Journal of Science and Transport Technology
Journal homepage: https://jstt.vn/index.php/en

# Predicting the Maximum Load Capacity of Circular RC Columns Confined with Fibre-Reinforced Polymer (FRP) Using Machine Learning Model

Indra Prakash[1], Thuy Anh Nguyen[2],*
[1]DDG (R), Geological Survey of India, Gandhinagar 382010, India
[2]University of Transport Technology, Hanoi 100000, Vietnam

**Article info**
**Type of article:**
Original research paper

**DOI:**
https://doi.org/10.58845/jstt.utt.2023.en.3.4.25-42

***Corresponding author:**
E-mail address:
anhnt@utt.edu.vn

**Received:** 21/11/2023
**Revised:** 19/12/2023
**Accepted:** 20/12/2023

**Abstract:** This article conducts an exhaustive investigation into the utilization of machine learning (ML) methods for forecasting the maximum load capacity (MLC) of circular reinforced concrete columns (CRCC) using Fiber-Reinforced Polymer (FRP). Extreme Gradient Boosting (XGB) algorithm is combined with novel metaheuristic algorithms, namely Sailfish Optimizer and Aquila Optimizer, to fine-tune its hyperparameters. The robustness and generalizability of these optimized hyperparameters are ensured through 200 Monte Carlo simulations (MCS). The model is constructed based on a database of 207 experimental results. Its performance is evaluated using three criteria: root mean squared error, mean absolute error, and the coefficient of determination. This study includes a performance comparison of the XGB4 model with eight other ML models, namely CatBoost (CAT), Gradient Boosting (GB), Hist Gradient Boosting (HGB), default XGB, Light Gradient Boosting (LGB), Linear Regression (LR), and Random Forest (RF). This comparison identifies the most effective model for predicting the MLC of columns. Additionally, this study explores the interpretability of the XGB model by SHAP values. This analysis illuminates the significance and interactions of various input features in predicting the FRP-confined CRCC's MLC. It offers insights into the primary elements influencing structural behavior by displaying a graphical depiction of the impact of specific characteristics on the model's output. This study culminates in developing an interactive Graphical User Interface (GUI) based on the XGB model. This tool allows users to investigate the influence of input parameters on the predicted MLC values, thereby enhancing their understanding and application of the model.

**Keywords:** FRP-confined RC circular columns, maximum load capacity, Machine Learning, XGB model.

## 1. Introduction

Using fiber-reinforced polymer (FRP) composites instead of traditional materials has greatly aided in strengthening or retrofitting various concrete elements because of their exceptional mechanical properties [1,2]. FRP composites are durable and resistant to corrosion, making them well-suited for situations where conventional reinforcing materials are susceptible to deterioration [3,4]. Specifically, by reducing the necessary concrete thickness and eliminating the need for additional reinforcement materials, FRP

confinement offers a cost-effective means of enhancing reinforced concrete (RC) components' performance [5]. FRP is mainly used for reinforcing RC columns and lateral confinement since it significantly enhances the performance of concrete structures when subjected to axial stresses. Confinement, especially when wrapping a column with circumferentially oriented FRP, significantly enhances the load capacity and ductility of RC columns [6]. Moreover, the outer confinement provided by the FRP wrapping increases the members' durability and prevents corrosion of the core concrete. According to relevant studies, the restraining effect of FRP can enhance the bearing capacity and ductility of core concrete in FRP-confined concrete columns by 14–923% and 2–46%, respectively [7-9]. Therefore, their axial load capacity must be investigated to expand the use of FRP-confined CRCC for structural engineering purposes.

Reinforcing circular RC columns (CRCC) using a variety of FRP materials in an effort to increase their ductility and strength has been the subject of numerous experimental investigations. Works such as Rahai et al. [10] and Ranolia et al. [11] studied the effects of several factors on the functionality of RC columns reinforced with FRP wraps. These characteristics included jacket stiffness, fiber orientation, and the presence of FRP. They found that the load-bearing capacity (LBC) and deformation capacities of the laden columns were enhanced by FRP reinforcement. Furthermore, the LBC of the columns improved with FRP layers, mainly when columns were enclosed by a circular FRP wrap [6]. Fardis investigated the influence of concrete core's compressive strength (CS) and cross-sectional dimensions on the axial compressive behavior of FRP-confined circular concrete columns. Regarding experimental findings, a matching axial compressive behavior model was developed [12]. In another study, Samaan devised a model to analyze the axial compressive behavior of CRCC incorporating GFRP and examined the impact of FRP layers on this behavior [13]. The main subject of these investigations was concentric loads on FRP-confined CRCC. When used in engineering practice, eccentric loads are exerted on concrete columns. The influence of load eccentricity on the confinement behavior of FRP-jacketed concrete columns has been established in multiple studies [14-16]. Eccentric loading results in fluctuations in the confinement pressure throughout the section, which generates stress states distinct from those observed under concentric loading, particularly during specific axial deformations [17]. Maaddawy [18] investigated the effectiveness of reinforcing RC columns subjected to eccentric loading with carbon FRP (CFRP). An increase in column's LBC is observed, with a proportional decrease in enhancement as the eccentricity ratio increased. Furthermore, CFRP's confinement effect diminished as the eccentricity ratio increased [19]. Other studies [16,20,21] utilized finite element models (FEM) to examine the impact of eccentric loading on FRP-confined concrete columns. The findings indicate that the mechanical characteristics of FRP-confined concrete were significantly altered by eccentric loading. The behavior of FRP-reinforced columns under eccentric and axial loads has been the subject of a plethora of research investigations. Nevertheless, the current models available are just semi-empirical and semi-theoretical constructions that rely on restricted data. These models fail to account for the impact of factors on the final compressive strain and CS of columns. To assess the precision of the forecast, it is critical to construct a model capable of precisely predicting the axial compressive constitutive relationship of FRP-confined CRCC [22].

Machine Learning (ML) applications in science and engineering are expanding [23,24]. Applying ML techniques to predict the maximum load capacity (MLC) of FRP-confined CRCC is a promising area of active research and

development. ML techniques and deep learning can be employed to analyze and predict numerous engineering problems rapidly and efficiently [25,26]. Specifically, ML approach can reduce time and costs compared to traditional experimental methods for assessing column LBC. However, using ML to determine the MLC of FRP-confined CRCC is still in the research and development phase, with very few scientific studies being conducted. In a specific study by Obaidat et al. [6], an Artificial Neural Network (ANN) model was utilized to predict the ultimate load based on a dataset comprising 92 FRP-confined CRCC. The performance of the ANN model is assessed using $R^2$. Ghanem and Elgazzar [27] proposed an ANN model for predicting axial compressive stress and deformation in FRP-confined concrete columns with internal steel reinforcement. The database used for the analysis included experimental results from 64 FRP-confined concrete columns under concentric compressive loading. However, a comprehensive examination of all factors influencing the LBC of FRP-confined CRCC remains a relatively unexplored area.

The recently proposed eXtreme Gradient Boosting (XGB) algorithm [28] is a state-of-the-art ML method. This algorithm is commonly employed to address supervised learning problems with high accuracy and has been effective in several studies. Bakouregui et al. [29] harnessed XGB to predict the LBC of RC columns containing FRP bars, achieving remarkable performance with average $R^2$ and MAPE values of 0.98 and 5.3%, respectively. Extending the applicability of XGB, Le et al. [30] constructed an XGB model to forecast the shear strength of FRP-RC beams, considering the presence or absence of stirrups. The model exhibited high predictive prowess, boasting an $R^2$ of 0.964 and an impressive RMSE of 24.134 kN, surpassing comparisons with other ML models. In a separate investigation, Amjad et al. [31] leveraged XGB to estimate the LBC of piles. The performance of the XGB model is compared with commonly used algorithms such as Decision Tree (DT), Random Forest (RF), AdaBoost, and Support Vector Machine (SVM), demonstrating its outstanding superiority with an $R^2$ of 0.995 and an RMSE of 80.653 kN. Collectively, these studies underscore the versatility and efficacy of the XGB model in solving diverse engineering problems. However, to date, no study has been conducted on the efficacy of the XGB method for estimating the MLC of FRP-confined CRCC. Therefore, this study utilizes the XGB method to predict the MLC of FRP-confined CRCC. Additionally, Shapley values are employed to quantify the importance and influence of input parameters on the MLC of FRP-confined CRCC. In this study, Python software is used for model development and also to generate a graphical user interface (GUI).

## 2. Database description

To guarantee the precision of the predictive model, a substantial amount of experimental data concerning MLC of FRP-confined circular RC columns is necessary for training and evaluating the model. This study employs a dataset comprising 207 FRP-confined unreinforced and RC columns under concentric and eccentric loads, featuring various slenderness ratios and material characteristics, as the primary data source. The dataset used in this study is derived from various sources in the literature [15,32-40]. Among them, most of the samples are reinforced with longitudinal and hoop steel bars, except for those in [37],[39],[15] and [32]. Table 1 presents a comprehensive database summary, including data and related proportions. Parameters affecting the MLC of FRP-confined RC columns representing geometric dimensions and material properties are collected, including the diameter of a circular column section, column height, load eccentricity, CS of concrete, number of layers of FRP hoop wraps, nominal thickness of an FRP hoop sheet, stress in FRP wraps, elastic modulus of the FRP, and ultimate tensile strain of FRP. For a more

intuitive grasp of these nine parameters, this study resorts to descriptive statistics and visual representation through box plots, as shown in Fig. 1.

To guarantee the prediction model's capacity for generalization, nine parameters are selected as the model's input parameters, and the MLC of FRP-confined CRCC is selected as the model's output parameter. This suggests that the interrelation between the nine factors and the dependent variable, which is the CS of concrete, was taken into account concurrently. Table 2 shows an overview of the numerical attributes of parameters.

The XGB model construction process adheres to the classic dataset partition approach. Specifically, 70% of the entire experimental dataset is allocated to the training set, with the remaining 30% reserved for the testing set. This 70/30 data split ratio has effectively ensured data confidence and representativeness during ML models' training and testing phases [41,42].

**Table 1**. Summary of the collected database

| Ref. | Data points | Proportion (%) |
|---|---|---|
| Al-Nimry and Rahadi [33] | 20 | 9.66 |
| Al-Nimry and Soman [34] | 32 | 15.50 |
| Bisby and Ranger [35] | 8 | 3.86 |
| Fitzwilliam and Bisby [36] | 18 | 8.70 |
| Jiang et al. [37] | 16 | 7.73 |
| Siddiqui et al. [38] | 7 | 3.38 |
| Wang et al. [32] | 8 | 3.86 |
| Wu and Jiang [15] | 36 | 17.40 |
| Wang et al. [39] | 24 | 11.60 |
| Kaeseberg et al. [40] | 38 | 18.40 |
| **Total** | **207** | **100** |

**Table 2.** Information of the considered parameters

| Parameter | Symbol | Range | Mean | Standard Deviation | Type |
|---|---|---|---|---|---|
| Diameter of a circular column section - mm | $X_1$ | 150.00 ÷ 305.00 | 181.12 | 38.00 | Input |
| Column height - mm | $X_2$ | 300.00 ÷ 1200.00 | 639.53 | 351.98 | Input |
| Load eccentricity - mm | $X_3$ | 0.00 ÷ 65.00 | 19.30 | 20.21 | Input |
| CS of concrete - MPa | $X_4$ | 21.20 ÷ 58.95 | 35.03 | 10.51 | Input |
| Number of layers of FRP hoop wraps | $X_5$ | 0.00 ÷ 2.00 | 1.16 | 0.75 | Input |
| Nominal thickness of an FRP hoop sheet - mm | $X_6$ | 0.11 ÷ 1.00 | 0.17 | 0.16 | Input |
| Stress in FRP wraps - MPa | $X_7$ | 846.00 ÷ 4900.00 | 4238.0 | 750.04 | Input |
| Elastic modulus of the FRP - MPa | $X_8$ | 77.30 ÷ 300.00 | 251.96 | 42.20 | Input |
| Ultimate tensile strain of FRP - % | $X_9$ | 1.10 ÷ 2.10 | 1.82 | 0.21 | Input |
| Maximum load capacity (MLC) - kN | $P_u$ | 35.00 ÷ 2041.00 | 658.04 | 526.10 | Output |

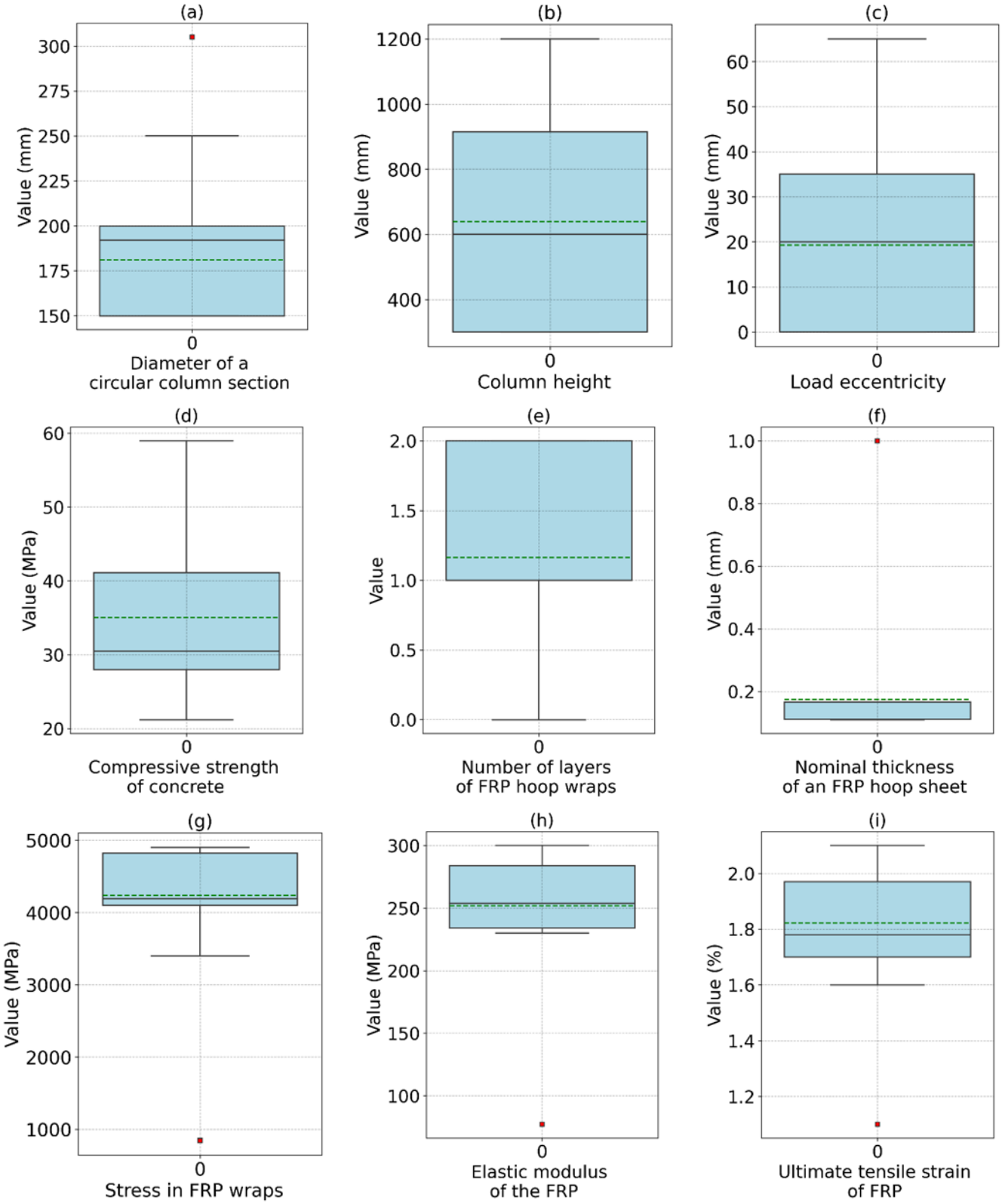


**Fig. 1.** Descriptive statistics and visual representation through box plots of variables

## 3. Methods

### 3.1. Machine learning methods

#### 3.1.1. Extreme Gradient Boosting

Extreme gradient boosting [28], commonly called XGB, is a powerful and highly effective ML algorithm under ensemble learning techniques. It has gained widespread recognition and popularity because of its exceptional performance in various

ML tasks, making it a prominent choice for data scientists and practitioners.

XGB, in its essence, is a boosting technique that amalgamates many weak learners, often decision trees, to construct a robust and precise prediction model. The "eXtreme" in XGB signifies its ability to handle complex relationships and patterns in data through an iterative approach. During each iteration, the algorithm focuses on the mistakes made by the previous models and works to correct them, ultimately improving the model's overall predictive accuracy.

One of the key features that sets XGB apart is its built-in support for regularization techniques, including L1 (Lasso) and L2 (Ridge) regularization. This helps prevent overfitting, ensuring the model generalizes well to unseen data. XGB also excels in handling missing data, as it learns the best imputation strategy during training.

By providing insights into feature significance, the method enables users to discern which factors have the most substantial influence on the predictions generated by the model. It is well-suited for parallel and distributed computing, making it capable of handling large datasets and benefiting from multiple cores and distributed computing environments.

### 3.1.2. Hybrid algorithms

*a. Sailfish Optimizer Algorithm (SFO)*

The SFO [43] is an innovative metaheuristic optimization technique that draws inspiration from the collective hunting behavior of sailfish, a species of fast-swimming marine fish. This algorithm has been designed to address various optimization problems across various domains, offering an efficient and natural approach to finding optimal solutions.

SFO capitalizes on the hunting strategy of sailfish as they work together to target and capture schools of sardines in the ocean. This collaborative approach among sailfish serves as a model for the optimization process, emphasizing the advantages of cooperation and coordination in achieving common goals.

In the SFO algorithm, sailfish represent the current optimal solution or the best-known solution, while the role of sardines is to explore the search space and seek improvements. The algorithm involves the random movement of sailfish and sardines within the solution space, where their positions correspond to the optimized problem variables.

The SFO algorithm can be divided into two phases: exploration and exploitation. During the exploration phase, sailfish and sardines move randomly to explore diverse candidate solutions. This phase generates a comprehensive understanding of the solution space. In the subsequent exploitation phase, the algorithm utilizes the information gathered during exploration to refine the search for the optimal solution. Sailfish positions are updated based on the best solution found thus far, ensuring that the algorithm converges toward the optimal solution.

*b. Aquila Optimizer Algorithm (AO)*

The AO is a novel metaheuristic algorithm inspired by the hunting behaviors of the Aquila, a genus of eagles in nature. This algorithm was introduced in 2021 by Abualigah et al. [44]. The name "Aquila" comes from the Latin term for eagle, considered a highly intelligent and skilled bird of prey, known for its powerful legs and sharp talons that it employs with speed and agility to capture its prey.

The AO algorithm is designed to mimic and leverage four primary hunting strategies employed by the Aquila: high soaring with a vertical stoop, contour fighting with a short glide, low flying with a gradual descent, and ground-based hunting and prey capture. These strategies serve as the foundation for the algorithm, which optimizes various processes by emulating the efficient hunting mechanisms of the Aquila.

The AO algorithm has demonstrated superior

computational efficiency and effectiveness compared with other optimization algorithms, making it a promising tool for solving various optimization problems across different domains. Its nature-inspired approach, drawing from the natural world, makes it an intriguing addition to optimization algorithms.

### 3.2. Model performance assessment

To precisely and impartially assess the precision of XGB model in predicting the MLC of FRP-confined CRCC, three criteria are utilized: $R^2$, MAE, and RMSE. $R^2$ measures the extent of the linear association between the model's predicted and desired values. Typically, a model is deemed reliable when its $R^2$ score exceeds 0.8, indicating accuracy. As the $R^2$ value approaches 1, the model's predictive precision becomes increasingly superior. Simultaneously, MAE and RMSE illustrate the disparity between the predicted and observed values. Smaller MAE and RMSE values correspond to enhanced predictive accuracy of the model. Overall, these evaluation indices are chosen to comprehensively evaluate the predictive model. These indices offer complementary information about the model's accuracy, precision, and goodness of fit, which is crucial for the objectives of the current study. The mathematical formulas for these three evaluation criteria can be found in the following equations:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Y_i - X_i\right)^2} \quad (1)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|Y_i - X_i\right| \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(Y_i - X_i\right)^2}{\sum_{i=1}^{N}\left(Y_i - \bar{X}\right)^2} \quad (3)$$

where $Y_i$, $X_i$ are the predicted and actual MLC values, respectively, $\bar{X}$ is experimental MLC values' average, and N is the number of instances.

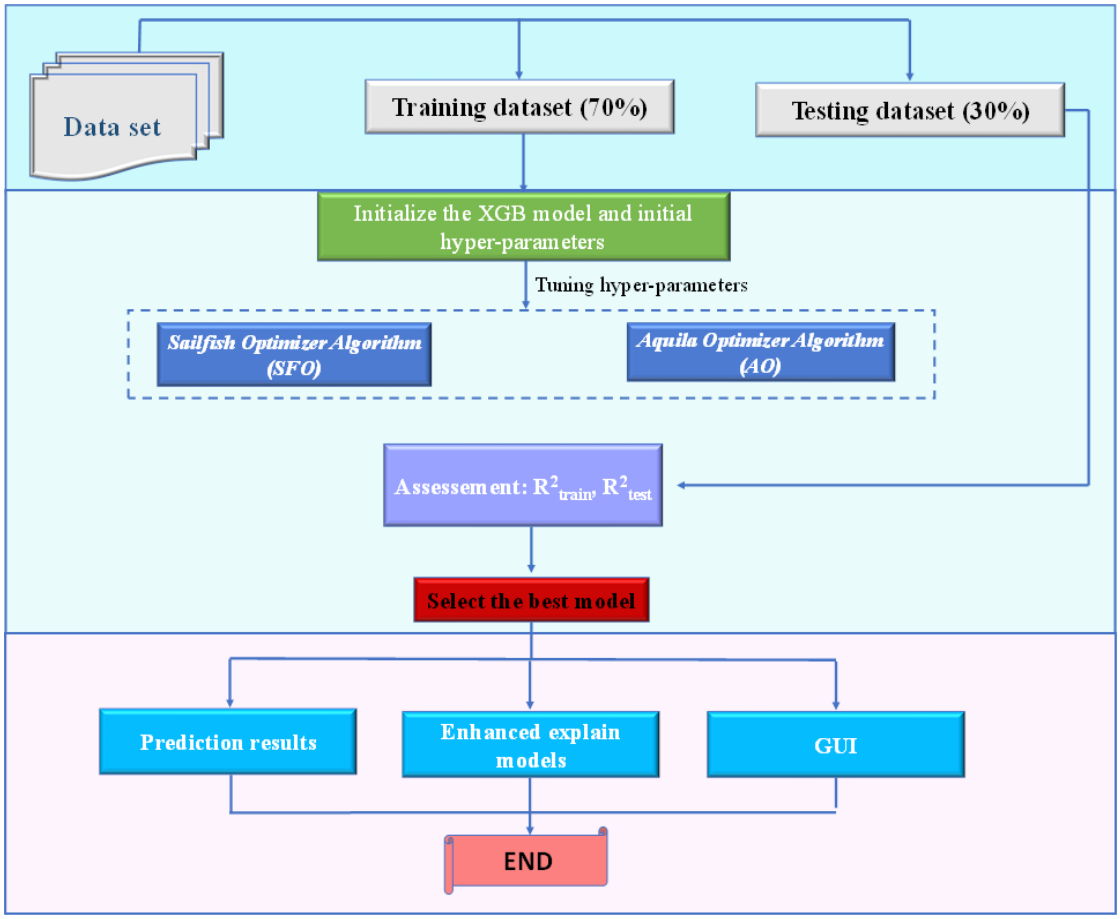


**Fig. 2.** Methodological flowchart

### 3.3. Methodological diagram of the study

The process of using the XGB model includes four main steps (Fig. 2):

**Step 1:** The dataset includes 207 experimental results from 10 published documents. It is randomly split into 2 parts at a 70/30 ratio, in which the training set accounts for 70%, and the testing one accounts for 30% of the remaining data set.

**Step 2:** Tune the hyperparameters of the XGB model: In this step, two optimization algorithms with population sizes of 10, 20, 30, and 40, respectively, are employed to fine-tune the 7 hyperparameters of the XGB model. The chosen objective function is $R^2$, which aims to optimize its value on both datasets.

**Step 3:** Select the best model: Compare the performance of the 8 established XGB models based on the highest $R^2$, then deduce the best model.

**Step 4:** In this step, representative prediction results regarding the MLC of FRP-confined CRCC, enhanced interpretability of the best-performing XGB model, and instructions for a user-friendly GUI are presented

## 4. Results and Discussion

### 4.1. Hyperparameter fine-tuning process

Fine-tuning the hyperparameters of ML models is paramount for achieving optimal performance. Adjusting these hyperparameters significantly influences a model's predictive accuracy and generalization capabilities. There are different methods for hyperparameter tuning, each with strengths and weaknesses. In this study, the SFO and the AO algorithm are selected as the optimization algorithms to fine-tune the hyperparameters of the XGB model. Seven critical hyperparameters of the XGB model have been chosen for fine-tuning: "n_estimators", "learning rate", "max_depth", "min_child_weight", "subsample", "colsample-bytree," and "gamma." These parameters are chosen based on their significant impact on the overall model performance [45,46]. The search space for these hyperparameters is outlined in Table 3, with the remaining parameters utilizing the default values in Python.

To further enhance the optimization process, the parameters of the optimization algorithm must be scrutinized. Specifically, careful consideration is given to the values for "epoch" and "population size" ($N_p$) to ensure the effectiveness and efficiency of the optimization process. The "epoch" is set at 500 to guarantee convergence, allowing for thorough solution space exploration while maintaining computational feasibility. Simultaneously, four different values of $N_p$, namely 10, 20, 30, and 40, are proposed during the optimization process. As mentioned earlier, the seven selected hyperparameters of the XGB model will serve as inputs for both the SFO and AO algorithms. With these chosen parameters, the algorithms will sweep through the ranges of the 7 hyperparameters to optimize the objective function. The selected objective function is $R^2$, aiming to maximize its value. The simulation results under different conditions are illustrated in Fig. 3. It is evident that as the "epoch" changes from 1 to 500, the $R^2$ function converges around the 200th iteration. This implies that the optimal hyperparameters for XGB models can be found after 200 epochs, and no superior hyperparameters are discovered after the exhaustive search of up to 500 epochs. Table 4 details the optimal hyperparameter and $R^2$ values of the eight XGB models. Comparing the performance of SFO and AO, it is evident that the SFO algorithm yields superior results. Further analysis reveals that the optimized XGB4 model using SFO with $N_p$ = 40 outperforms the other population sizes on the training and validation datasets. Therefore, the XGB4 model, optimized with the SFO algorithm and a population size of 40, is chosen as the final model for predicting the MLC of FRP-confined CRCC.

To assess the generalizability of the optimized hyperparameter set for the XGB4 model across diverse training and testing datasets, 200 MCS are conducted. The primary purpose of running MCS is to account for the inherent variability and uncertainty in the datasets and model outcomes. By repeatedly sampling the distributions of input parameters, MCS provide a comprehensive understanding of how the model behaves under different conditions and variations in the dataset. The convergence curves for the two evaluation criteria, RMSE and $R^2$, are illustrated in Fig. 4. These curves depict the normalized convergence of the model's predictive performance for the simulations. Analyzing the convergence curves for the training dataset reveals that the RMSE criterion converges within a 1% range, and the $R^2$ criterion converges within a 0.05% range after approximately 14 simulations. Conversely, these criteria converge after approximately 100 simulations for the testing dataset. This indicates that the proposed 200 MCS are sufficient to thoroughly examine the convergence of the model and validate the generalizability of the optimized hyperparameter set. The convergence behavior, mainly the stabilization of the evaluation criteria, demonstrates that the model's performance stabilizes and generalizes well across various datasets. In essence, MCS provide a robust means of assessing the reliability and generalizability of the model's hyperparameters, ensuring that the model's performance is consistently effective across various conditions and datasets.

**Table 3.** The XGB hyperparameters and their upper and lower bound values

| n_estimators | learning _rate | max_depth | min_child _weight | subsample | colsample_by tree | gamma |
|---|---|---|---|---|---|---|
| 1-500 | 0.005-0.3 | 1-10 | 1-10 | 0-1 | 0-1 | 0-2 |

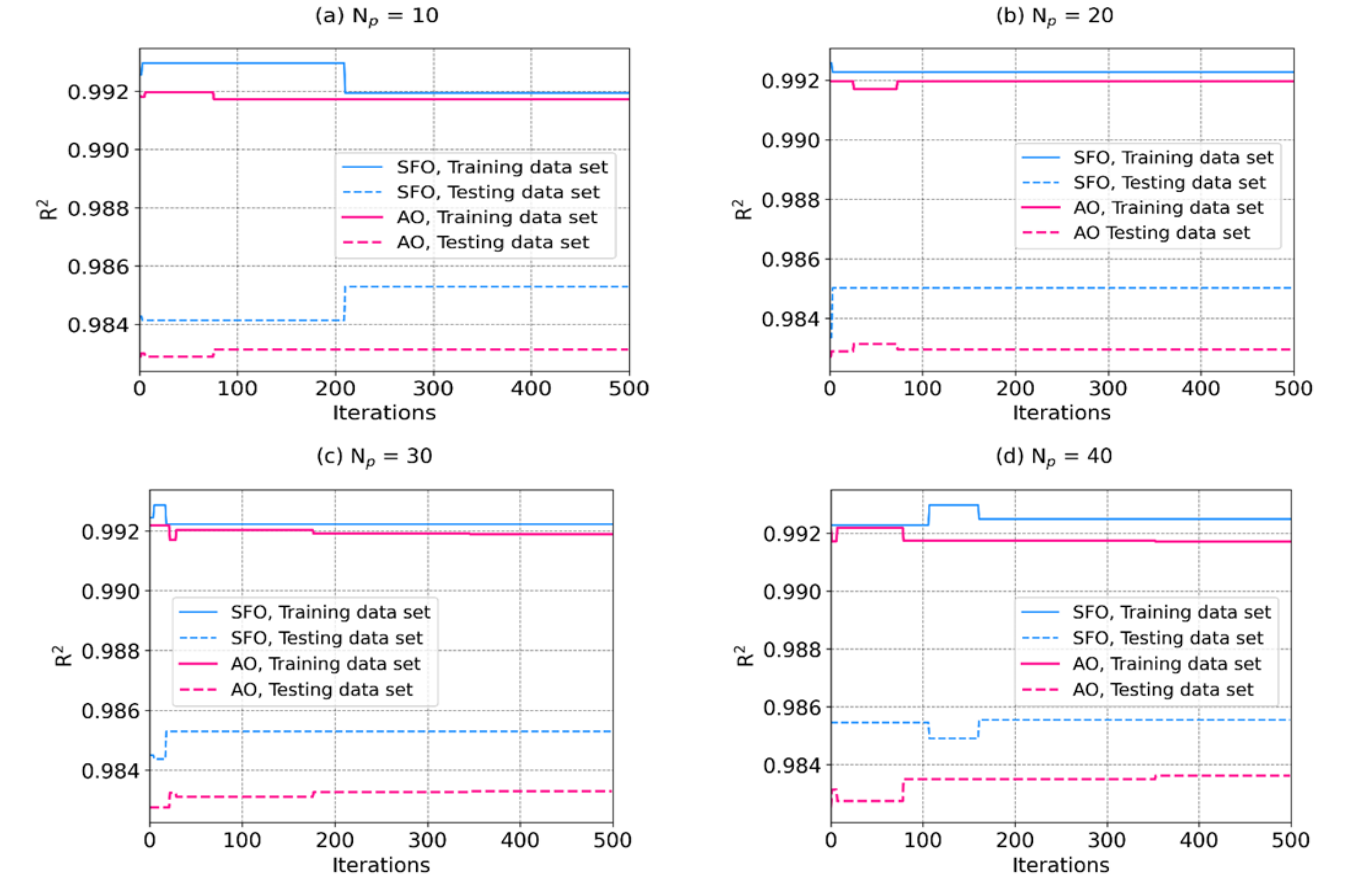


**Fig. 3.** Optimizing the XGB model using $R^2$ criteria

**Table 4.** Prediction results of 8 XGB models

| Optimization algorithm | SFO Algorithm | | | | AO Algorithm | | | |
|---|---|---|---|---|---|---|---|---|
| **Model XGB** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** |
| **$N_p$** | 10 | 20 | 30 | 40 | 10 | 20 | 30 | 40 |
| **n_estimators** | 81 | 79 | 81 | 83 | 192 | 132 | 106 | 62 |
| **learning_rate** | 0.259 | 0.248 | 0.3 | 0.270 | 0.3 | 0.3 | 0.3 | 0.3 |
| **max_depth** | 4 | 4 | 4 | 4 | 10 | 7 | 5 | 9 |
| **min_child_weight** | 2 | 1 | 2 | 2 | 10 | 7 | 6 | 6 |
| **subsample** | 0.864 | 0.869 | 0.837 | 0.899 | 1 | 0.991 | 1 | 1 |
| **colsample_bytree** | 0.864 | 0.848 | 0.837 | 0.899 | 1 | 0.991 | 1 | 1 |
| **gamma** | 1.727 | 1.444 | 1.201 | 1.862 | 2 | 1.700 | 2 | 2 |
| **$R^2_{train}$** | 0.9919 | 0.9923 | 0.9922 | 0.9925 | 0.9917 | 0.9919 | 0.9919 | 0.9917 |
| **$R^2_{test}$** | 0.9853 | 0.9850 | 0.9853 | 0.9856 | 0.9831 | 0.9830 | 0.9833 | 0.9836 |

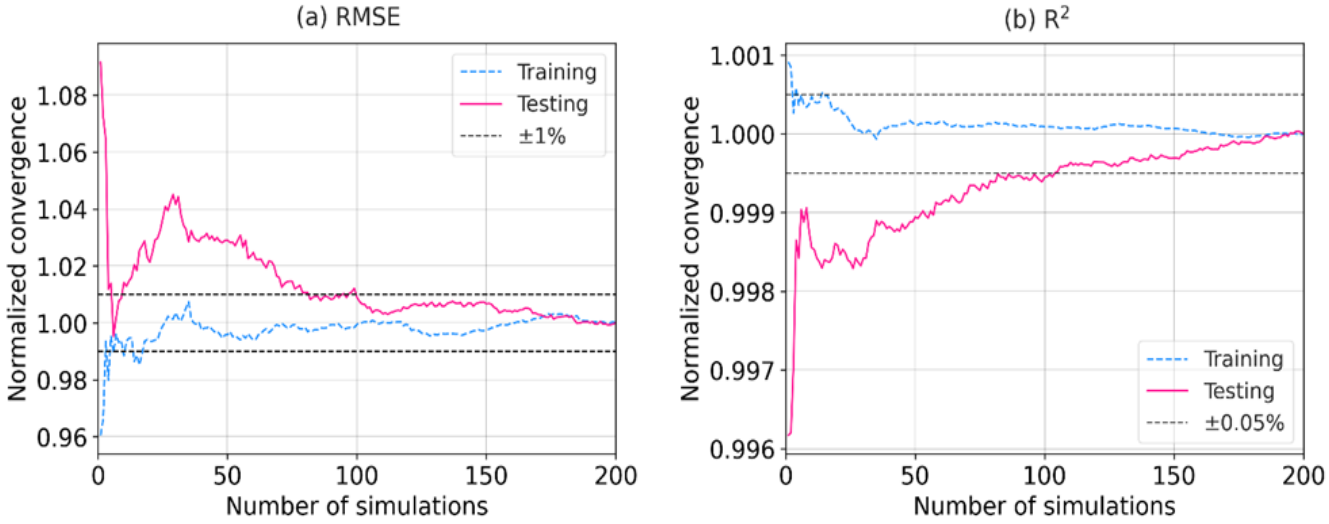


**Fig. 4.** Normalized convergence curves of RMSE and $R^2$ after 200 MCS

### 4.2. Model prediction results

This part shows and discusses the outcomes of using the XGB4 model to forecast the MLC of FRP-confined CRCC. As seen in Fig. 5, the association between expected and actual values is presented on a regression plot. The x-axis represents the experimental LBC of the columns, while the y-axis represents the predicted values. The solid black line indicates a near-perfect match, whereas the black dashed and blue dotted lines represent 15% and 30% error margins, respectively. The XGB4 model exhibits exceptional performance on both the training and testing datasets, with most data points aligning closely along the y = x line, indicating a highly accurate prediction. Furthermore, Fig. 6 provides an error plot for the training and testing datasets. The significance of the error plot lies in visualizing the disparity between predicted and actual values. A concentrated cluster of points around zero implies low prediction errors, strengthening the model's reliability. The detailed performance of the XGB4 model is in Table 5.

### 4.3. Comparison with other ML models

In this section, the performance comparison of the XGB4 model with eight other ML models is performed, including CatBoost (CAT), Gradient Boosting (GB), Hist Gradient Boosting (HGB), default XGB, Light Gradient Boosting (LGB), Linear Regression (LR), and RF, to identify the most

effective model for predicting the LBC of columns. Model performance evaluation is conducted using three statistical criteria: RMSE, MAE, and $R^2$. Fig. 7 provides a bar chart illustrating the performance values of the ML models on both the training and testing datasets, while Table 6 details these values. Regarding the training dataset, the XGB_def model exhibits the best performance, followed by the CAT and XGB4 models. Notably, the XGB4 model does not achieve the highest accuracy on the training set. However, the XGB4 model outperforms all the other models on the testing set. The performance of the XGB4 model on the testing set improved the error by 43.97% compared with XGB_def and by 28.96% compared with CAT. The LR performs the poorest on the training and testing sets. This could be attributed to its linear assumption, which might not capture the complex relationships in the data as effectively as other models.

In conclusion, the XGB4 model demonstrates the highest efficiency in predicting the LBC of columns, especially evident in the validation set. Despite not being the top performer on the training set, the XGB4 model's improved performance on the validation set emphasizes its effectiveness in generalization and its superior capability in predicting the MLC of FRP-confined CRCC.

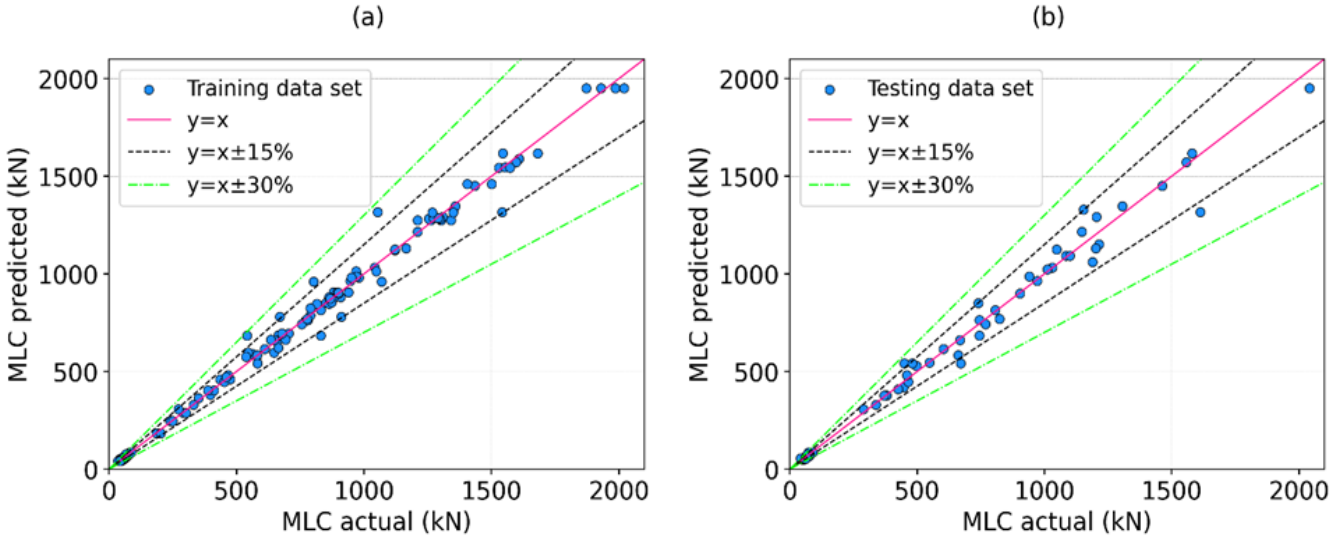


**Fig. 5.** Regression graphs between MLC prediction results and actual values: (a) training, and (b) testing

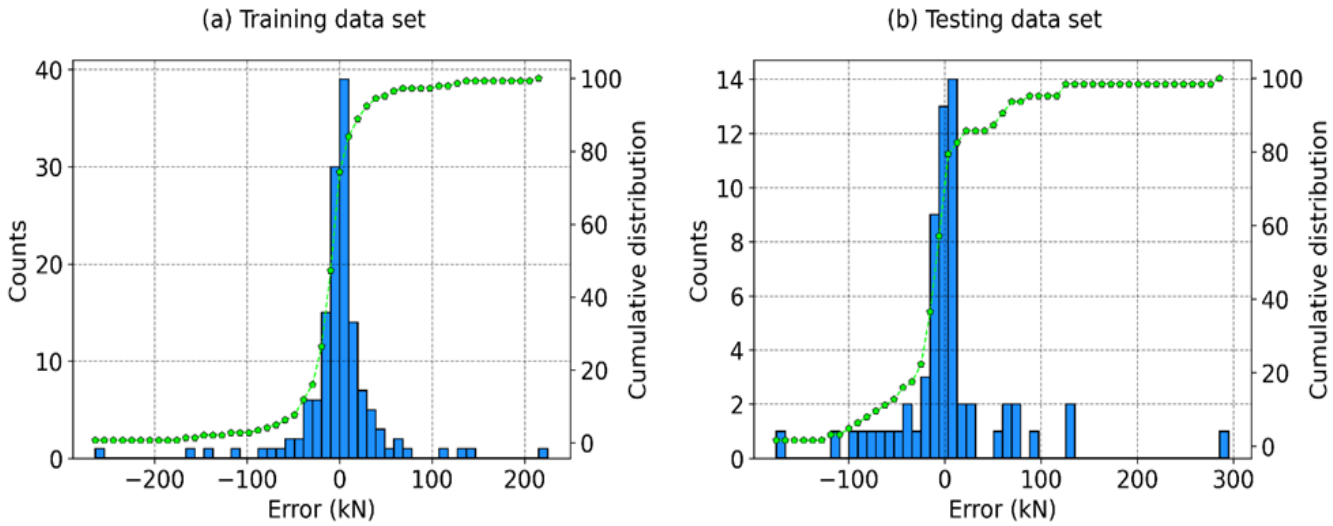


**Fig. 6.** Error between MLC prediction results and actual values: (a) training, and (b) testing

**Table 5.** Values of statistical criteria of XGB4 model

| | RMSE (kN) | MAE (kN) | $R^2$ |
|---|---|---|---|
| **Training data set** | 45.982 | 24.020 | 0.992 |
| **Testing data set** | 61.100 | 34.394 | 0.986 |
| **All data** | 51.059 | 27.177 | 0.991 |

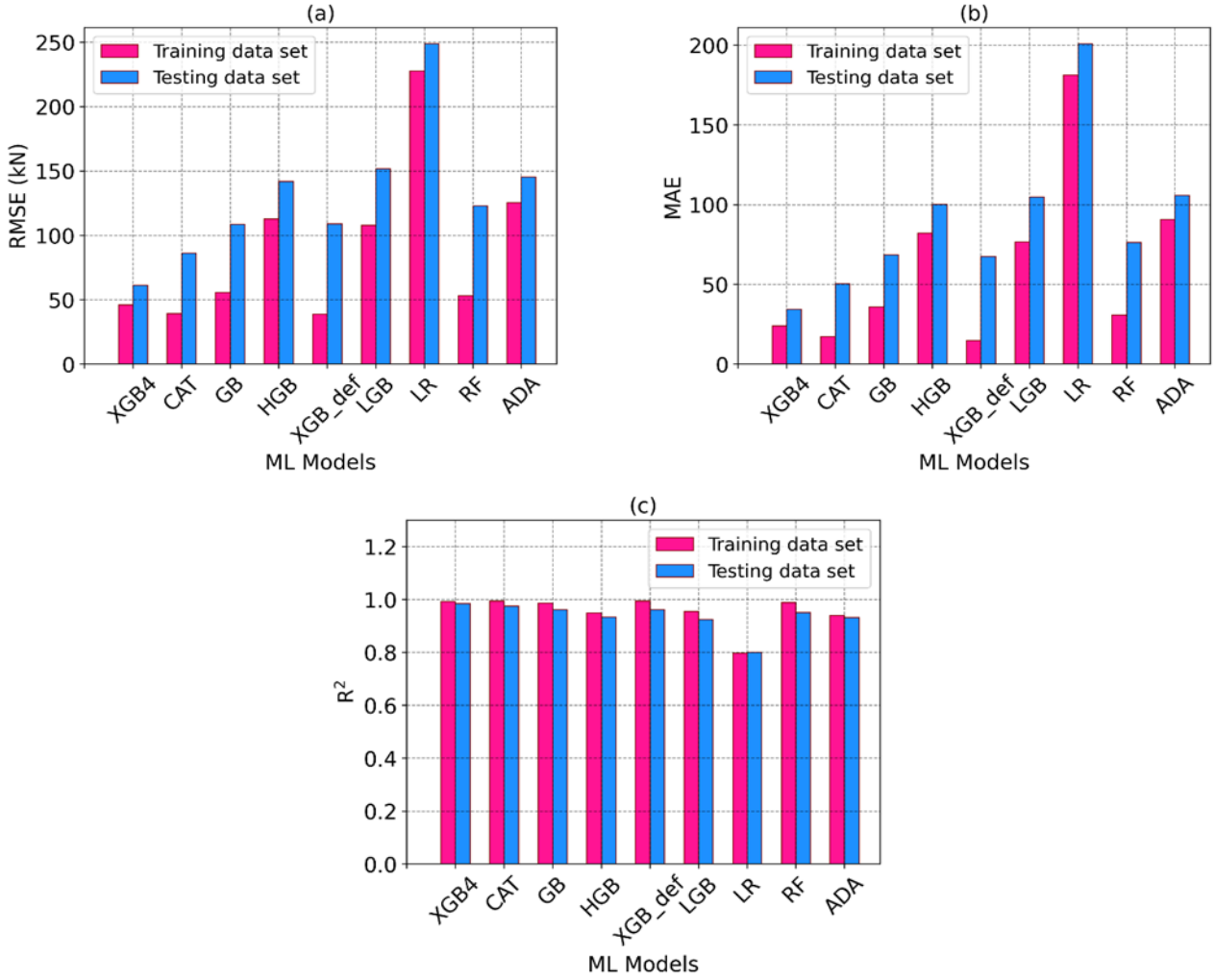


**Fig. 7.** Detailed accuracy of different ML models

**Table 6.** Statistical criteria of different ML models

| | RMSE (kN) | MAE (kN) | $R^2$ | RMSE (kN) | MAE (kN) | $R^2$ |
|---|---|---|---|---|---|---|
| | Training data set | | | Testing data set | | |
| **XGB4** | 45.982 | 24.020 | 0.992 | **61.100** | **34.394** | **0.986** |
| **CAT** | 39.131 | 17.385 | 0.994 | 86.002 | 50.498 | 0.976 |
| **GB** | 55.604 | 35.948 | 0.988 | 104.265 | 65.454 | 0.965 |
| **HGB** | 112.804 | 82.184 | 0.950 | 141.862 | 100.262 | 0.935 |
| **XGB_def** | 38.706 | 14.922 | 0.994 | 109.053 | 67.374 | 0.962 |
| **LGB** | 107.722 | 76.641 | 0.955 | 151.842 | 104.814 | 0.925 |
| **LR** | 227.624 | 181.457 | 0.800 | 248.935 | 200.951 | 0.799 |
| **RF** | 54.673 | 32.937 | 0.988 | 119.194 | 76.812 | 0.954 |
| **ADA** | 133.131 | 101.524 | 0.931 | 159.323 | 119.159 | 0.918 |

### 4.4. Enhanced Explaining Ability of ML Models

This study comprehensively depicts the ML model and the dependence/interaction of all considered features. A technique closely approximating SHAP, termed *TreeExplainer* and tailored for tree-based models, has been applied. The model's predictions are interpreted using this technique, which assigns importance to each feature based on the decision tree's structure. Each feature's contribution to the model's output for a specific prediction is quantified by TreeExplainer, enabling a more profound understanding of the model's decision-making process.

The superior predictive accuracy of the XGB4 model for the MLC of FRP-confined CRCC is demonstrated. The interpretation of this model for the MLC of FRP-confined circular RC columns using SHAP values is discussed in this section. Fig. 8a illustrates the correlation of various features with SHAP values for MLC of FRP-confined CRCC derived from XGB4 model.

Stress in FRP wraps emerges as the parameter that most crucially influences the MLC of FRP-confined CRCC. The highest SHAP values for this parameter indicate a substantial impact. Stress in FRP wraps measures the stress experienced by the FRP wraps, suggesting that the behavior of these wraps significantly affects the structural integrity of the column. The emergence of stress in FRP wraps as the parameter most crucially influencing MLC can be attributed to the material properties of FRP. When FRP wraps are subjected to stress, they contribute significantly to the confinement effect on the concrete core. Higher stress levels in the FRP wraps indicate a greater capacity to withstand external loads and provide enhanced confinement to the concrete, leading to increased MLC. Following this, the CS of concrete garners the second-highest Shap value, underscoring its pivotal role in determining MLC. The concrete CS plays a pivotal role in determining MLC due to its direct impact on the load-carrying capacity of the unconfined concrete core. A higher CS means the concrete can withstand larger axial loads before experiencing failure. Load eccentricity ranks next in importance, significantly impacting the model's predictions. When the applied load is eccentric (not perfectly aligned with the column's centerline), it can induce bending moments in addition to axial forces. This bending effect can lead to localized stresses and influence the failure mode of the column, ultimately affecting the MLC. Subsequently, features such as the diameter of the column, the number of layers of FRP hoop wraps, the ultimate tensile strain of FRP, the elastic modulus of the FRP, the nominal thickness of an FRP hoop sheet, and column height are considered. Finally, the column height is observed to have a relatively minor impact.

The SHAP summary plot for the XGB4 model, depicted in Fig.8b, aligns with the previous observations. Model outcomes are visualized through a bee-swarm plot, where the color spectrum spans from blue to red, indicating values from low to high. In the case of stress in FRP wraps and CS of concrete, positive influences are evident on the right side of the axis, indicating a direct relationship between both features and the MLC of FRP-confined CRCC. A similar analysis reveals that the number of layers of FRP hoop wraps, the ultimate tensile strain of FRP, and the nominal thickness of an FRP hoop sheet positively affect the MLC of FRP-confined CRCC. Conversely, load eccentricity, diameter of the circular column section, elastic modulus of the FRP, and column height exhibit adverse effects. This can be explained by the enhanced LBC when the number or thickness of FRP layers increases. Conversely, a larger eccentric load diminishes the impact of the load on the column's structural strength. These observations are contingent on the database used in this study. More accurate results could be obtained with a larger dataset. The findings offer valuable insights into the nuanced interplay between input features and the model's predictions

for the MLC of FRP-confined CRCC, providing a foundation for further refinement and practical applications.

### 4.5. XGB model-based interactive GUI

The XGB model-based interactive GUI holds significant importance; it serves as an intuitive and accessible platform for users to leverage the power of the XGB algorithm in predictive modeling and analysis. This GUI facilitates interaction with the model, making it easier for users to experiment with different parameters, input data, and interpret results without advanced programming skills. Fig. 9 visually represents the primary interface of the GUI, showcasing input fields, sliders, and the submission button. Users can input specific values or use sliders to define and specify the parameters of interest. Users can customize the values of input parameters related to the MLC of FRP-confined CRCC. After adjusting the parameters, users can click the "Submit" button. The GUI displays the corresponding MLC value for FRP-confined CRCC.

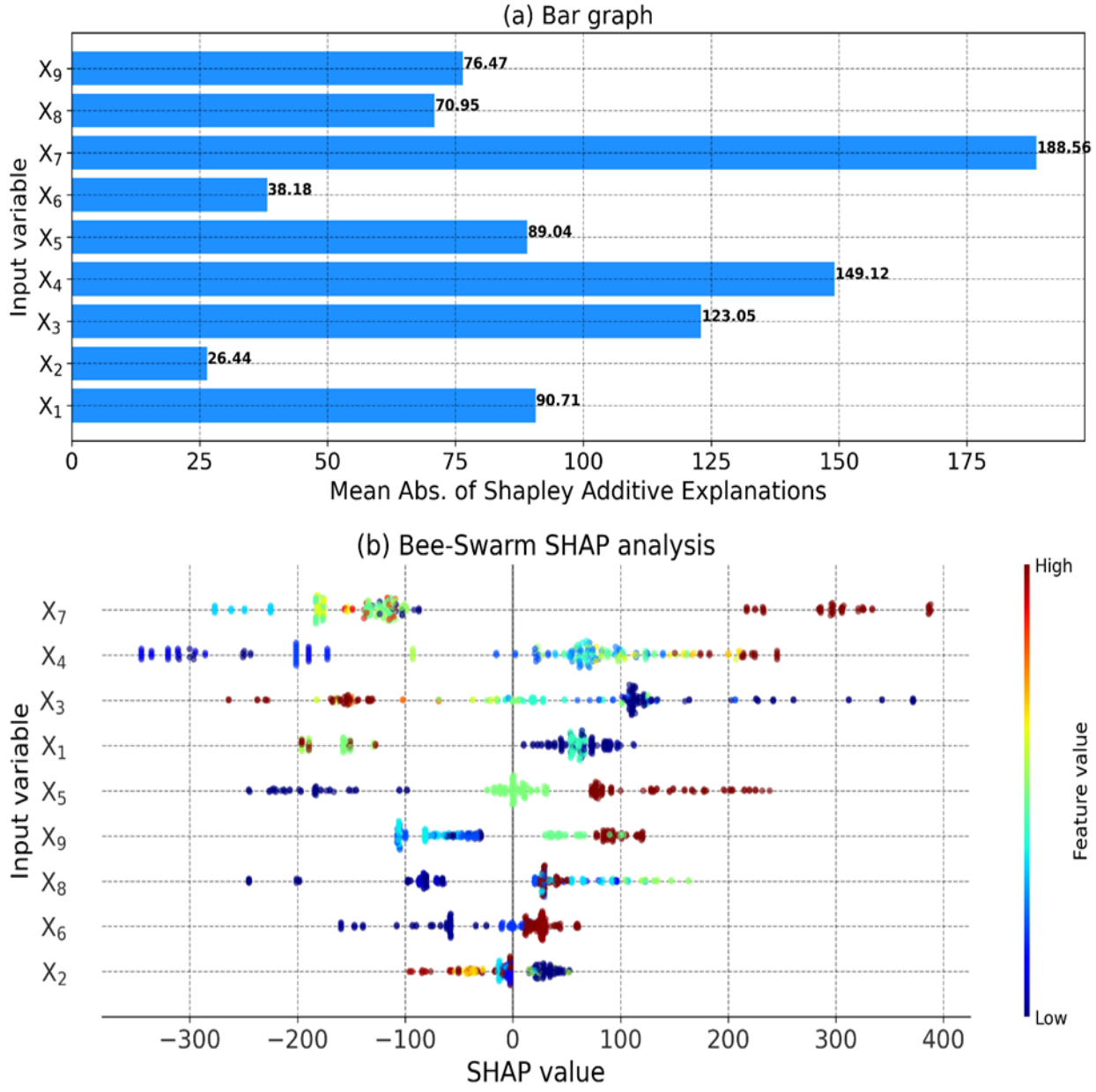


**Fig. 8.** (a) Mean SHAP values used to determine feature importance, (b) Visualization of SHAP impact values on model outputs

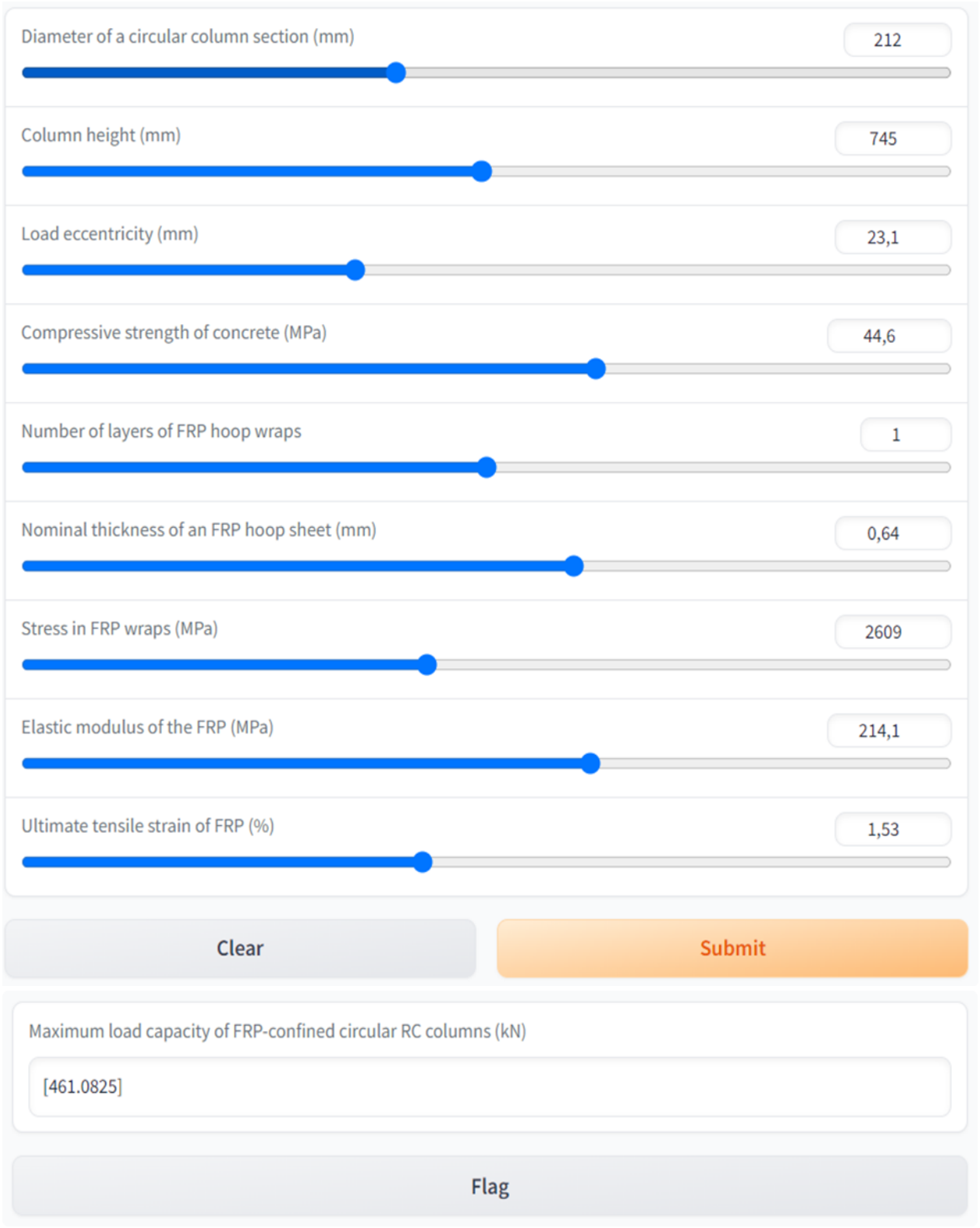


**Fig. 9.** GUI interface for users

## 5. Conclusion

In conclusion, this research explores applying ML techniques for predicting the MLC of FRP-confined CRCC. Utilizing the XGB algorithm, combined with the innovative SFO and AO for hyperparameter tuning, has proven highly effective in enhancing predictive accuracy. Incorporating MCS involving 200 runs ensures the robustness and generalizability of the optimized hyperparameters. An extensive evaluation of

model performance, employing metrics such as RMSE, MAE, and $R^2$ across training and testing datasets, provides a comprehensive understanding of the model's capabilities. Additionally, the interpretability of the XGB model through SHAP values uncovers intricate relationships and underscores the significance of various input features in predicting the MLC of FRP-confined CRCC. A notable contribution of this study is the development of an interactive GUI based on the XGB model, providing users with an intuitive platform to explore and visualize the impact of input parameters on predicted MLC values. Implemented on the Python platform, this GUI facilitates user-friendly interaction, empowering researchers and engineers to make informed decisions in designing and assessing FRP-confined CRCC.

Nevertheless, it is critical to acknowledge that while this study offers significant perspectives, it is not without constraints. The model's performance depends on the dataset's quality and comprehensiveness. Therefore, the accuracy of predictions may vary when applied to different datasets or real-world scenarios. Additionally, while the SHAP values provide a valuable model interpretation, they do not completely understand the complex interactions and non-linear relationships between features. Future research could address these limitations and improve the model's predictive accuracy and interpretability.

**References**

[1]. Y. Zhou, J. Hu, M. Li, L. Sui, F. Xing. (2016). FRP-confined recycled coarse aggregate concrete: Experimental investigation and model comparison. *Polymers*, 8(10), 375.

[2]. P. Li, Y.-F. Wu, Y. Zhou, F. Xing. (2019). Stress-strain model for FRP-confined concrete subject to arbitrary load path. *Composites Part B: Engineering*, 163, 9-25.

[3]. Y. Zhou, X. Liu, F. Xing, H. Cui, L. Sui. (2016). Axial compressive behavior of FRP-confined lightweight aggregate concrete: An experimental study and stress-strain relation model. *Construction and Building Materials*, 119, 1-15.

[4]. E. Fathalla, R. Rajapakse, B.I. Mihaylov. (2022). Modeling the shear behavior of deep beams strengthened with FRP sheets. *Engineering Structures*, 260, 114232.

[5]. Y.A. Sayed, A.A. Ibrahim, A.G. Tamrazyan, M.F. Fahmy. (2023). Machine-learning-based models versus design-oriented models for predicting the axial compressive load of FRP-confined rectangular RC columns. *Engineering Structures*, 285, 116030.

[6]. Y.T. Obaidat, W. Barham, N.A.A. Al-Khazaaleh. (2022). Modeling of confined circular RC columns using artificial neural network and finite element method. *Structures, Elsevier*, 40, 74-87.

[7]. A. Kashi, A.A. Ramezanianpour, F. Moodi. (2017). Durability evaluation of retrofitted corroded reinforced concrete columns with FRP sheets in marine environmental conditions. *Construction and Building Materials*, 151, 520-533.

[8]. H. Toutanji, M. Saafi. (2001). Durability studies on concrete columns encased in PVC–FRP composite tubes. *Composite Structures*, 54(1), 27-35.

[9]. A. Ahmed, S. Guo, Z. Zhang, C. Shi, D. Zhu. (2020). A review on durability of fiber reinforced polymer (FRP) bars reinforced seawater sea sand concrete. *Construction and Building Materials*, 256, 119484.

[10]. A.R. Rahai, P. Sadeghian, M.R. Ehsani. (2008). Experimental behavior of concrete cylinders confined with CFRP composites. *The 14th World Conference on Earthquake Engineering*, pp 12-17.

[11]. K.V. Ranolia, B.K. Thakkar, J.D. Rathod. (2013). Effect of different patterns and cracking in FRP wrapping on compressive strength of confined concrete. *Procedia Engineering*, 51, 169-175.

[12]. J.B. Mander, M.J.N. Priestley, R. Park.

(1988). Theoretical Stress-Strain Model for Confined Concrete. *Journal of Structural Engineering*, 114(8), 1804-1826.

[13] J.G. Teng, L. Lam. (2004). Behavior and Modeling of Fiber Reinforced Polymer-Confined Concrete. *Journal of Structural Engineering*, 130(11), 1713-1723.

[14] T. El Maaddawy. (2009). Strengthening of eccentrically loaded reinforced concrete columns with fiber-reinforced polymer wrapping system: Experimental investigation and analytical modeling. J*ournal of Composites for Construction*, 13(1), 13-24.

[15] Y.-F. Wu, C. Jiang. (2013). Effect of load eccentricity on the stress–strain relationship of FRP-confined concrete columns. *Composite Structures*, 98, 228-241.

[16] B. Hu, J. Wang, G. Li. (2011). Numerical simulation and strength models of FRP-wrapped reinforced concrete columns under eccentric loading. *Construction and Building Materials*, 25(5), 2751-2763.

[17] C. Jiang, Y.-F. Wu. (2020). Axial strength of eccentrically loaded FRP-confined short concrete columns. *Polymers*, 12(6), 1261.

[18] T. El Maaddawy. (2008). Behavior of corrosion-damaged RC columns wrapped with FRP under combined flexural and axial loading. *Cement and Concrete Composites*, 30(6), 524-534.

[19] Z. Gu, C. Wei, C. Wu, D. Gao, P. You, D. Wei. (2023). Analysis of eccentrically loaded FRP partially wrapped reinforced concrete columns subjected to combined environmental erosion. *Engineering Structures*, 280, 115720.

[20] B. Csuka, L.P. Kollár. (2011). FRP-confined circular columns subjected to eccentric loading. *Journal of Reinforced Plastics and Composites*, 30(14), 1167-1178.

[21] G. Lin, J.G. Teng. (2017). Three-Dimensional Finite-Element Analysis of FRP-Confined Circular Concrete Columns under Eccentric Loading. *Journal of Composites for Construction*, 21(4), 04017003.

[22] P. Chen, H. Wang, S. Cao, X. Lv. (2022). Prediction of mechanical behaviours of FRP-confined circular concrete columns using artificial neural network and support vector regression: Modelling and performance evaluation. *Materials*, 15(14), 4971.

[23] Q.H. Nguyen, H.-B. Ly, T.-A. Nguyen, V.-H. Phan, L.K. Nguyen, V.Q. Tran. (2021). Investigation of ANN architecture for predicting shear strength of fiber reinforcement bars concrete beams, *Plos One*, 16(4), e0247391.

[24] T.-A. Nguyen, H.-B. Ly, H.-V.T. Mai, V.Q. Tran. (2021). Using ANN to Estimate the Critical Buckling Load of Y Shaped Cross-Section Steel Columns. *Scientific Programming*, 1-8, 5530702.

[25] H.-B. Ly, M.H. Nguyen, B.T. Pham. (2021). Metaheuristic optimization of Levenberg–Marquardt-based artificial neural network using particle swarm optimization for prediction of foamed concrete compressive strength. *Neural Computing and Applications*, 33(24), 17331-17351.

[26] H.-V.T. Mai, T.-A. Nguyen, H.-B. Ly, V.Q. Tran. (2021). Investigation of ANN model containing one hidden layer for predicting compressive strength of concrete with blast-furnace slag and fly ash. *Advances in Materials Science and Engineering*, 5540853.

[27] S.Y. Ghanem, H. Elgazzar. (2021). Predicting the behavior of reinforced concrete columns confined by fiber reinforced polymers using data mining techniques. *SN Applied Sciences*, 3, 170.

[28] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, K. Chen, R. Mitchell, I. Cano, T. Zhou. (2015). Xgboost: extreme gradient boosting. *R Package Version*, pp 1-4.

[29] A.S. Bakouregui, H.M. Mohamed, A. Yahia, B. Benmokrane. (2021). Explainable extreme gradient boosting tree-based prediction of load-carrying capacity of FRP-RC columns. *Engineering Structures*, 245(93), 112836.

[30] H.-A. Le, D.-A. Le, T.-T. Le, H.-P. Le, T.-H. Le,

H.-G.T. Hoang, T.-A. Nguyen. (2022). An Extreme Gradient Boosting approach to estimate the shear strength of FRP reinforced concrete beams. *Structures, Elsevier*, 45, 1307-1321.

[31] M. Amjad, I. Ahmad, M. Ahmad, P. Wróblewski, P. Kamiński, U. Amjad. (2022). Prediction of pile bearing capacity using XGBoost algorithm: modeling and performance evaluation. *Applied Sciences*, 12(4), 2126.

[32] Z. Wang, D. Wang, S.T. Smith, D. Lu. (2012). Experimental testing and analytical modeling of CFRP-confined large circular RC columns subjected to cyclic axial compression. *Engineering Structures*, 40, 64-74.

[33] H.S. Al-Nimry, R.A. Al-Rabadi (2019). Axial-flexural interaction in FRP-wrapped RC columns. *International Journal of Concrete Structures and Materials,* 13(1), 13:53.

[34] H. Al-Nimry, A. Soman. (2018). On the slenderness and FRP confinement of eccentrically-loaded circular RC columns. *Engineering Structures, Elsevier,* 164, 92-108.

[35] L. Bisby, M. Ranger. (2010). Axial-flexural interaction in circular FRP-confined reinforced concrete columns. *Construction and Building Materials,* 24(9), 1672-1681.

[36] J. Fitzwilliam, L.A. Bisby. (2010). Slenderness Effects on Circular CFRP Confined Reinforced Concrete Columns. *Journal of Composites for Construction*, 14(3), 280-288.

[37] Y. Cao, Y.-F. Wu, C. Jiang. (2018). Stress-strain relationship of FRP confined concrete columns under combined axial load and bending moment, *Composites Part B: Engineering*, 134, 207-217.

[38] N.A. Siddiqui, S.H. Alsayed, Y.A. Al-Salloum, R.A. Iqbal, H. Abbas. (2014). Experimental investigation of slender circular RC columns strengthened with FRP composites. *Construction and Building Materials*, 69, 323-334.

[39] W. Wang, P.R. Martin, M.N. Sheikh, M.N.S. Hadi. (2018). Eccentrically Loaded FRP Confined Concrete with Different Wrapping Schemes. *Journal of Composites for Construction*, 22(6), 04018056.

[40] J. Li, M.N.S. Hadi. (2003). Behaviour of externally confined high-strength concrete columns under eccentric loading. *Composite Structures*, 62(2), 145-153.

[41] A. Sharma, M.K. Goyal. (2015). Bayesian network model for monthly rainfall forecast. *2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), IEEE*, pp 241-246.

[42] S. Salcedo-Sanz, R.C. Deo, L. Carro-Calvo, B. Saavedra-Moreno. (2016). Monthly prediction of air temperature in Australia and New Zealand with machine learning algorithms. *Theoretical and Applied Climatology*, 125, 13-25.

[43] S. Shadravan, H.R. Naji, V.K. Bardsiri. (2019). The Sailfish Optimizer: A novel nature-inspired metaheuristic algorithm for solving constrained engineering optimization problems, *Engineering Applications of Artificial Intelligence*, 80, 20-34.

[44] L. Abualigah, D. Yousri, M. Abd Elaziz, A.A. Ewees, M.A. Al-Qaness, A.H. Gandomi. (2021). Aquila Optimizer: A novel meta-heuristic optimization algorithm. *Computers & Industrial Engineering*, 157, 107250.

[45] J. Elith, J.R. Leathwick, T. Hastie. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4), 802-813.

[46] G.G. Moisen, E.A. Freeman, J.A. Blackard, T.S. Frescino, N.E. Zimmermann, T.C. Edwards Jr. (2006). Predicting tree species presence and basal area in Utah: a comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. *Ecological Modelling*, 199(2), 176-187.