Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 7001710, 19 pages
https://doi.org/10.1155/2021/7001710

Research Article

An Extra Tree Regression Model for Discharge Coefficient Prediction: Novel, Practical Applications in the Hydraulic Sector and Future Research Directions

Mohammed Majeed Hameed,[1] **Mohamed Khalid AlOmar**,[1] **Faidhalrahman Khaleel**,[1] **and Nadhir Al-Ansari**[2]

[1]*Department of Civil Engineering, Al-Maaref University College, Ramadi, Iraq*
[2]*Civil Engineering Department, Environmental and Natural Resources Engineering, Lulea University of Technology, Lulea 97187, Sweden*

Correspondence should be addressed to Mohamed Khalid AlOmar; mohd.alomar@yahoo.com and Nadhir Al-Ansari; nadhir.alansari@ltu.se

Received 28 May 2021; Revised 13 August 2021; Accepted 9 September 2021; Published 21 September 2021

Academic Editor: Danial Armaghani

Despite modern advances used to estimate the discharge coefficient (C_d), it is still a major challenge for hydraulic engineers to accurately determine C_d for side weirs. In this study, extra tree regression (ETR) was used to predict the C_d of rectangular sharp-crested side weirs depending on hydraulic and geometrical parameters. The prediction capacity of the ETR model was validated with two predictive models, namely, extreme learning machine (ELM) and random forest (RF). The quantitative assessment revealed that the ETR model achieved the highest accuracy in the predictions compared to other applied models, and also, it exhibited excellent agreement between measured and predicted C_d (correlation coefficient is 0.9603). Moreover, the ETR achieved 6.73% and 22.96% higher prediction accuracy in terms of root mean square error in comparison to ELM and RF, respectively. Furthermore, the performed sensitivity analysis shows that the geometrical parameter such as b/B has the most influence on C_d. Overall, the proposed model (ETR) is found to be a suitable, practical, and qualified computer-aid technology for C_d modeling that may contribute to enhance the basic knowledge of hydraulic considerations.

1. Introduction

Weirs are one of the most common hydraulic structures used for water engineering projects such as irrigation, hydropower systems, drainage, and sewage networks. Weirs transit floods in reservoirs of dams, redirect water from channels, and can be utilized as a device for measuring the discharge in channels. The safety of water-transferring channels and dams is strongly linked to the weir's capacity [1, 2]. One of the main types of weirs is called the side weirs, which is a hydraulic structure constructed on the channel side and used as a protective system by discharging the excessive flow and controlling the level of water. Considering different channels, and different flow characteristics, the relationship between the hydraulic behavior and the coefficient of discharge (C_d) in weirs has been well investigated in the literature [3–7]. The researchers have carried out various analytical and experimental approaches regarding the flow characteristics of side weirs. The analytical approach involves using a mathematical model to obtain new equations that require empirical coefficients taking into consideration the other effective parameters. On the other hand, the experimental approach used multiple empirical relationships to present the flow characteristics in side weirs. Lastly, artificial intelligence (AI) has shown a great potential to be a reliable alternative for the experimental and empirical approaches due to its ability to capture the nonlinear behavior and extract the most effective parameters that are used to predict C_d. In what follows we discuss some of the main works in the literature on the above approaches.

1.1. Related Works. Singh et al. [8] found that using a multiple regression approach under subcritical flow the coefficient of discharge is depending on the Froud number (F_1), and the ratio between sill height (s) and depth of upstream flow (y_1). Borghei et al. [9] found that the De-Marchi coefficient of discharge is a function of Froud number (F_1), the ratio between weir height (w) and the depth of upstream (y_1), and the ratio between weir length (L) and channel width (B). Cheong et al. [7] found that, after a laboratory research, C_d is a function of Froud number (F_1) under subcritical flow in the trapezoidal channel.

However, there are certain limitations both experimentally and mathematically, like high-cost laboratory equipment, which needs to be used to obtain the properties of the side weir [10, 11], lack in the accuracy of the results, and the limited number of tests that depends on the empirical relationships [12]. Nevertheless, the mathematical models used in the analytical approach are unable to predict the coefficient of discharge with the required and desired accuracy. Due to these limitations, the researchers were encouraged to approach this issue in numerical methods [10, 11]. Furthermore, one of the attractive numerical methods is artificial intelligence and soft computing approaches, which are used to model complex phenomena in different sectors like, hydraulic, hydrology, and the water management science [13–15]. The approach of soft computing achieves a great performance compared to multiple nonlinear equations systems-based models and classical approaches [16–18].

Ebtehaj et al. [19] used Group Method of Data Handling (GMDH) by taking Froud number (F_1), the dimensionless ratio of weir length (L) to channel width (B), the ratio of weir length (L) to the depth of upstream (y_1), and the ratio of weir height (w) to weir length (L) as input parameters to establish a predictive model for estimating the C_d of the side weir. The results of the suggested model were more precise in the prediction of C_d than conventional nonlinear regression equations. Azimi et al. predicted the C_d in a trapezoidal channel by using Gene Expression Programming (GEP) and obtained satisfactory accuracy [20]. Furthermore, Ebtehaj et al. [21] used a hybrid approach that contains Group Method of Data Handling (GMDH) and Adaptive Neuro-Fuzzy Inference System (ANFIS) merged with Genetic Algorithm (GA) and Singular Value Decomposition (SVD) to predict the C_d in the side weir. The dimensionless ratio of weir length, Froud number (F_1), the ratio of weir length (L) to the depth of upstream (y_1), and the ratio of weir height (w) to weir length (L) were taken as input parameters in both models. The reviewed studies have demonstrated that the Artificial intelligence (AI) models could accurately provide a desirable prediction and show a high ability to select the most significant parameters that have a major impact on the discharge coefficient. Moreover, two AI modeling approaches called ANN and ANFIS have been employed to predict the discharge coefficient of triangular labyrinth weirs [12]. The study found that ANFIS presents more accurate estimates than ANN. However, the study faced a limitation regarding the proper selection of input variables; therefore, gamma test (GT) was used to address this issue. Azimi et al. [22] presented a study for predicting the coefficient of discharge of a side weir in a trapezoidal channel using support vector regression (SVR) model. The study used different parameters, including geometrical and hydraulic parameters for developing the SVR model. The study also declared that the ratio of side weir length to trapezoidal channel bottom width (L/b) parameter has a major impact on the discharge coefficient. Furthermore, Salmasi et al. [23] developed six AI models, namely, ANN, SVR, random forest (RF), Gaussian process (GP), generalized regression neural network (GRNN), and random forest tree (RT), to estimate the discharge coefficient of oblique sluice gates. Among the prediction models, the ANN model was noticed to provide excellent estimates with less prediction error.

While AI modeling approaches were adopted, a range of limitations such as the tuning of hyperparameters for the model, case studies stochasticity, and model stability have been observed [24]. However, machine learning models such as extreme learning machine (ELM), random forest (RF), and extra tree regression (ETR) provide a way to overcome these limitations and have recently become increasingly popular [25–29]. Machine learning models essentially build on data extraction and pattern recognition between data, through the development of algorithms using a dataset subset known as training data and the prediction accuracy verification using the separate dataset subset known as the testing set.

1.2. Motivation and Contribution. The accurate estimation of C_d is considered a problematic issue to the hydraulic engineers. However, the complexity of hydraulic properties of the weirs, variability in flow characteristics, and stochasticity of discharge urge the researchers to apply novel means to address this issue in order to increase the accuracy of estimations. Besides, it is critical to assess the most important parameters that have a major impact on the discharge coefficient. To the best of our knowledge, few studies used AI approaches to estimate the discharge coefficient of rectangular side weir.

The main objective of this research is to develop three robust AI approaches for estimating the discharge coefficient (C_d), namely, random forest (RF), extreme learning machine (ELM), and extra tree regression (ETR). Additionally, the ETR is a developed approach of the RF method, which is used to address the generalization (overfitting) problem associated with the RF. The ETR model as a novel and advanced tool is used for the first time to address the hydraulic issue according to researchers' knowledge. The models have been developed using hydraulic parameters related to discharge properties and other geometrical characteristics. The models were assessed using different statistical criteria (error metrics and agreement measures) and graphical presentations such as boxplot, line-graph, Taylor diagram, violin plot, and histogram. Lastly, sensitivity analysis has been conducted for the selection of the most influential parameters that affect the C_d.

2. Methodology

2.1. Data Collection and Description. The dataset used in this study involves geometrical parameters of the rectangular sharp-crested side weirs and other parameters related to the discharge characteristics. There were 84 samples collected from different studies in the literature [28, 29]. Emiroglu et al. [30] performed their experiments in the hydraulic laboratory using a rectangular channel with a length, depth, width, and gradient of 12 m, 0.5 m, 0.5 m, and 0.01, respectively. In order to provide a circular free-surface condition, the collection channel is made with a 1.3 m width. The side weirs were manufactured from steel plates with very sharp crests, and they were placed and aerated at the same level as the main channel side. It is worth mentioning that the pipe and the sluice gate were used for regulating and controlling the water. According to Emiroglu et al. study, the applied discharge between 0.01 and 0.15 m^3/sec was determined by using an electromagnetic flow meter. Then, the researchers calibrated the discharge (Q) results of the electromagnetic flow meter by utilizing a V-notched weir placed at the beginning of the system. The Q passing over the side weir has been eventually calibrated via standard rectangular weir, which is placed downstream from the discharge collection channel.

Bagheri et al. [31] carried out several experiments including rectangular sharp-crested weirs for different widths and heights, within a horizontal rectangular channel with 0.40 m width, 8 m long, and 0.60 m depth. It is important to mention that all experiments were conducted under subcritical flow conditions. The profiles of the free surface have been computed along the central axis of the channel and along the side of the weir using utilizing points meter (±0.5 mm accuracy) placed on a mobile carriage. Furthermore, the electromagnetic flowmeter has been used for measuring upstream discharge. The slice gate has been calibrated at the maximum value of weir error within ±5 percent used for controlling the downstream discharge and the flow depth. Moreover, the diverted flow from the weirs was measured by the difference between the upstream and downstream discharge.

The statistical descriptions of all variables are reported in Table 1, where the C_d, F_1, $p/y_1, b/y_1$, and b/B denote the coefficient of discharge, Froud number, weir height to its length, ratios of weir length to its depth of upstream flow, and weir dimensionless length, respectively. Figure 1 presents the correlation matrix between the used variables in this study and the histogram distribution for each variable. The variables are in general have a poor linear correlation with C_d, meaning that the process of developing an accurate prediction is a tough task. However, there is a slight similarity between the distribution of C_d, F_1, and p/y_1, respectively.

2.2. Random Forest. Random forest (RF) is a new approach that belongs to the data mining field, and it is structured to give a precise prediction without overfitting the data [32]. By using bootstrapping, RF creates several trees (500–2000), where each tree is trained by randomizing the predictors' subsets, which reduces the correlation between the trees and, simultaneously, adds randomness to bagging. Unlike the conventional trees, RF uses a subset of predictors at each node to split it instead of selecting the best splits among all variables.

In contrast to other machine learning classification systems such as support vector regression and neural networks, RF delivers high-quality results [32, 33]. The tree-based models are generally growing by using regression and classification tree techniques during the process of training data. Figure 2 shows the basic structure of the RF model.

The theory of the RF process is summed up as follows:

Step 1: the n subsets of the training sample $D_1, D_2, \ldots, D_n$ are extracted using the bootstrap sampling process from the full training sample set D. The sample size of the E_n subsets is similar to the overall sample set for training D.

Step 2: n decision trees are built in line with the n subsets and results of the n classification are obtained.

Step 3: each decision tree casts one voting unit for the most common class, which decides optimal outcomes.

2.3. Extreme Learning Machine. The extreme learning machine (ELM) is a technique based on the algorithm of optimization learning least square (LS) and proposed by Huang et al. [34]; it is used for generalizing the single-layer feed-forward network (SLFN). Since the classical machine learning algorithms suffer from latency due to the iterative approach they employ, ELM is proposed as an effective approach to overcome the latency problems [34, 35]. The structure of ELM contains three main layers known as the input, output, and the hidden layer, where the latter (hidden layer) is considered as the most important layer in the ELM approach. Unlike the traditional neural networks, ELM has the ability to overcome tuning obstacles; therefore, there is no need to tune the hidden layer, furthermore, by keeping the parameters in the hidden layer fixed and using LS to resolve the weights in the output layers ELM tends to get minor training errors. In other words, hidden layer weights are separated from the training data and can be randomly initialized, while the weights in the output layer need to be optimized, which can be done by using the pseudoinverse [36, 37]. Due to its basic existence, fast learning, and universal approximation capacity, ELM provides huge advantages regarding mapping the relationship of variables between the input and the output [38].

The relationship between SLNF and m hidden nodes in the ELM model can be expressed mathematically as shown in the following equation:

$$Z_t = \sum_{k=1}^{m} B_k g_k(\alpha_k \cdot x_k + \beta_k), \quad k = 1, 2, \ldots, j. \tag{1}$$

B_k are weight values connected to the output node of the kth hidden node (s), Z_t is the target of ELM, m represents the hidden nods' number, $g_k(\alpha_k \cdot x_k + \beta_k)$ is the hidden output

Table 1: Statistical parameters of the used variable.

Variable	Mean	StDev	Median	Minimum	Maximum	IQR	Skewness	Kurtosis
F_1	0.365	0.1623	0.34	0.088	0.804	0.2507	0.47	−0.56
p/y_1	0.6348	0.1804	0.641	0.181	0.91	0.2663	−0.31	−0.73
b/y_1	2.188	1.417	2.083	0.347	6.554	2.457	0.57	−0.45
b/B	0.9244	0.5033	1.008	0.3	1.513	1.2	−0.06	−1.69
Cd	0.48454	0.097	0.4925	0.281	0.846	0.13425	0.53	0.89

Here, StDev and IQR are, respectively, referring to the standard deviation and interquartile range.

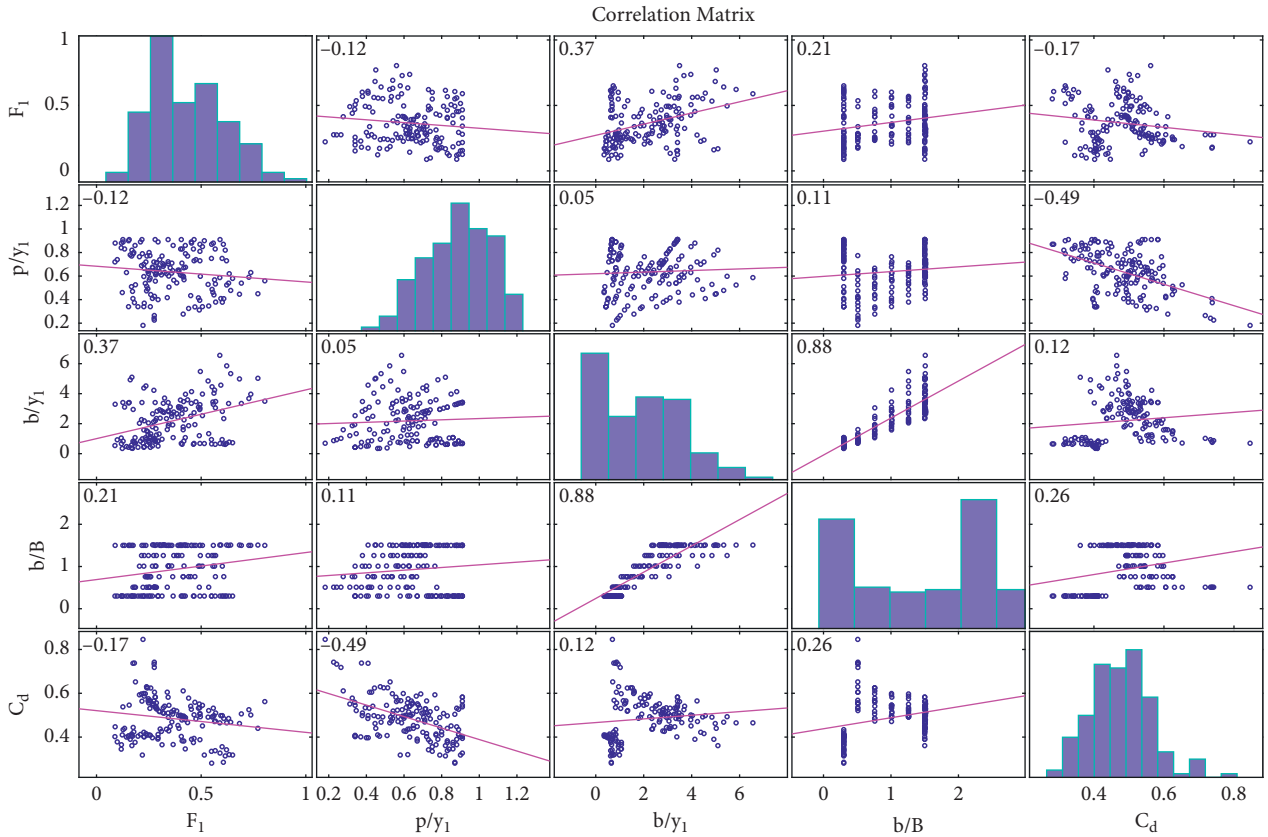

Figure 1: Correlation matrix and histogram plots for parameters.

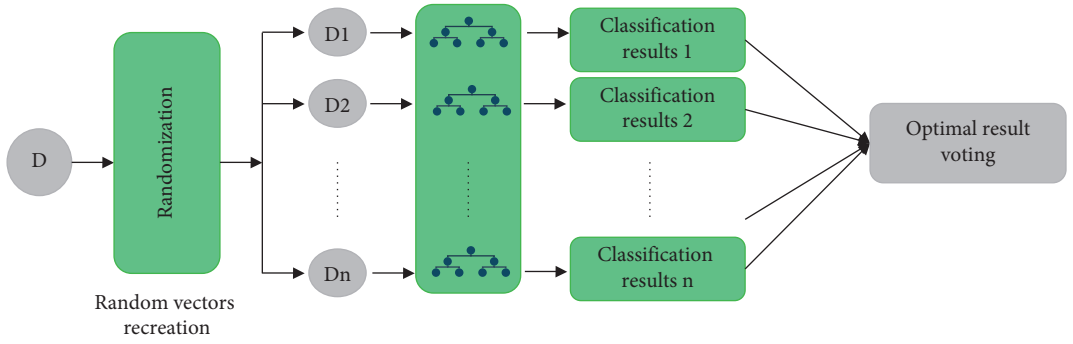

Figure 2: Random forest structure.

function, and (α_k, β_k) are the parameters of the hidden node randomly initialized. It is important to mention that the Sigmoid transfer function is used in the development process of ELM. The j equation above can be written as compact:

$$\mathbf{R}C = T, \tag{2}$$

where $\mathbf{R}$ is the hidden layer output matrix of the neural network,

$$\mathbf{R}(\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m, x_1, \ldots, x_m) = \begin{bmatrix} g_k(\alpha_1 \cdot x_1 + \beta_1) & \cdots & g_m(\alpha_m \cdot x_1 + \beta_m) \\ \vdots & & \vdots \\ g_m(\alpha_m \cdot x_m + \beta_1) & \cdots & g_m(\alpha_m \cdot x_1 + \beta_m) \end{bmatrix}_{N \times m}, \quad (3)$$

$$C = \begin{bmatrix} C_1^T \\ \vdots \\ C_m^T \end{bmatrix}, \quad (4)$$

$$T = \begin{bmatrix} t_1^T \\ \vdots \\ t_m^T \end{bmatrix}, \quad (5)$$

where $(\cdot)^T$ is the transpose of the matrix. Figure 3 shows the structure of ELM.

2.4. Extra Tree Regression. The Extra Tree Regression (ETR) approach is a developed approach derived originally from Random Forest (RF) model and suggested by Geurts et al. [39]. According to the conventional top-down technique, the Extra Tree Regression (ETR) algorithm constructs a collection of unpruned decisions or regression trees [39].

To perform the regression, Random Forest (RF) model utilizes two steps, bootstrapping and bagging, respectively. In the bootstrapping step, a set of decision trees is produced by the growth of each individual tree that uses a random training dataset sample. The bagging step, which operates in two steps, is used to divide the decision tree nodes after achieving the ensemble, where several random subsets of training data are selected during the initial bagging process. The decision-making process is accomplished by selecting the best subset and its value [26].

Breiman [32] considered the RF model to be a series of decision trees, in which $G(x, \theta_r)$ indicates the G^{th} predicting tree, where θ indicates a uniform independent distribution vector assigned before the growth of the tree. All the trees are combined and averaged in an ensemble of trees (forming forest) of $G(x)$ constructed using Breiman [32] equation:

$$G(x, \theta_1, \ldots \theta_r) = \frac{1}{R} \sum_{r=1}^{R} G(x, \theta_r). \quad (6)$$

There are two key distinctions between the ETR and the RF systems. First, the ETR uses all the cutting points and divides nodes by the random selection from these points. Second, it uses the entire learning samples to cultivate the trees [39] in order to minimize bias. The splitting process in the ETR approach is controlled by two parameters, namely, k and $n_{\min}$, where k refers to the number of features that are randomly picked up in the node, and $n_{\min}$ parameter refers to the minimum sample size expected to separate nodes. Furthermore, the strength of the selection of attributes and the average output noise strength are determined by k and $n_{\min}$ respectively. These two parameters improve the precision and reduce overfitting in the ETR model [40, 41]. Figure 4 shows the structure of the ETR.

2.5. Model Development and Performance Evaluation. The development of C_d predictive models in this current paper (ELM, RF, and ETR) was developed based on geometrical and flow characteristics dataset such as $p/y_1, b/y_1$, b/B, and F_1. All of the inputs and outputs used variables are divided into two sets, training (70%) and testing (30%), and then normalized within the range 0 and 1 [42]. Moreover, out of 162 samples, 122 samples were used for training the three proposed models and the rest (40 samples) were used as the testing set. The normalization step is very important in enhancing the performance of machine learning models, and all variables are receiving the same attention and considerations. Figure 5 shows the development step of the proposed models in this study. It is worth mentioning that the hyperparameters of ETR models are randomly selected, because there is no robust method nor related study that could become a benchmark for calculating these parameters in the literature. In this study, a variety of statistical error plots were reported to measure the error and the similarity between the predicted and experimental value of C_d. In addition to the novelty of using ETR as a new application in the hydraulic field, we also performed a sensitivity analysis to identify the most important parameters that have a major impact on the C_d. Another contribution of this study is the conduct of reliable analysis to compare the results of this study with previous studies in order to ensure the efficiency of the used model.

2.6. Performance Evaluation. Twelve statistical parameters have been used to assess the accuracy performance of different AI models used in this study for predicting the C_d. The statistical parameters include different matrices to assess the prediction error and other parameters used for evaluating the agreement between actual and predicted values of C_d

(i) Mean absolute error (MAE) is used as an indicator of how similar the estimated values to the observed ones, and it is given by [41, 42]

$$\text{MAE} = \frac{1}{n} \sum_{r=1}^{n} \left| C_d^{r,m} - C_d^{r,c} \right|. \quad (7)$$

(ii) Root mean square error (RMSE) [43] is a common measure for comparing prediction error for different models, and the lower the RMSE value, the higher the prediction ability of the model in terms of its absolute deviation, and it is given by [42]

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{r=1}^{n} \left(C_d^{r,m} - C_d^{r,c} \right)^2}. \quad (8)$$

RMSE and MAE values are utilized in order to present the fitting error, and these values are always nonnegative. The zero value of MAE and RMSE indicates the perfect fit to data, which has almost never been achieved. Lower RMSE and MAE values generally are better than the higher ones.

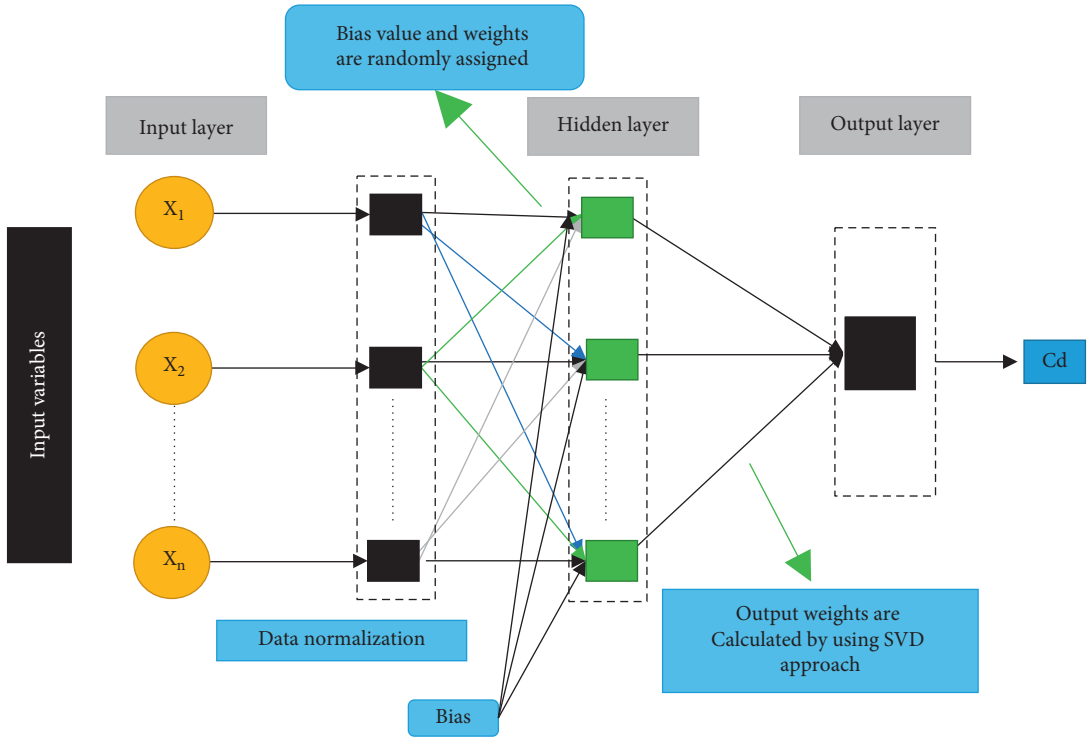

Figure 3: ELM structure.

(iii) Root mean squared relative error (RMSRE) is used to test the training and testing datasets performance and given by [44]

$$\mathrm{RMSRE} = \sqrt{\frac{1}{n}\sum_{r=1}^{n}\left(\frac{C_d^{r,m} - C_d^{r,c}}{C_d^{r,m}}\right)^2}. \tag{9}$$

(iv) Relative root mean square relative error (RMSRE) is obtained by dividing RMSE by the average data value and given by

$$RRMSE = \frac{\sqrt{1/n\sum_{r=1}^{n}\left(C_d^{r,m} - C_d^{r,c}\right)^2}}{\sum_{r=1}^{n} C_d^{r,m}} \times 100. \tag{10}$$

The accuracy of a model is considered excellent when the values of RRMSE are less than 10%, good if the values are between 10% and 20%, fair if the values are between 20% and 30%, and poor if the values of RRMSE are greater than 30%.

(v) The coefficient of determination (R^2) is used to estimate the model efficiency and given by

$$R^2 = 1 - \frac{\sum_{r=1}^{n}\left(C_d^{r,m} - C_d^{r,c}\right)^2}{\sum_{r=1}^{n}\left(C_d^{r,m} - \overline{C_d}^{r,m}\right)^2}. \tag{11}$$

Values of R^2 near 1 indicate a more effective model.

(vi) Relative error (RE):

$$\mathrm{RE} = \frac{C_d^{r,m} - C_d^{r,c}}{C_d^{r,m}} * 100\%. \tag{12}$$

(vii) Maximum absolute relative error (erMAX) is the maximum absolute relative error:

$$\mathrm{erMAX} = \max\left(\left|\frac{C_d^{r,m} - C_d^{r,c}}{C_d^{r,m}}\right|\right). \tag{13}$$

(viii) The correlation coefficient (CC) is used to evaluate the correspondence between actual and predicted values [45].

$$\mathrm{CC} = \frac{\sum_{r=1}^{n}\left(C_d^{r,m} - \overline{C_d}^{r,m,}\right)\left(C_d^{r,c} - \overline{C_d}^{r,c}\right)}{\sqrt{\sum_{r=1}^{n}\left(C_d^{r,m} - \overline{C_d}^{r,m}\right)^2 \sum_{r=1}^{n}\left(C_d^{r,c} - \overline{C_d}^{r,c}\right)^2}}. \tag{14}$$

(ix) Nash–Sutcliffe efficiency (NSE) [46] and the mean squared error (MSE) are considered the most significant criteria in hydrological model calibration and evaluation with observed data. The NSE can show an excellent match between the observed values and predicted values when it equals one.

$$\mathrm{NSE} = 1 - \frac{\sum_{r=1}^{n}\left|C_d^{r,m} - C_d^{r,c}\right|}{\sum_{r=1}^{n}\left|C_d^{r,c} - \mu\right|}, \tag{15}$$

where μ is the observed value mean deviation.

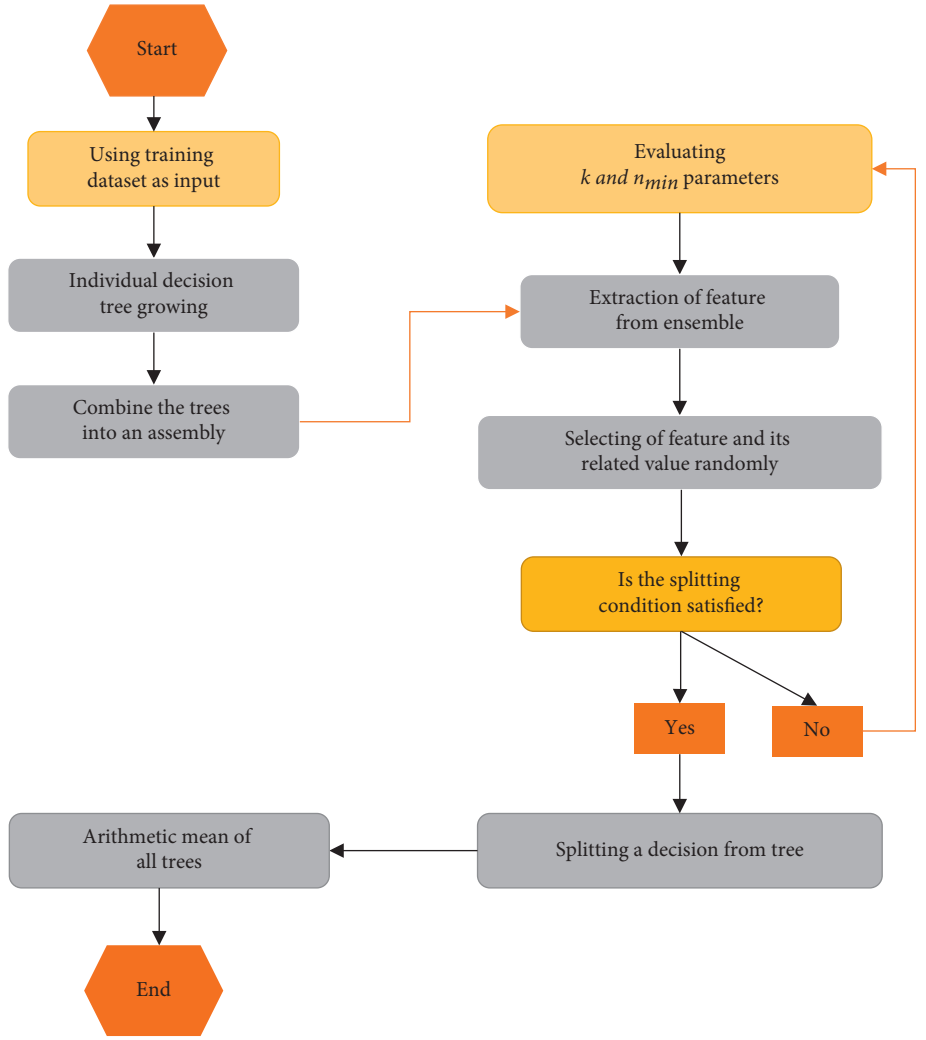

Figure 4: ETR structure.

(x) Willmott index (WI) [47, 48]: the WI can show a good agreement between the observed values and predicted values when it equals one:

$$d = 1 - \frac{\sum_{r=1}^{n} \left(C_d^{r,c} - \overline{C_d}^{r,m}\right)^2}{\sum_{r=1}^{n} \left(\left|C_d^{r,c} - \overline{C_d}^{r,m}\right| + \left|C_d^{r,m} - \overline{C_d}^{r,m}\right|\right)^2}. \quad (16)$$

(xi) Mean bias error (MBA): this indicator shows the trend of the model toward underestimating or overestimating the coefficient of discharge (C_d). It is preferred when the MBA value is closest to zero.

$$\text{MBA} = \frac{1}{n} \sum_{r=1}^{n} \left(C_d^{r,m} - C_d^{r,c}\right). \quad (17)$$

(xii) t-statistic test (t-stat): it is used to confirm whether or not the calculated (C_d) values differ significantly from their measured counterparts. Besides, for a more complete evaluation of (C_d), the estimation model t-stat is used in conjunction with RMSE and MBE. This indicator is used widely to assess the reliability of the predicted model [49, 50].

$$t - \text{stat} = \sqrt{\frac{(n-1)\text{MBA}^2}{\text{RMSE}^2 - \text{MBA}^2}}, \quad (18)$$

where $\overline{C_d}^{r,m}$ and $\overline{C_d}^{r,c}$ are the average of the measured and calculated C_d values, $C_d^{r,m} - C_d^{r,c}$ are representing the measured and predicted values, and n is the total number of observations.

3. Results and Discussion

The key purpose of the current study is to implement the viability of a rigorous method of artificial intelligence called ETR for C_d estimation. The collected data are split into two partitions: 75% of the data are used for the training phase, and the remaining 25% are used for validation purposes. In order to check the predictability of the proposed model during the training and testing steps, two comparable models (RF and ELM) were used in this study for estimating C_d. Different statistical measure indexes are utilized for assessing the performance of the proposed modeling

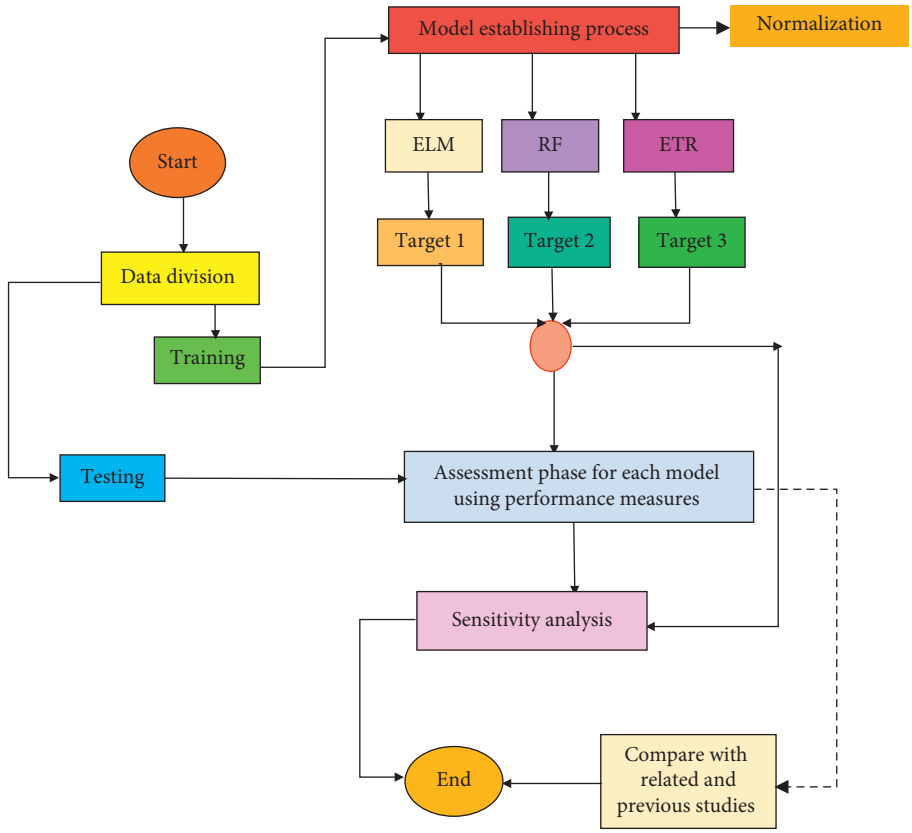

Figure 5: Flow chart of the developed model.

techniques including agreement criteria (CC, NSE, and WI) and error criteria such as RMSE, MAE, RMSRE, RRMSE, and erMax. Furthermore, visual assessment tools like histogram, boxplot presentation, Taylor diagram, Violin presentation, and scatter plots are used for evaluating the proposed models in order to select the best predictive one.

The prediction capacity of all proposed models is statistically assessed using agreement and error criteria through the calibration phase. Based on Figure 6, there are no significant differences in terms of prediction precision between ELM and RF models. More specifically, the performance of ELM is noticed to be slightly better than that of the RF model, where RF has lower values of forecasting errors (MAE = 0.0173, RMSE = 0.0245, RMSRE = 0.0555, RRMSE = 5.064, erMAX = 0.2603) and higher level of agreement (CC = 0.9715, NSE = 0.9438, and WI = 0.9854). Moreover, the RF model reported a desirable estimate but slightly lower than ELM, producing acceptable statistical metrics with MAE of 0.0228, RMSE of 0.0316, RMSRE of 0.0669, RRMSE of 6.521, CC of 0.9541, erMAX of 0.2471, NSE of 0.9068, and WI of 0.9735. However, the proposed ETR model managed efficiently to predict C_d with higher level of accuracy than other predicted models, reporting the lowest value of forecasted errors (MAE = 0.0051, RMSE = 0.0067, RMSRE = 0.0151, RRMSE = 1.379, and erMAX = 0.0575) and highest accuracy indexes (CC = 0.998, NSE = 0.9958, and WI = 0.9989). In terms of RMSE, the proposed model (ETR) showed a prediction enhancement compared to RF and ELM models by 78.80% and 72.65%, respectively.

The visualization assessment is vital to examine the efficiency of the proposed models in predicting each sample separately. A scatter plot is one of the most significant figures used frequently to assess the performance efficiency of the predicted models. It provides significant information about the diversion of each point from the actual observation. Besides, the equation of the regression line can be obtained from the scatter plot ($y = ax + b$). The efficiency of the tested model can be determined according to the value of the slope (a), which is preferred to be close to one. The other significant criterion is the interception value (b). When the b value goes to zero, and a value is close to one, this means that the suggested model is generating more accurate predictions. The measured and predicted values of C_d can be seen in Figure 7. As seen from Figure 7, better accuracy of the ETR model is clearly observed. Furthermore, the ETR is found to have the highest value of R^2 of 0.996. However, the ELM model also provided more accurate estimations but slightly less than ETR, having excellent R^2 (0.944). The other key observation from Figure 7 is that the RF model appeared to produce the lowest prediction accuracy than the two other models, achieving the lower value of R^2 (0.910). Moreover, the visualized assessment is very efficient in the selection process of the most reliable predictive models.

In Figure 8, a relative error diagram is shown to visually examine the performance of each predictive model. The

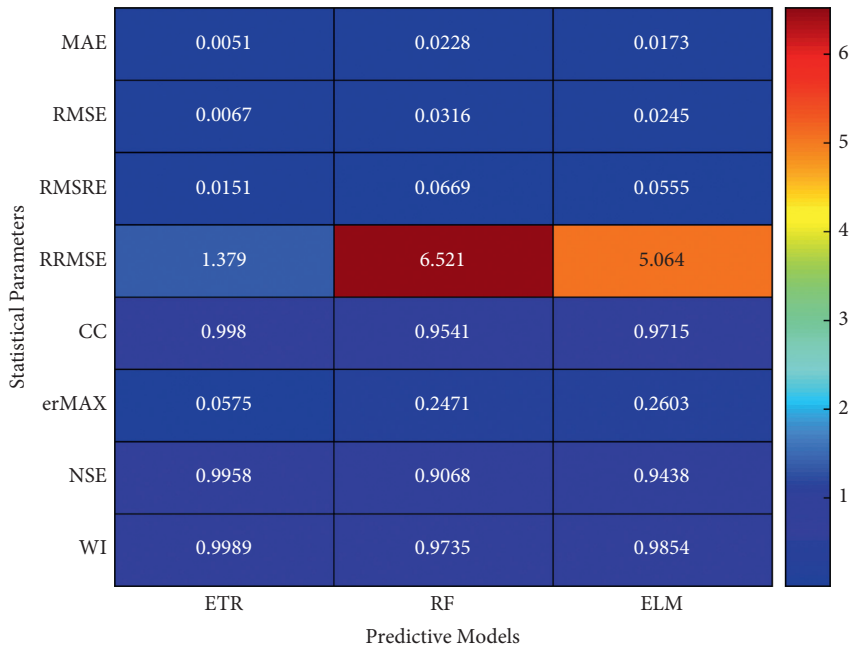

Figure 6: Performance of different models in modeling C_d during the training phase.

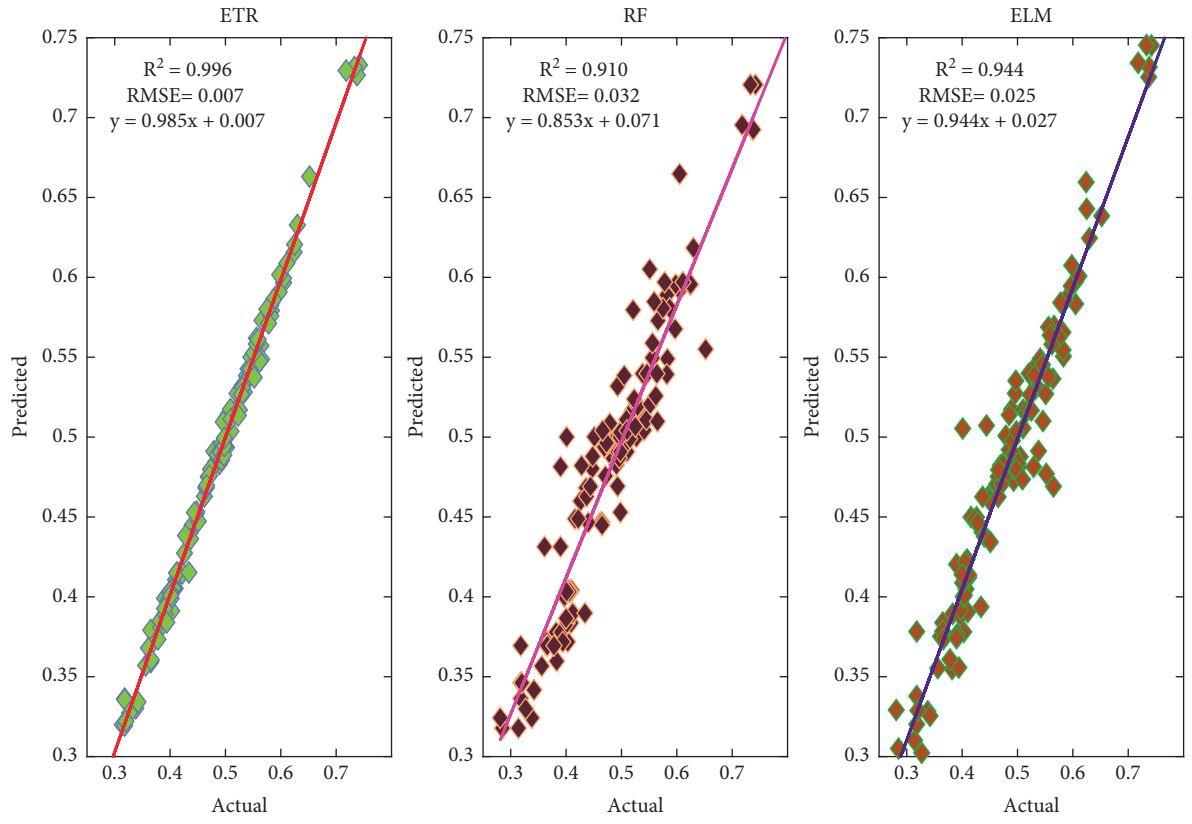

Figure 7: Scatter plots showing the comparison between the observed and predicted value of C_d: training phase.

figures illustrated that the ETR-model outperforms the other models. However, the ELM and RF techniques produced higher relative errors for almost all observations. Furthermore, both techniques (ELM and RF) showed a significant bias for negative values, meaning that these applications tend to overestimate. Consequently, the proposed model (ETR) does not exhibit a significant bias, and it also produces the lowest RE% values.

For further comparison, the boxplot diagram is created in Figure 9 to visually examine the prediction accuracy of the

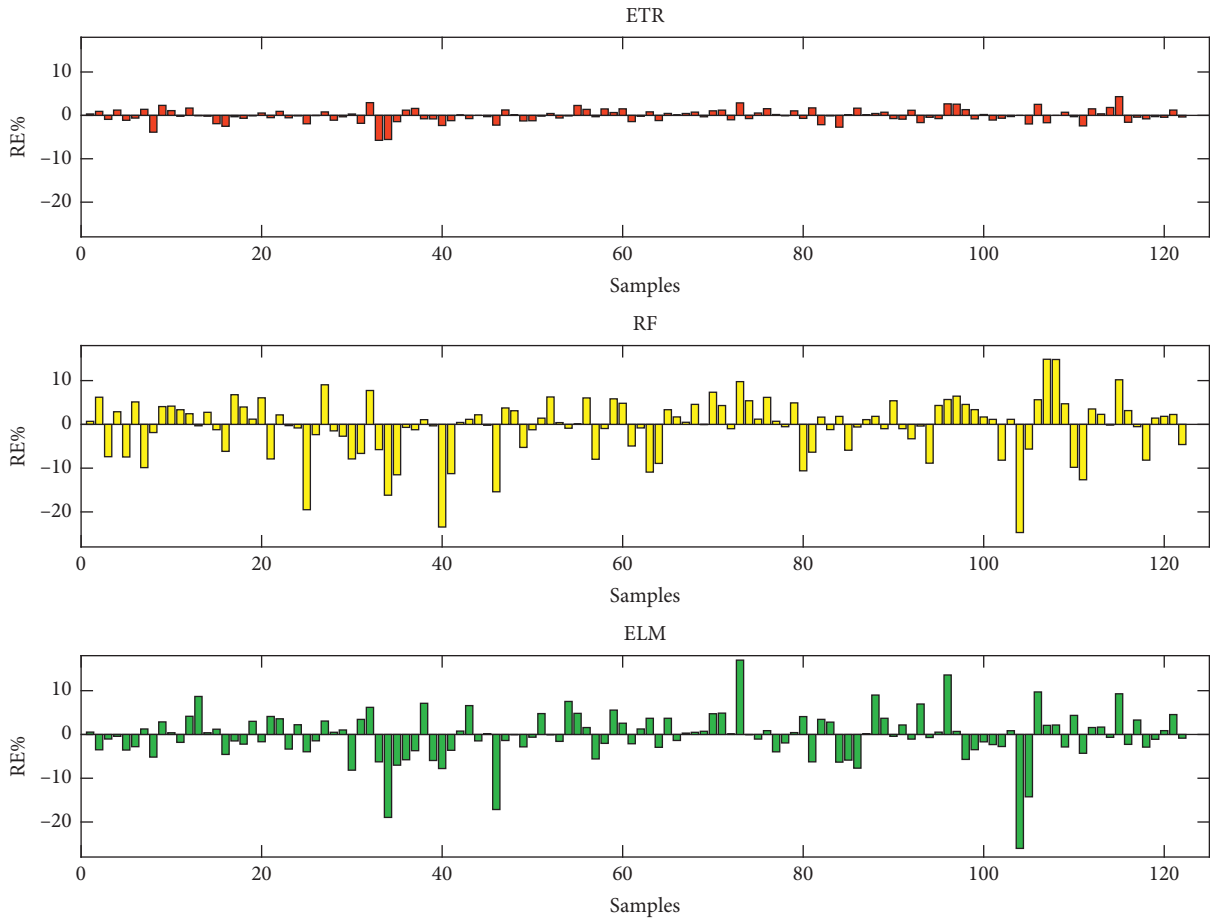

Figure 8: Relative error diagrams to show the performance capacity of each predicted model: training phase.

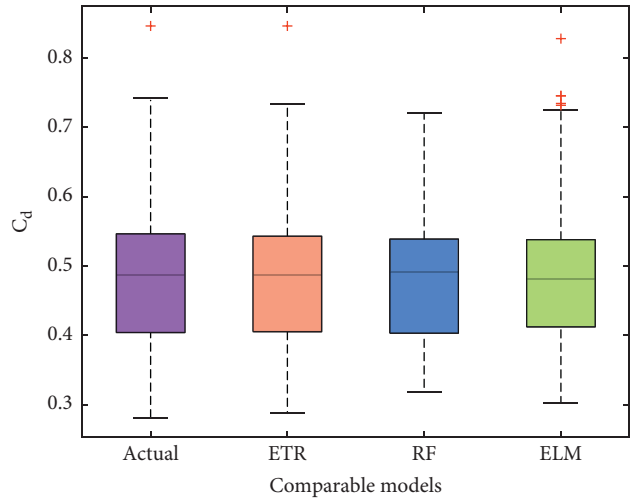

Figure 9: Boxplot diagram used for comparing the outcomes of the proposed models with observer data: training step.

proposed models against the observed values. Boxplot diagram is a tool for graphically representing numerical data groups through their quartiles. In the boxplot diagram, lines are extending from the box called (Whiskers) representing the variability outside the lower and upper quartiles. Moreover, the outliers are shown in the boxplot figure as points, and the line plotted inside the box represents the median of the data. Based on Figure 9, there is a significant similarity between actual observations and the ones predicted by the ETR. Furthermore, the median criteria are shown as a line inside the box and found to be very close to the actual median. It is worth noting that the ETR models

managed efficiently to simulate the outlier values of C_d. One of the most difficult tasks for any AI model is the prediction accuracy of outlier observations, and, however, the ETR models could address this obstacle efficiently. On the other hand, the RF model showed poor prediction performance as shown in Figure 6 and showed less accuracy when capturing the outlier values and giving very far values of median to the actual one. Although the ELM performance was much better than RF, it generated relatively excessive value of outlier C_d estimates. Thus, among the three proposed models, the ETR was superior and found to produce more accurate and consistent predictions with their corresponding observations.

Advanced analysis techniques are important to graphically select the best and more efficient model for the prediction of C_d. In this regard, as an advanced graphical presentation tool, a violin plot is used to produce Figure 10 in order to visually clarify the similarity between actual and predicted values of C_d. From a statistical perspective, the violin plot is considered an efficient diagram, because it combines a boxplot and a density plot that is rotated and placed on each side to show the distribution shape of data. This figure provided more informative information about the performance of each predictive model, where it can be seen that the RF model did not mimic the shape of the dataset as well as desired. On the other hand, both ELM and ETR showed accurate estimates when compared to the actual observations. Moreover, it is noticed that there was a desirable agreement between the actual data set and the predicted ones by ETR. Besides, the ETR models could mimic the distribution function of the measured dataset more efficiently than ELM models.

After presenting the quantitative and visualized assessment during the training phase, it is important to take a step forward to present the capacity of each suggested model through the testing phase. This phase is very important in assessing the qualification of predicted models, because, in the testing phase, the models deal with a new dataset that was excluded during the training phase. Besides, in the testing phase, the issues related to the predictive models such as generalization (overfitting) and the ability to address outlier values can be disclosed more clearly than in the training phase [29, 51–53].

The performance of the predictive models during the testing phase according to statistical measures is provided in Figure 11. All statistical matrices show better performance of the suggested ETR model than other models in the prediction of C_d. Besides, the ETR model showed lower values of error measures (MAE = 0.0166, RMSE = 0.0208, RMSRE = 0.044, RRMSE = 4.284, erMAX = 0.1031) and higher agreement with the actual values (CC = 0.9603, NSE = 0.9175, WI = 0.9793). However, the prediction skill of the RF was found to be poor, thereby generating higher prediction error (MAE = 0.023, RMSE = 0.027, RMSRE = 0.059, RRMSE = 5.565, and erMAX = 0.1382), as well as lower agreement matrices (CC = 0.9331, NSE = 0.8609, and, WI = 0.9651). Furthermore, the performance of ELM generates more accurate predictions than RF but lower than ETR with MAE of 0.0178, RMSE of 0.0223, RMSRE of 0.0463, RRMSE of 4.599, CC of 0.9556, erMAX of 0.1186, NSE of 0.9049, and WI of 0.9767. The ETR model's superiority was measured in terms of its RMSE reduction capacity during the testing phase. The obtained results showed prediction augmentation of 78.80% and 72.65% by using the ETR model compared to the RF and ELM models, respectively. The results confirmed that a more accurate and reliable prediction of C_d could be achieved through the proposed model (ETR).

The observed and estimated C_d in the testing phase are illustrated in Figure 12. According to this Figure 12, the better accuracy of the ETR model in comparison with other models is clearly observed. Furthermore, ETR estimates are closer to the actual values than RF and ELM, specifically for the peak values of C_d. The figure also shows that the RMSE parameter, which is used frequently for evaluating the models, gives a very close evaluation for all proposed models. Thus, the quantitative assessments, which mainly depend on the prediction errors, may sometimes be misleading, and hence, the visualized assessments are very important as complementary to quantitative assessments.

For attaining a robust evaluation of the performance of each suggested model, relative error (RE) and absolute relative error (ARE%) figures are established to graphically provide more evaluation information about the prediction capacity of each predicted sample separately. The relative error (RE) for each sample is presented in Figure 13. Among all AI approaches used in this paper, ETR provides fewer values of relative error percentages. The key observation is that all the predicted observations have very small relative error (Maximum RE is less than ±10%). However, the RF model produces predictions with higher RF compared to ETR and ELM models. ELM also provides some estimates with higher values of relative error (more than ±10%). Furthermore, a histogram of ARE distribution is presented in Figure 14. This figure provided more information on how many percentages of estimated samples have a specific range of error. The RF model was found to produce relatively fewer observations with ARE% of 0%–5% than ETR and ELM models. Nevertheless, ETR was shown to be the best model in terms of predicting samples with fewer ARE% compared to RF and ELM. More specifically, when it comes to using ETR model, more than 72% of the data points (samples) showed an error less than 5%.

The boxplot and violin figures have also been developed to further evaluate the relative performance of the proposed models. Furthermore, they graphically presented more information on the effectiveness of each model separately. The results acquired during the model testing phase were used to create the boxplots as shown in Figure 15. Furthermore, the interquartile range (IQR) and median of the ETR model were found to be very close to the observed IQR and median. Overall, the ETR models generated high accuracy of prediction, much similar to the observed ones. On the other hand, the ELM and models did not successfully provide estimations as close as the observed ones. For instance, ELM models tended to overestimate, while the RF model underestimates the discharge coefficient. According to Figure 16, there is a desirable corresponding between the distributions of

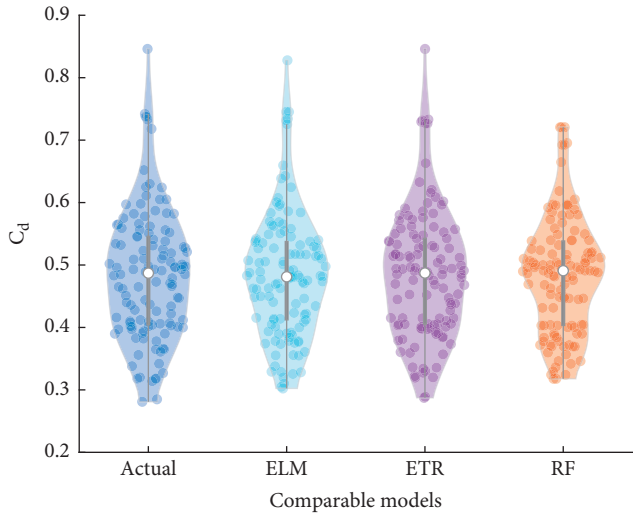

Figure 10: Violin presentation for showing the similarity between observed and predicted for C_d during the training phase.

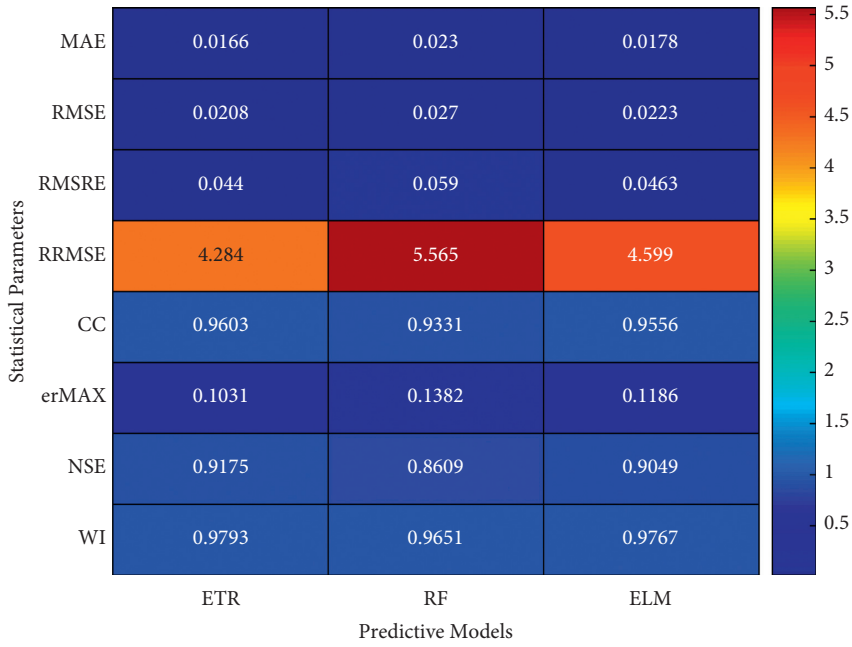

Figure 11: Performance of different models in modeling C_d during the testing phase.

actual C_d and predicted values by ETR. Nevertheless, it can be observed that the ELM could efficiently capture the higher values of C_d, and the skewness toward the largest values was noticed. It can be seen that, based on the two figures (Figures 16 and 17), the ELM model is weak in predicting the higher value of the coefficient of discharge.

The performance of suggested models in simulating the measured C_d is visually presented using Taylor diagram. In Figure 17, Taylor diagram affords a measure of association, error, and variability in simulated C_d compared to measure one and, hence, provides a detailed appraisal of model performance. Figure 17 clearly illustrated that, during the testing step, the ETR simulated C_d very closer to the actual C_d than ELM and RF, respectively. Besides, the obtained results exhibited significantly higher prediction accuracy of the ETR model than other predictive models used in this study.

For further assessment of the reliability of the applied model in this study (ETR) model against the comparable models such as ELM and RF, t-stat and MBA as statistical parameters are used to verify the efficiency of the used model through training and testing phases. Figure 18 shows the superiority of the ETR models in providing more accurate estimates compared to benchmark models (i.e., ELM, RF) through the training phase and testing phase, respectively.

To examine the performance of the proposed model, three benchmark models are considered for the comparison in terms of the coefficient of determination (R^2). R^2 obtained from the proposed ETR model through the testing phase is

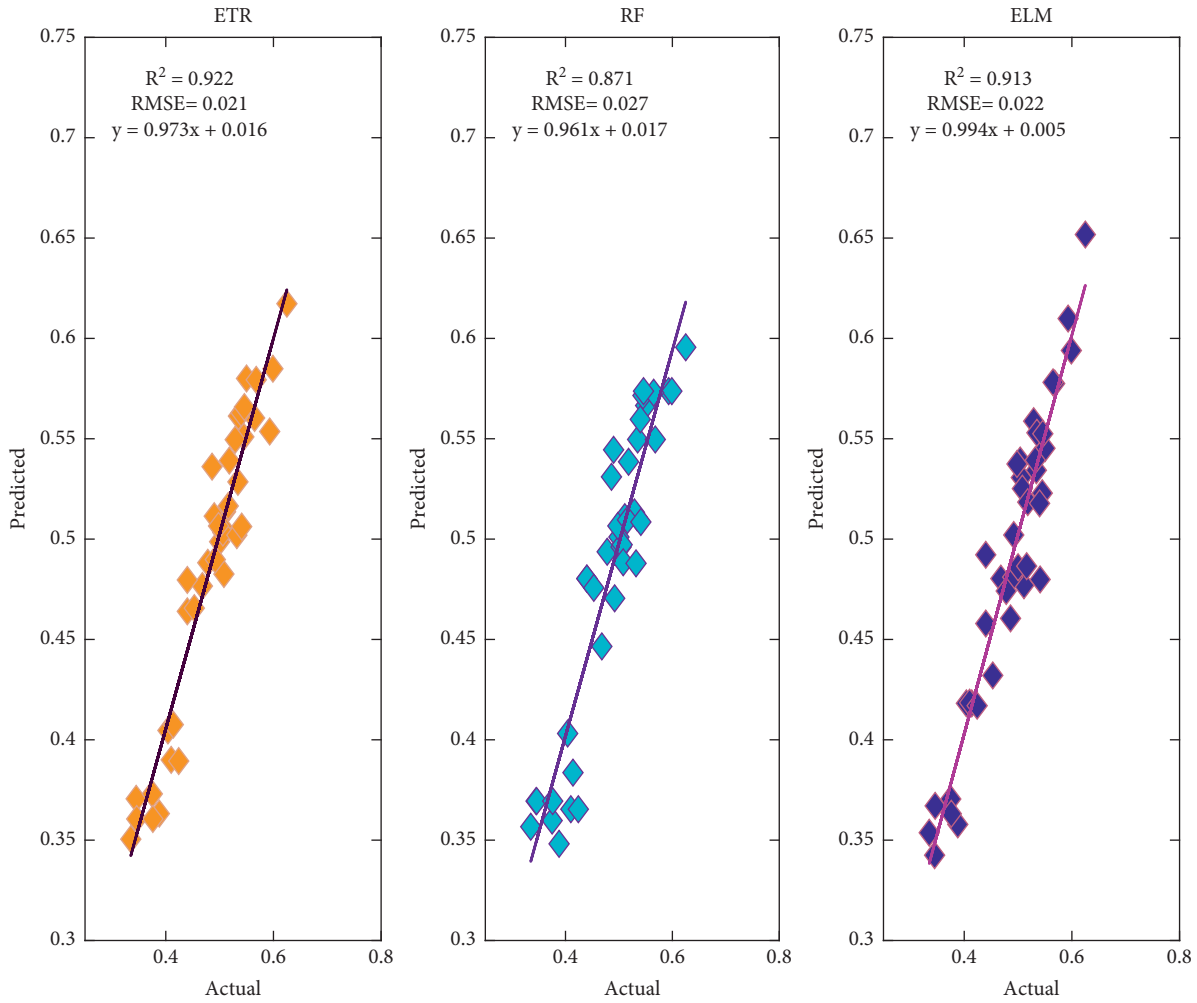

Figure 12: Scatter plots showing the comparison between the observed and predicted value of C_d: testing phase.

compared with Ebtehaj et al. [19], where the Group Method of Data Handling (GMDH) is used by the benchmark model to predict the coefficient of discharge. By comparing R^2 for both the benchmark and proposed models, the latter showed a better accuracy, where R^2 of 0.779 and 0.922 is obtained from the benchmark and proposed models, respectively, for the same data of [19]. The second benchmark model is proposed by Bilhan et al. [54] and uses the artificial neural network (ANN) to predict the coefficient of discharge of triangular labyrinth side wear in a curved channel. The obtained R^2 from ANN model was 0.8836. The third benchmark model is proposed by Parsaie et al. [55] with R^2 of 0.7396, where, based on principal component analysis, a developed adaptive neuro-fuzzy method is proposed to predict the coefficient of discharge of side weirs.

3.1. Sensitivity Analysis. In this study, the sensitivity analysis is performed to explore the effect of each input variable on the coefficient of discharge (C_d). The procedure of conducting sensitivity analysis is carried out by training the model of this study (ETR) based on several combinations of input variables and then checking the reliability of the performance model during the testing set. This strategy can provide more useful information regarding the proposed model and examining its performance in a difficult situation when the inputs are insufficient. To conduct the sensitivity analyses, different input combinations are established in Table 2. In this table, the first model (MO) corresponds to the proposed model of this study (ETR) using all provided input parameters for predicting the discharge coefficient. Figure 19 presents the results of the sensitivity analysis using different sorts of input variables. According to attained outcomes, model M3 produced the highest prediction accuracy among the other models in comparison with the benchmark model (M0), based on the fact that the p/y_1 has the lowest impact on the discharge coefficient followed. On the other hand, b/B is generally considered the most important variable, where the absence of that variable reduces the prediction accuracy significantly than other parameters. Figure 20 presents the relative importance of each variable based on comparing the outcomes proposing input combination (M1, M2, M3, and M4) with the benchmark one (M0) according to the relative differences of RMSE. It is significant noting that the b/B parameter has the highest impact on the coefficient of discharge (29.69%), followed by b/y_1 (25.58%), F_1 (25.08%), and p/y_1 (20.66%).

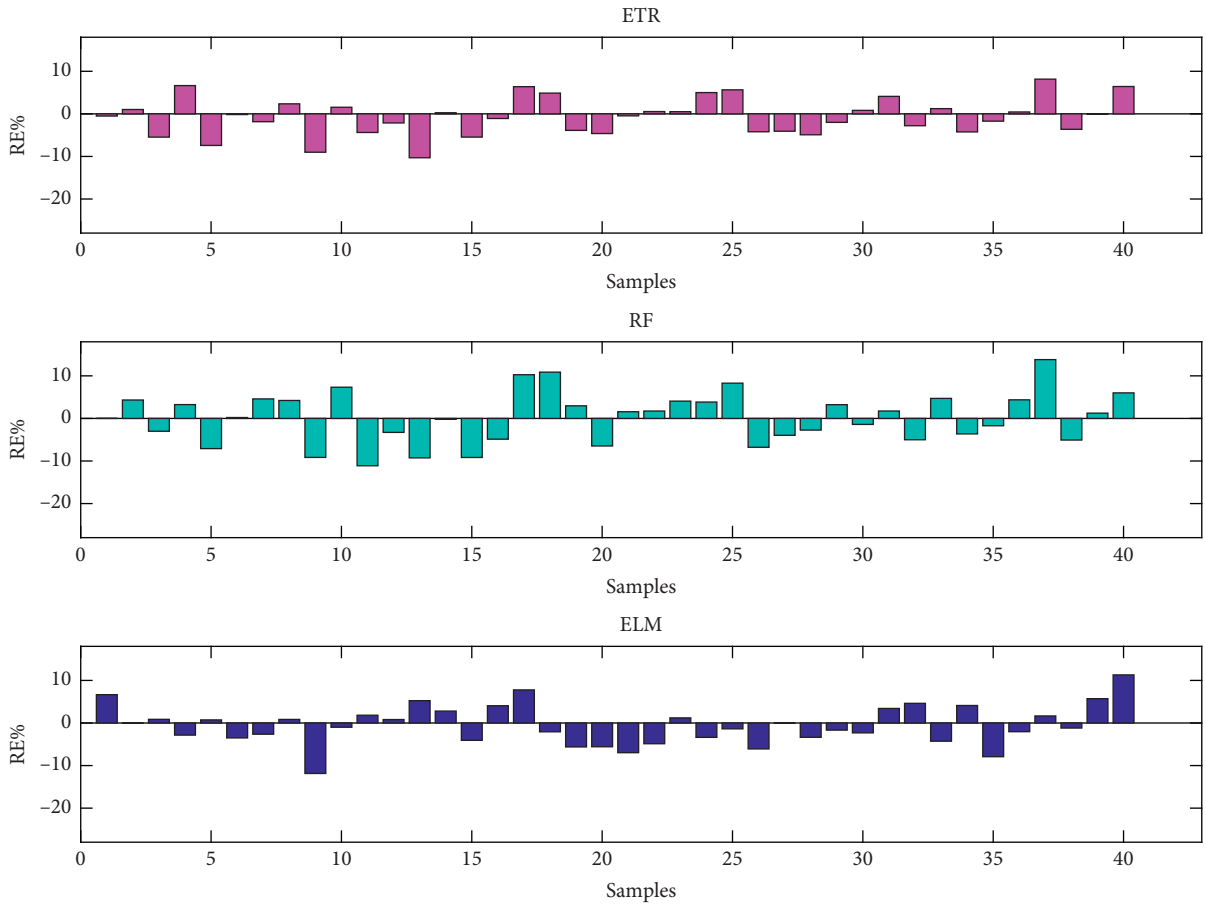

Figure 13: Relative error diagrams to show performance capacity of each predicted model: testing phase.

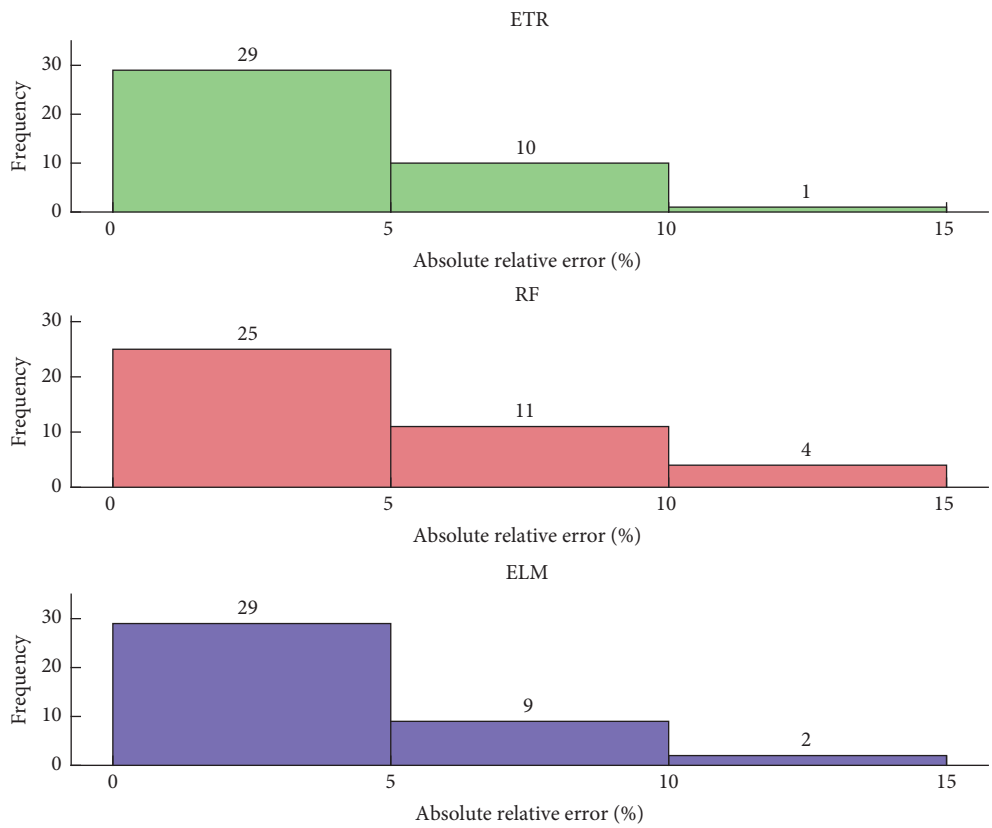

Figure 14: Histogram of absolute relative error.

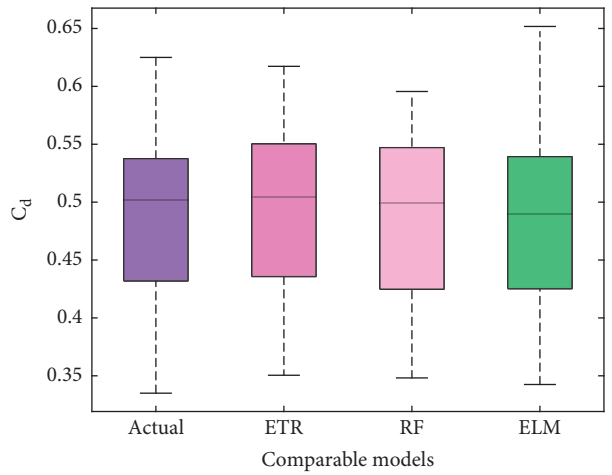

Figure 15: Boxplot diagram used for comparing the outcomes of the proposed models with observer data: testing step.

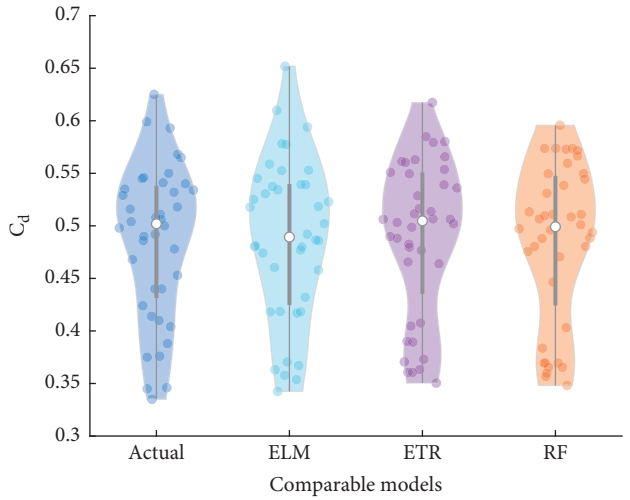

Figure 16: Violin presentation for showing the similarity between observed and predicted for C_d during the training phase.

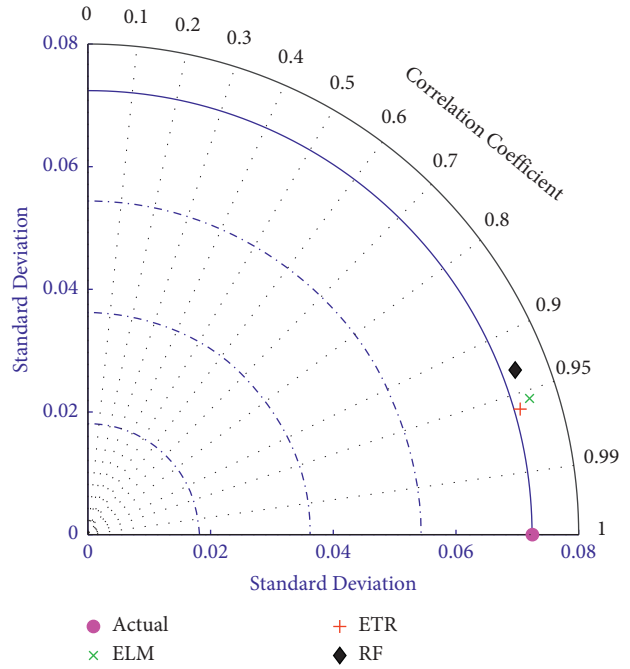

Figure 17: Graphical Taylor diagram for comparing the performances of the proposed models during the testing phase.

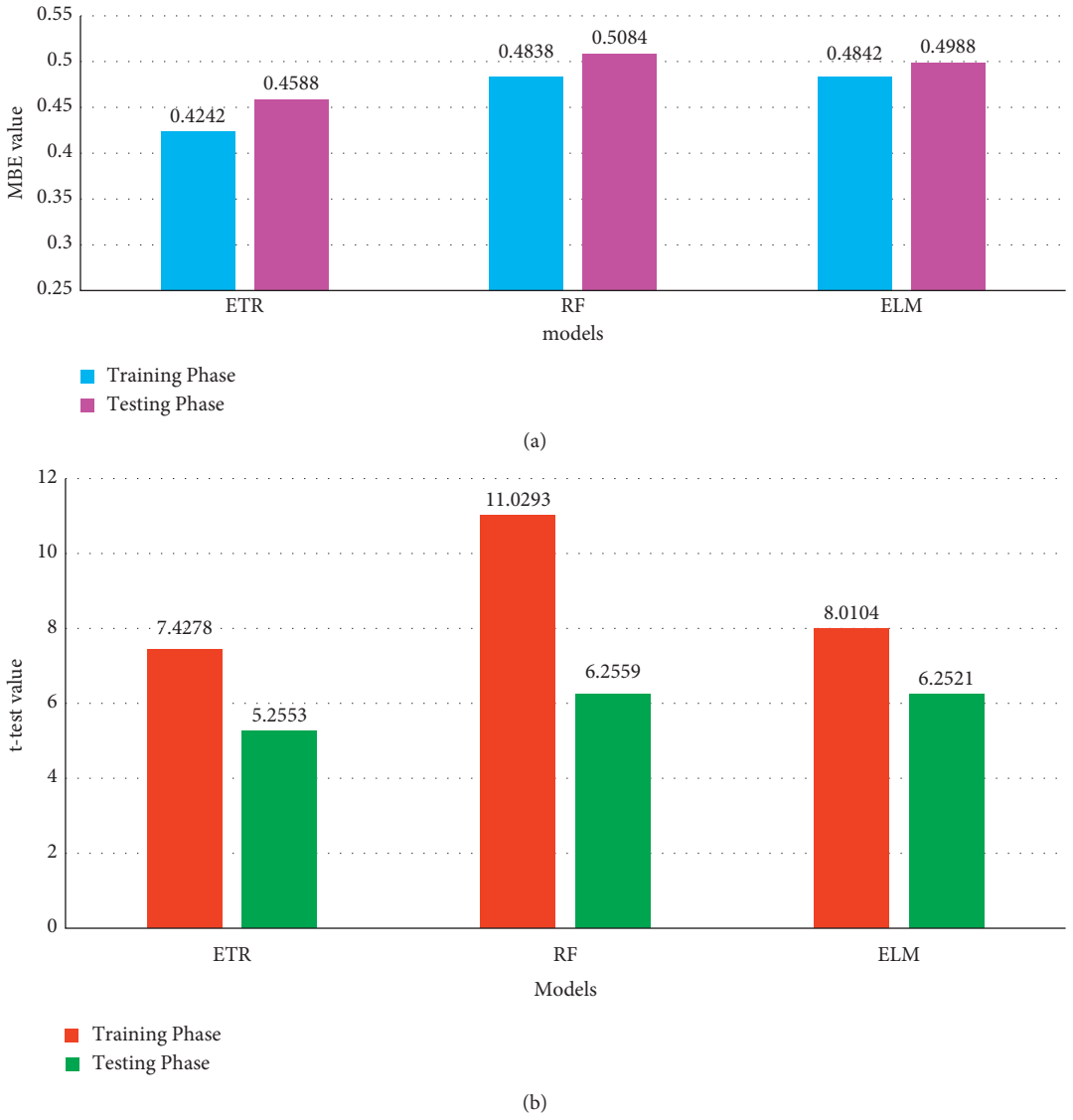

Figure 18: A comparison between the performances of the applied models: mean bias error (MBA) (a) and the comparison according to *t*-statistic test (*t*-stat) (b).

Table 2: Different input combinations were used to conduct sensitivity analysis.

Model	Absent	Function
M0	No	$C_d = f(F1, p/y_1, b/y_1, b/B)$
M1	b/B	$C_d = f(F1, p/y_1, b/y_1)$
M2	b/y_1	$C_d = f(F1, p/y_1, b/B)$
M3	p/y_1	$C_d = f(F1, b/y_1, b/B)$
M4	F_1	$C_d = f(p/y_1, b/y_1, b/B)$

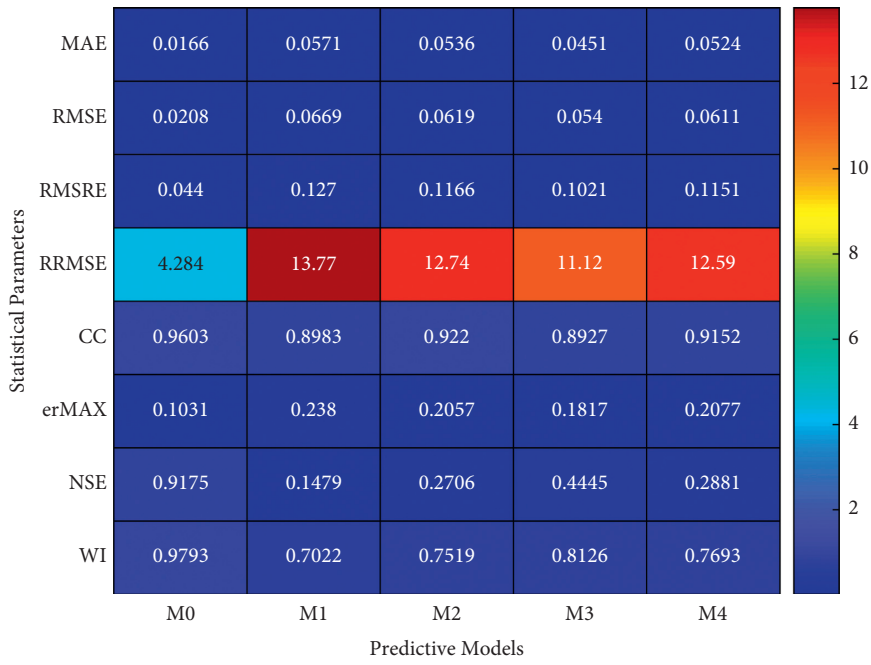

Figure 19: Results of the sensitivity analysis for the ETR model.

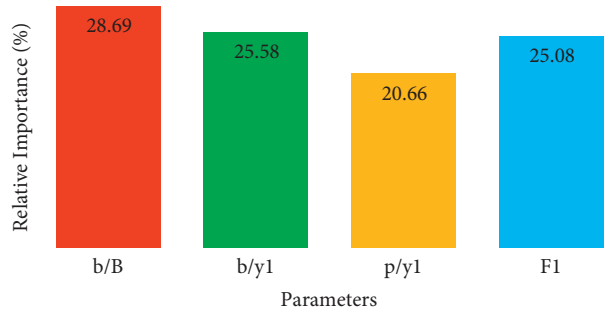

Figure 20: Relative importance presentation showing the percentage effect of each variable on the coefficient discharge.

4. Conclusion

In this study, three different machine learning approaches, namely, ETR, RF, and ELM, have been utilized to predict the coefficient of discharge C_d by using the geometrical and discharge parameters (F_1, p/y_1, b/y_1, and b/B) as the models' inputs. The performance of each model was assessed using different statistical parameters (error and agreement metrics) and novel graphical presentations like Taylor, boxplot, violin, and line-graph diagrams. The outcomes of the applied models are summarized as follows:

(1) ETR model outperformed other models in both training and testing phases. For instance, in terms of RMSE (during the testing phase), the ETR model showed a prediction improvement of 22.96% and 6.73% compared to the RF and ELM models, respectively.

(2) The obtained results in this study confirmed that, unlike the ETR, the RF suffers from an overfitting issue.

(3) Sensitivity analyses were carried to select the most influential parameter on the coefficient of discharge. According to sensitivity analysis, the b/B parameter has the highest effects on the discharge coefficient followed by b/y_1, F_1, and p/y_1.

(4) It can be concluded that the proposed model (ETR) is very effective in modeling complex hydraulic systems. Accordingly, this method can be used more comprehensively in future studies for solving complex issues regarding the hydraulic sector.

For future researches, this study recommends applying the preprocessing technique to handle the dataset before starting to train the models. One of the most efficient approaches is Principle Component Analysis (PCA) technique, which is widely used to remove the correlation between input vectors and reduce the input uncertainty as well [56]. Thus, the presence of PCA may improve the performance of ETR model by providing to the model clean and nonredundant information.

Data Availability

The data are available upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to express their gratitude to Al-Maaref University College for supporting and funding this study.

References

[1] A. El-Khashab, K. V. H. Smith, Experimental investigation of flow over side weirs," *Journal of the Hydraulics Division*, vol. 102, no. 9, pp. 1255–1268, 1976.

[2] G. De Marchi, "Essay on the performance of lateral weirs," *L'Energia Elettrica Milan, Italy*, vol. 11, pp. 849–860, 1934.

[3] A. Abdollahi, A. Kabiri-Samani, K. Asghari, H. Atoof, and S. Bagheri, "Numerical modeling of flow field around the labyrinth side-weirs in the presence of guide vanes," *ISH Journal of Hydraulic Engineering*, vol. 23, no. 1, pp. 71–79, 2017.

[4] K. Subramanya and S. C. Awasthy, "Spatially varied flow over side-weirs," *ASCE Journal of the Hydraulics Division*, vol. 98, no. 1, 1972.

[5] K. G. Ranga Raju, B. Prasad, and S. K. Gupta, "Side weir in rectangular channel," *ASCE Journal of the Hydraulics Division*, vol. 105, no. 5, 1979.

[6] W. H. Hager, "Lateral outflow over side weirs," *Journal of Hydraulic Engineering*, vol. 113, no. 4, 1987.

[7] H. Cheong, "Discharge coefficient of lateral diversion from trapezoidal channel," *Journal of Irrigation and Drainage Engineering*, vol. 117, no. 4, 1991.

[8] R. Singh, D. Manivannan, and T. Satyanarayana, "Discharge coefficient of rectangular side weirs," *Journal of Irrigation and Drainage Engineering*, vol. 120, no. 4, 1994.

[9] S. M. Borghei, M. R. Jalili, and M. Ghodsian, "Discharge coefficient for sharp-crested side weir in subcritical flow"," *Journal of Hydraulic Engineering*, vol. 125, no. 10, 1999.

[10] A. Parsaie, A. H. Haghiabi, and A. Moradinejad, "CFD modeling of flow pattern in spillway's approach channel," *Sustainable Water Resources Management*, vol. 1, no. 3, pp. 245–251, 2015.

[11] A. Parsaie, H. A. Yonesi, and S. Najafian, "Predictive modeling of discharge in compound open channel by support vector machine technique," *Modeling Earth Systems and Environment*, vol. 1, no. 1-2, 2015.

[12] A. H. Haghiabi, A. Parsaie, and S. Ememgholizadeh, "Prediction of discharge coefficient of triangular labyrinth weirs using adaptive neuro fuzzy inference system," *Alexandria Engineering Journal*, vol. 57, no. 3, pp. 1773–1782, 2018.

[13] T. Saba and A. Rehman, J. S. AlGhamdi, Weather forecasting based on hybrid neural model," *Applied Water Science*, vol. 7, no. 7, pp. 3869–3874, 2017.

[14] S. Shaghaghi, H. Bonakdari, A. Gholami, I. Ebtehaj, and M. Zeinolabedini, "Comparative analysis of GMDH neural network based on genetic algorithm and particle swarm optimization in stable channel design," *Applied Mathematics and Computation*, vol. 313, pp. 271–286, 2017.

[15] H. Azimi, H. Bonakdari, I. Ebtehaj, B. Gharabaghi, and F. Khoshbin, "Evolutionary design of generalized group method of data handling-type neural network for estimating the hydraulic jump roller length," *Acta Mechanica*, vol. 229, no. 3, pp. 1197–1214, 2018.

[16] B. Ebrahimi, M. Rahmani, and S. H. Ghodsypour, "A new simulation-based genetic algorithm to efficiency measure in IDEA with weight restrictions," *Measurement*, vol. 108, pp. 26–33, 2017.

[17] J. A. Ali, M. A. Hannan, A. Mohamed, and M. G. M. Abdolrasol, "Fuzzy logic speed controller optimization approach for induction motor drive using backtracking search algorithm," *Measurement*, vol. 78, pp. 49–62, 2016.

[18] H. Bonakdari, A. H. Zaji, S. Shamshirband, R. Hashim, and D. Petkovic, "Sensitivity analysis of the discharge coefficient of a modified triangular side weir by adaptive neuro-fuzzy methodology," *Measurement*, vol. 73, pp. 74–81, 2015.

[19] I. Ebtehaj, H. Bonakdari, A. H. Zaji, H. Azimi, and F. Khoshbin, "GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs," *Engineering Science and Technology, an International Journal*, vol. 18, no. 4, pp. 746–757, 2015.

[20] H. Azimi, H. Bonakdari, and I. Ebtehaj, "A highly efficient Gene expression programming model for predicting the discharge coefficient in a side weir along a trapezoidal canal," *Irrigation and Drainage*, vol. 66, no. 4, pp. 655–666, 2017.

[21] I. Ebtehaj, H. Bonakdari, and B. Gharabaghi, "Development of more accurate discharge coefficient prediction equations for rectangular side weirs using adaptive neuro-fuzzy inference system and generalized group method of data handling," *Measurement*, vol. 116, pp. 473–482, 2018.

[22] H. Azimi, H. Bonakdari, and I. Ebtehaj, "Design of radial basis function-based support vector regression in predicting the discharge coefficient of a side weir in a trapezoidal channel," *Applied Water Science*, vol. 9, no. 4, p. 78, 2019.

[23] F. Salmasi, M. Nouri, P. Sihag, and J. Abraham, "Application of SVM, ANN, GRNN, RF, GP and RT models for predicting discharge coefficients of oblique sluice gates using experimental data," *Water Supply*, vol. 21, no. 1, pp. 232–248, 2021.

[24] P. Refaeilzadeh, L. Tang, and H. Liu, "Cross-validation," *Encyclopedia of Database Systems*, vol. 5, pp. 532–538, 2009.

[25] O. Bilhan, M. E. Emiroglu, C. J. Miller, and M. Ulas, "The evaluation of the effect of nappe breakers on the discharge capacity of trapezoidal labyrinth weirs by ELM and SVR approaches," *Flow Measurement and Instrumentation*, vol. 64, pp. 71–82, 2018.

[26] A. Sharafati, S. B. H. S. Asadollah, and M. Hosseinzadeh, "The potential of new ensemble machine learning models for effluent quality parameters prediction and related uncertainty," *Process Safety and Environmental Protection*, vol. 140, pp. 68–78, 2020.

[27] S. B. H. S. Asadollah, A. Sharafati, D. Motta, and Z. M. Yaseen, "River water quality index prediction and uncertainty analysis: a comparative study of machine learning models," *Journal of Environmental Chemical Engineering*, vol. 9, Article ID 104599, 2020.

[28] S. Shamshirband, E. Jafari Nodoushan, J. E. Adolf, A. Abdul Manaf, A. Mosavi, and K. Chau, "Ensemble models with uncertainty analysis for multi-day ahead forecasting of chlorophyll a concentration in coastal waters," *Engineering Applications of Computational Fluid Mechanics*, vol. 13, no. 1, pp. 91–101, 2019.

[29] M. K. Alomar, M. M. Hameed, N. Al-Ansari, and M. A. Alsaadi, "Data-driven model for the prediction of total

dissolved gas: robust artificial intelligence approach," *Advances in Civil Engineering*, vol. 2020, 20 pages, 2020.

[30] M. E. Emiroglu, H. Agaccioglu, and N. Kaya, "Discharging capacity of rectangular side weirs in straight open channels," *Flow Measurement and Instrumentation*, vol. 22, no. 4, pp. 319–330, 2011.

[31] S. Bagheri, A. R. Kabiri-Samani, and M. Heidarpour, "Discharge coefficient of rectangular sharp-crested side weirs part II: Domínguez's method," *Flow Measurement and Instrumentation*, vol. 35, pp. 116–121, 2014.

[32] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[33] A. Liaw and M. Wiener, "Classification and regression by randomforest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.

[34] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.

[35] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 42, no. 2, pp. 513–529, 2011.

[36] J. M. Martínez-Martínez, P. Escandell-Montero, E. Soria-Olivas, J. D. Martín-Guerrero, R. Magdalena-Benedito, and J. Gómez-Sanchis, "Regularized extreme learning machine for regression problems," *Neurocomputing*, vol. 74, no. 17, pp. 3716–3721, 2011.

[37] S. Lu, X. Wang, G. Zhang, and X. Zhou, "Effective algorithms of the Moore-Penrose inverse matrices for extreme learning machine," *Intelligent Data Analysis*, vol. 19, no. 4, pp. 743–760, 2015.

[38] M. N. A. Raja and S. K. Shukla, "An extreme learning machine model for geosynthetic-reinforced sandy soil foundations," *Proceedings of the Institution of Civil Engineers: Geotechnical Engineering*, pp. 1–21, 2020.

[39] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 63, no. 1, pp. 3–42, 2006.

[40] V. John, Z. Liu, C. Guo, S. Mita, and K. Kidono, "Real-time lane estimation using deep features and extra trees regression," in *Image and Video Technology*, T. Bräunl, B. McCane, M. Rivera, and X. Yu, Eds., vol. 9431, pp. 721–733, Springer, Cham, Switzerland, 2016.

[41] G. Mishra, D. Sehgal, D. Sehgal, and J. K. Valadi, "Quantitative structure activity relationship study of the anti-hepatitis peptides employing random forest and extra tree regressors," *Bioinformation*, vol. 13, no. 3, pp. 60–62, 2017.

[42] M. K. AlOmar, M. M. Hameed, and M. A. AlSaadi, "Multi hours ahead prediction of surface ozone gas concentration: robust artificial intelligence approach," *Atmospheric Pollution Research*, vol. 11, no. 9, pp. 1572–1587, 2020.

[43] C. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Research*, vol. 30, pp. 79–82, 2005.

[44] H. Tao, A. W. Ewees, A. O. Al-Sulttani et al., "Global solar radiation prediction over North Dakota using air temperature: development of novel hybrid intelligence model," *Energy Report*, vol. 1174, pp. 136–157, 2021.

[45] M. M. Hameed and M. K. AlOmar, "Prediction of compressive strength of high-performance concrete: hybrid artificial intelligence technique," *Communications in Computer and Information Science*, pp. 323–335, 2020.

[46] J. E. Nash and J. V. Sutcliffe, "River flow forecasting through conceptual models part I—a discussion of principles," *Journal of Hydrology*, vol. 10, no. 3, pp. 282–290, 1970.

[47] C. J. Willmott and D. E. Wicks, "An empirical method for the spatial interpolation of monthly precipitation within California," *Physical Geography*, vol. 1, no. 1, pp. 59–73, 1980.

[48] C. J. Willmott, "On the validation of models," *Physical Geography*, vol. 2, no. 2, pp. 184–194, 1981.

[49] R. J. Stone, "Improved statistical procedure for the evaluation of solar radiation estimation models," *Solar Energy*, vol. 51, no. 4, pp. 289–291, 1993.

[50] M. Despotovic, V. Nedic, D. Despotovic, and S. Cvetanovic, "Review and statistical analysis of different global solar radiation sunshine models," *Renewable and Sustainable Energy Reviews*, vol. 52, pp. 1869–1880, 2015.

[51] M. M. Hameed, M. K. AlOmar, W. J. Baniya, and M. A. AlSaadi, "Prediction of high-strength concrete: high-order response surface methodology modeling approach," *Engineering with Computers*, 2021.

[52] G. Zhang, Z. H. Ali, M. S. Aldlemy et al., "Reinforced concrete deep beam shear strength capacity modelling using an integrative bio-inspired algorithm with an artificial intelligence model," *Engineering with Computers*, 2020.

[53] J. A. Khalaf, A. A. Majeed, M. S. Aldlemy et al., "Hybridized deep learning model for perfobond rib shear strength connector prediction," *Complexity*, vol. 202121 pages, 2021.

[54] O. Bilhan, M. E. Emiroglu, and O. Kisi, "Use of artificial neural networks for prediction of discharge coefficient of triangular labyrinth side weir in curved channels," *Advances in Engineering Software*, vol. 42, no. 4, pp. 208–214, 2011.

[55] A. Parsaie and A. H. Haghiabi, "Prediction of discharge coefficient of side weir using adaptive neuro-fuzzy inference system," *Sustainable Water Resources Management*, vol. 2, no. 3, pp. 257–264, 2016.

[56] M. M. Hameed, M. K. AlOmar, W. J. Baniya, and M. A. AlSaadi, "Incorporation of artificial neural network with principal component analysis and cross-validation technique to predict high-performance concrete compressive strength," *Asian Journal of Civil Engineering*, vol. 22, no. 6, pp. 1019–1031, 2021.