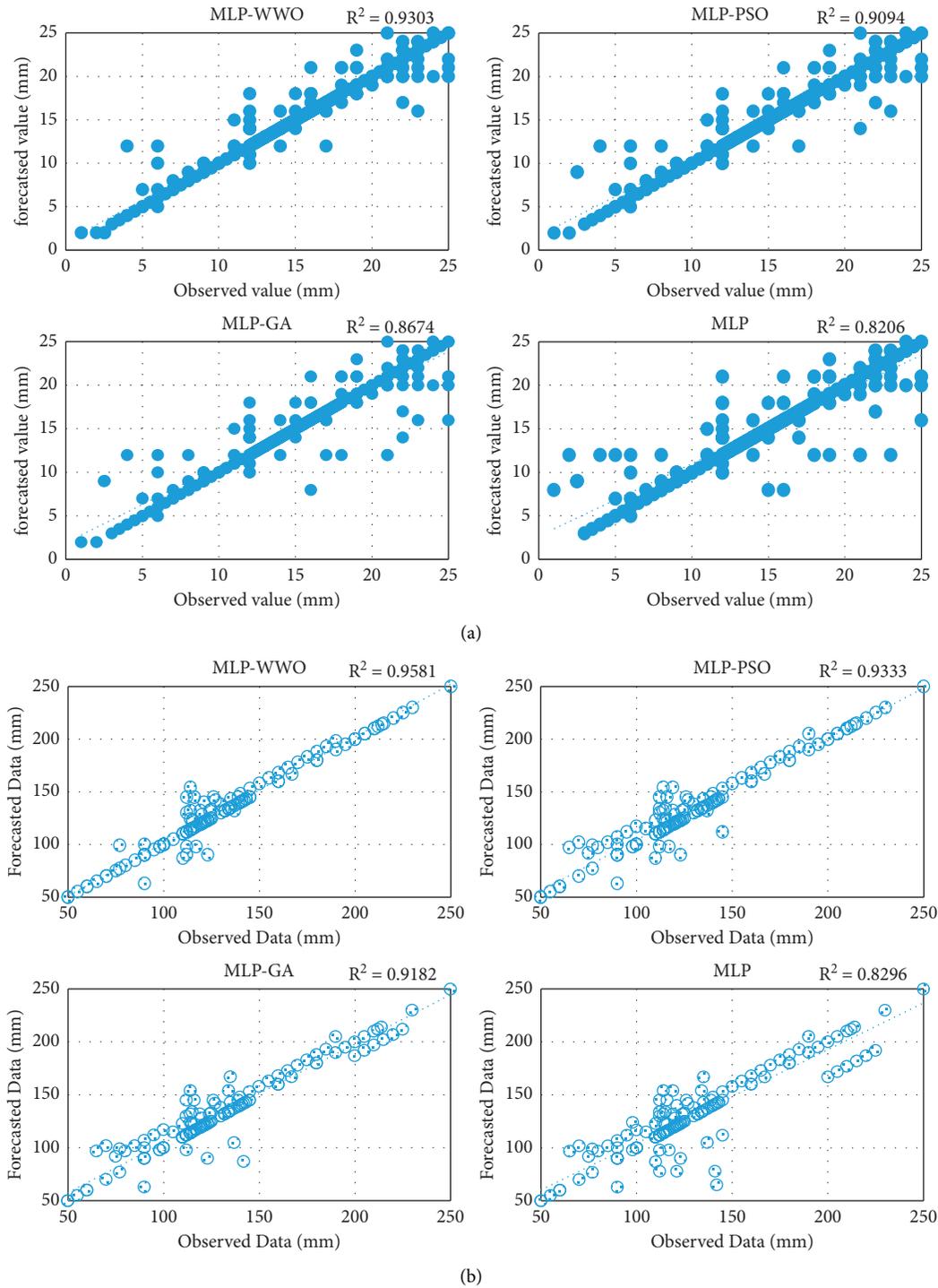


FIGURE 9: Scatter plots for (a) daily values and (b) monthly values.

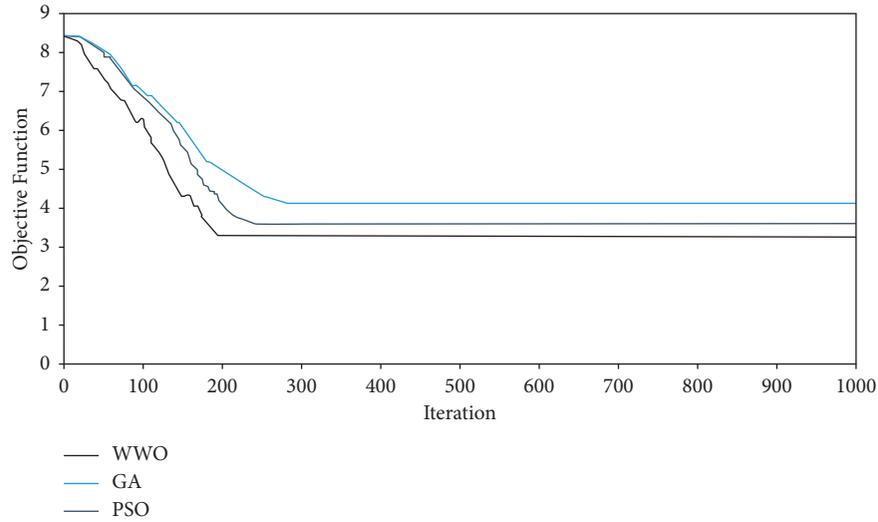
TABLE 6: Classification of hydrological model [40].

Goodness of fit	The accuracy of models
$0.90 \leq NSE \leq 1$	Very good
$0.80 \leq NSE \leq 0.90$	Good
$0.65 \leq NSE \leq 0.80$	Satisfactory
$NSE \leq 0.65$	Unsatisfactory

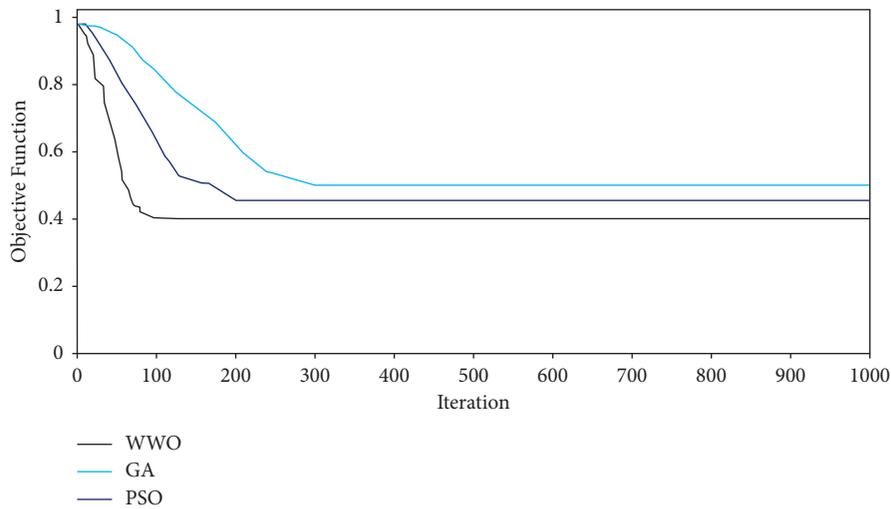
### 3.3. Scatter Plots

3.3.1. *Daily Scale.* Figure 9 shows the scatter plots for the soft computing models.

The results indicated that the outputs related to the MLP-WWO were closer to the observed data indicating the



(a)



(b)

FIGURE 10: Convergence curve of algorithms for (a) daily scale and (b) monthly scale.

accurate performance of the combined model. Figure 9(a) demonstrates that the MLP model has the worst performance among the MLP models. Furthermore, as displayed in Figure 9(a), the MLP-PSO scatter points are closer to the 45° line in comparison to the MLP-GA scattered points.

**3.3.2. Monthly Scale.** The computed  $R^2$  for soft computing models indicated that the MLP-WWO had the best performance as compared to the other models, and the PSO outperformed the GA. To sum up, the hybrid MLP models outperformed the standalone MLP model (Figure 9(b)).

**3.4. Probability Distribution of NSE.** The training data were randomly sampled  $M$  times with replacement to build a model and evaluate its NSE for each resample.  $M$  trained models were used to compute the NSE based on the validation data. This approach was used to perform the goodness of fit of predicted data and observed data.

The procedure may require high computational time, depending on the number of patterns. After approximating the probability distribution of NSE, its significance was evaluated based on the 95% confidence interval (CI) (Table 6). The results were analyzed to predict the daily and monthly ET as follows:

**3.4.1. Daily Scenario (NSE).** Figure 10(a) displays how the probability that  $NSE > 0.80$  is as high as 93% in the MLP-WWO model. Thus, the MLP-GA model did not achieve 0.90 NSE and decreased to 0.50 to 0.89 (Figure 11(a)). The obtained results for the MLP-PSO show that more than 60% of the CIs are above 0.80 NSE (Figure 11(a)). Based on the results, the MLP-PSO is better than the MLP-GA and MLP-WWO.

**3.4.2. Monthly Scenario (NSE).** Regarding the MLP-WWO, the results indicated that more than 93% of the CIs are above

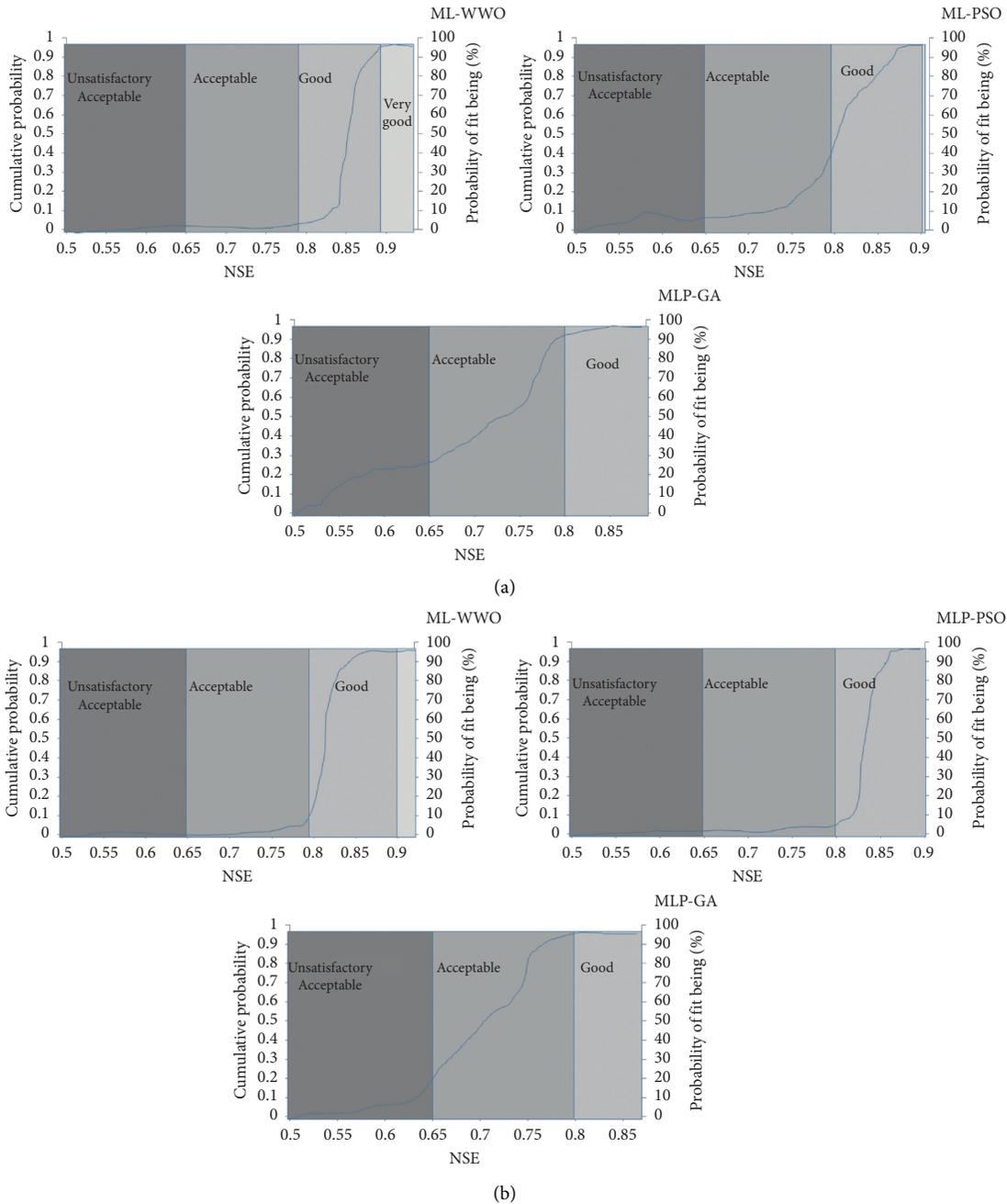


FIGURE 11: The probability distribution for NSE: (a) daily ET and (b) monthly ET.

0.80 NSE (Figure 11(b)). Therefore, the MLP-GA model failed to reach 0.90 NSE and decreased to 0.5 to 0.87 (Figure 11(b)).

Based on Figure 10 showing the convergence curves, the WWO converted earlier than other methods.

#### 4. Conclusion

In this study, the soft computing models were used as the input for the soft computing models. The forecasting models were generated using the lagged ET values for

Aidoghmoush basin (Iran). Then, the outputs of the soft computing models were compared, which indicated that the MLP-WWO outperformed the other MLP models. In addition, the MAE of the MLP-WWO is 1.3%, 2.5%, and 3.3% lower than those of the MLP-PSO, MLP-GA, and MLP models in the daily scale models. The MAE of the MLP-WWO was 7.2%, 14%, and 17% lower than those of the MLP-PSO, MLP-GA, and MLP models in the monthly scale. In addition, the outputs related to the MLP-WWO were closer to the observed data. Finally, it must be stated that the appropriate optimization algorithms affect the accuracy of standalone soft computing models. Thus, the

selection of a robust optimization algorithm is an important issue for developing the soft computing models. Future investigations can develop the performance of the models of this study. The proposed models can be used for predicting other hydrological variables such as rainfall, temperature, and runoff. Furthermore, the next studies can investigate the effect of uncertainty on the accuracy of the models.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon reasonable request.

## Conflicts of Interest

The authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest in the subject matter in this article.

## References

- [1] V. Ramaswamy and F. Saleh, "Ensemble based forecasting and optimization framework to optimize releases from water supply reservoirs for flood control," *Water Resources Management*, vol. 34, pp. 989–1004, 2020.
- [2] M. Mohammadi, S. Farzin, S. Farhad Mousavi, and H. Karami, "Investigation of a new hybrid optimization algorithm performance in the optimal operation of multi-reservoir benchmark systems," *Water Resources Management*, vol. 33, pp. 4767–4782, 2019.
- [3] A. Pashazadeh and M. Javan, "Comparison of the gene expression programming, artificial neural network (ANN), and equivalent Muskingum inflow models in the flood routing of multiple branched rivers," *Theoretical and Applied Climatology*, vol. 139, no. 3-4, pp. 1349–1362, 2020.
- [4] F. Banadkooki, M. Ehteram, A. N. Ahmed et al., "Enhancement of groundwater-level prediction using an integrated machine learning model optimized by whale algorithm," *Natural Resources Research*, vol. 29, pp. 3233–3252, 2020.
- [5] S. Farzin, N. C. Farideh, M. V. Anaraki, and F. Mahmoudian, "Introducing a framework for modeling of drug electrochemical removal from wastewater based on data mining algorithms, scatter interpolation method, and multi criteria decision analysis (DID)," *Journal of Cleaner Production*, vol. 266, 2020.
- [6] G. M. Kovoor and L. Nandagiri, "Sensitivity analysis of FAO-56 penman-monteith reference evapotranspiration estimates using Monte Carlo simulations," in *Hydrologic Modeling. Water Science and Technology Library*, V. Singh, S. Yadav, and R. Yadava, Eds., Springer, Singapore, 2018.
- [7] G. Huang, L. Wu, X. Ma et al., "Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions," *Journal of Hydrology*, vol. 574, pp. 1029–1041, 2019.
- [8] B. A. Kimball and K. J. Boote, "Simulation of maize evapotranspiration: an inter-comparison among 29 maize models," *Agricultural and Forest Meteorology*, vol. 271, pp. 264–284, 2019.
- [9] J. Shiri, "Improving the performance of the mass transfer-based reference evapotranspiration estimation approaches through a coupled wavelet-random forest methodology," *Journal of Hydrology*, vol. 561, pp. 737–750, 2018.
- [10] H. Farzanpour, J. Shiri, A. A. Sadraddini, and S. Trajkovic, "Global comparison of 20 reference evapotranspiration equations in a semi-arid region of Iran," *Hydrology Research*, vol. 50, no. 1, pp. 282–300, 2019.
- [11] J. Shiri, "Modeling reference evapotranspiration in island environments: assessing the practical implications," *Journal of Hydrology*, vol. 570, pp. 265–280, 2019.
- [12] M. K. Saggi and S. Jain, "Reference evapotranspiration estimation and modeling of the Punjab northern India using deep learning," *Computers and Electronics in Agriculture*, vol. 156, pp. 387–398, 2019.
- [13] J. Shiri, Ö. Kişi, G. Landeras et al., "Daily reference evapotranspiration modeling by using genetic programming approach in the Basque country (northern Spain)," *Journal of Hydrology*, vol. 415, pp. 302–316, 2012.
- [14] H. Tabari, C. Martinez, A. Ezani, and P. Hosseinzadeh Talaei, "Applicability of support vector machines and adaptive neurofuzzy inference system for modeling potato crop evapotranspiration," *Irrigation Science*, vol. 31, pp. 575–588, 2013.
- [15] Y. Luo, S. Traore, X. Lyu et al., "Medium range daily reference evapotranspiration forecasting by using ann and public weather forecasts," *Water Resources Management*, vol. 29, 2015.
- [16] A. P. Patil and P. Chandra Deka, "An Extreme learning machine approach for modeling evapotranspiration using extrinsic inputs," *Computers and Electronics in Agriculture*, vol. 121, 2016.
- [17] A. P. Patil and P. Chandra Deka, "Performance evaluation of hybrid wavelet-ANN and wavelet-ANFIS models for estimating evapotranspiration in arid regions of India," *Neural Computing & Applications*, vol. 28, 2017.
- [18] A. Seifi and H. Riahi, "Estimating daily reference evapotranspiration using hybrid gamma test-least square support vector machine, gamma test-ann, and gamma test-anfis models in an arid area of Iran," *Journal of Water and Climate Change*, vol. 11, 2020.
- [19] M. A. Ghorbani, R. Kazempour, K. W. Chau, S. Shamsirband, and T. G. Pezhman, "Forecasting Pan evaporation with an integrated artificial neural network quantum-behaved particle swarm optimization model: a case study in Talesh, Northern Iran," *Engineering Applications of Computational Fluid Mechanics*, vol. 12, no. 1, pp. 724–737, 2018.
- [20] Y. Guan, B. Mohammadi, Q. B. Pham et al., "A novel approach for predicting daily pan evaporation in the coastal regions of Iran using support vector regression coupled with krill herd algorithm model," *Theoretical and Applied Climatology*, vol. 142, 2020.
- [21] S. Mohamadi, M. Ehteram, and A. El-Shafie, "Accuracy enhancement for monthly evaporation predicting model utilizing evolutionary machine learning methods," *International journal of Environmental Science and Technology*, vol. 17, pp. 1–24, 2020.
- [22] F. B. Banadkooki, M. Ehteram, A. N. Ahmed et al., "Suspended sediment load prediction using artificial neural network and ant lion optimization algorithm," *Environmental Science and Pollution Research*, vol. 27, no. 30, pp. 38094–38116, 2020.
- [23] M. Ehteram, S. Q. Salih, and Z. M. Yaseen, "Efficiency evaluation of reverse osmosis desalination plant using hybridized multilayer perceptron with particle swarm

- optimization,” *Environmental Science and Pollution Research*, vol. 27, pp. 5278–15291, 2020.
- [24] S. Mohamadi, S. Sh Sammen, F. Panahi et al., “Zoning map for drought prediction using integrated machine learning models with a nomadic people optimization algorithm,” *Natural Hazards*, vol. 104, pp. 537–579, 2020.
- [25] M. Ehteram, F. Y. Teo, A. N. Ahmed et al., “Performance improvement for infiltration rate prediction using hybridized adaptive neuro-fuzzy inferences system (ANFIS) with optimization algorithms,” *Ain Shams Engineering Journal*, vol. 12, pp. 1665–1676, 2020.
- [26] H. Nguyen, H. Moayedi, W. A. W. Jusoh, and A. Sharifi, “Proposing a novel predictive technique using M5Rules-PSO model estimating cooling load in energy-efficient building system,” *Engineering with Computers*, vol. 35, 2020.
- [27] Y. Zhou, J. Zhang, X. Yang, and Y. Ling, “Optimal reactive power dispatch using water wave optimization algorithm,” *Operational Research*, vol. 20, pp. 1–17, 2018.
- [28] A. Hematabadi and A. A. Foroud, “Optimizing the multi-objective bidding strategy using min–max technique and modified water wave optimization method,” *Neural Computing & Applications*, vol. 31, 2019.
- [29] A. Liu, L. Peng, W. Sun et al., “Prediction of mechanical properties of micro-alloyed steels via neural networks learned by water wave optimization,” *Neural Computing & Applications*, vol. 32, 2020.
- [30] Y. Han, J. Wu, B. Zhai et al., “Coupling a bat algorithm with XGBoost to estimate reference evapotranspiration in the arid and semiarid regions of China,” *Advances in Meteorology*, vol. 2019, Article ID 9575782, 2019.
- [31] F. Alasali, H. Stephen, and W. Holderbaum, “Stochastic optimal energy management system for RTG cranes network using genetic algorithm and ensemble forecasts,” *Journal of Energy Storage*, vol. 24, 2019.
- [32] M. Li, H. Chen, X. Wang, N. Zhong, and S. Lu, “An improved particle swarm optimization algorithm with adaptive inertia weights,” *International Journal of Information Technology and Decision Making*, vol. 18, no. 03, pp. 833–866, 2019.
- [33] Y.-J. Zheng, “Water wave optimization: a new nature-inspired metaheuristic,” *Computers & Operations Research*, vol. 55, pp. 1–11, 2015.
- [34] A. Ashrafzadeh, M. A. Ghorbani, S. M. Biazar, and Z. M. Yaseen, “Evaporation process modelling over northern Iran: application of an integrative data-intelligence model with the krill herd optimization algorithm,” *Hydrological Sciences Journal*, vol. 64, 2019.
- [35] A. A. Heidari, Y. Yin, M. Mafarja, S. M. J. Jalali, J. S. Dong, and S. Mirjalili, “Efficient moth-flame-based neuroevolution models,” *Algorithms for Intelligent Systems. In Evolutionary Machine Learning Techniques*, Springer, Singapore, 2020.
- [36] P.-S. Ashofteh, O. Bozorg-Haddad, and H. A. Loáiciga, “Logical genetic programming (LGP) development for irrigation water supply hedging under climate change conditions,” *Irrigation and Drainage*, vol. 66, no. 4, pp. 530–541, 2017.
- [37] R. Noori, A. R. Karbassi, A. Moghaddamnia et al., “Assessment of input variables determination on the SVM model performance using PCA, gamma test, and forward selection techniques for monthly stream flow prediction,” *Journal of Hydrology*, vol. 401, 2011.
- [38] R. Noori, A. Khakpour, B. Omidvar, and A. Farokhnia, “Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic,” *Expert Systems with Applications*, vol. 37, 2010.
- [39] J. M. Kim and J. Lee, “Time series analysis for evaluating hydrological responses of pore-water pressure to rainfall in a slope,” *Hydrological Sciences Journal*, vol. 62, 2017.
- [40] D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith, “Model evaluation guidelines for systematic quantification of accuracy in watershed simulations,” *Transactions of the ASABE*, vol. 50, no. 3, pp. 885–900, 2007.