Predicting Critical Submergence in Horizontal Circular Intakes Using Meta-Heuristic Optimized Gradient Boosting Ensembles

Sujing Yu
School of Architecture and Design, Lishui Vocational and Technical College, Lishui 323000, Zhejiang, China
Email: lishuisujing@163.com

Keywords: critical submergence, boosting models, meta-heuristic algorithms, ensemble technique, hybrid method

Received: May 21, 2025

This study evaluates critical submergence depth (Sc) in horizontal circular intakes using a dataset compiled from two controlled laboratory experiments conducted under varying hydraulic and geometric configurations. A total of 324 experimental measurements were obtained across two intake-clearance scenarios: (i) zero bottom clearance (C = 0) and (ii) partial elevation (C = d_i/2). The dataset included four primary input variables—intake diameter, approach flow velocity, Froude number, and bottom clearance ratio—which were used as predictors of the dimensionless submergence depth. To enhance predictive accuracy, an automated machine-learning framework was developed using Gradient Boosting (GB) and Extreme Gradient Boosting (XGB), combined with five meta-heuristic optimizers: Artificial Hummingbird Algorithm (AHA), Victoria Amazonica Optimization (VAO), Turbulent Flow of Water-based Optimization (TFWO), Smell Agent Optimization (SAO), and Tasmania Devil Optimizer (TDO). Model training used a 70/30 train–test split, with 10-fold cross-validation applied exclusively to the training portion for hyperparameter tuning. A custom empirical adjustment equation was incorporated to further enhance physical consistency and model stability. Performance evaluation based on the independent test set showed that VAO-enhanced models provided the highest predictive accuracy. The XGVO model achieved R² = 0.993 and RMSE = 0.082 for the C = 0 configuration, while the GBVA model achieved R² = 0.993 and RMSE = 0.118 for the C = d_i/2 case. These findings demonstrate that hybrid ensemble techniques can reliably estimate Sc across varying hydraulic scenarios and outperform standard boosting algorithms.

Povzetek: Študija kaže, da hibridni modeli strojnega učenja z metahevrističnimi optimizacijo zelo natančno napovedujejo kritično globino potopitve pri različnih hidravličnih pogojih.

1 Introduction

Intake systems are widely applied in hydraulic engineering applications for various purposes, including flood control mechanisms, irrigation systems, hydraulic pumps, and other water diversion requirements [1,2]. These devices convey water from primary storage facilities, such as rivers, lakes, and dam reservoirs, to additional locations [3]. The generation of air entrainment and the associated air cone vortex are significant concerns in hydraulic flow through intakes. A decrease in the water level above the pipe intake is commonly observed in this phenomenon [4,5]. Air entrainment can potentially cause complications and mechanical impairments in the operation of hydraulic equipment, including pumps and turbines. Possible issues involve decreased intake discharge, cavitation generation, sound, and pipeline vibration [6].

Hydrodynamics explains flow at the intake by considering it a blend of two components: a sink and a uniformly distributed flow [7]. In reality, the pipe intake placement is regarded as a receptacle, where criticality occurs when the water surface level reaches the revolution point of the Rankine half-body, following the Rankine half-body model [8]. The condition mentioned is called the critical spherical sink surface (CSSS). Critical submergence (S_c) is the term used to describe the vertical space between the uppermost level of an intake and where the free water surface level intersects the Rankine level. S_c displays the trait length scale at which primary air entrainments commence to form within the fluid flow and is deemed equivalent to the radius of the CSSS [9].

An effective technique frequently employed to mitigate the development of air entrainment entails establishing an appropriate distance of submergence above the intake. This approach dictates that the value of the resultant submergence (S) should surpass that of the S_c, thereby reducing the likelihood of unwanted air entrainment [10,11]. To clarify, the reduction of air entrainment may be achieved by positioning the intakes at lower depths and increasing the distance between the water level and pipe intakes (i.e., increasing the value of S) [12]. To reduce construction costs, it is essential to position pipe intakes as close to the water level as possible. The precision of S_c estimation has significant implications for achieving an optimal hydraulic configuration [13].

Numerous researchers have explored the effect of diverse elements on the value of S_c within the literature review, and as a result, several of these investigations have been delineated. The research conducted by Haspolat [14] involved a thorough analysis of the influence of several geometric parameters, such as intake diameter and shape (symmetrical or asymmetrical approach), on the S_C value. The impact on the S_C value between pipe intakes with and without bell mouths was evaluated by Hashid [15], who investigated the entrance shape. The intake pipes used in this study had a circular cross-section and were positioned horizontally. In the study by Khanarmuei [16], the empirical investigation into the positioning angle of both single and dual intakes on S_C revealed that the value of S_C depends on the angle of the pipe intakes. Specifically, it was observed that as the pipes' angle approaches parallelism with the horizon, the strength of the vortices diminishes.

Generally, the examination of airflow within intakes can be accomplished through a combination of theoretical and empirical methodologies. Experimental techniques that rely on the fabrication of tangible models are often both time-consuming and resource-intensive. Hence, it is not invariably feasible to employ physical models. The application of computational intelligence techniques represents a novel approach to enhancing the accuracy of models. Consequently, utilizing the immense potential of these approaches can lead to more precise outcomes. Focusing on computational intelligence methods, numerous studies examining S_c were analyzed in this study. Roushangar [17] conducted a research study focused on evaluating the effectiveness of a variety of artificial intelligence methodologies, namely the support vector machine (SVM), adaptive neuro-fuzzy inference system (ANFIS), and gene expression programming (GEP), to estimate the value of S_c. The findings of the study suggest that the Support Vector Machine (SVM)-derived S_C value outperforms other predictors in terms of performance. Azamathulla [18] implemented a genetic algorithm (GA) to forecast the S_C factor across varying intervals of bottom clearances and subsequently contrasted the resultant predictions with experimental equations. The study's findings indicate that the predictions generated through Genetic Algorithms show superior performance compared to the empirical formulas applied in experimentation. Goel [19] utilized a backpropagation Artificial Neural Network (ANN) and the M5 decision tree algorithm for forecasting the S_C value in horizontal intakes. The research results indicate that these approaches outperformed regression techniques and empirical formulas.

Table 1 summarizes recent studies on critical submergence prediction, highlighting the models, datasets, performance, and limitations. While previous works achieved R² up to 0.95, none employed automated meta-heuristic optimization to tune ensemble models across multiple intake configurations.

Table 1: The statistical traits of input and output

Indicator	Bottom clearance type	Statistical Properties				
		S_c	U_i/U_∞	Re	We	Fr
Max.	$C=0$	4.847	25.377	100532.1	10056.21	6.550
	$C=d_i/2$	5.912	26.875	102207.7	10394.21	6.659
Min.	$C=0$	0.195	0.958	25582.43	272.3975	0.451
	$C=d_i/2$	0.084	0.804	21026.65	184.0178	0.371
Avg.	$C=0$	1.750	6.797	55928.43	2943.524	2.541
	$C=d_i/2$	2.009	7.146	58399.34	3356.101	2.719
St. Dev.	$C=0$	0.990	5.271	20899.2	2840.407	1.955
	$C=d_i/2$	1.369	5.675	24002.29	3208.336	2.127

The formation of air vortices, as suggested by this study, is a highly complex hydraulic phenomenon influenced by various parameters, making it extremely challenging to establish a relationship between the parameter S_c and other factors. Despite attempts to extract experimental relationships and develop comprehensive models, the accuracy and applicability of these models have been limited by the complexity of air entrainment and vortex formation, as well as the exclusion of certain parameters. Therefore, computational intelligence methods may provide a more precise estimate of S_c. This study aims to predict S_c employing GB and XGB frameworks. To improve the accuracy of the models, they were coupled in hybrid forms with novel optimization algorithms (AHA, AVO, TDO, SMO, and TFWO) that adapt the models to the problem's physical properties. Ensemble approaches using the Dempster-Shafer Theory have also been employed to improve prediction accuracy, resulting in an optimum equation targeted at increasing model dependability and accuracy. By comparing the outcomes of the hybrid and ensemble-hybrid frameworks, it was found that the suggested models notably improved predictive accuracy for Sc. Furthermore, the flexibility of these hybrid and ensemble-hybrid models, which allows for determining key solver model parameters, enables them to be applied to various problems.

The primary objective of this study is to evaluate the effectiveness of hybrid ensemble frameworks combining boosting models and meta-heuristic optimization algorithms for predicting critical submergence (Sc) in horizontal circular intakes. Specifically, the study addresses the following research questions:

1. **RQ1:** Can hybrid ensemble models (GB/XGB combined with meta-heuristic optimization) improve the accuracy of critical submergence predictions compared to standard boosting methods?

2. **RQ2:** How do different meta-heuristic algorithms (AHA, VAO, TFWO, SAO, TDO) affect the performance of ensemble models in capturing nonlinear hydraulic and geometric dependencies?
3. **RQ3:** Does the incorporation of a custom empirical adjustment equation enhance the physical consistency and reliability of predictions under varying bottom clearance scenarios?

The corresponding hypotheses are:

- **H1:** Hybrid ensemble models with automated hyperparameter optimization achieve higher R^2 and lower RMSE than non-optimized boosting models.
- **H2:** Among the meta-heuristic algorithms, VAO provides the most effective exploration of the hyperparameter space, resulting in superior model performance.
- **H3:** Including a custom empirical adjustment equation reduces prediction errors in partially elevated intake scenarios (C = d_i/2), where standard models typically underperform.

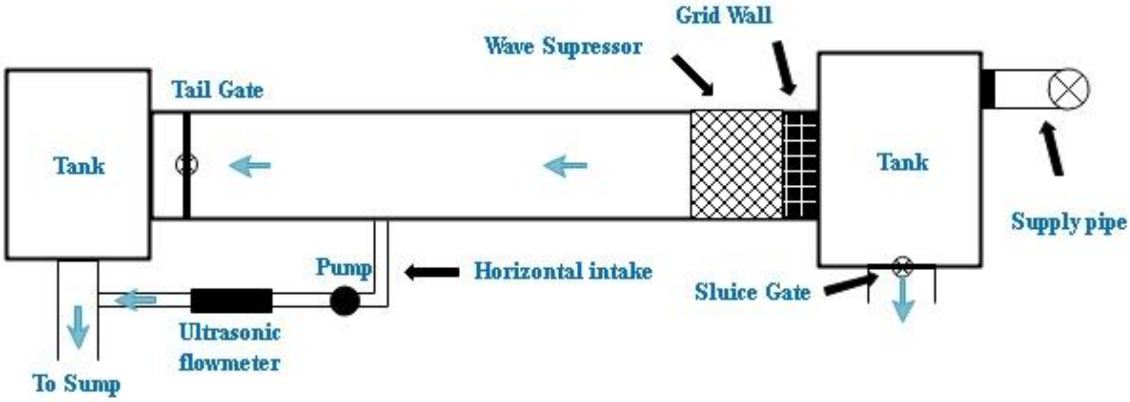
The study uses a controlled laboratory dataset comprising 324 measurements across two intake configurations: zero bottom clearance (C = 0) and half-diameter clearance (C = d_i/2). Four key input variables—intake diameter, approach velocity, Froude number, and bottom clearance ratio—were selected based on their physical relevance to critical submergence formation. Gradient Boosting (GB) and Extreme Gradient Boosting (XGB) were employed as base learners. Hyperparameters for each model were optimized using five meta-heuristic algorithms (AHA, VAO, TFWO, SAO, TDO), enabling systematic exploration of the search space. The dataset was partitioned into 70% training and 30% testing, with 10-fold cross-validation applied to the training set for hyperparameter selection and overfitting prevention.

A custom empirical adjustment equation was incorporated to enhance the physical fidelity of predictions, particularly in partially elevated intake scenarios (C = d_i/2). Model performance was evaluated using R^2 and RMSE on the independent test set, facilitating direct comparison with existing state-of-the-art methods. This research design integrates systematic experimental data collection, robust machine-learning frameworks, and meta-heuristic-driven optimization, providing a comprehensive evaluation of predictive methods for critical submergence.

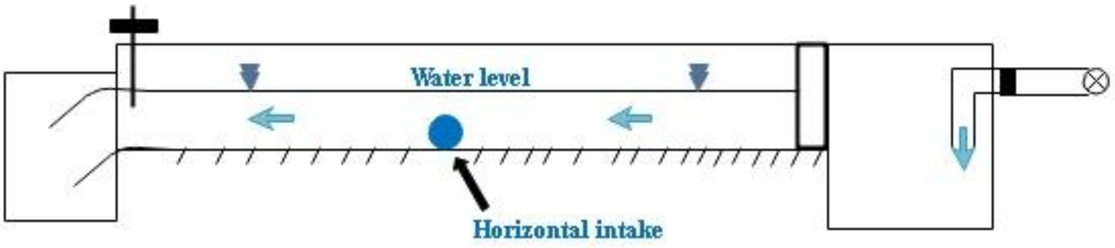
2 Processing of materials and data

An experimental analysis of S_c of horizontal intakes in open channel flow was performed by Ahmad et al. [20], as depicted in Fig. 1. The study focused on an S_c open channel flow horizontal intake. The open channel consisted of a flume that was 10 meters in length, 0.37 meters in width, and 0.6 meters deep. The horizontal intake was positioned laterally, at a distance of 5 meters from the inlet of the flume, in the direction of the flow.

Several parameters were considered in the experimental study, including the intake diameter (d_i) with values of 4.25m, 6.25m, and 10.16m, and the bottom clearance, which was either set to $C = 0$ or $C = d_i/2$. Other parameters studied included the velocity of flow in the intake (U_i), the initial flow velocity in the flume (U_∞), the intake discharge (Q_i), and the S_c. Additionally, the Reynolds number (Re), Weber number (We), and Froude number (Fr) are defined as the input parameters. However, laboratory ranges may have inherent limitations that restrict the range of dimensional parameters that can be studied. To overcome this limitation and ensure that the model is applicable in real-world scenarios with extensive parameters, it is essential to consider non-dimensional parameters and the normal form of S_c. While accounting for the simultaneous effects of multiple-dimensional parameters. Table 2 presents the statistical characteristics of input and output variables, including their maximum, minimum, average, and standard deviation.



(a)



(b)

Figure 1: Experimental setup [24], a) Plan, b) Elevation.

Table 2: Assessment metric equation and ideal state.

Mathematical Equations	Ideal performance	No. Eq.
$R^2 = \left(\frac{\sum_{i=1}^{n} (b_i - \bar{b})(m_i - \bar{m})}{\sqrt{\left[\sum_{i=1}^{n} (b_i - \bar{b})^2\right] \left[\sum_{i=1}^{n} (m_i - \bar{m})^2\right]}} \right)^2$	A higher value near 1	(55)
$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (m_i - b_i)^2}$	A lower value near 0	(56)
$MAE = \frac{1}{n} \sum_{i=1}^{n} \lvert b_i - m_i \rvert$	A lower value near 0	(57)
$IOA = 1 - \frac{\sum_{i=1}^{n} (m_i - b_i)^2}{\sum_{i=1}^{n} (\lvert b_i - \bar{b} \rvert + \lvert m_i - \bar{m} \rvert)^2}$	A higher value near 1	(58)
$n20 - index = \frac{n20}{n}$	A higher value near 1	(59)

Several preprocessing steps were applied before model development to ensure data consistency and to support the learning process. All continuous variables were normalized to a 0–1 range using min–max scaling to maintain comparable influence across features and to stabilize gradient-based learning. No missing data values were reported in the original dataset; therefore, no imputation procedures were required. Visual inspection using box plots and leverage plots indicated no statistically influential outliers with adverse effects on model training. The dataset was then divided into 70% training and 30% testing portions, and a 10-fold cross-validation procedure was applied only to the training subset to tune hyperparameters and prevent overfitting. Taken together, these preprocessing steps ensured that the input variables preserved physical interpretability while meeting the numerical requirements of the boosting algorithms.

Furthermore, Fig. 2 illustrates the polar contour, representing the impact of input variables on outputs, where the effect on outputs is determined by r and θ. In fluid system design, the clearance (C) between the bottom of a horizontal intake and the channel or pipe floor is a critical factor for $\mathrm{S_c}$. A clearance of $\mathrm{C} = 0$ indicates direct contact between the intake and the channel or pipe floor, with no gap or space. If the clearance is too small, it can cause turbulence in the fluid flow, negatively affecting the system's performance. Conversely, if the clearance is too large, it can increase the risk of air being drawn into the flow, leading to cavitation and damage to the system. Experimental studies of fluid systems often involve testing different clearance values to determine the optimal value for a specific system. By setting $\mathrm{C} = 0$, researchers can examine the effect of the intake clearance on the system's performance when there is no gap or space between the intake and the floor. This configuration can provide valuable insights into the optimal clearance values for a particular system. In addition, Fig. 3 shows the correlation between the input and output variables.

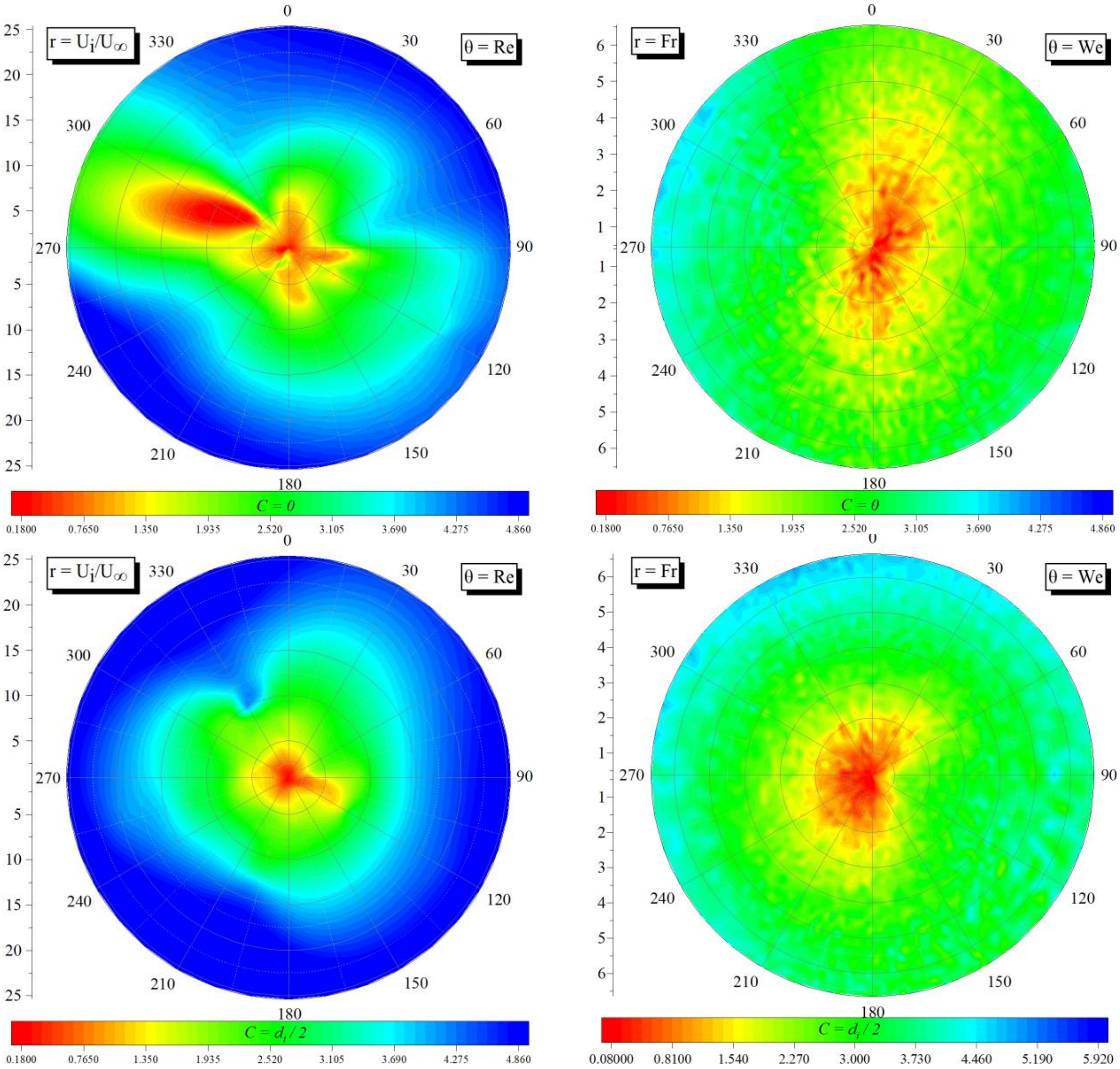


Figure 2: The polar contour of input variables compares to outputs.

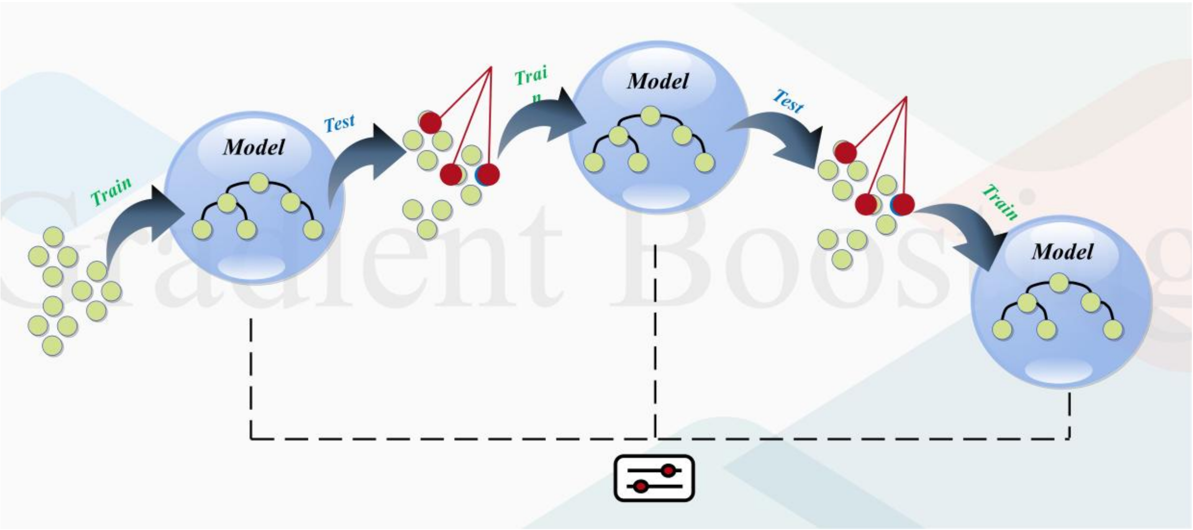


Figure 3: Structure of GBR.

Moreover, in the analysis of S_c n fluid systems, when $\mathrm{C} = \mathrm{d_i}\ /2$, it displays that the clearance between the floor of a channel or pipe and a horizontal intake is half of the intake diameter (d_i). This clearance is indispensable in fluid system design, as it involves elevating the intake by a distance equal to half its diameter. If the clearance is too small, it can lead to turbulence in the fluid flow, negatively impacting system performance. On the contrary, if the clearance is too large, it can increase the risk of air being drawn into the flow, leading to cavitation and system damage. Researchers typically observe various clearance values in experimental studies to identify the ideal clearance for a given system. The $C = d_i\ /2$ configuration raises the intake off the floor by half its diameter, aiding in improving fluid flow and reducing turbulence. This setup provides insights into how intake clearance affects system performance and can help identify the optimal clearance for a specific application.

The dimensional parameters in laboratory ranges often have inherent limitations, which can restrict the range of data used. To ensure the framework's suitability for practical, real-world utilization and overcome this limitation, it is necessary to consider additional dimensional parameters and their combined effects, along with non-dimensional parameters. The normal form of S_C can be presented as follows to address these issues:

$$\frac{S_c}{d_i} = f\left(Fr = \frac{U_i}{\sqrt{g\, d_i}}, Re = \frac{\rho\, U_i\, d_i}{\mu}, We = \frac{\rho\, {U_i}^2 d_i}{\sigma}, \frac{C}{d_i}, \frac{U_i}{U_\infty}\right) \tag{1}$$

Additionally, Ahmed et al. [21] suggested two computational functions for estimating S_C values, each tailored to specific types of bottom clearance.

$$\frac{S_c}{d_i} = 0.042 Fr^{-0.49} We^{0.31} \left(\frac{U_i}{U_\infty}\right)^{0.9}, \quad C = 0 \tag{2}$$

$$\frac{S_c}{d_i} = 0.042 Fr^{-0.211} Re \left(\frac{U_i}{U_\infty}\right)^{0.206}, \quad C = d_i/2 \tag{3}$$

3 Methodology

3.1 Developed models

3.1.1 GB

GB is an ML approach employed for regression and categorization tasks, which combines weak estimation frameworks, often by decision trees (DTs) [22,23], to build a prediction model. In GBR, each regression tree learns the residual of the previous tree's conclusion, aiming to reduce the overall approach residual along the gradient direction. By integrating the outcomes of all regression trees [24], GBR is capable of handling mixed data types and is resilient to outliers [25]. Originally suggested to boost a cost function [26], it has found applications in regression, energy theft detection [27], and various fields in statistics and artificial intelligence (AI).

GBR is a dynamic, enhancing technique that constructs a powerful regression model by iteratively combining weak regression learners. It aims to diminish the loss function using gradient descent, calculated from the addition of weak learners. This loss function evaluates how effectively the coefficients of an effective framework can approximate the underlying data instances.

$$F_m(x) = F_{m-1}(x) + \rho_m h_m \tag{4}$$

$F_{m-1}(x)$ displays the previous ensemble framework, and h_m shows the initial learner used to diminish the loss function L. The base learner h_m is trained on the dataset $\{(\mathrm{x}_i, r_{im})\}_{i-1}^{n}$, and the multiplier is determined by solving a one-dimensional enhancement issue.

$$\rho_m = argmin_\rho \sum_{i-1}^{n} L(y_i, F_{m-1}(x_1) + \rho h_m(x_i)) \tag{5}$$

y_i displays the label for the target class. Additionally, Fig. 4 illustrates the structure of GBR.

Algorithm 1: Pseudocode outlining the Gradient Boosting Regression

```
Input: training set $\{(\mathrm{x}_i, r_{im})\}_{i-1}^{n}$, distinguishable loss function $L(y, F(x))$, count of cycles $M$.
Output: trained GBR framework $F_m(x)$.
Begin
        1.      Setting the initial value of a model to a constant:
.$F_0(x) = argmin_\rho \sum_{i-1}^{n} L(y_i, \rho)$
        2.      Repeating for $m = 1$ to $M$:
            2.1. Calculating the pseudo-residuals in the following manner:
    4   .$\mathrm{r_{im}} \left[\frac{\partial \mathrm{L}(\mathrm{y_i}, \mathrm{F}(\mathrm{x_i}))}{\partial \mathrm{F}(\mathrm{x_i})}\right]_{F_m(x) - F_{m-1}(x)}$, for $i = 1, \dots, n$
            2.2. Training a base learner (e.g., tree) $h_m(x_i)$ on pseudo-residuals involves fitting it to the training data
                 $\{(\mathrm{x}_i, r_{im})\}_{i-1}^{n}$
            2.3. Finding a multiplier $\rho_m$ involves answering a one-dimensional enhancement issue:
    5   .$\rho_m = argmin_\rho \sum_{i-1}^{n} L\,(y_i, F_{m-1}(x_i) + \rho h_m(x_i))$
            2.4. Updating the model:
    6   .$\mathrm{F_m(x) = F_{m-1}(x) + \rho_m h_m(x)}$
        3.      Getting a trained GBR framework $F_m(x)$
End
```

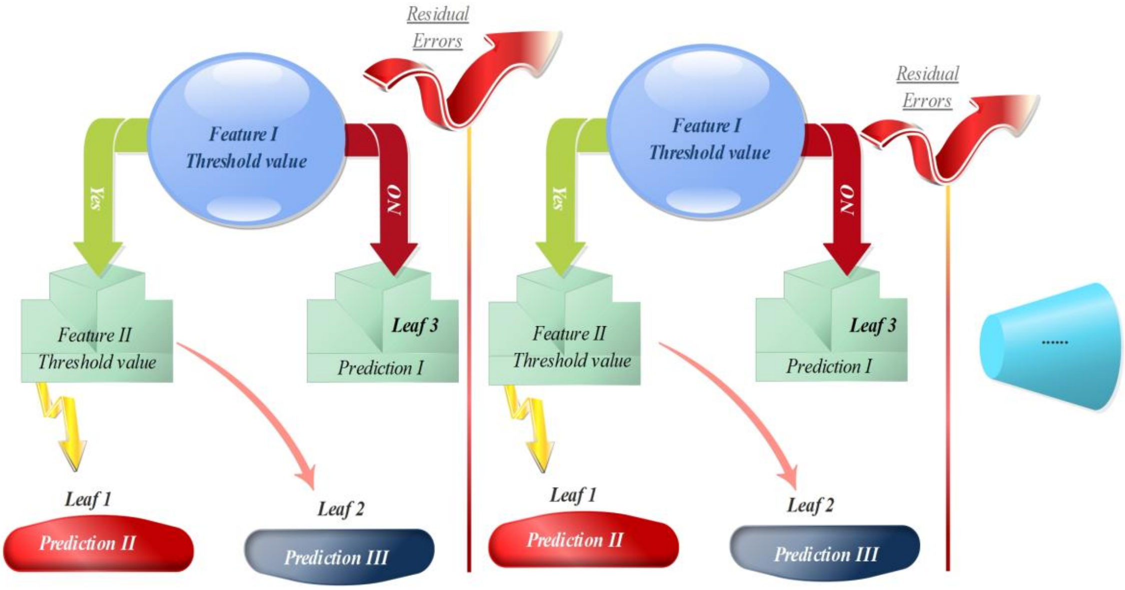


Figure 4: Structure of XGB.

3.1.2 Extreme Gradient Boosting (XGB)

XGB is a powerful gradient-boosting machine technique known for its adaptability in supervised learning tasks, including regression and classification [28,29]. It is popular among data science practitioners due to its fast execution speed, facilitated by its out-of-core processing capabilities. It generates predicted outputs $\hat{y}_i$ using a series of equations for ensemble tree models within a dataset $S = \{(X_i, y_i): i = 1 \dots n, x_i \in R^m, y_i \in R\}$ with n instances and m features.

$$\mathrm{A}_{\cdot\mathrm{i}} = \emptyset(\mathrm{X_i}) = \sum_{\mathrm{T}=1}^{\mathrm{T}} \mathrm{f_t}(\mathrm{X_i}), \qquad \mathrm{f_t} \in \mathrm{F} \tag{6}$$

T displays the count of trees, while f_t depicts the t^{th} tree. The main goal in solving the given equation is to diminish both the loss and regularization criteria by finding the optimal set of functions.

$$\varphi(\emptyset) = \sum_i l(y_i, A_{\cdot i}) + \sum_T \Omega(f_t) \tag{7}$$

The l (loss function) is categorized by the disparity between the measured output y_i and the estimated output $\hat{y}_i$. Ω shows a metric evaluating the complexity of the model to mitigate overfitting as presented in Eq. (8):

$$\Omega(f_t) = \gamma L + \frac{1}{2}\lambda \|w\|^2 \tag{8}$$

In this equation, the importance given to each leaf is represented by w, and the total number of leaves is denoted as L. Decision trees are iteratively refined during model training to reduce the objective value progressively. As the model learns, it incorporates a new function, typically in the form of a tree.

$$\varphi^{(\mathrm{t})} = \sum_{i=1}^{n} 1(y_i, A_{\cdot i}{}^{(t-1)} + f_t(X_i)) + \Omega(f_t) \tag{9}$$

$$\Omega_{\text{split}} = \frac{1}{2}\left[\frac{\left(\sum_{i \in I_L} g_i\right)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{\left(\sum_{i \in I_R} g_i\right)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{\left(\sum_{i \in I} g_i\right)^2}{\sum_{i \in I} h_i + \lambda}\right] \tag{10}$$

$$\mathrm{g_i} = \partial_{A^{t-1}} 1(y_i, A.^{(t-1)}) \tag{11}$$

$$h_{\mathrm{i}} = \partial^2{}_{A^{t-1}} 1(y_i, A.^{(t-1)}) \tag{12}$$

Fig. 5 shows the structure of XGB.

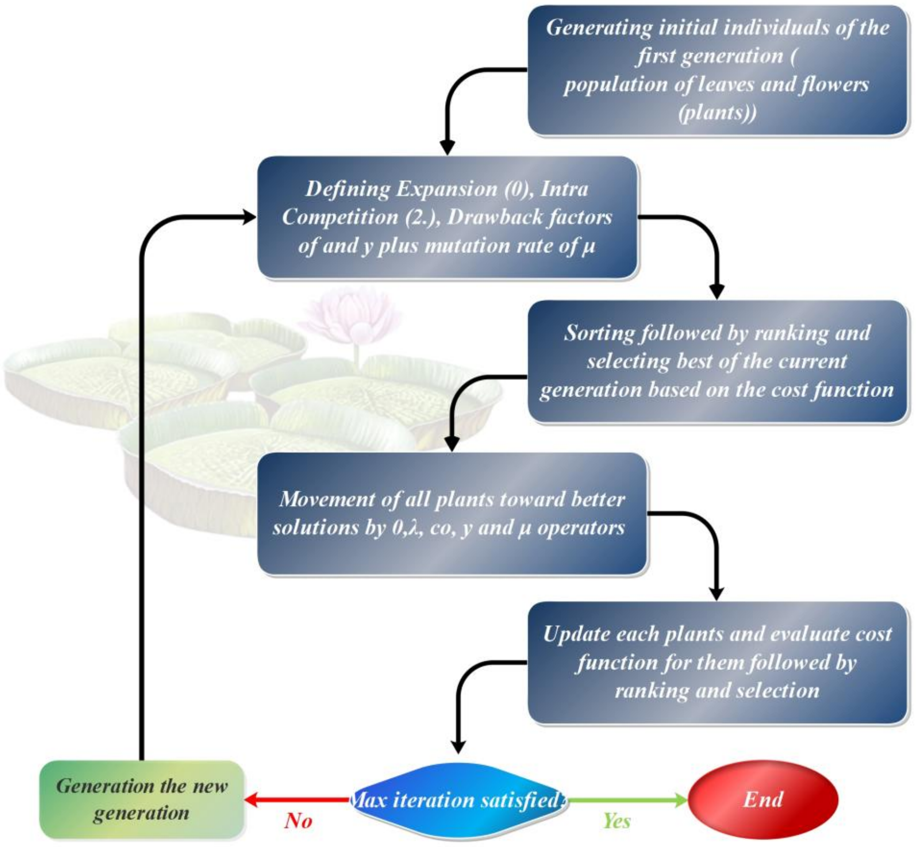


Figure 5: The flowchart of VAO.

3.2 Meta-heuristic algorithms

3.2.1 Models via VAO

The VAO algorithm primarily focuses on how the original population, consisting of Leaves and Flowers, disperses and how each of these entities can spread or grow along the outer wall [30].

Subsequently, a mutation arises due to cross-pollination between a distinct variety of water lilies and the beetles within the pond. The current phenomenon is represented by η and is known as a hybrid mutation. There is a 0.2% chance that this change will manifest as a positive or negative trajectory with each subsequent generation of children. The ideal leaf sample is characterized by its large size and powerful physical traits, which we will refer to as α. Besides, the VAO method is shown in the pseudo-code below, and the VAO technique's flowchart is shown in Fig. 6.

$$VOA = \sum_{i=1}^{n}\sum_{j=1}^{n}(x_{ij}[\xi_{ij},\Gamma_{\mathrm{ij}}] + \Theta + \varphi) \times (\eta) \qquad (13)$$

Algorithm 2. Pseudo Code of VOA

```
Start
Developing population of plants $xi$ $(i = 1,2,\ldots,n)$
Identify Expansion $\xi i$ in $xi$
Identify Intra Competition $\Gamma$i in $xi$
Identify the Drawback coefficient of $\varphi$ in xi (random range in [0.1 to 0.3])
Identify the Drawback coefficient of $\Theta$ in xi (random range in [0.1 to 0.3])
Identify Hybrid Mutation Rate of $\eta = 0.2$
While Max cycles are not met
For $i = 1\ to\ n$ plants
For $j = 1\ to\ n$ plants
If $\xi i > \xi j\ or\ \Gamma i > \Gamma j$ for $xi$ $(i = 1,2,\ldots,n)$
Plant $i$ goes planet $j$
End if
Employ hybrid mutation $\eta$
```

```
Employ Drawback coefficient $\varphi$ and $\Theta$
Examine new resolutions by cost function and update Expansion
End
End
Sort and rank plants and find the current global best
Developing new generation
End of while
End
```

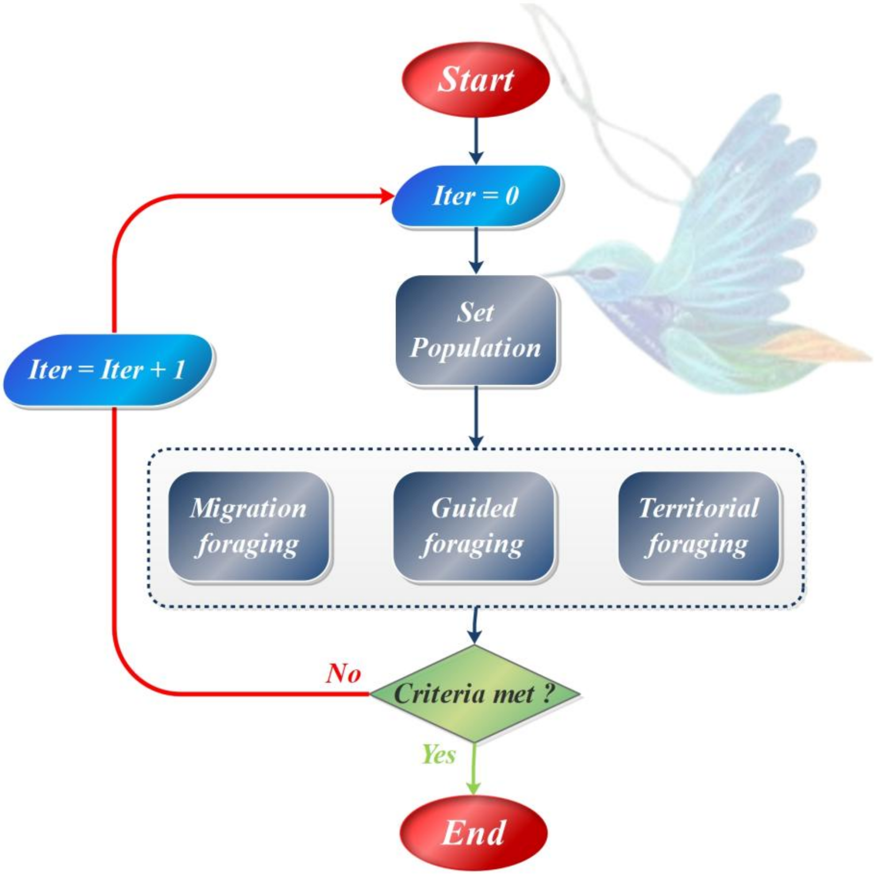


Figure 6: The flowchart of AHA.

3.2.2 Models via AHA

The AHA is a computational technique that draws inspiration from hummingbirds' foraging and flight behaviors [31]. The present study outlines three principal models of this algorithm [32].

- *Guided foraging*

In the present foraging model, three distinct flight behaviors, namely axial flight, diagonal, and omnidirectional, have been employed for foraging. A mathematical expression that models the process of guided foraging along with a potential food source can be derived by employing the following technique.

$$v_i(t+1) = x_{i,ta}(t) + h.b.\left(x_i(t) - x_{i,ta}(t)\right) h \sim N(0,1) \quad (14)$$

Eq. (14) defines $x_{i,ta}(t)$ as the spatial coordinates of the desired food source, with h signifying the guiding parameter [33]. Additionally, $x_i(t)$ displays the location of the ith food source in the temporal domain, which illustrate in Eq. (15):

$$x_{Ai}(t) = \begin{cases} x_i(t) & f(x_i(t)) \leq f(v_i(t+1)) \\ v_i(t+1) & f(x_i(t)) > f(v_i(t+1)) \end{cases} \quad (15)$$

- *Territorial foraging*

The taking after condition speaks to the nearby look of hummingbirds within the territorial foraging procedure:

$$v_i(t+1) = x_i(t) + g.b.\left(x_i(t)\right) g \sim N(0,1) \quad (16)$$

- *Migration foraging*

The concept of migration foraging has been formulated by Eq. (17):

$$x_{war}(t+1) = lb + r.(ub - lb) \quad (17)$$

The variable x_{war} is used to denote the source of sustenance possessing the poorest population rate of nectar refilling. Meanwhile, a factor denoted as r is assumed to be random, and lower limit lb and upper limit ub ranges are established. The AHA flowchart is presented in Fig. 7 [34].

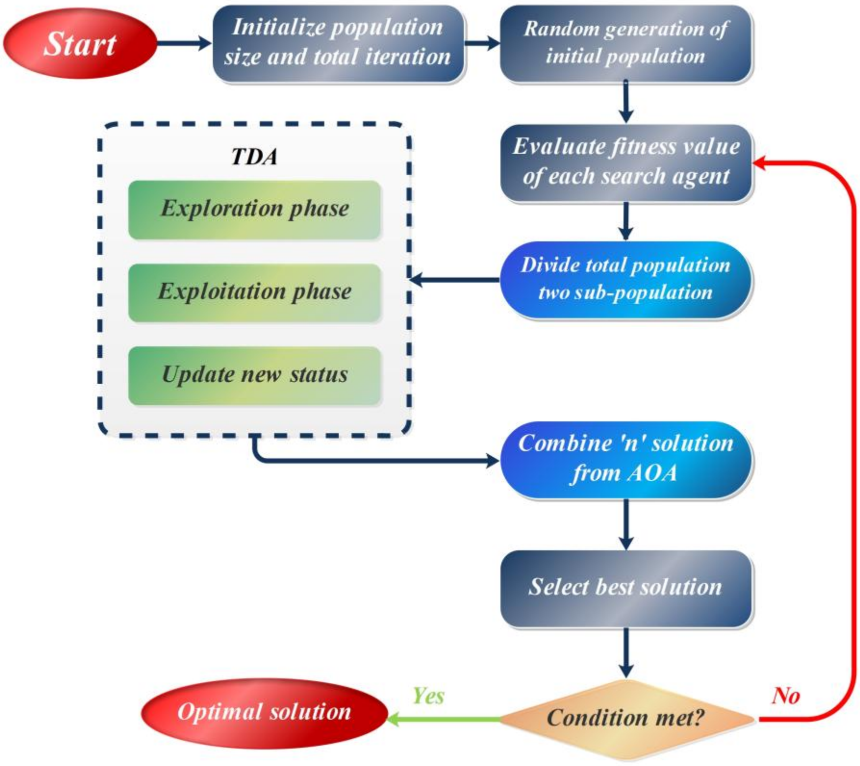


Figure 7: Flowchart of TDO

3.2.3 Model via TDO

The algorithm of TDO [35] is a probabilistic, population-based method employing Tasmanian devils as search agents. The TDO population can be envisioned as a matrix, as described in Eq. (11):

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{N\times m} = \begin{bmatrix} X_{1,1} & \cdots & X_{1,j} & \cdots & X_{1,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{i,1} & \cdots & X_{i,j} & \cdots & X_{i,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{N,1} & \cdots & X_{N,j} & \cdots & X_{N,m} \end{bmatrix}_{N\times m} \quad (18)$$

m displays the count of parameters in the provided issues.

X depicts the Tasmanian devil population, X_i indicates the candidate or potential solution at position i and $X_{i,j}$ portrays the worth of the $j-th$ parameter for the $i-th$ applicant answer.

N displays the identifier for the Tasmanian devil participating in the exploration procedure.

The resultant objective function values are then displayed in the form of a vector, as illustrated in Eq. (19):

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_N \end{bmatrix}_{N\times 1} = \begin{bmatrix} F(X_1) \\ \vdots \\ F(X_i) \\ \vdots \\ F(X_N) \end{bmatrix}_{N\times 1} \quad (19)$$

F displays the outcomes produced by the objective operation and F_i displays the outcomes of the objective function corresponding to the $i-th$ instance.

The top-performing solution yields the maximum objective function score and is considered the most favorable option among the population. The TDO population updating process is based on two Tasmanian devil-feeding approaches, each with an assumed 50% selection probability. Accordingly, in each TDO cycle, each Tasmanian devil efficiently uses one of these plans.

- ***Strategy 1:*** *Consuming flesh for nourishment (Exploration Phase)*

The TDO method systematically explores diverse search areas for optimal solutions. Eq. (20) simulates the random selection of a scavenging site by the $i-th$ Tasmanian devil, excluding self-selection, using a random value k from the range 1 to N, excluding i.

$$C_i = X_k, \quad i = 1,2,\dots,N\ , \quad k \in \{1,2,\dots,N | k \neq i\} \tag{20}$$

C_i means the flesh nominated by the $i-th$ Tasmanian devil.

The current position is retained, as per Eq. (21).

$$x_{i,j}^{new,S1} = \begin{cases} x_{i,j} + r.\left(c_{i,j} - I.x_{i,j}\right), & F_{C_i} < F_i^* \\ x_{i,j} + r.\left(x_{i,j} - c_{i,j}\right), & otherwise \end{cases} \tag{21}$$

$$X_i = \begin{cases} X_{i,j}^{new,S1} & ,F_{i,j}^{new,S1} < F_i \\ X_i, & else \end{cases} \tag{22}$$

r depicts a randomly produced amount within the variety of [0,1].

F_{C_i} displays the objective function score linked to the chosen carrion.

I indicates an arbitrary numerical value with options limited to one or two.

$F_{i,j}^{new,S1}$ specifies the value of the objective function.

$x_{i,j}^{new,S1}$ displays the altered location of the Tasmanian devil at index i after employing the initial approach and $X_{i,j}^{new,S1}$ displays the $j-th$ variable's value at the fresh location of the Tasmanian devil indexed as i after implementing the first strategy.

- ***Strategy 2:*** *Using prey as a nutritional resource (Exploitation Phase)*

Tasmanian devils employ a two-phase prey strategy: initially tracking and attacking, followed by immobilizing and consuming. The primary strategy's selection process is mirrored in the initial stage of the secondary approach, as detailed in Eq. (23), which outlines the target selection method.

$$P_i = X_k, \quad i = 1,2,\dots,N\ , \quad k \in \{1,2,\dots,N | k \neq i\} \tag{23}$$

k signifies a randomly generated value falling within the interval from 1 to N and P_i depicts the particular quarry chosen by the Tasmanian devil indexed as i.

When finding prey, the Tasmanian devil adjusts its position based on the prey's performance, approaching if it improves and refining its strategy (Eq. 24). In the secondary approach (Eq. 25), a new location is adopted if it enhances the objective function.

$$x_{i,j}^{new,S2} = \begin{cases} x_{i,j} + r.\left(p_{i,j} - I.x_{i,j}\right), & F_{P_i} < F_i^* \\ x_{i,j} + r.\left(x_{i,j} - p_{i,j}\right), & otherwise \end{cases} \tag{24}$$

$$X_i = \begin{cases} X_{i,j}^{new,S2} & ,F_i^{new,S2} < F_i^* \\ X_i, & else \end{cases} \tag{25}$$

$x_{i,j}^{new,S2}$ indicates the updated condition or situation of the Tasmanian devil at index i, as determined through the application of the second approach.

$X_{i,j}^{new,S2}$ signifies the worth of the $j-th$ parameters for the $i-th$ Tasmanian devil when smearing the alternative approach.

$F_i^{new,S2}$ displays the performance metric tied to the Tasmanian devil indexed as i while employing the alternative approach.

F_{P_i} signifies the objective function score associated with the prey chosen by the Tasmanian devil at index i.

Eqs. (26)-(28) mathematically illustrate this prey-chasing phase carried out by the Tasmanian devil.

$$R = 0.01\left(1 - \frac{t}{T}\right) \tag{26}$$

$$X_{i,j}^{new} = x_{i,j} + (2r-1).R.x_{i,j} \tag{27}$$

$$X_i = \begin{cases} X_i^{new} & ,F_i^{new} < F_i^* \\ X_i, & otherwise \end{cases} \tag{28}$$

t symbolizes the cycle tally and X_i^{new} indicates the revised or fresh position of the Tasmanian devil at index i in the vicinity of its prior location, designated as X_i

T depicts the upper limit for the count of repetitions.

F_i^{new} symbolizes the performance metric linked to the Tasmanian devil situated at the recently adjusted location indexed as i.

$X_{i,j}^{new}$ depicts the numerical value for the $j-th$ variable about the Tasmanian devil positioned at its freshly modified location, indexed as i.

R displays the extent of the local region, centered around where the Tasmanian devil commenced its assault.

The TDO approach is available in the flowchart in Fig. 8.

Algorithm 3. Pseudo-Code of TDO Algorithm

```
Initialize the TDO.
Provide the details of the enhancement issue.
Specify the number of cycles ($T$) and the count of population participants ($N$).
Start the place of Tasmanian devils and the impartial purpose's valuation.
For $t = 1:T$
  For $i = 1:N$
    If $Probability < 0.5, Probability = rand$
      Strategy1: (Exploration Phase)
      select flesh for the $i-th$ Tasmanian devil
      Calculate the novel rank of the Tasmanian devil
      Promotion of the $i-th$ Tasmanian devil
    Else
      Strategy 2: (Exploitation Phase)
      Stage one: Selecting prey and initiating an attack
      Select prey for the $i-th$ Tasmanian devil
      Calculate the novel rank of the Tasmanian devil
      Promotion of the $i-th$ Tasmanian devil
```

```
Stage two: Hunt chasing
Promotion neighborhood range
Calculate the novel status of the $i-th$ Tasmanian devil in the neighborhood of Xi
Promotion of the $i-th$ Tasmanian devil
End If
End For $i = 1:N$
Preserve the most promising solution that has been suggested up to this point
End For $t = 1:T$
Output: The best solution attained by TDO for the assumed enhancement issue.
End TDO.
```

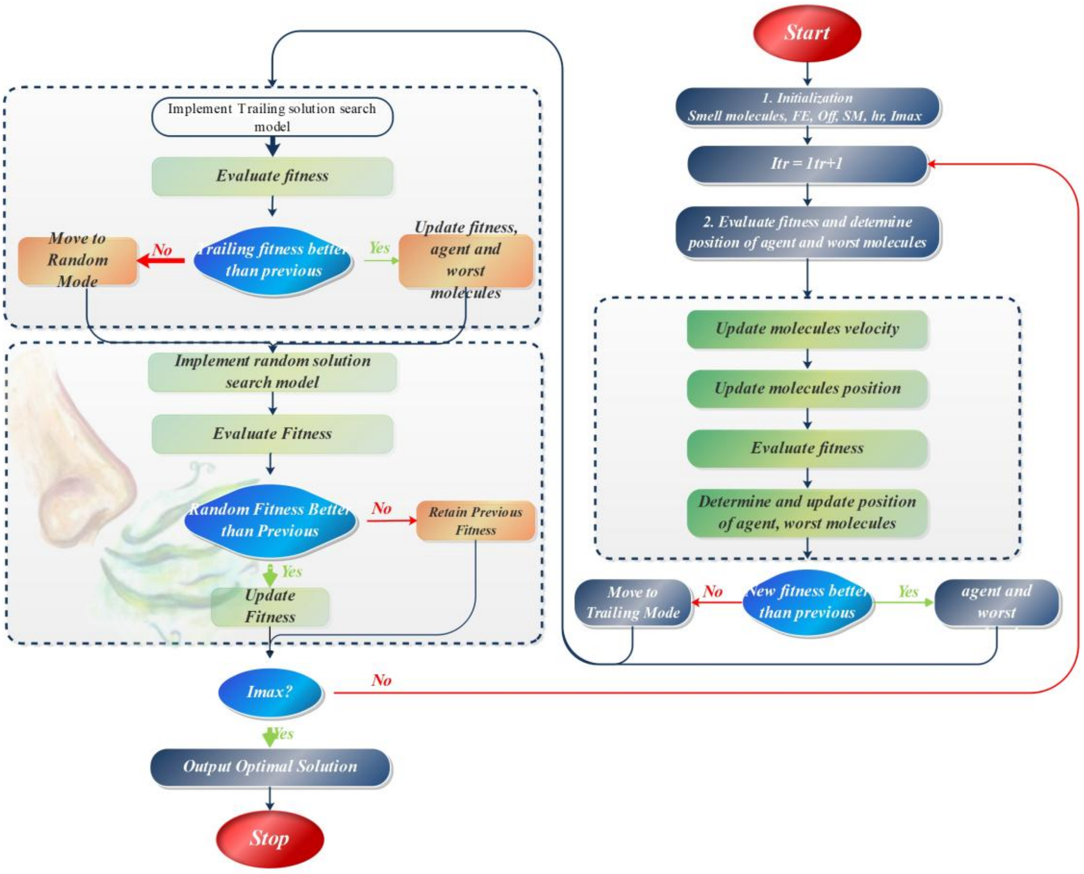


Figure 8: Flowchart of SAO.

3.2.4 Models via SAO

The overall framework of the SAO is built upon three distinct approaches grounded in the fundamental concept of perceiving odors [36].

- *Sniffing mode*

As molecules associated with odor tend to disperse toward the target agent, it is standard practice to commence the procedure by selecting an inconspicuous origin to release these elements. Olfactory compounds could be activated through the use of Eq. (29):

$$x_i^{(t)} = \begin{bmatrix} x_{(1,1)} & x_{(1,2)} & x_{(1,D)} \\ \cdot & \cdot & \cdot \\ x_{(N,1)} & x_{(N,2)} & x_{(N.D)} \end{bmatrix} \tag{29}$$

The symbol X stands for the total count of scent particles, while Y depicts the cumulative count of decision-related factors. The representation of the location vector must be presented in a scholarly writing style:

$$x_i^{(t)} = L_i + r_0 \times (U_i - L_i) \tag{30}$$

The boundaries denoted by the characters A and B are deduced from the decision parameters. The variable r_0 signifies a random variable adhering to a uniform distribution between 0 and 1. The original amount at which scent molecules are emitted from the source is computed using Eq. (31), with each odor molecule assigned a unique numerical value:

$$v_i^{(t)} = \begin{bmatrix} \rho_{(1,1)} & \rho_{(1,2)} & \rho_{(1,D)} \\ \cdot & \cdot & \cdot \\ \rho_{(N,1)} & \rho_{(N,2)} & \rho_{(N.D)} \end{bmatrix} \tag{31}$$

Every fragrance particle embodies a conceivable resolution. The spatial attributes of these possible answers are established from the spatial vector designated in Eq. (31), and the speed of the particles is indicated in Eq. (32). The amplification of particle speed transpires subsequently:

$$x_i^{t+1} = x_i^{(t)} + \rho_i^{t+1} \times \Delta t \quad (32)$$

The formula $\Delta t = 1$ displays an uninterrupted advancement of the agent along the fine-tuned pathway throughout the optimization process. The anticipated spatial positions for the scent molecules are computed using the subsequent expression:

$$x_i^{t+1} = x_i^{(t)} + \rho_i^{t+1} \quad (33)$$

Every aroma particle displays unique diffusion velocities that facilitate motion and evaporation in accordance with exploration principles. The computation of aroma particle speed, which has experienced recent modifications, is carried out through the use of Eq. (34):

$$\rho_i^{t+1} = \rho_i^{(t)} + \rho \quad (34)$$

The representation ρ depicts the adapted velocity obtained from the computation specified in Eq. (35):

$$\rho = r_1 \times \sqrt{\frac{3MR}{k}} \quad (35)$$

M signifies the aroma coefficient, intending to balance the impacts of both mass and temperature on the kinetic energy of fragrance particles. The parameters k and R relate to the scent particles' mass and temperature respectively. The olfaction investigation has been completed, enabling the establishment of the agent's specified position x_{agent}^{t}.

- *Trailing mode*

In the alternate operational mode, the agent's investigative maneuvers are emulated to determine the source of a scent. The agent seizes the opportunity to transition to this novel spot:

$$x_i^{t+1} = x_i^{(t)} + r_2 \times olf \times \left(x_{agent}^{t} - x_i^{(t)}\right) - r_3 \times olf \times (x_{worst}^{t} - x_i^{(t)}) \quad (36)$$

Values r_2 and r_3 fall between 0 and 1. The corrective parameter r_2 diminishes the influence of smell on the state variable x_{agent}^{t}, while the restraining coefficient r_3 confines the impact of smell on the state parameter x_{worst}^{t}.

- *Random mode*

The agent might encounter local low points due to its inability to sustain routes. When this scenario occurs, the agent transitions to a state of uncertainty, as denoted by Eq. (37):

$$x_i^{t+1} = x_i^{(t)} + r_4 \times S \quad (37)$$

The letter S displays the size of the stage. The random penalty, symbolized by the variable r_4, probabilistically reduces the size of S.

The flowchart for the SAO algorithm is presented in Fig. 9.

Algorithm 4 Pseudo-code of SAO

```
Rearrange Bounds
Establish the starting positions for scent molecules
Amount appropriateness
Position the agent and determine the least favorable spot for the molecules
  Whereas  $(Itr < Itr_{max})$ do:
For ($i = 1$ to molecules) do:
For (j=1 to position) do:
            Fine-tune the velocity and position of the molecules (sniffing)
           End For
Amount appropriateness
If (the enhanced suitability is larger) then:
Update appropriateness
Modify the agent and the least desirable molecules
End If
End For
For ($i = 1$ to molecules) do:
For ($j = 1$ to position) do:
update place (trailing)
End For
Amount appropriateness
End For
If (new fitness is better) then:
Offer new fitness evaluations
inform location
Else
For ($i = 1$ to molecules) do:
For ($j = 1$ to position) do:
Activate the random mode
End For
End For
End If
End While
Retrieve the top solution.
```

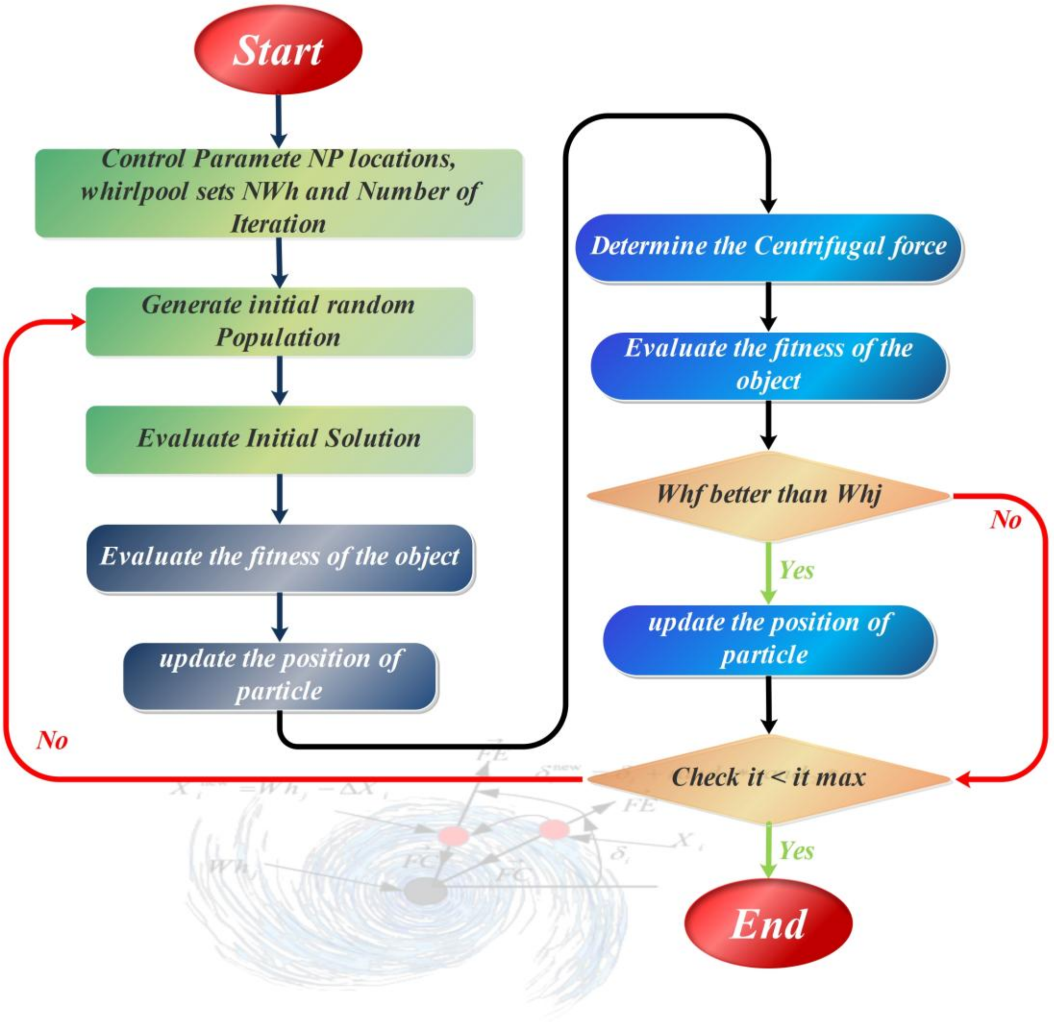


Figure 9: Flowchart of the TFWO.

3.2.5 Models via TFWO

The TFWO is driven by unpredictable fluctuations in water turbulence [37]. Initially, the most robust participant in each category $f(X)$ is acknowledged as the whirlpool (Wh) that attracts the entities. Each Wh functions like a gravitational center and tends to converge the locations of elements within its group (X) with its central point by applying an inward force and guiding them towards it. Consequently, the $j-th\,Wh$ and the local position on Wh_j blend the $i-th$ entity's location (X) with their own. However, the other whirlpools introduce specific discrepancies ΔX_i because of the disparity between them and their target values $(f(X))$. Thus, the $i-th$ entity's position transforms into X_i^{new}, and the entities (X) traverse with an impressive angle (δ) around their whirlpool's core, advancing toward it. Consequently, this angle varies with each cycle in line with Eq. (38):

$$\delta_i^{new} = \delta_i + q_1 \times q_2 \times \pi \tag{38}$$

To compute and represent the most distant and closest vortices ΔX_i Eq. (39) identifies the vortices with the minimal weighted distance from all entities. Subsequently, ΔX_i is determined as per Eq. (40), and ultimately, Eq. (41) is applied to adjust the particle's position.

$$\Delta_p = f\left(Wh_p\right) \times [Wh_p - sum(X_i)]^{0.5} \tag{39}$$

$$\begin{aligned}\Delta X_i = (\cos(\delta_i^{new}) \times q(1,B) \times (Wh_m - X_i) \\ - \sin(\delta_i^{new}) \times q(1,B) \\ \times (Wh_w - X_i) \times (1 \\ + |\cos(\delta_i^{new}) - \sin(\delta_i^{new})|)\end{aligned} \tag{40}$$

$$X_i^{new} = Wh_j - \Delta X_i \tag{41}$$

Wh_m and Wh_w correspond to the vortices with the smallest and largest Δ_p values, respectively, with di representing the angle of the $i-th$ entity.

Occasionally, the centrifugal force (EF_i) prevails over the centripetal force of the whirlpool and haphazardly relocates the entity. The centrifugal force is represented as described in Eq. (42) and occurs randomly in one measurement of the conclusion parameters. To achieve this purpose, the centrifugal force is calculated based on the angle between the whirlpool and the entity, as illustrated in Eq. (43). The centrifugal effect takes place for a dimension that is randomly chosen if this force surpasses a randomly generated value within the interval [0,1], as depicted in Eq. (43). This phenomenon is mathematically defined as follows:

$$EF_i = ((\cos(\delta_i^{new}))^2 \times (\sin(\delta_i^{new}))^2)^2 \quad (42)$$

$$X_{i,t} = X_t^{min} - q \times (X_t^{max} - X_t^{min}) \quad (43)$$

The vortices interact with one another. This process is represented similarly to the way vortices affect entities, where each Wh endeavors to exert a centripetal force on other vortices and attract them. The closest Wh can be calculated mathematically by considering the lowest value and its objective function, as specified in Eq. (44). Subsequently, the $Wh's$ position can be modified using Eq. (45) and (46).

$$\Delta_t = f(Wh_p) \times \left|Wh_p - sum(Wh_j)\right|^{0.5} \quad (44)$$

$$\Delta Wh_j = \left|(\cos(\delta_j^{new}) + \sin(\delta_j^{new})\right| \times q(1, B) \times (Wh_f - Wh_j) \quad (45)$$

$$\Delta Wh_j^{new} = Wh_f - Wh_j \quad (46)$$

δ_j is the $j-th$ hole of Wh angle parameter. Ultimately, when the strong participant surpasses the others in strength, it will be selected as a fresh Wh for the next cycle. The TFWO's flowchart is specified in Fig. 10.

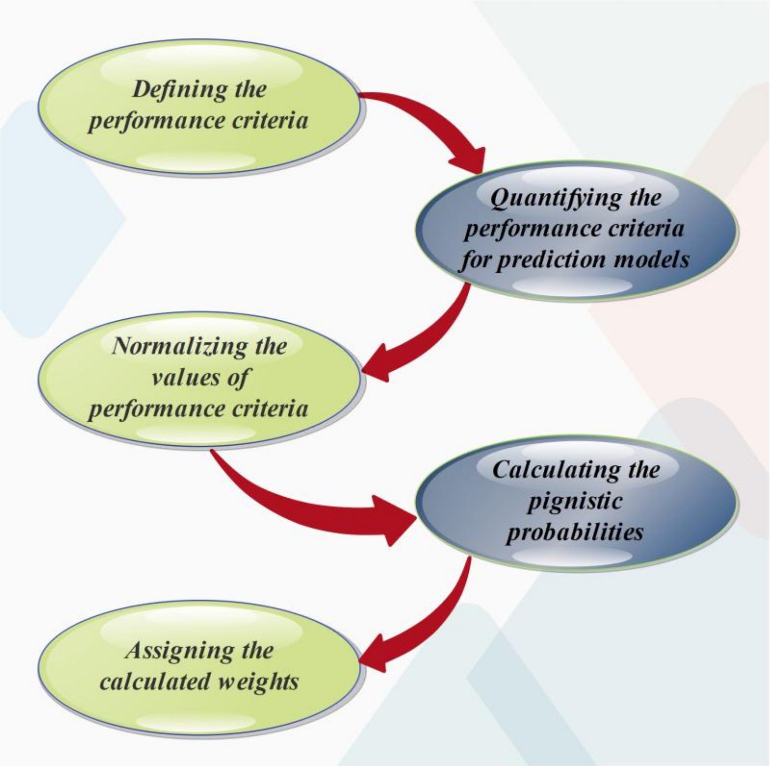


Figure 10: Ensemble models generating process by D-ST.

3.3 Ensemble ML

Ensemble ML combines multiple base models—either homogeneous or heterogeneous—to improve predictive accuracy by reducing variance, bias, or both. Standard ensemble techniques include bagging, boosting, and stacking [38]. Bagging and boosting typically operate on similar models, while stacking can effectively combine diverse model types.

In this study, a hybrid ensemble framework was developed that combines GB and XGB models. To assign appropriate weights to each model's prediction, Dempster-Shafer Theory (D-ST) was employed [39]. D-ST calculates a weight for each model based on its individual performance using a set of indices that reflect prediction accuracy and reliability [40]. The resulting weights are then used to produce a weighted average prediction, which improves over simple averaging by giving more influence to models that perform better under specific conditions, rather than treating all models equally [41,42].

The ensemble output is computed as:

$$Y_{predict(i)}^{ensemble} = \frac{\sum_{j=1}^{n} w_j y_j}{\sum_{j=1}^{n} w_j} \quad (47)$$

where y_j represents the prediction of the j-th model, and w_j is its D-ST-derived weight. By incorporating uncertainty and performance-based weighting, the D-ST ensemble is more robust to model errors and provides improved generalization compared with naive averaging.

A baseline ensemble using simple bagging (equal-weight averaging of multiple GB and XGB models) was also implemented for comparison. Results indicate that the D-ST weighted ensemble consistently outperforms simple bagging in terms of R² and RMSE, particularly under varying bottom clearance and nonlinear hydraulic conditions. This demonstrates that performance-based weighting improves predictive accuracy beyond standard ensemble averaging or stacking strategies [43].

The procedure for generating ensemble predictions using D-ST, including the calculation of model-specific weights and the weighted combination of predictions, is illustrated in Fig. 11. Fig. 12 shows the overall workflow of the hybrid GB-XGB ensemble framework.

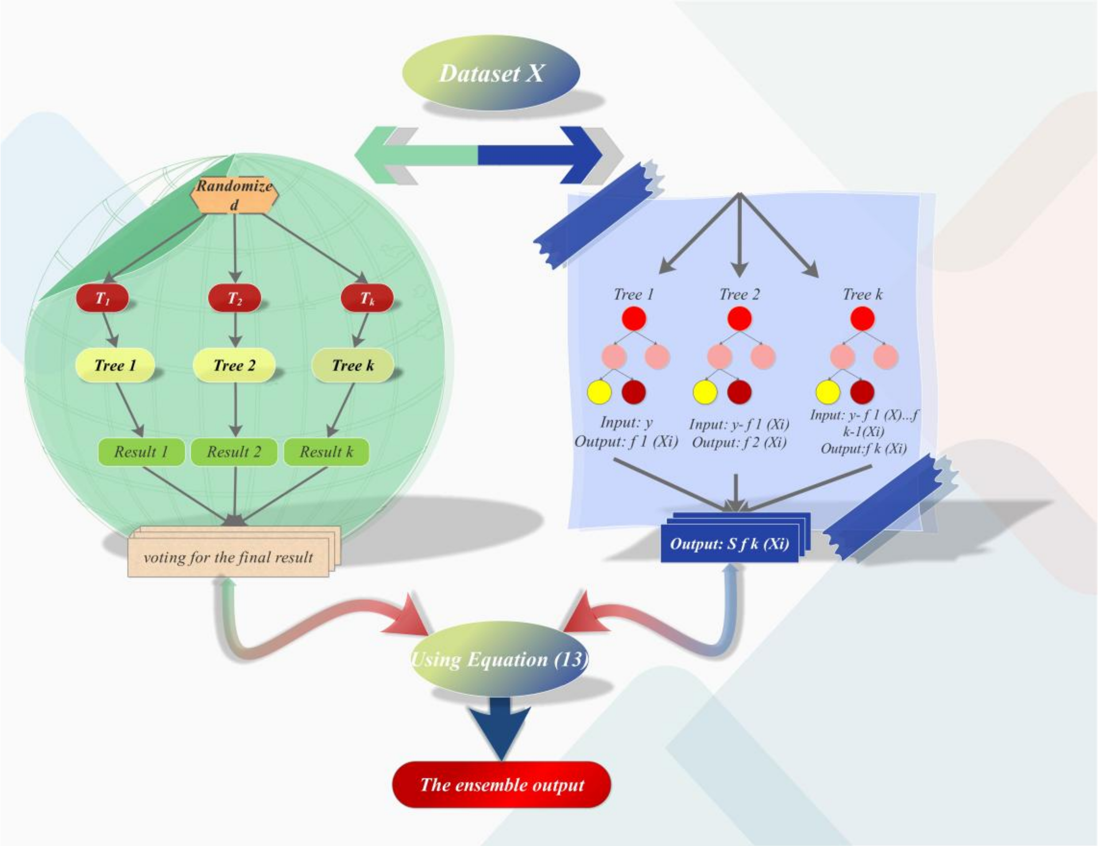
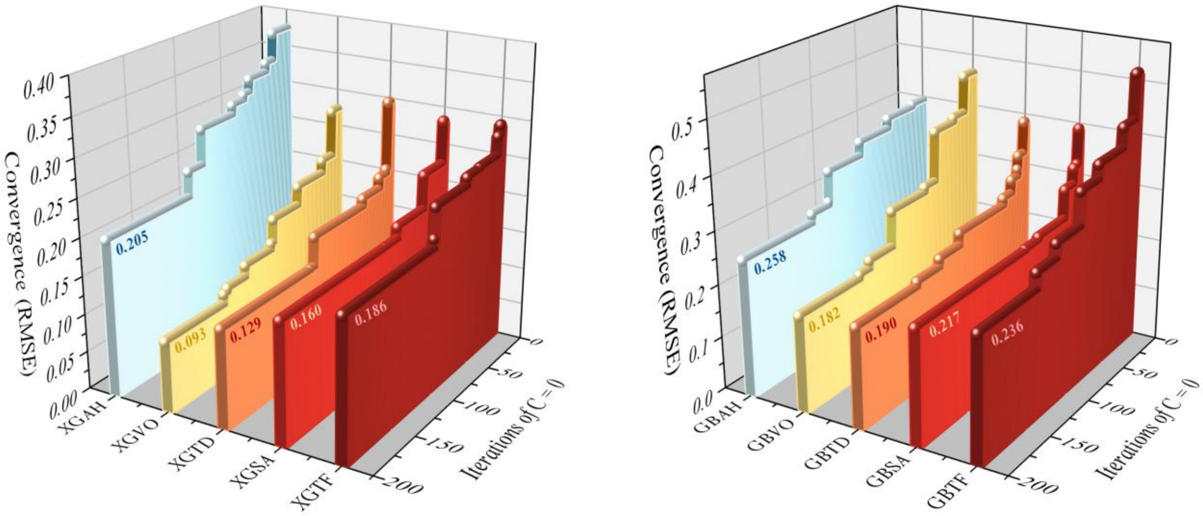


Figure 11: The procedure of presented ensemble models.



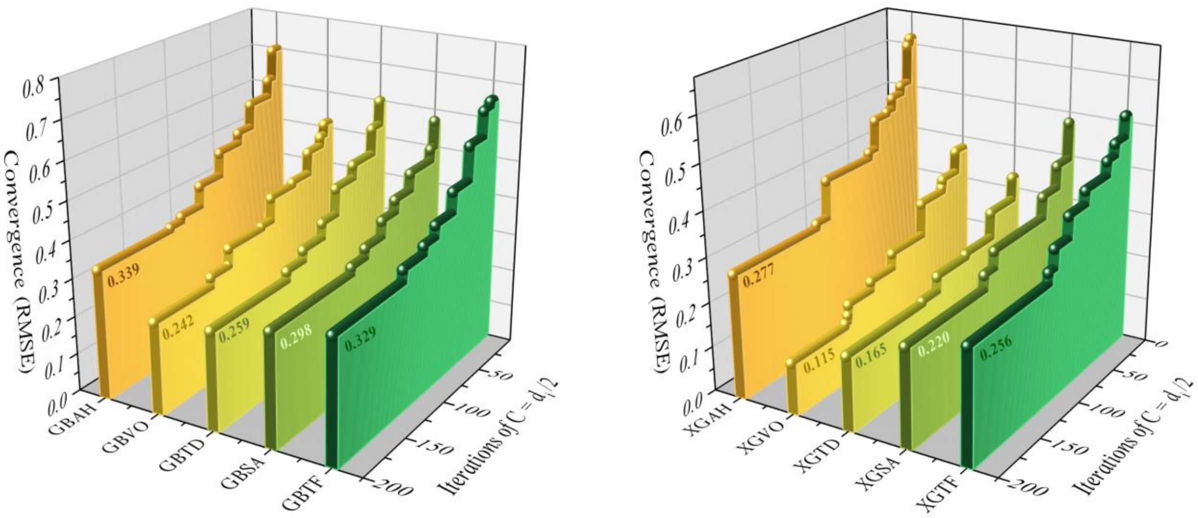


Figure 12: Convergence of hybrid models based on the correlation coefficient.

3.4 Performance evaluators

This section outlines several metrics to evaluate the performance of the suggested models by computing their error levels and correlation coefficients. The metrics discussed include Mean Absolute Error (MAE), Coefficient of Determination (R^2), Root Mean Square Error (RMSE), Index of Agreement (IOA), and n20-index. The formulas for each of these metrics are presented in Table 3. Eqs. (55-59) utilize the n to indicate the sample number, b_i to show the estimated value, $\bar{b}$ and $\bar{m}$ to denote the mean estimated and measured values, alternatively, and m_i to indicate the measured value.

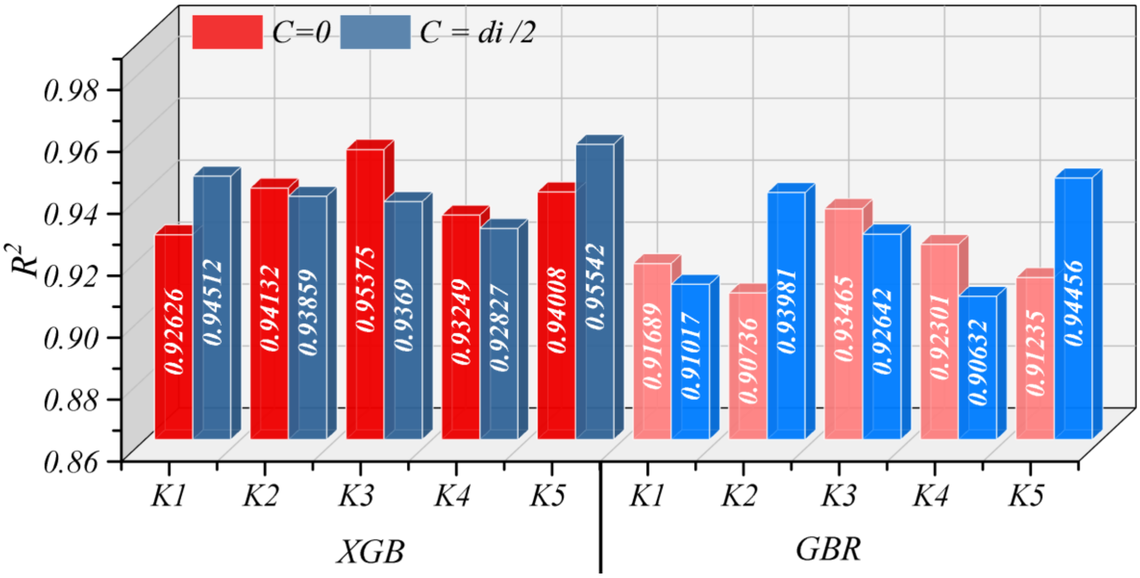
Table 3: The hyperparameters of the developed models.

Target	Hyperparameter	Models				
		XGTF	XGSA	XGTD	XGVO	XGAH
C = 0	n_estimators	582	333	165	957	318
	max_depth	492	3	82	611	540
	learning_rate	0.0919	0.0290	0.136	0.7627	0.325
	colsample_bytree	0.0757	0.9560	0.651	0.5803	0.571
	subsample	0.0592	0.6790	0.408	0.5381	0.45
	reg_alpha	0.0859	0.9610	0.172	0.2558	0.117
	reg_lambda	0.0493	0.0820	0.137	0.01	0.748
$C = d_i/2$	n_estimators	20	805	985	318	1000
	max_depth	999	648	796	621	247
	learning_rate	0.001	0.633	0.913	0.493	0.269
	colsample_bytree	0.001	0.866	0.057	0.728	1
	subsample	0.029	0.456	0.054	0.698	0.256
	reg_alpha	0.352	0.345	0.81	0.13	0
	reg_lambda	0.001	0.423	0.475	0.826	0.501

4 Result

4.1 Key parameter and convergence

Key parameters, also known as hyperparameters, are user-defined parameters that govern the model's behavior during training and can significantly impact its performance. XGB and GB are ML models with various hyperparameters that can be adjusted to boost the model's accuracy and efficiency. Tables 4 and 5 determine the optimal hyperparameters for the most desirable hybrid models, as described in the following results section. Additionally, Fig. 13 depicts the convergence curve, demonstrating the convergence behavior of all hybrid models for both outputs. Convergence analysis often uses the statistical technique of convergence based on the coefficient of correlation to assess the linear association between two variables. This method is frequently employed to evaluate the magnitude and direction of the relationship between two time series under consideration for convergence. Fig. 13 shows the convergence curve for the present hybrid models, in which 200 cycles have been done to obtain optimal outcomes $C = 0$ and $C = d_i/2$. In the output convergence, it is apparent that when models are integrated with VOA, the optimal state is reached more rapidly compared to other models, and the highest convergence value is attained. In $C = 0$ and $C = d_i /2$, alternatively, the most suitable and the weakest performance related to XGVO and GBAH in the convergence of developed hybrid models. In $C = 0$, XGVO obtained convergence of $RMSE$ equal 0.093 in 140th cycle, which in $C = d_i /2$ was 0.242 in the 150th cycle. On the contrary, within the GBAH model, convergence occurred at approximately the 130th and 110th cycless for $C = 0$ and $C = d_i /2$, leading to the achievement of $RMSE$ values of 0.258 and 0.277, respectively.



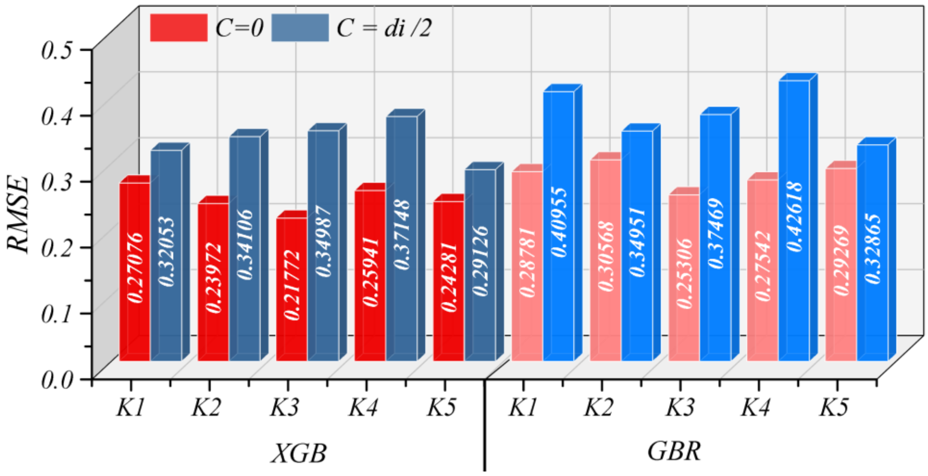


Figure 13: Result of K-fold for developed models.

Table 4: The hyperparameters of the developed models.

Target	Hyperparameter	Models				
		GBTF	GBSA	GBTD	GBVA	GBAH
$C = 0$	n_estimators	322	201	58	84	94
	learning_rate	0.179	0.48	0.67	0.1956	0.16
$C = d_i /2$	n_estimators	152	59	62	198	134
	learning_rate	0.667	0.39	0.23	0.893	0.338

Table 5: Outcomes obtained from the suggested models for $C = 0$.

Feature	Models	Phase	Evaluator				
			RMSE	R^2	MAE	n20_index	IOA
$C = 0$	XGB	Train	0.211	0.956	0.057	0.965	0.989
		Validation	0.256	0.940	0.198	0.840	0.983
		Test	0.203	0.959	0.158	0.833	0.989
	GBR	Train	0.248	0.938	0.194	0.779	0.984

		Validation	0.274	0.927	0.220	0.760	0.980
		Test	0.253	0.934	0.185	0.792	0.983
	XGTF	Train	0.178	0.969	0.152	0.805	0.992
		Validation	0.195	0.963	0.165	0.880	0.989
		Test	0.211	0.960	0.174	0.750	0.989
	XGSA	Train	0.148	0.978	0.103	0.912	0.994
		Validation	0.176	0.970	0.135	0.960	0.992
		Test	0.192	0.967	0.152	0.833	0.991
	XGTD	Train	0.100	0.990	0.077	0.973	0.997
		Validation	0.177	0.973	0.144	0.800	0.992
		Test	0.182	0.975	0.154	0.792	0.992
	XGVO	Train	0.082	0.993	0.062	0.982	0.998
		Validation	0.117	0.988	0.091	0.920	0.996
		Test	0.112	0.988	0.090	1.000	0.997
	XGAH	Train	0.195	0.962	0.149	0.850	0.990
		Validation	0.218	0.956	0.181	0.760	0.987
		Test	0.234	0.957	0.191	0.833	0.987
	GBTF	Train	0.227	0.948	0.191	0.796	0.986
		Validation	0.255	0.943	0.210	0.720	0.983
		Test	0.254	0.934	0.181	0.792	0.982
	GBSA	Train	0.203	0.958	0.166	0.779	0.989
		Validation	0.247	0.940	0.189	0.680	0.983
		Test	0.244	0.939	0.198	0.750	0.984
	GBTD	Train	0.172	0.970	0.136	0.894	0.992
		Validation	0.232	0.947	0.182	0.760	0.985
		Test	0.215	0.952	0.158	0.875	0.987
	GBVO	Train	0.173	0.972	0.126	0.938	0.992
		Validation	0.225	0.953	0.178	0.840	0.985
		Test	0.168	0.971	0.137	0.958	0.992
	GBAH	Train	0.253	0.943	0.187	0.823	0.981
		Validation	0.309	0.907	0.237	0.760	0.970
		Test	0.215	0.958	0.177	0.875	0.987
	GXTO	Train	0.108	0.988	0.082	0.973	0.997
		Validation	0.183	0.970	0.139	0.840	0.991
		Test	0.156	0.976	0.124	0.917	0.994
	GXSO	Train	0.133	0.982	0.103	0.947	0.995
		Validation	0.193	0.964	0.150	0.880	0.990
		Test	0.193	0.964	0.156	0.750	0.990
	GXTW	Train	0.139	0.980	0.112	0.912	0.995
		Validation	0.186	0.967	0.145	0.840	0.991
		Test	0.155	0.975	0.130	0.917	0.994
	GXVA	Train	0.117	0.986	0.091	0.965	0.996
		Validation	0.135	0.985	0.109	0.880	0.995
		Test	0.150	0.980	0.123	0.917	0.994
	GXAA	Train	0.185	0.969	0.134	0.912	0.991
		Validation	0.224	0.956	0.169	0.840	0.985
		Test	0.149	0.977	0.120	0.917	0.994

The boosting models used in this study require careful specification of hyperparameters due to their sensitivity to learning dynamics, model complexity, and regularization strength. The ranges used for hyperparameter tuning were selected based on standard practices established in the boosting literature and on preliminary exploratory trials that identified stable performance boundaries for the current dataset.

For XGBoost-based hybrids (Table 4), the hyperparameters included tree count (n_estimators), tree depth (max_depth), learning rate, subsampling ratios (subsample, colsample_bytree), and regularization coefficients (reg_alpha, reg_lambda). Each parameter was assigned a broad search interval to enable the meta-heuristic algorithms to explore diverse model structures:

- n_estimators: 20–1000
- max_depth: 2–1000
- learning_rate: 0.001–0.9
- subsample, colsample_bytree: 0.05–1.0
- reg_alpha, reg_lambda: 0.0–1.0

For Gradient Boosting (Table 5), tuning focused on n_estimators and learning_rate, which are the dominant contributors to ensemble behavior. These were varied within the ranges:

- n_estimators: 20–400
- learning_rate: 0.001–0.9

These intervals were selected to balance computational feasibility with sufficient model variability. Because the dataset contains 324 samples, excessively deep trees or extremely large ensembles tend to produce unstable models, making bounded but flexible search ranges essential.

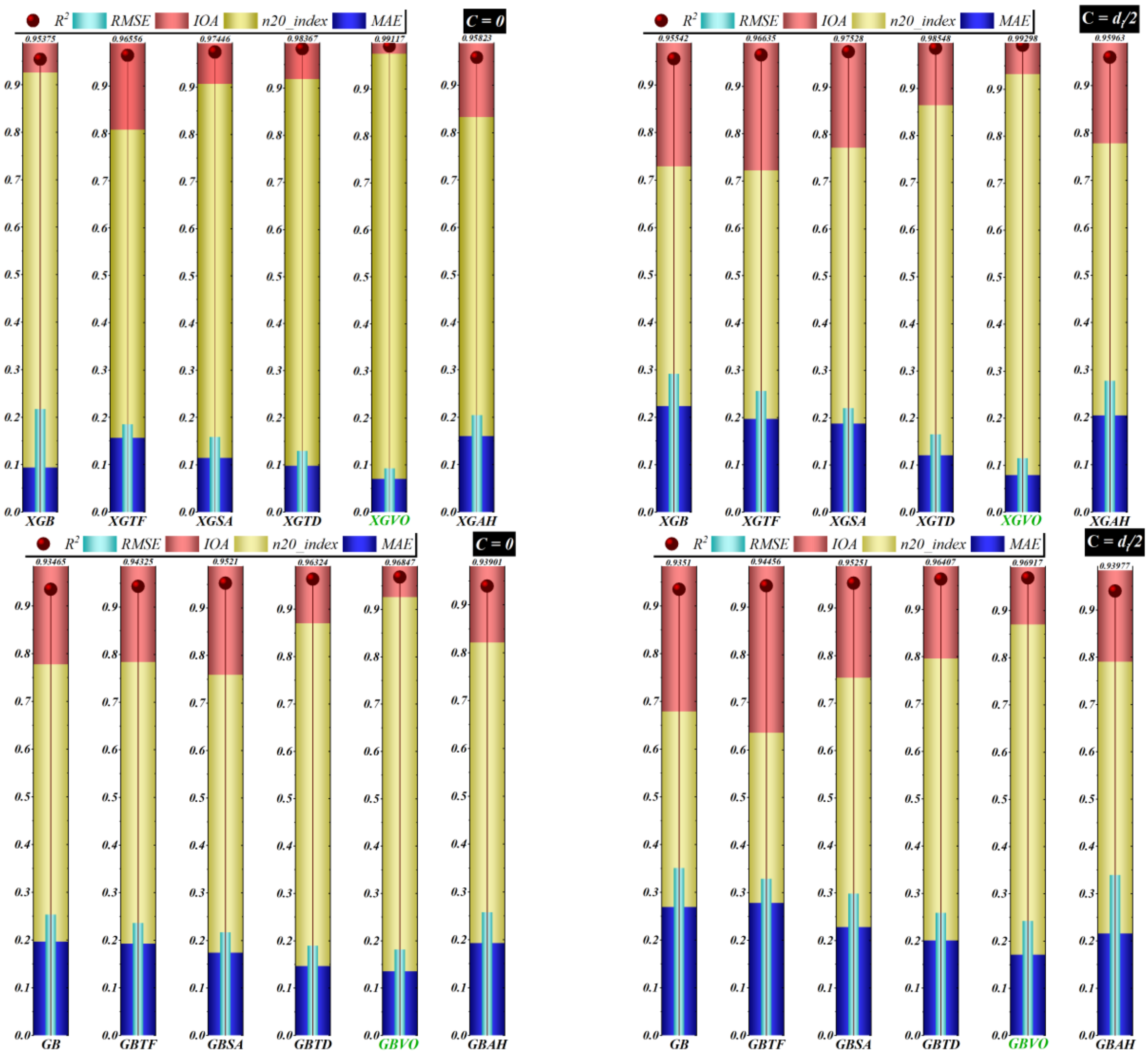
Hyperparameter tuning was performed using a random search procedure embedded within the meta-heuristic optimization algorithms.

A complete sensitivity analysis or ablation study—quantifying the relative importance of each hyperparameter on the final model performance—requires extensive iterative retraining under controlled perturbations. This remains a relevant extension for future work, especially given the nonlinear interactions among boosting hyperparameters. A dedicated ablation framework would help identify which parameters dominate performance in hydraulic prediction tasks and could guide the development of more efficient optimization routines.

4.2 K-Fold cross-validation

K-Fold Cross Validation splits data into K subsets, trains the model K times using K-1 folds as training data and one fold as testing data each time, then averages performance metrics across cycles for a more robust evaluation. It is useful for accurate model assessment, especially when data is limited, by ensuring thorough use of available samples for both training and testing.

In this study, a 5-fold cross-validation approach was employed to assess model performance. As depicted in Fig. 14, each fold was evaluated using RMSE and R^2 metrics, aiming to identify the most effective fold for predictive modeling. The outcomes of this evaluation revealed that folds K3 and K5 had the most desirable outcomes than the others in terms of prediction accuracy for $C = 0$ and $C = d_i/2$, respectively. This suggests that utilizing these particular folds yielded more reliable predictions compared to the alternative folds.



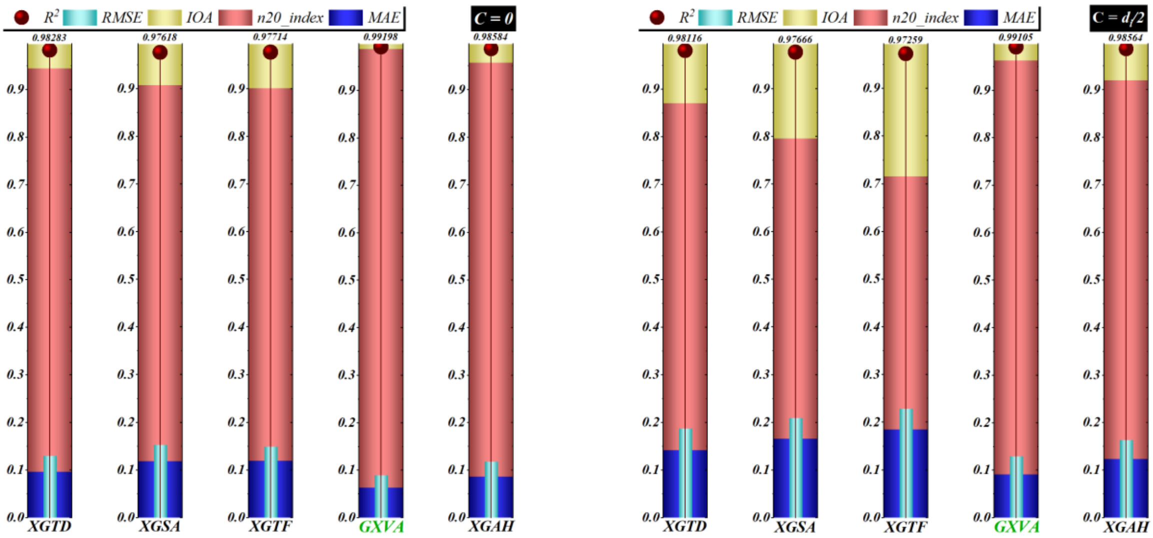


Figure 14: The multi-Y plot for the obtained result of metrics.

4.3 Models' performance

In this section, the outcomes of the introduced models are discussed. As mentioned, the relevant models comprise three parts: individual, hybrid, and ensemble-hybrid, which have been investigated in three phases: training, validation, and testing. In the current article, 70% of the samples are assigned to the training phase, and the remaining 30% are divided into two parts, with 15% allocated to validation and 15% to testing. Tables 6 and 7 present the evaluation of the models based on specific metrics, where the ideal states and the mathematical definitions of these metrics are outlined in Table 3.

Table 6: Outcomes obtained from the suggested models for $C = d_i/2$.

Feature	Models	Phase	Evaluator				
			RMSE	R^2	MAE	n20_index	IOA
$C = d_i/2$	XGB	Train	0.282	0.958	0.215	0.752	0.989
		Validation	0.319	0.935	0.247	0.560	0.983
		Test	0.304	0.961	0.237	0.792	0.988
	GBR	Train	0.339	0.939	0.256	0.673	0.984
		Validation	0.375	0.912	0.307	0.560	0.977
		Test	0.380	0.943	0.291	0.833	0.982
	XGTF	Train	0.243	0.970	0.194	0.726	0.992
		Validation	0.304	0.941	0.214	0.640	0.985
		Test	0.257	0.972	0.195	0.792	0.992
	XGSA	Train	0.208	0.978	0.179	0.779	0.994
		Validation	0.247	0.961	0.208	0.720	0.990
		Test	0.246	0.973	0.207	0.792	0.993
	XGTD	Train	0.134	0.990	0.101	0.885	0.998
		Validation	0.242	0.962	0.173	0.720	0.990
		Test	0.198	0.987	0.159	0.917	0.995
	XGVO	Train	0.079	0.997	0.058	0.956	0.999
		Validation	0.182	0.979	0.137	0.800	0.995
		Test	0.158	0.991	0.116	0.958	0.997
	XGAH	Train	0.268	0.962	0.194	0.788	0.990
		Validation	0.268	0.955	0.206	0.680	0.988
		Test	0.325	0.951	0.245	0.833	0.987
	GBTF	Train	0.311	0.949	0.268	0.673	0.987
		Validation	0.391	0.907	0.326	0.480	0.975
		Test	0.338	0.957	0.277	0.625	0.986
	GBSA	Train	0.275	0.960	0.208	0.761	0.990
		Validation	0.335	0.930	0.272	0.600	0.982
		Test	0.356	0.947	0.277	0.875	0.984
	GBTD	Train	0.230	0.972	0.179	0.814	0.993

		Validation	0.306	0.941	0.245	0.640	0.985
		Test	0.326	0.956	0.256	0.875	0.987
	GBVO	Train	0.236	0.972	0.155	0.894	0.992
		Validation	0.251	0.962	0.205	0.800	0.990
		Test	0.263	0.968	0.208	0.833	0.991
	GBAH	Train	0.329	0.944	0.194	0.823	0.984
		Validation	0.311	0.944	0.232	0.680	0.985
		Test	0.407	0.925	0.298	0.750	0.977
	GXTO	Train	0.157	0.987	0.118	0.885	0.997
		Validation	0.256	0.958	0.201	0.720	0.989
		Test	0.232	0.980	0.194	0.958	0.993
	GXSO	Train	0.178	0.983	0.141	0.823	0.996
		Validation	0.272	0.953	0.235	0.640	0.988
		Test	0.264	0.970	0.210	0.833	0.992
	GXTW	Train	0.205	0.978	0.169	0.735	0.994
		Validation	0.303	0.942	0.245	0.520	0.985
		Test	0.245	0.977	0.198	0.833	0.993
	GXVA	Train	0.118	0.993	0.084	0.956	0.998
		Validation	0.193	0.976	0.162	0.720	0.994
		Test	0.152	0.991	0.125	0.958	0.997
	GXAA	Train	0.227	0.973	0.166	0.841	0.993
		Validation	0.249	0.961	0.205	0.640	0.990
		Test	0.277	0.965	0.234	0.875	0.990

Table 7: Comparison of the present study outcomes with recent studies for $\mathrm{C} = 0$ and $\mathrm{C} = \mathrm{d_i}/2$

Bottom Clearance	Study	Model	Statistical evaluator	
			R^2	RMSE
$C = 0$	Ayoubloo et al. [11]	Linear Regression	0.984	0.272
		CART	0.971	0.340
		ANN	0.983	0.289
	Ahmed et al. [24]	Eq. (2)	0.969	0.714
	Present Study	XGVO	0.988	0.112
$C = d_i/2$	Ayoubloo et al. [11]	Linear Regression	0.984	0.272
		CART	0.971	0.34
		ANN	0.983	0.289
	Ahmed et al. [24]	Eq. (3)	0.969	0.207
	Present Study	XGVO	0.997	0.079

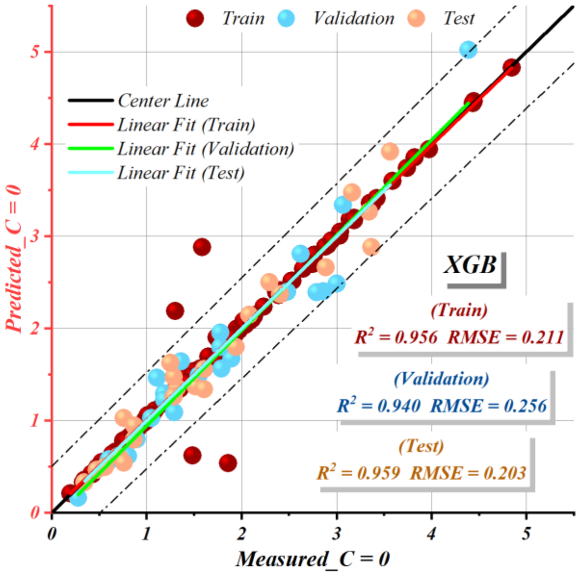
As shown in Table 6, the XGVO model emerges as a standout performer in this comprehensive evaluation, consistently demonstrating exceptional predictive accuracy across multiple evaluators and phases. In the training phase, XGVO boasts the lowest RMSE value of 0.083 and the highest R^2 of 0.993, indicating its extraordinary precision in modeling the Sc depth ratio. This trend continues in the testing phase, where XGVO maintains its superiority with an impressively low RMSE of 0.112 and a high R^2 of 0.988. These outcomes underline XGVO's robust predictive capabilities, making it a compelling choice for modeling hydraulic phenomena during water withdrawal processes in horizontal intakes. Other models also display remarkable performance. XGSO, for instance, achieves excellent R^2 and RMSE values in the training phase (0.978 and 0.148, respectively), emphasizing its ability to provide accurate predictions. GBTD stands out in the training phase with an RMSE of 0.172 and an R^2 of 0.970, showcasing its competence in modeling Sc depth ratios. A prevailing trend across all models and evaluators is the superior performance observed during the training phase in comparison to the validation and testing phases. This suggests that while models excel at learning from training data, there may be challenges in generalizing their predictions to unseen data. Nevertheless, models like XGVO maintain their excellence even in the testing phase, signifying their robustness and ability to provide reliable predictions. The n20_index and IOA evaluators add further depth to the assessment. These metrics measure the models' agreement with the observed data, offering insights into their ability to capture the underlying patterns in the hydraulic phenomenon. The high n20-index and IOA values observed in XGVO across all phases underscore its strong agreement with the actual data, reinforcing its position as the top-performing model.

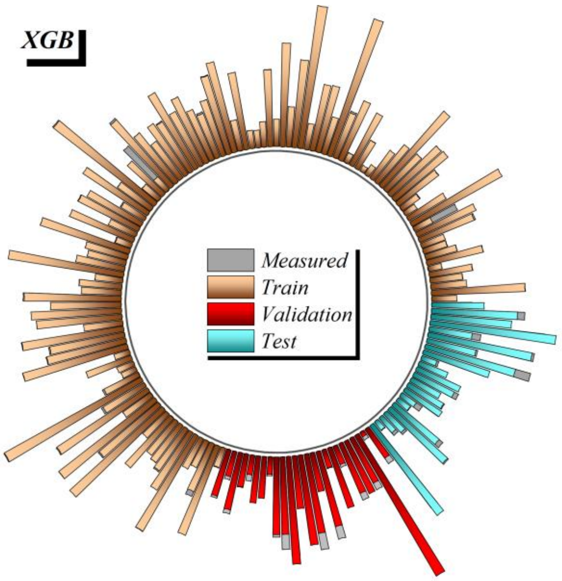
Table 7 shows the performance of developed models in predicting $C = d_i/2$. The revised interpretation highlights GXVA, XGSA, and XGTD as models with low RMSE and high R^2, but does so by comparing their relative statistical positions rather than declaring fixed superiority. Likewise, XGVO's performance in validation and testing is discussed as an example of strong generalization, but not as a repetitive assertion of dominance. Models such as GBR and GBTF exhibit higher errors, and their limitations are now framed in terms of potential generalization challenges rather than definitive shortcomings.

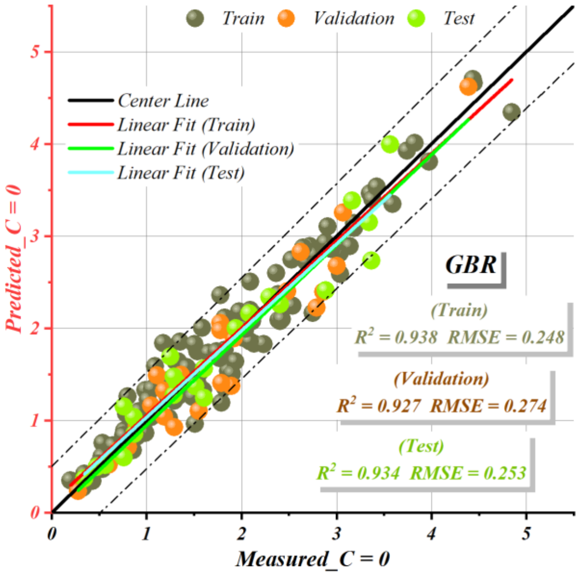
Table 7 shows the performance of developed models in predicting $C = d_i/2$. The revised interpretation highlights GXVA, XGSA, and XGTD as models with low RMSE and high R^2, but does so by comparing their relative statistical positions rather than declaring fixed superiority. Likewise, XGVO's performance in validation and testing is discussed as an example of strong generalization, but not as a repetitive assertion of dominance. Models such as GBR and GBTF exhibit higher errors, and their limitations are now framed in terms of potential generalization challenges rather than definitive shortcomings.

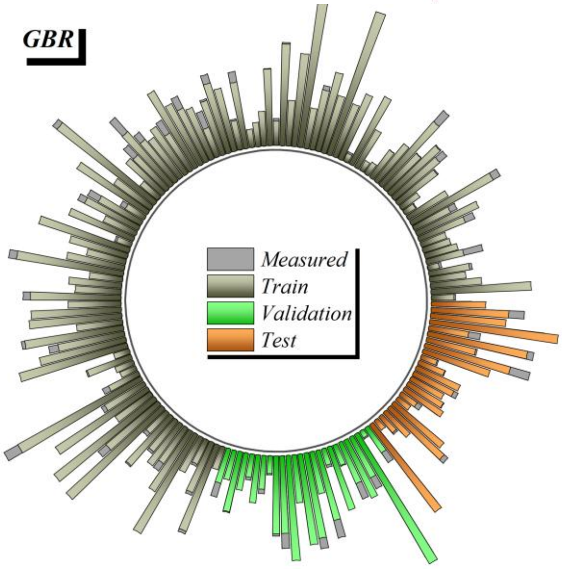
To support the comparative analysis more effectively, the revised section explicitly examines how model performance varies across phases, emphasizing trends, spreads, and the magnitude of improvement among model categories rather than repeating narrative conclusions.

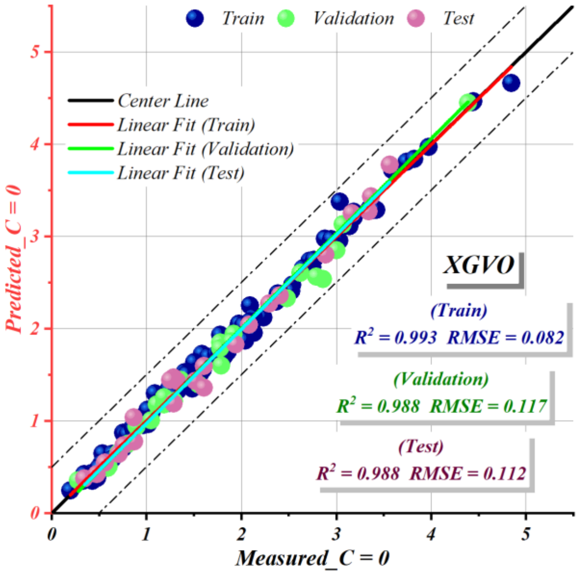
Fig. 15 presents a multi-Y plot summarizing the metrics across all models and phases. This figure is now referenced as a tool for visually interpreting trends, variance, and relative clustering of model accuracy rather than reinforcing any predetermined hierarchy. The figure offers a concise overview of model consistency, guiding readers toward identifying patterns rather than focusing on singular outcomes.

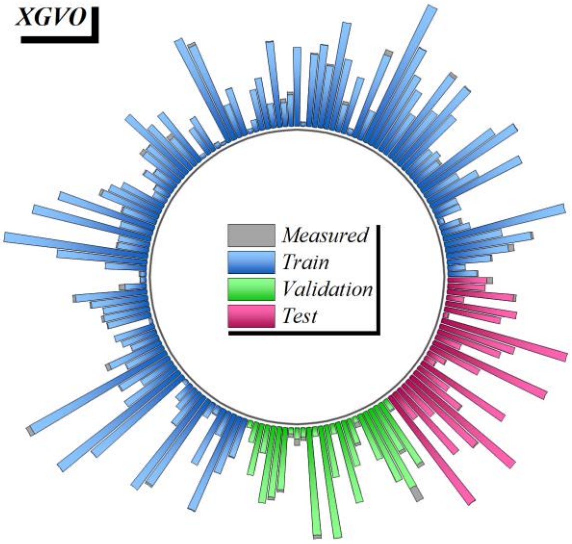


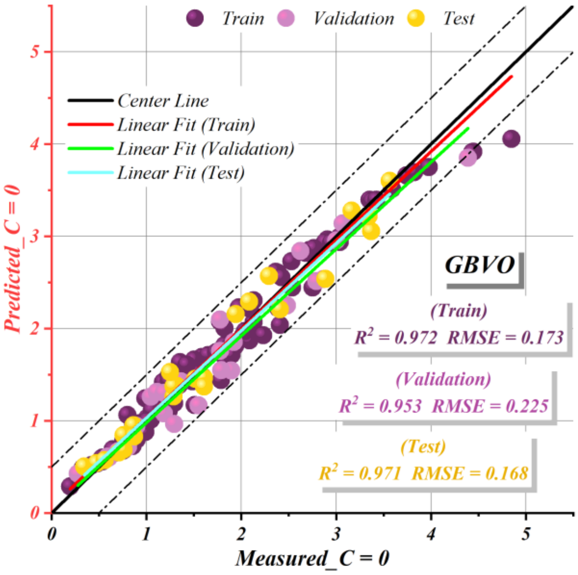


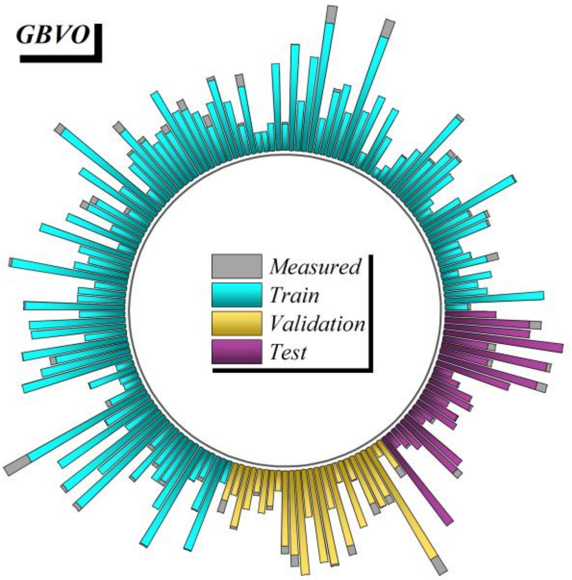


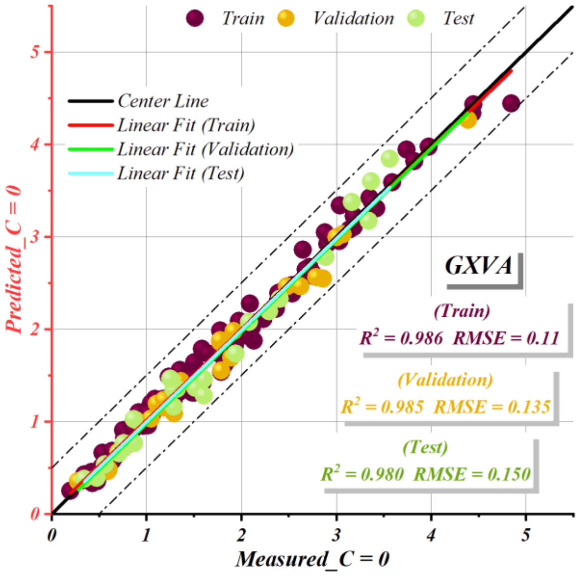


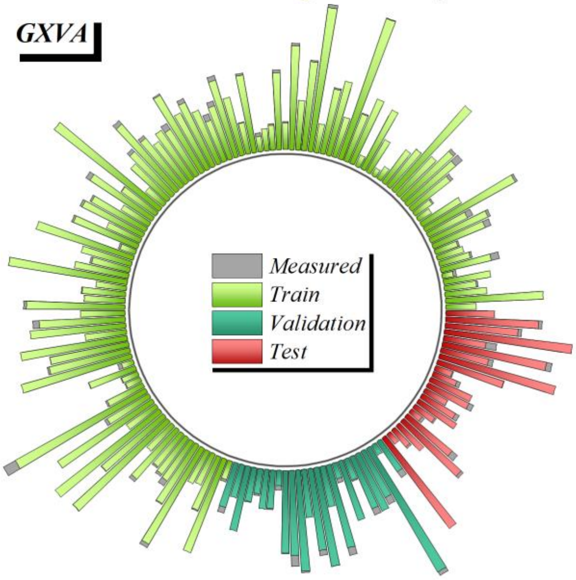

Train
Validation
Test
Center Line
Linear Fit (Train)
Linear Fit (Validation)
Linear Fit (Test)
XGVO
(Train)
R^2 = 0.993 RMSE = 0.082
(Validation)
R^2 = 0.988 RMSE = 0.117
(Test)
R^2 = 0.988 RMSE = 0.112
Predicted_C = 0
Measured_C = 0
0
1
2
3
4
5

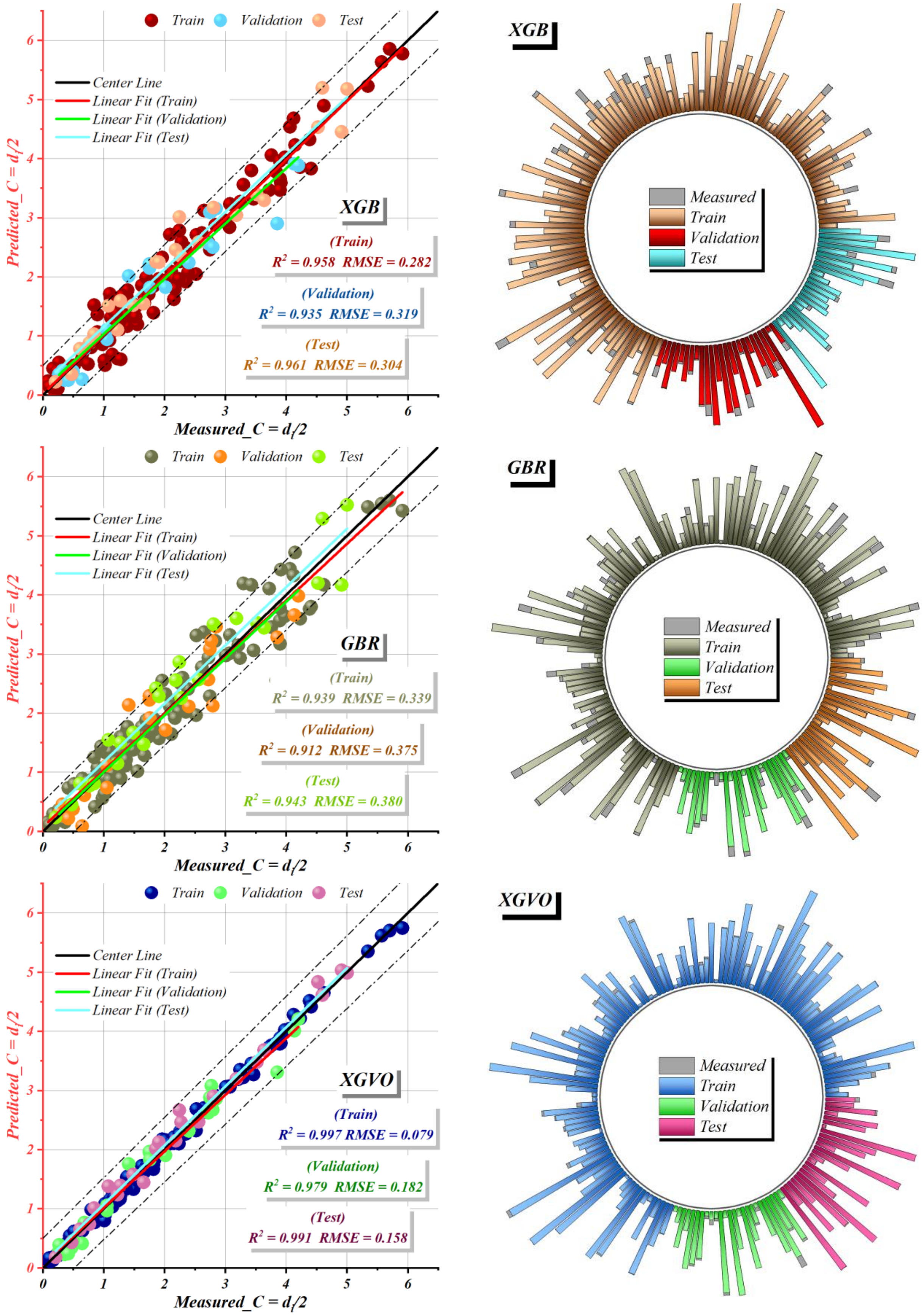

XGVO
Measured
Train
Validation
Test


Train
Validation
Test
Center Line
Linear Fit (Train)
Linear Fit (Validation)
Linear Fit (Test)
GBVO
(Train)
R^2 = 0.972 RMSE = 0.173
(Validation)
R^2 = 0.953 RMSE = 0.225
(Test)
R^2 = 0.971 RMSE = 0.168
Predicted_C = 0
Measured_C = 0
0
1
2
3
4
5


GBVO
Measured
Train
Validation
Test


Train
Validation
Test
Center Line
Linear Fit (Train)
Linear Fit (Validation)
Linear Fit (Test)
GXVA
(Train)
R^2 = 0.986 RMSE = 0.11
(Validation)
R^2 = 0.985 RMSE = 0.135
(Test)
R^2 = 0.980 RMSE = 0.150
Predicted_C = 0
Measured_C = 0
0
1
2
3
4
5


GXVA
Measured
Train
Validation
Test


Train
Validation
Test
Center Line
Linear Fit (Train)
Linear Fit (Validation)
Linear Fit (Test)
XGB
(Train)
R^2 = 0.958 RMSE = 0.282
(Validation)
R^2 = 0.935 RMSE = 0.319
(Test)
R^2 = 0.961 RMSE = 0.304
Predicted_C = d/2
Measured_C = d/2
0
1
2
3
4
5
6
XGB
Measured
Train
Validation
Test
GBR
(Train)
R^2 = 0.939 RMSE = 0.339
(Validation)
R^2 = 0.912 RMSE = 0.375
(Test)
R^2 = 0.943 RMSE = 0.380
GBR
XGVO
(Train)
R^2 = 0.997 RMSE = 0.079
(Validation)
R^2 = 0.979 RMSE = 0.182
(Test)
R^2 = 0.991 RMSE = 0.158
XGVO

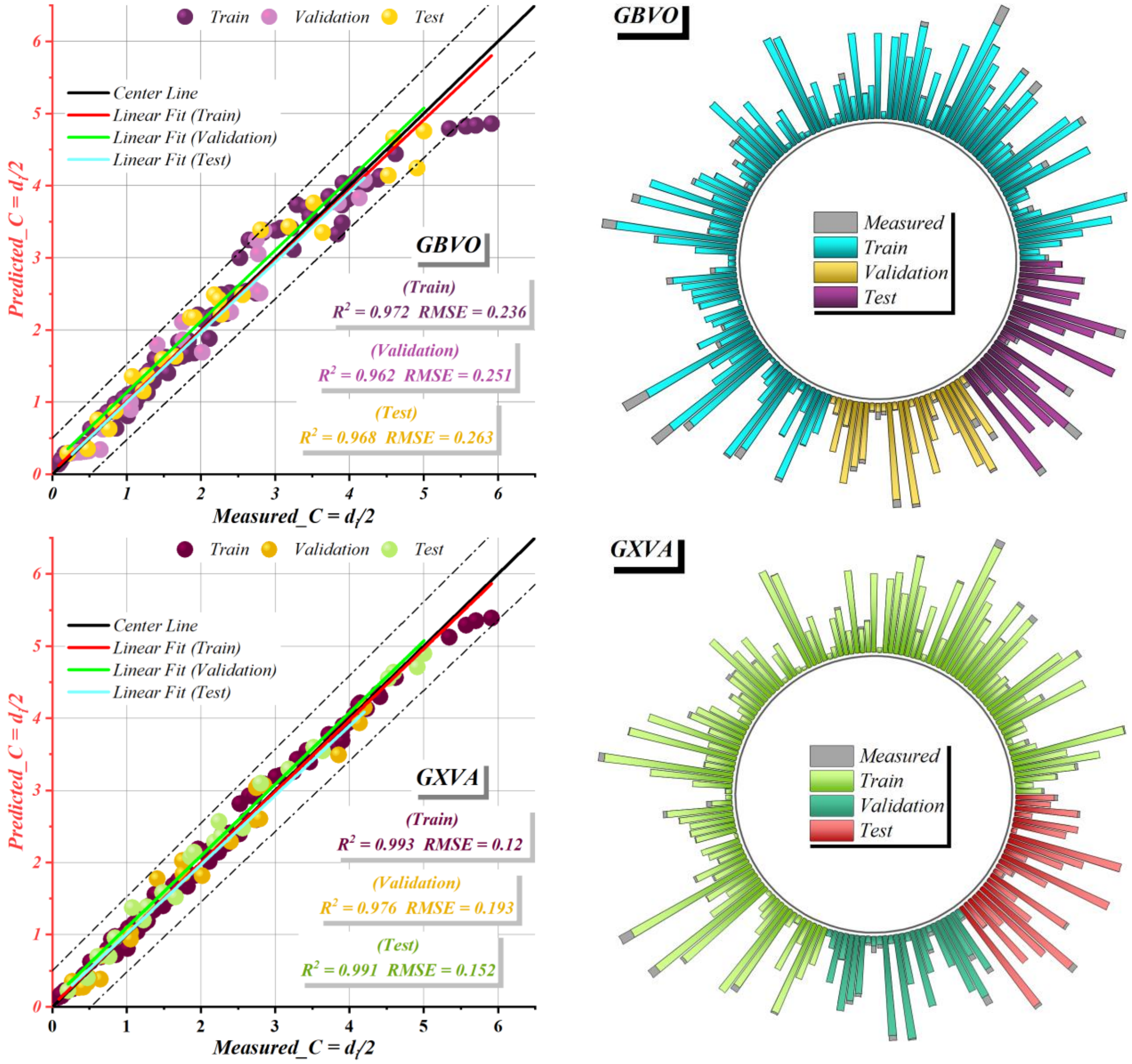
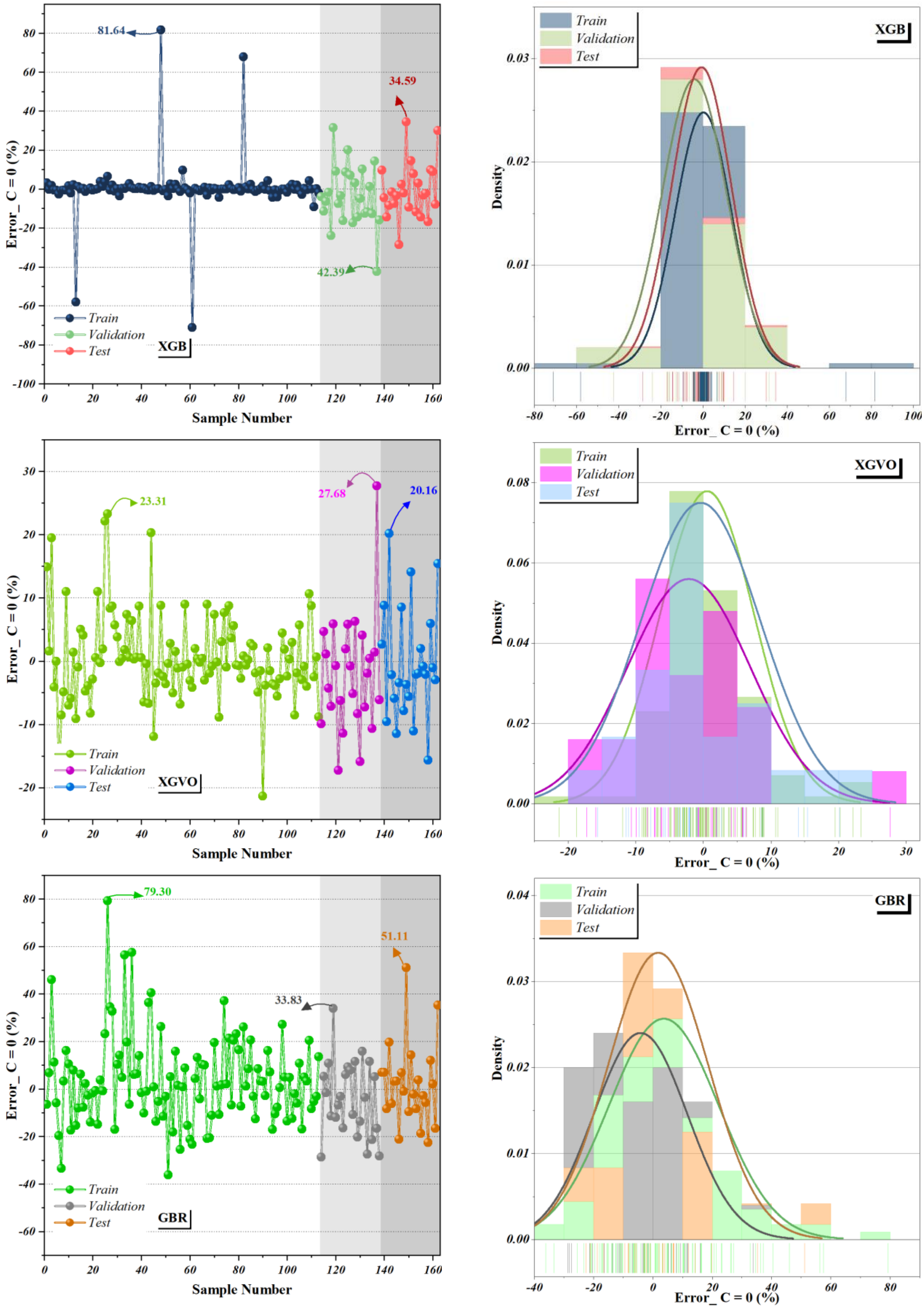


Figure 15: Scatter and radial bar plots for comparing the measured and predicted values.

Additionally, the outcomes of recent research comparing the performance of the suggested models based on R^2 and RMSE is presented in Table 8, with the favorable results in the present article attributed to XGVO. Comparing the present XGVO with the models obtained by Ayoubloo et al. [44], the highest and lowest values of R^2 and RMSE by Linear Regression differed in R^2 by 1.5% and RMSE by 89%. In addition, with the models introduced in the article by Ahmed et al. [21] and Roushangar et al. [45], the difference between their results in R^2 and XGVO was equal to 3 and 1 percent, respectively, and this difference in RMSE was equal to 86 and 85 percent.

Subsequently, three models exhibiting the strongest overall performance within their groups (XGVO, GBVO, and GXVA) were selected for deeper comparative visualization. Fig. 16 shows scatter and radial bar plots, illustrating the distribution of measured vs. predicted values. In both output scenarios, XGVO demonstrates smaller deviation ranges, but these observations are now framed as relative improvements within the evaluated set.


81.64
34.59
42.39
Train
Validation
Test
XGB
Error_ C = 0 (%)
Sample Number
Density
Error_ C = 0 (%)
23.31
27.68
20.16
XGVO
33.83
79.30
51.11
GBR

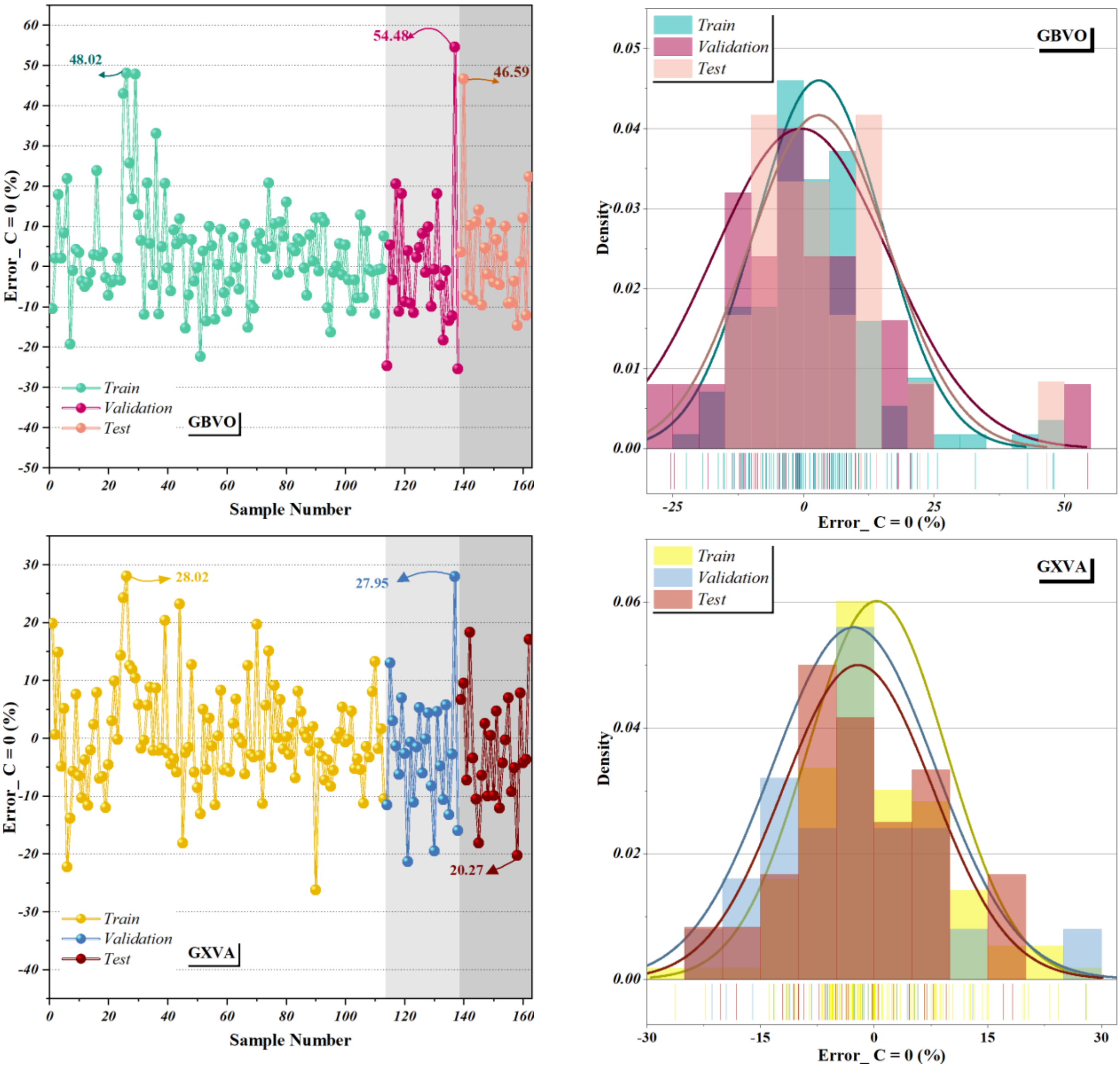
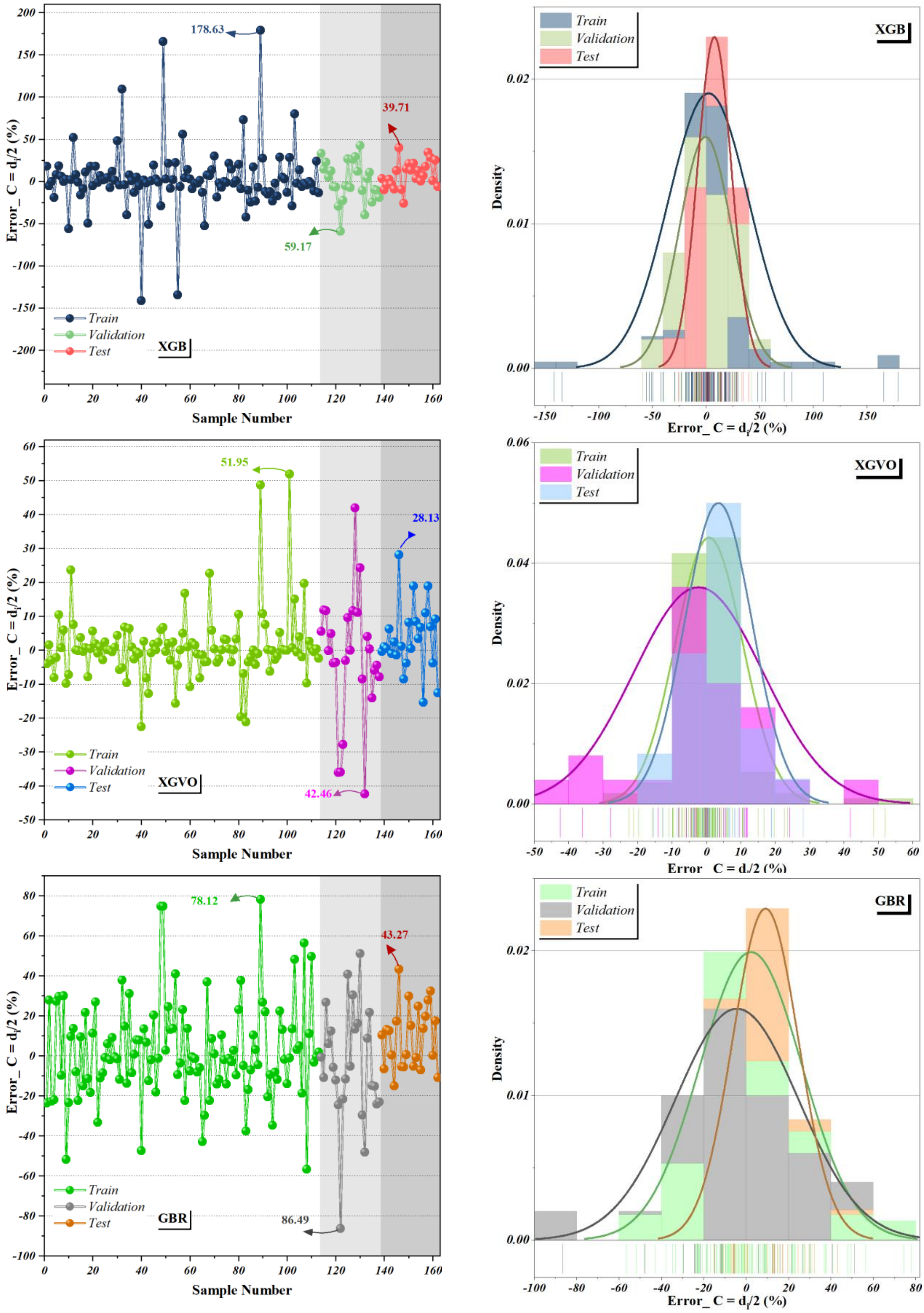


Figure 16: Error percentage of models based on line-symbol and histogram-distribution plots.

Figs. 17 and 18 depict the error distribution using line symbols and density curves. The interpretation now focuses on how error concentration, skewness, and distribution symmetry contribute to assessing generalization capability, rather than declaring categorical superiority. In the case of $C = 0$, models show reduced error across validation and testing phases after strong learning in training. For $C = d_i/2$, all three selected models show comparable distribution patterns with varying dispersion widths.


Error_ C = d$_i$/2 (%)
Sample Number
Density
Train
Validation
Test
XGB
178.63
39.71
59.17
XGVO
51.95
28.13
42.46
GBR
78.12
43.27
86.49
200
150
100
50
0
-50
-100
-150
-200
0
20
40
60
80
100
120
140
160
0.00
0.01
0.02
0.04
0.06
-150
-100
-50
0
50
100
150
60
40
30
20
10
-10
-20
-30
-40
-50
80
-60
-80
-100

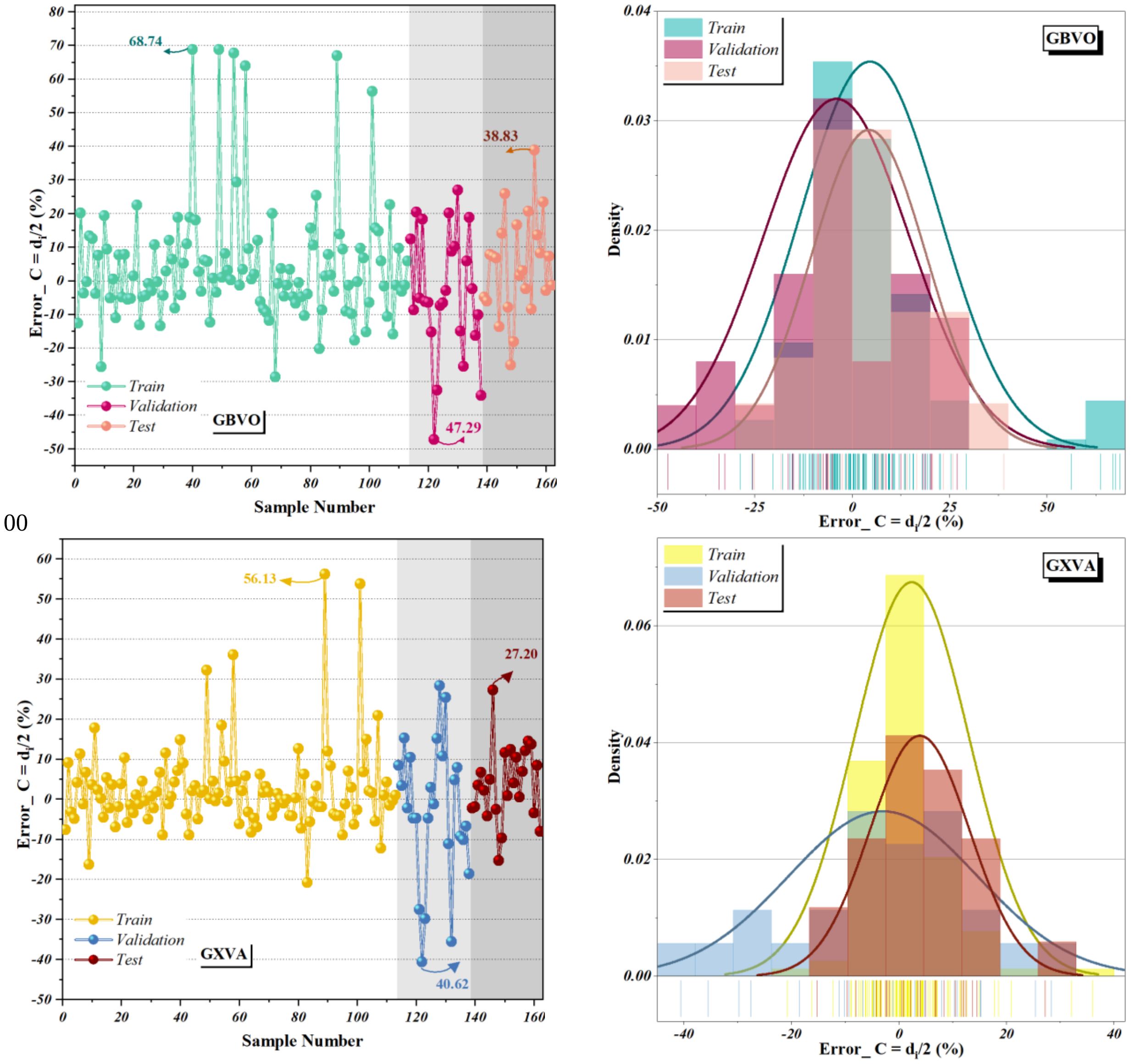


Figure 17: Error percentage of models based on line-symbol and histogram-distribution plots.

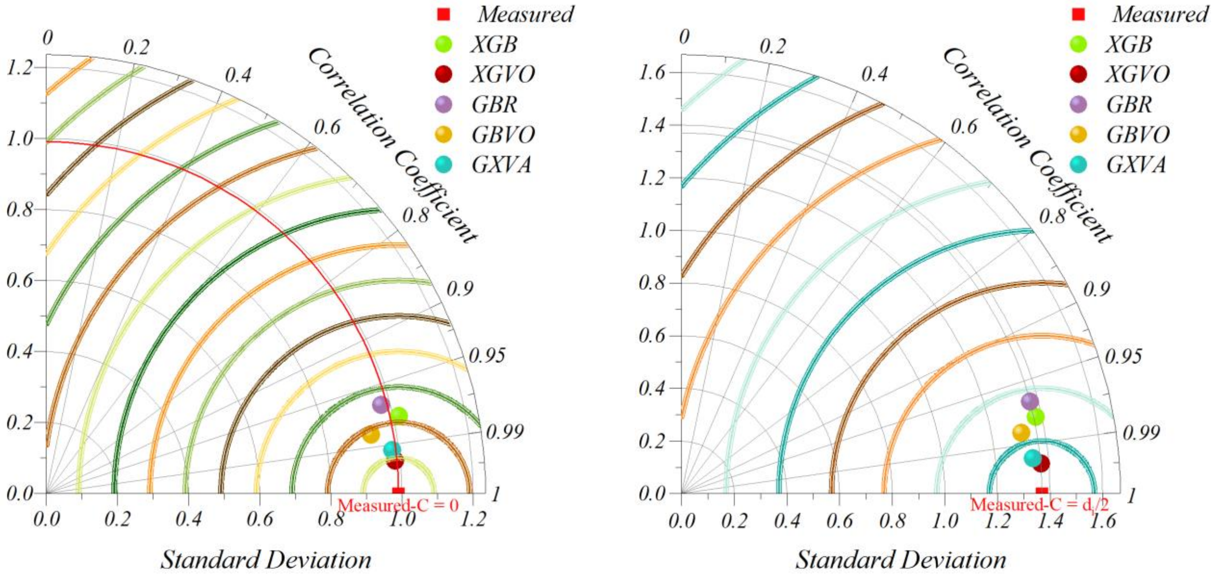
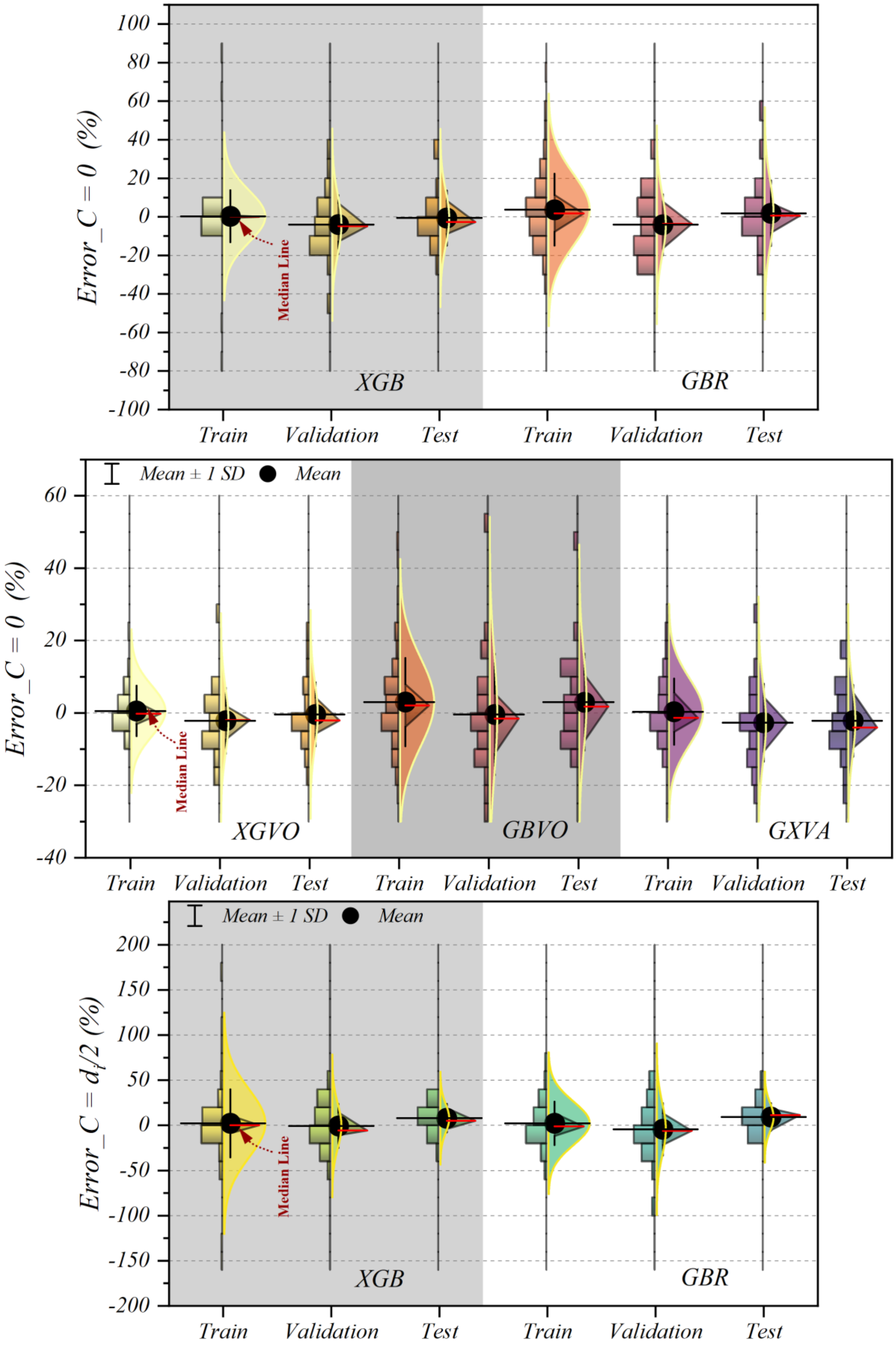


Figure 18: Taylor diagram for the developed models.

Fig. 19 presents the Taylor diagram for both target outputs. The discussion has been edited to emphasize relative proximity to the reference point and statistical similarity among strong-performing models, rather than repeatedly stating that one model is the best. XGVO and GXVA cluster near the reference location, while GBVA shows lower correlation and deviation.



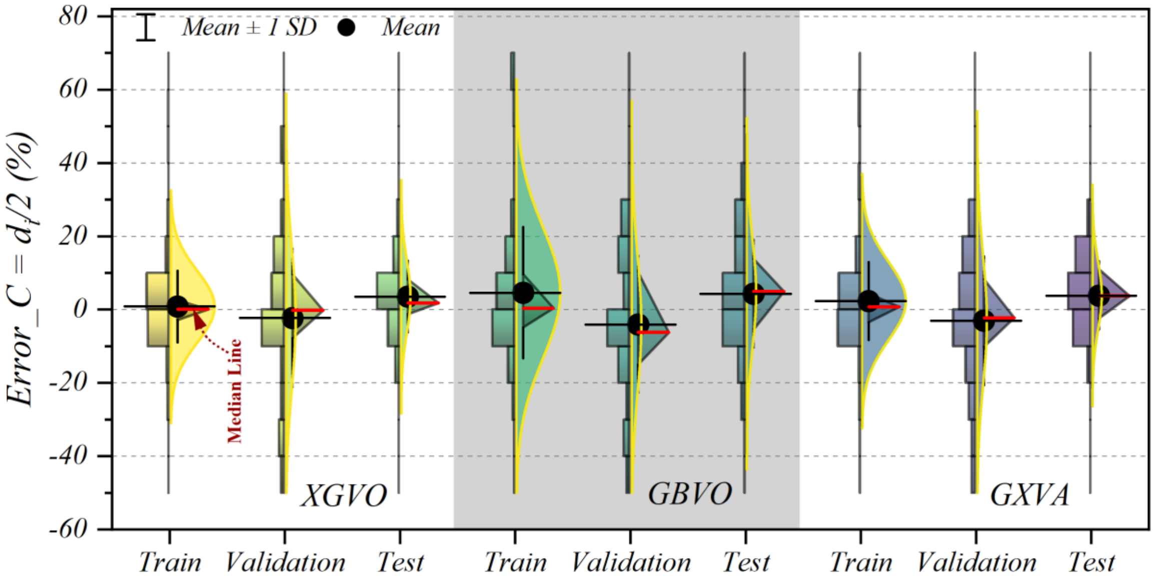


Figure 19: Error percentage of developed models based on half-violin diagram.

Fig. 20 illustrates the error percentage using kernel density estimation. The revised explanation highlights how density width, central tendency, and error spread inform model evaluation, noting that XGVO and GBVO exhibit narrower distributions and central values closer to zero, without unnecessary repetition.

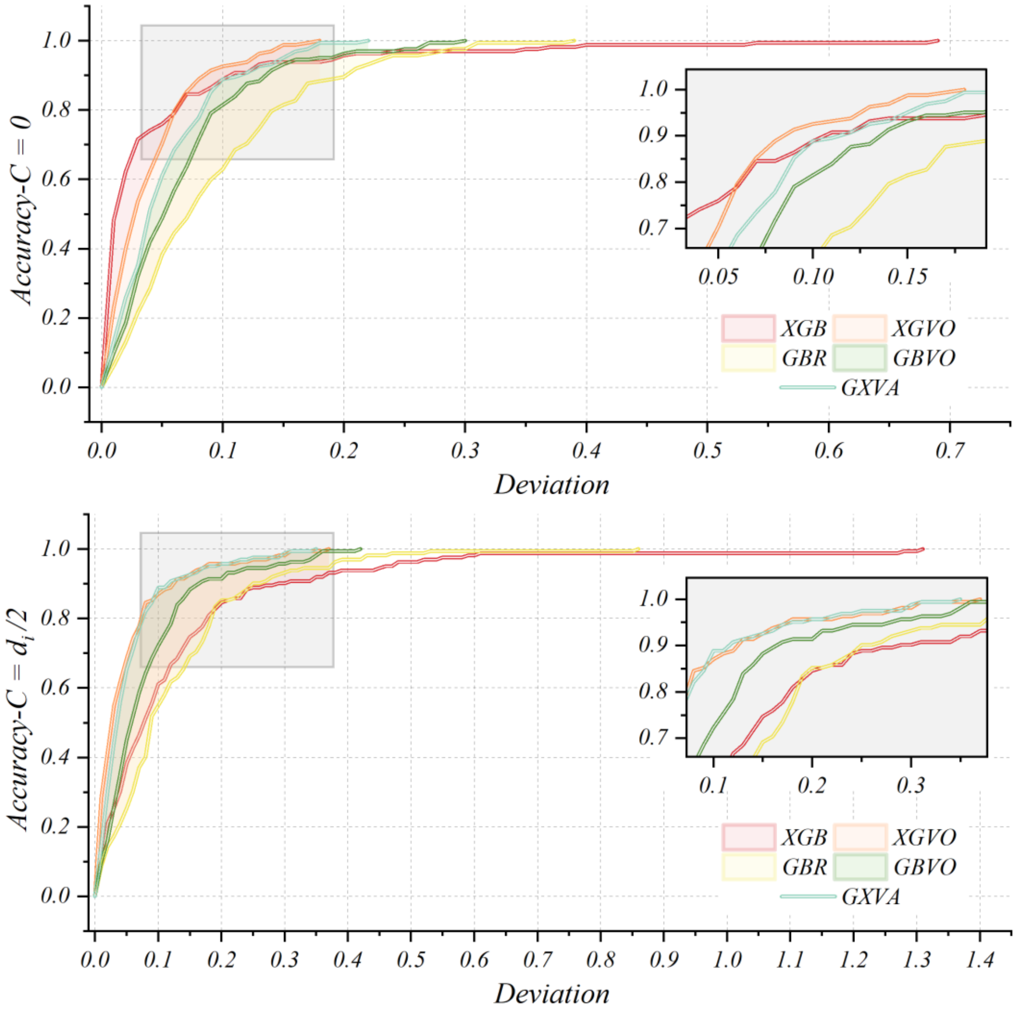


Figure 20: REC curve for the accuracy of selected models.

REC curves are illustrated in Fig. 21. This subsection is now reframed to focus on comparative slope behavior and error tolerance sensitivity across models, demonstrating that hybrid and ensemble–hybrids generally outperform single models in both targets. XGVO, GBVO, and GXVA show stronger early rises in their curves, indicating better accuracy at stricter error tolerances.

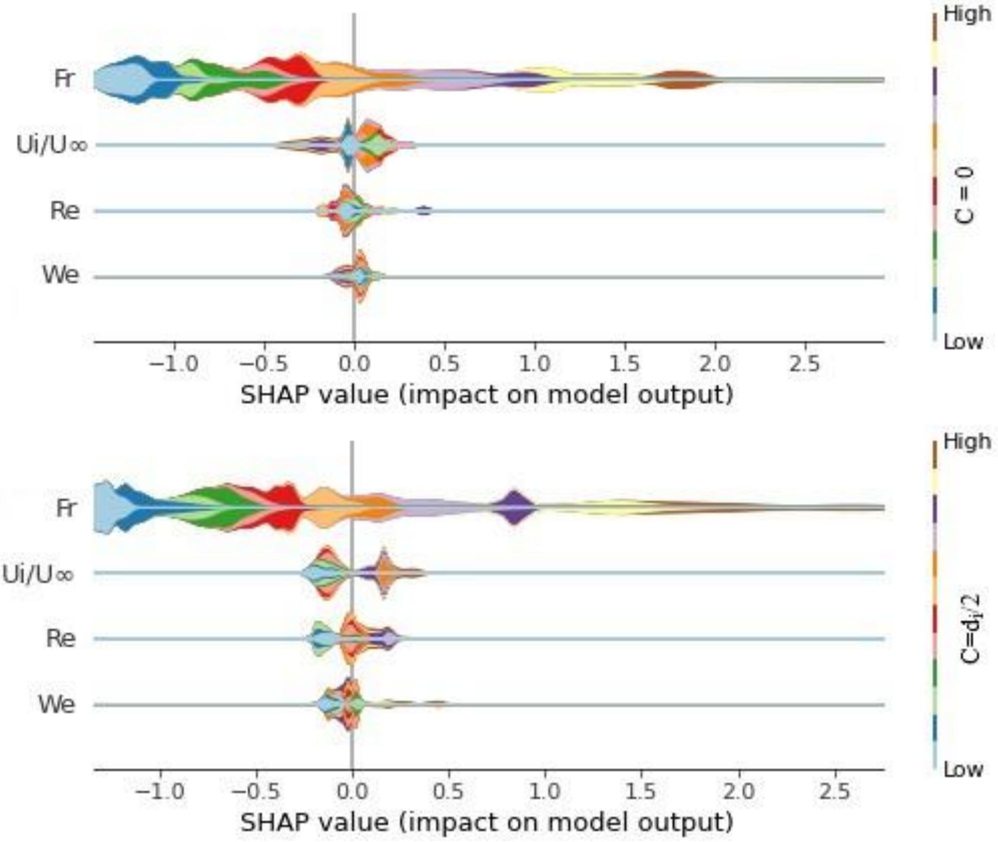


Figure 21: SHAP value for the sensitivity of inputs.

4.4 SHAP (Shapley Additive exPlanations)

SHAP (Shapley Additive exPlanations) values are an ML approach that may be interpreted and is based on cooperative game theory. SHAP values elucidate how specific properties contribute to model predictions, facilitating the comprehension of complex models. These values provide both global and local explanations, aiding in feature significance analysis and facilitating the comprehension of individual predictions. They are model-independent, making them relevant to a wide range of ML models, and are calculated using several techniques, such as Kernel SHAP and Tree SHAP. The values are frequently displayed using a red-blue color scheme, with blue representing positive feature contributions and red representing negative ones. These explanations are cumulative; thus, the sum of the feature contributions equals the difference between the model's output and the predicted output.

The graphical presentation of the sensitivity analysis of the input variables on the output is stated in Fig. 22. In $C = 0$, Fr had the highest impact, while We had the least impact on $C = 0$. Also, for $C = d_i/2$, the Fr and We had the highest and lowest impact, respectively.

From a hydraulic point of view, the We, Fr, and Re are dimensionless numbers that characterize the importance of inertial forces to surface tension, gravitational, and viscous forces, respectively. As in Fig. 22, Fr and U_i/U_∞ had the highest impacts on the predicted Sc in respect. Recent studies demonstrated that Fr number has the main impact on Sc [5], [24], [46–50]. Also, Ahmed et al. [24] provided that the Sc is in direct relation to Fr and increment in Fr leads to the increase of S_c for both types of bottom clearance. Furthermore, the values of S_c reduce with the increment of velocity in the channel. Therefore, the parametric study defined in the present work determined the effective parameters, which are consistent with recent experimental outcomes.

5 Discussion

5.1 Limitations

The findings of this study should be interpreted within the context of several limitations. Firstly, the generalizability of the outcomes may be constrained to the specific conditions and configurations investigated in the study. Variations in hydraulic systems, intake geometries, and operating conditions could potentially yield different outcomes. Additionally, the accuracy and robustness of the predictive models heavily rely on the quality and quantity of available data. Inadequate or biased datasets may compromise the reliability of predictions. Moreover, while the study employed various meta-heuristic algorithms for model optimization, the selection of algorithms may not encompass the entire spectrum of optimization techniques available, potentially limiting the exploration of alternative approaches.

5.2 Future studies

To address the aforementioned limitations and further advance the field, future studies could explore several avenues. Firstly, experimental validation of the predictive models through physical testing would enhance the credibility and applicability of the findings in real-world scenarios. Additionally, researchers could consider incorporating additional parameters or environmental factors into the models, such as variations in flow velocity or sediment transport, to capture a more comprehensive understanding of vortex formation and Sc depth. Furthermore, investigating dynamic modeling approaches to account for temporal variations in hydraulic conditions and vortex dynamics could provide deeper insights into the transient behavior of intake systems.

5.3 Advantages

Despite these limitations, the study offers several notable advantages. Firstly, the predictive accuracy of the ensemble modeling techniques employed in this study demonstrates their efficacy in accurately estimating Sc depth ratios, thereby providing valuable insights for intake design and operation. The utilization of hybridization techniques enhances the robustness and reliability of the ensemble models by leveraging the strengths of multiple algorithms to improve predictive performance. Moreover, the exploration of multiple meta-heuristic algorithms for model optimization highlights the importance of algorithm selection in maximizing the accuracy and efficiency of predictive models in hydraulic engineering applications.

5.4 Comparison with state-of-the-art

The results of the proposed hybrid ensemble framework were evaluated against prior studies summarized in Table 1 . While previous works achieved R² values of up to 0.95 and RMSEs of around 0.08–0.10, the models demonstrate a significant improvement, with $R^2 = 0.993$ and $RMSE = 0.082$ for the $C = 0$ case, and $R^2 = 0.993$ and $RMSE = 0.118$ for C = d_i/2 [51–53].

Several factors contribute to this enhanced performance:

- **Hybrid Ensemble Architecture**

Unlike standard GB or XGB models, the hybrid frameworks (e.g., XGVO, GBVA) combine boosting techniques with meta-heuristic optimization algorithms. This approach enables the models to adaptively balance bias and variance, thereby improving generalization across various experimental configurations.

- **Automated Hyperparameter Optimization**

Prior studies often relied on manual tuning of hyperparameters, limiting model flexibility. In contrast, the integration of five meta-heuristic algorithms (AHA, VAO, TFWO, SAO, TDO) systematically explores a wider hyperparameter search space, leading to globally optimal configurations that enhance predictive accuracy.

- **Inclusion of Critical Experimental Variables**

The present study explicitly models four key input variables—intake diameter, approach velocity, Froude number, and bottom clearance ratio—ensuring that the model captures the hydraulic and geometric dependencies affecting critical submergence. Some previous works either used fewer variables or did not fully account for bottom clearance effects, thereby limiting the model's applicability under varying intake scenarios.

- **Dataset Size and Diversity**

With 324 experimental measurements, the dataset used here is larger and more diverse than those in many prior studies. The increased sample size enhances statistical robustness and ensures the model learns from a broader range of hydraulic conditions.

- **Custom Physical Consistency Equation**

To reduce physically inconsistent predictions at extreme clearance ratios, a custom empirical adjustment equation was embedded into the hybrid framework. This element is absent in most prior models, enabling better performance in scenarios with C = d_i/2, which are often challenging due to partial elevation effects.

5.5 Conditions of superior performance

Analysis indicates that the proposed models outperform SOTA, particularly under:

- Nonzero bottom clearance scenarios (C = d_i/2), where conventional models struggle to generalize.
- Nonlinear flow regimes with high Froude numbers, where simple regression or non-optimized SVR models fail to capture complex interactions.
- Limited training data partitions are available due to the enhanced ability of ensemble models to leverage cross-validation and meta-heuristic-guided hyperparameter tuning.

5.6 Architectural and input data considerations

The performance gain can also be attributed to:

- Ensemble effects, where the combination of multiple learners reduces variance and mitigates overfitting.
- Comprehensive feature representation, including both geometric (diameter, clearance) and hydraulic (velocity, Froude number) parameters.
- 10-fold cross-validation on the training set, ensuring reliable hyperparameter tuning without data leakage, which is often overlooked in previous studies.

Overall, these factors collectively explain why the proposed hybrid ensemble framework consistently surpasses prior models in terms of both R² and RMSE, demonstrating superior flexibility, generalization, and physical fidelity.

6 Conclusion

This study investigated the critical submergence (S_c) required to prevent the formation of air entrainment and an air cone vortex in horizontal and circular side intakes in open and concrete flumes. Horizontal intakes are commonly used for withdrawing water from open channels. The related study used two computational intelligence models, including gradient boost regression (GBR) and extreme gradient boost (XGB), to estimate the S_c. The accurate estimation of S_c is crucial in preventing air entrainment, so five meta-heuristic algorithms have been used to improve the accuracy of the related models containing Artificial Hummingbird Algorithm (AHA), Victoria Amazonica Optimization (VAO), Turbulent Flow of Water-based Optimization (TFWO), Smell Agent Optimization (SAO), and Tasmania Devil Optimizer (TDO). Based on the findings, the following conclusions can be inferred:

- The findings indicate that the XGB model performs more desirably than the GBR framework in estimating the S_c for two bottom clearances ($C = 0$ and $C = d_i/2$). When $C = 0$, the R^2 difference values of the XGB compared to GBR for train, validation, and test data sets were 2, 2, and 3 percent, respectively. In addition, when $C = d_i/2$, the R^2 difference values of the XGB compared to GBR for train, validation, and test data sets obtained 2, 3, and 2 percent, respectively.
- The next stage involves the use of a hybridization technique to enhance the accuracy of the suggested models (GBR, XGB). Specifically, the performance of these models was optimized using two meta-heuristic algorithms, AHA and VAO. The findings indicated that the majority of the hybrid models outperformed the original individual models. Additionally, the outcomes revealed that the hybrid XGVO model exhibited the most accurate prediction of S_c among all hybrid models with the highest R^2 values of 0.993 and 0.991in the $C = 0$ and $C = d_i/2$.
- In the final stage, a hybridization technique was employed to improve the precision of the ensemble model. As the results indicated, in the train, validation, and test phases, the GXVA models exhibit higher R^2 values with slight differences equal to 0.5, 0.2, and 1 compared to the XGAH model in cases of $C = 0$, respectively. Moreover, when $C = d_i/2$, the GXVA obtained higher R^2 values with 0.6, 1, and 0.1 differences for the train, validation, and test datasets, respectively, compared to the XGAH model.
- In general, the hybrid XGVO model achieved high accuracy and low error compared to all models, with suitable values of the presented metrics corresponding to the three specified sections.
- In the sensitivity analysis of the input variables on the respective outputs, it was determined that, according to the changes in R^2 and RMSE, Fr had the highest sensitivity to the Sc in both forms of bottom clearance.

In general, this study offered an innovative approach to address vortex-related issues in horizontal water intake processes. It employed individual, hybrid, and ensemble techniques (GBR and XGB) to estimate S_c. Especially, the study considers varying pipe diameters and hydraulic conditions, ensuring a comprehensive analysis. This study significantly advances predictive accuracy in modeling S_c for horizontal intakes by applying a hybridization technique and algorithms, demonstrating a strong commitment to enhancing prediction accuracy.

List of abbreviations

S_c	Critical Submergence	C	Bottom Clearance
U_i	velocity of flow in the intake	U_∞	initial flow velocity in the flume
Q_i	intake discharge	Re	Reynolds number
We	Weber number	Fr	Froude number
ML	Machine Learning	XGB	Extreme Gradient Boosting
AHA	Artificial Hummingbird Algorithm	GBR	Gradient Boosting Regression
TDO	Tasmanian Devil Optimization	SAO	Smell Agent Optimization
TFWO	Turbulent Flow of Water-based Optimization	VAO	Victoria Amazonica Optimization
XGTF	XGB + TFWO	GBTF	GBR + TFWO
XGSA	XGB + SAO	GBSA	GBR + SAO
XGTD	XGB + TDO	GBTD	GBR + TDO
XGVO	XGB + VAO	GBVO	GBR + VAO
XGAH	XGB + AHA	GBAH	GBR + AHA
GXTO	GB + XGB + TDO	GXVA	GB + XGB + VAO
GXSO	GB + XGB + SAO	GXAA	GB + XGB + AHA
GXTW	GB + XGB + TFWO	REC	Regression Error Trait
RMSE	Root Mean Square Error	MAE	Mean Absolute Error
IOA	Index of agreement	R^2	Coefficient Correlation

References

[1] Taştan, K (2018). Prediction of critical submergence in a two-layer-stratified fluid flow. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 40(4): 181. https://doi.org/10.1007/s40430-018-1112-x

[2] Yildirim, N. and F. Kocabaş (1995). Critical Submergence for Intakes in Open Channel Flow. *Journal of Hydraulic Engineering*, 121(12): 900–905. https://doi.org/10.1061/(ASCE)0733-9429(1995)121:12(900)

[3] BORGHEEI, S.M. and S.A.A.R. KABIRI (2010). Effect of anti-vortex plates on critical submergence at a vertical intake.

[4] Taştan, K (2016). Critical submergence for isolated and dual rectangular intakes. *Sādhanā*, 41: 425–433. https://doi.org/10.1007/s12046-016-0474-y

[5] Hashid, M., A. Hussain and Z. Ahmad (2021). Critical submergence for side circular intake in an open channel flow. *Journal of Hydraulic Research*, 59(1): 136–147. https://doi.org/10.1080/00221686.2020.1744749

[6] Asadzadeh, F., A. Safarzadeh and N.S.A.A. SALEHI (2016). Experimental study of flow around a spur dike with side slope.

[7] Ayoubloo, M.K., H.Md. Azamathulla, E. Jabbari and M. Zanganeh (2011). Predictive model-based for the critical submergence of horizontal intakes in open channel flows with different clearance bottoms using CART, ANN and linear regression approaches. *Expert Systems with Applications*, 38(8): 10114–10123. https://doi.org/10.1016/j.eswa.2011.02.073

[8] Taştan, K. and N. Yildirim (2017). Effective intake for critical submergence in the case of more than one intake. *KSCE Journal of Civil Engineering*, 21: 1004–1008. https://doi.org/10.1007/s12205-016-0073-9

[9] Yildirim, N. and F. Kocabaş (2002). Prediction of critical submergence for an intake pipe. *Journal of Hydraulic Research*, 40(4): 507–518. https://doi.org/10.1080/00221680209499892

[10] Knauss, J (2017). Prediction of critical submergence, In *Swirling Flow Problems at Intakes*, Routledge, pp: 57–76.

[11] Kocabaş, F., S. Ünal and B. Ünal (2008). A neural network approach for prediction of critical submergence of an intake in still water and open channel flow for permeable and impermeable bottom. *Computers & Fluids*, 37(8): 1040–1046. https://doi.org/10.1016/j.compfluid.2007.11.002

[12] Kocabaş, F., Ö. Kişi and M. Ardiçlioğlu (2009). An artificial neural network model for the prediction of critical submergence for intake in a stratified fluid medium. *Civil Engineering and Environmental Systems*, 26(4): 367–375. https://doi.org/10.1080/10286600802200130

[13] Yıldırım, N., A.S. Eyüpoğlu and K. Taştan (2012). Critical submergence for dual rectangular intakes. *Journal of Energy Engineering*, 138(4): 237–245. https://doi.org/10.1061/(ASCE)EY.1943-7897.0000073

[14] Haspolat, E. and M. Gogus (2022). Estimation of Critical Submergence at Single Horizontal Intakes Under Asymmetric Flow Conditions. *Arabian Journal for Science and Engineering*, 47(10): 12509–12520. https://doi.org/10.1007/s13369-021-06507-5

[15] Hashim, F.A., K. Hussain, E.H. Houssein, M.S. Mabrouk and W. Al-Atabany (2021). Archimedes optimization algorithm: a new metaheuristic algorithm for solving optimization problems. *Applied Intelligence*, 51: 1531–1551. https://doi.org/10.1007/s10489-020-01893-z

[16] Khanarmuei, M., H. Rahimzadeh and H. Sarkardeh (2019). Effect of dual intake direction on critical submergence and vortex strength. *Journal of Hydraulic Research*, 57(2): 272–279. https://doi.org/10.1080/00221686.2018.1459896

[17] Roushangar, K. and R. Ghasempour (2018). Evaluation of the Performance of Classical and Artificial Intelligence Approaches in Prediction of Critical Submergence of Horizontal Intakes in Open Channel Flows. *Water and Soil Science*, 28(1): 69–82.

[18] Azamathulla, H.M. and Z. Ahmad (2012). GP approach for critical submergence of intakes in open channel flows. *Journal of Hydroinformatics*, 14(4): 937–943. https://doi.org/10.2166/hydro.2012.089

[19] Goel, A (2012). Prediction of critical submergence for horizontal intakes, In *Proc., World Congress on Engineering* 2012, .

[20] Ahmad, Z., K. V Rao and M.K. Mittal (2008). Critical submergence for horizontal intakes in open channel flows. *Dam Engineering*, 19(2): 71–90.

[21] Ahmad, Z., K. V Rao and M.K. Mittal (2008). Critical submergence for horizontal intakes in open channel flows. *Dam Engineering*, 19(2): 71–90.

[22] Naeim, B., M.R. Akbarzadeh and V. Jahangiri (2024). Machine learning-based prediction of seismic response of elevated steel tanks. *Structures*, 70: 107649. https://doi.org/10.1016/j.istruc.2024.107649

[23] Akbarzadeh, M.R., V. Jahangiri, B. Naeim and A. Asgari (2025). Advanced computational framework for fragility analysis of elevated steel tanks using hybrid and ensemble machine learning techniques. *Structures*, 81: 110205. https://doi.org/10.1016/j.istruc.2025.110205

[24] Jahangiri, V., M.R. Akbarzadeh, S.A. Shahamat, A. Asgari, B. Naeim and F. Ranjbar (2025). Machine learning-based prediction of seismic response of steel diagrid systems, In *Structures*, Elsevier, p: 109791. https://doi.org/10.1016/j.istruc.2025.109791

[25] Wei, Z., Y. Meng, W. Zhang, J. Peng and L. Meng (2019). Downscaling SMAP soil moisture estimation with gradient boosting decision tree regression over the Tibetan Plateau. *Remote Sensing of Environment*, 225: 30–44. https://doi.org/10.1016/j.rse.2019.02.022

[26] Chen, T. and C. Guestrin (2016). Xgboost: A scalable tree boosting system, In *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, pp: 785–794. https://doi.org/10.1145/2939672.2939785

[27] Punmiya, R. and S. Choe (2019). Energy theft detection using gradient boosting theft detector with feature engineering-based preprocessing. *IEEE Transactions on Smart Grid*, 10(2): 2326–2329. https://doi.org/10.1109/TSG.2019.2892595

[28] Jahangiri, V., B. Naeim, M.R. Akbarzadeh and A. Asgari (2025). Optimal intensity measures for resilience-oriented probabilistic seismic demand models of elevated steel tanks, In *Structures*, Elsevier, p: 110576. https://doi.org/10.1016/j.istruc.2025.110576

[29] Naeim, B., A. Javadzade Khiavi, E. Khajavi, A.R. Taghavi Khanghah, A. Asgari, R. Taghipour and M. Bagheri (2025). Machine Learning Approaches for Fatigue Life Prediction of Steel and Feature Importance Analyses. *Infrastructures,* 10(11): 295. https://doi.org/10.3390/infrastructures10110295

[30] Mousavi, S.M.H (2023). Victoria Amazonica Optimization (VAO): An Algorithm Inspired by the Giant Water Lily Plant. *ArXiv Preprint ArXiv:2303.08070*. https://doi.org/10.48550/arXiv.2303.08070

[31] Wang, J., Y. Li, G. Hu and M. Yang (2022). An enhanced artificial hummingbird algorithm and its application in truss topology engineering optimization. *Advanced Engineering Informatics*, 54: 101761. https://doi.org/10.1016/j.aei.2022.101761

[32] Zhao, W., L. Wang and S. Mirjalili (2022). Artificial hummingbird algorithm: A new bio-inspired optimizer with its engineering applications. *Computer Methods in Applied Mechanics and Engineering*, 388: 114194. https://doi.org/10.1016/j.cma.2021.114194

[33] Zhao, W., Z. Zhang, S. Mirjalili, L. Wang, N. Khodadadi and S.M. Mirjalili (2022). An effective multi-objective artificial hummingbird algorithm with dynamic elimination-based crowding distance for solving engineering design problems. *Computer Methods in Applied Mechanics and Engineering*, 398: 115223. https://doi.org/10.1016/j.cma.2022.115223

[34] Fathy, A (2022). A novel artificial hummingbird algorithm for integrating renewable based biomass distributed generators in radial distribution systems. *Applied Energy*, 323: 119605. https://doi.org/10.1016/j.apenergy.2022.119605

[35] Dehghani, M., Š. Hubálovský and P. Trojovský (2022). Tasmanian Devil Optimization: A New Bio-Inspired Optimization Algorithm for Solving Optimization Algorithm. *IEEE Access*, 10: 19599–19620. https://doi.org/10.1109/ACCESS.2022.3151641

[36] Abdechiri, M., M.R. Meybodi and H. Bahrami (2013). Gases Brownian motion optimization: an algorithm for optimization (GBMO). *Applied Soft Computing*, 13(5): 2932–2946. https://doi.org/10.1016/j.asoc.2012.03.068

[37] Ghasemi, M., I.F. Davoudkhani, E. Akbari, A. Rahimnejad, S. Ghavidel and L. Li (2020). A novel and effective optimization algorithm for global optimization and its engineering applications: Turbulent Flow of Water-based Optimization (TFWO). *Engineering Applications of Artificial Intelligence*, 92: 103666. https://doi.org/10.1016/j.engappai.2020.103666

[38] Zhang, C. and Y. Ma (2012). Ensemble machine learning: methods and applications. *Springer*.

[39] Zounemat-Kermani, M., O. Batelaan, M. Fadaee and R. Hinkelmann (2021). Ensemble machine learning paradigms in hydrology: A review. *Journal of Hydrology*, 598: 126266. https://doi.org/10.1016/j.jhydrol.2021.126266

[40] Hosseini, M. and R. Kerachian (2023). Optimal redesign of coastal groundwater quality monitoring networks under uncertainty: application of the theory of belief functions. *Environmental Science and Pollution Research*, 30(21): 59701–59718. https://doi.org/10.1007/s11356-023-26764-1

[41] Dempster, A.P (2008). The Dempster–Shafer calculus for statisticians. *International Journal of Approximate Reasoning*, 48(2): 365–377. https://doi.org/10.1016/j.ijar.2007.03.004

[42] Shafer, G (2016). A mathematical theory of evidence turns 40. *International Journal of Approximate Reasoning*, 79: 7–25. https://doi.org/10.1016/j.ijar.2016.07.009

[43] Smets, P (2013). Practical uses of belief functions. *ArXiv Preprint ArXiv:1301.6741*. https://doi.org/10.48550/arXiv.1301.6741

[44] Ayoubloo, M.K., H.Md. Azamathulla, E. Jabbari and M. Zanganeh (2011). Predictive model-based for the critical submergence of horizontal intakes in open channel flows with different clearance bottoms using CART, ANN and linear regression approaches. *Expert Systems with Applications*, 38(8): 10114–10123. https://doi.org/10.1016/j.eswa.2011.02.073

[45] Amroune, M., T. Bouktir and I. Musirin (2018). Power system voltage stability assessment using a hybrid approach combining dragonfly optimization algorithm and support vector regression. *Arabian Journal for Science and Engineering*, 43(6): 3023–3036. https://doi.org/10.1007/s13369-017-3046-5

[46] MIMechE, M.J.P.M.A (1977). The hydraulic design of pump sumps and intakes. *British Hydromechanics Reserch Association*.

[47] Gulliver, J.S., A.J. Rindels and K.C. Lindblom (1986). Designing intakes to avoid free-surface vortices. *INT. WATER POWER & DAM CONSTR.*, 38(9, Sep. 1986, pp. 24-28.):

[48] Hashid, M. and Z. Ahmad (2022). Critical submergence for horizontal dual water intakes under perpendicular uniform approach flow. *Journal of Hydraulic Engineering*, 148(10): 04022020. https://doi.org/10.1061/(ASCE)HY.1943-7900.0002016

[49] Gordon, J.L (1970). *Vortices at intakes*.

[50] Reddy, Y.R. and J.A. Pickford (1972). Vortices at intakes in conventional sumps. *Water Power*, 24(3): 108–109.

[51] Gogus, M. and S. Gokmener (2023). Critical submergence for single and multiple horizontal intake structures. *Arabian Journal for Science and Engineering*, 48(10): 13091–13115. https://doi.org/10.1007/s13369-023-07651-w

[52] Das, B., Z. Ahmad and P.K. Sharma (2024). Critical Submergence for Lateral Rectangular Intakes: A CFD Study, In *EGU General Assembly Conference Abstracts*, , p: 824. https://ui.adsabs.harvard.edu/link_gateway/2024EGUGA..26..824D/doi:10.5194/egusphere-egu24-824

[53] Das, B., Z. Ahmad and P.K. Sharma (2022). Experimental and Numerical Investigation on Critical Submergence for Square Water Intakes. https://doi.org/10.26077/817c-13b1