

Case study

Data-driven axial compressive strength investigation of FRP-confined coral aggregate concrete

Chang Zhou^a, Kai-Di Peng^{b,*}, Yu-Lei Bai^c^a Department of Architecture and Civil Engineering, City University of Hong Kong, 999077, Hong Kong^b Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Kowloon, China^c State Key Laboratory of Bridge Safety and Resilience, Beijing University of Technology, Beijing 100124, China

ARTICLE INFO

Keywords:

Coral aggregate concrete

FRP

Confinement

Axial compressive strength

Machine learning

Model explanation

ABSTRACT

This study utilizes machine learning (ML) method to investigate the axial compressive strength of fiber-reinforced polymer (FRP)-confined coral aggregate concrete (CAC). A dataset comprising 115 samples is created, and eight input features are selected for developing and evaluating ML models. Besides, six empirical formulae are used to compare their performance against the ML models. The SHapley Additive exPlanation (SHAP) algorithm is employed to elucidate the prediction mechanisms of the ML models and to clarify the interactions between the eight input features and the axial compressive strength of FRP-confined CAC. A comparison of evaluation metrics indicates that the empirical model, which is developed for compressive strength of FRP-confined geopolymer-based CAC prediction, outperforms the other five empirical formulas in precision, boasting the highest R^2 value of 0.84. In comparison, with the exception of the KNN model, the remaining five data-driven ML models exhibit high precision in predicting the axial compressive strength of FRP-confined CAC, with metric R^2 values exceeding 0.93 on both the training and testing dataset. Besides, the axial compressive strength of confined CAC is primarily influenced by thickness of FRP layer and unconfined compressive strength of CAC, and the elastic modulus and ultimate strength of FRP are also critical factors. Furthermore, excessive FRP confinement will not further enhance the axial compressive strength of confined CAC, and CAC column with a larger diameter necessitates either a thicker FRP layer or a higher FRP strength to achieve desired compressive strength.

1. Introduction

In marine construction, the production of conventional concrete is typically reliant on inland resources that are geographically distant from the ocean, such as river sand, crushed stone, and fresh water. The transportation of these materials to remote marine construction sites is often associated with high costs and lengthy supply cycles, posing significant challenges to meeting construction demands in locations like islands and reefs [1]. To address these limitations, researchers have recently introduced a novel type of concrete material-coral aggregate concrete [2]. This innovative material utilizes locally available coral-based aggregates, such as coral debris and coral sand, in with combination seawater and cement. By leveraging these readily accessible materials, CAC not only reduces transportation costs but also effectively mitigates the shortage of aggregate in marine engineering projects [3]. However, the

* Corresponding author.

E-mail address: 18071319r@connect.polyu.hk (K.-D. Peng).<https://doi.org/10.1016/j.cscm.2025.e05065>

Received 11 June 2025; Received in revised form 7 July 2025; Accepted 15 July 2025

Available online 17 July 2025

2214-5095/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

naturally occurring coral gravel and coral sand possess inherent characteristics such as porosity, low density, and relatively low strength. When utilized directly in concrete formulations, these properties tend to undermine the mechanical performance and durability of the concrete [4]. A feasible method to address this shortcoming is to apply FRP as a confinement technique to bolster the structural integrity of coral aggregate concrete. FRP is renowned for its combination of light weight, high tensile strength, excellent resistance to corrosion, and superior bonding capabilities [5,6]. These features allow FRP to significantly improve the compressive strength, ductility, and crack resistance of coral aggregate concrete. Furthermore, the integration of FRP helps to reduce the brittle failure tendencies of coral aggregate concrete, thereby enhancing its overall durability in the harsh marine environments [7].

Recent advancements have spurred extensive research on the compressive strength of FRP-confined CAC [8–12]. In light of this, Zhang et al. [10] investigated the axial compressive behavior of BFRP-confined geopolymer-based CAC columns, focusing on the effects of specimen sizes and the number of BFRP layers, and found that BFRP confinement significantly enhanced the load-bearing capacity and deformability of columns, and the ultimate axial strength and strain of the specimens increased with the number of BFRP layers. Similarly, Wang et al. [11] investigated the axial compressive behavior of seawater coral aggregate concrete (SCAC) and SCAC-filled FRP tubes. They extended an existing model for FRP-confined lightweight concrete by incorporating the strength of the coral aggregates to predict the compressive strength of FRP-confined SCAC. In addition, Li et al. [12] concluded that smaller coral aggregate sizes resulted in better compressive strength for the FRP-confined CAC. Besides, the dilation behavior and ultimate FRP rupture strain were significantly influenced by the aggregate size, with smaller aggregates leading to better confinement efficiency. Based on these findings, they also proposed a modified stress-strain model to more accurately predict the behavior of FRP-confined CAC under axial compression. While the aforementioned models demonstrate high accuracy in forecasting compressive strength, their reliance on a single dataset for both model construction and verification introduces uncertainties in prediction when applied to generalized samples.

In recent decades, ML algorithms have been regarded as an effective method in material and structural engineering [13–16]. Their ability to deliver highly accurate predictions and classifications makes them a compelling solution for addressing challenges in practical application. In this regard, different ML methods are implemented to forecast the axial mechanical behavior of FRP-confined normal concrete, including Conditional Tabular Generative Adversarial Network [17], Random Forest [18], Support Vector Regression [19], Deep Learning [20], and Extreme Gradient Boosting Trees (XGBoost) [21]. All these studies demonstrated the exceptional precision of ML models. However, none of these models account for the reduction in compressive strength induced by coral aggregate. Another key challenge is that before these models can be fully implemented, their interpretability issues must be overcome.

To fill these knowledge gaps, this study compiles a dataset of 115 samples documenting the axial compressive strength of FRP-confined CAC along with their design parameters. Six existing empirical formulae, which include five formulae for FRP-confined normal concrete and one for FRP confined geopolymer-based CAC, are selected to check their applicability for FRP-confined CAC. In addition, ML algorithms are utilized to construct surrogate models for confined compressive strength, and their accuracy and generalization performance are systematically evaluated to identify the optimal model. Furthermore, SHAP method is employed to interpret the decision-making mechanism of the top-performing model and quantify the influence of key design variables. Finally, a

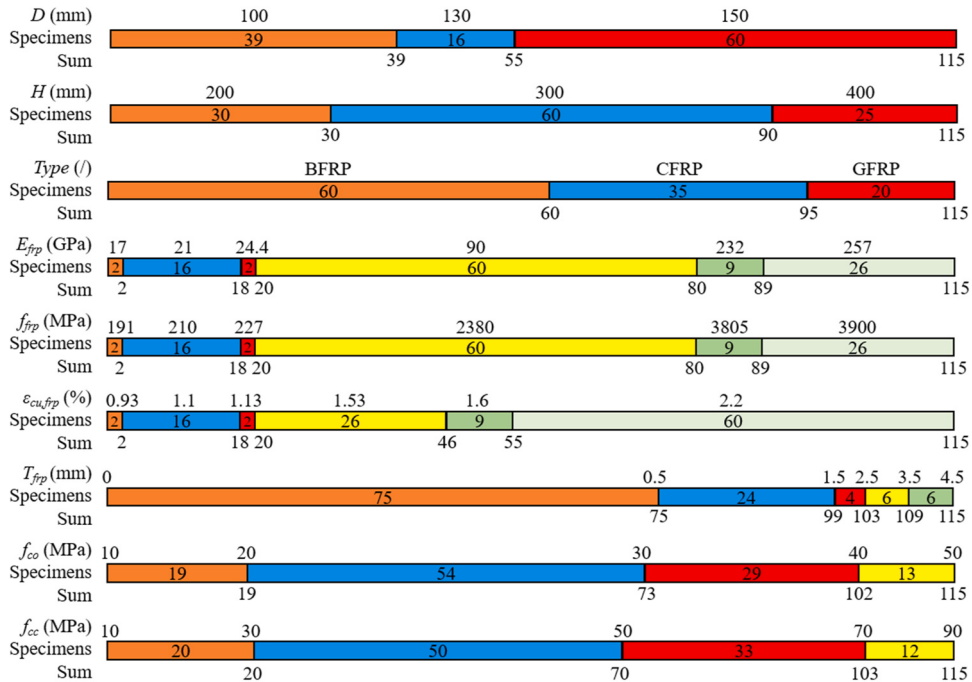


Fig. 1. Distribution of input features and axial compressive strength of FRP confined CAC.

comprehensive parameter analysis is conducted across the entire dataset to assess the positive or negative effects of each variable on compressive strength of FRP-confined CAC. The innovations of this research are as follows: 1) High-precision models are developed to predict the axial compressive strength of FRP-confined CAC; 2) The impact of each design variables on the compressive strength have been elucidated; 3) A novel parametric analysis method has been proposed. This method integrates all samples in the dataset, even if they differ from other samples in multiple parameters.

2. Dataset description

This study establishes a comprehensive database by integrating 115 samples for FRP-confined CAC specimen subjected to axial loading, as detailed in Appendix A. Five relevant publications are incorporated herein [10,12,22–24], and eight variables concerning the geometry of specimen and material strength are taken into account to examine their effects on the axial compressive strength (f_{cc}). These parameters encompass the diameter and height of coral aggregate concrete column (D and H), type of FRP material for concrete confinement (*Type*), elastic modulus of FRP material (E_{frp}), ultimate strength and strain of FRP material (f_{frp} and $\varepsilon_{cu,frp}$), diameter of FRP bars (d_b), surface treatment of FRP bars (*st*), FRP types (T_{frp}), thickness of FRP layer (T_{frp}) and compressive strength of unconfined coral aggregate concrete (f_{co}). The compressive strength of unconfined and FRP-confined coral aggregate concrete are all obtained through experimental test, and the size of specimens for these strength measurements are identical within each sample to quantify the influence of FRP confinement. Besides, all these specimens have a circular shape to eliminate the invalid constraint area appearing in square column [25]. Furthermore, to eliminate the influence of longitudinal or lateral reinforcement, this study only considers coral aggregate concrete cylinders that are free from steel bars. In addition, BFRP, CFRP, and GFRP are labeled as 1, 2, and 3, respectively, to streamline the building and assessing procedures of the ML models.

Fig. 1 and Table 1 presents the distribution diagrams and statistic properties of the eight input features and the axial compressive strength of FRP-confined CAC. The Pearson correlation coefficient matrix is plotted in Fig. 2 to reveal the relation between each input and output variables. The Pearson correlation coefficient is a powerful tool which helps to understand how closely two sets of data are related in a linear way, with an absolute value smaller than or equal to 1. A coefficient value of 1 indicates that the relationship between two variables can be perfectly represented by a linear equation, with all data points lying exactly on a straight line and increasing in tandem. In the provided correlation matrix, the color gradient from white to dark blue the strength of the correlation, with dark blue showing stronger correlations and light-colored indicating weaker correlations. Notably, two feature, i.e., elastic modulus of FRP and ultimate strength of FRP, has a strong correlation, but both of them are selected as input features to quantify their influence on the strength improvement of coral aggregate concrete.

3. Performance of existing empirical formulae

In order to quantify the contribution of FRP confinement to the compressive strength of FPR-confined CAC, several empirical formulae have been proposed. This study collects six formulae to check their prediction in terms of compressive strength forecasting [10,26–30]. In general, all these formulae are expressed as the following closed-form equation:

$$\frac{f_{cc}}{f_{co}} = \alpha + \beta \left(\frac{f_l}{f_{co}} \right)^\gamma \quad (1)$$

$$f_l = \frac{2k_e T_{frp} E_{frp} \varepsilon_{cu,frp}}{D} \quad (2)$$

where f_{cc} and f_{co} are the axial compressive strength of FRP-confined CAC and unconfined CAC, respectively; α , β and γ are coefficients of which the values are summarized in Table 2; f_l is confinement pressure of FRP; T_{frp} , E_{frp} and $\varepsilon_{cu,frp}$ are thickness, elastic modulus and ultimate tensile strain of FRP layer, respectively; k_e is effective rupture strain coefficient of which the value is 0.586, 0.580, and 0.624 for GFRP, BFRP, and CFRP, respectively [10]; D represents diameter of CAC column.

Evaluation metrics are crucial for assessing the performance of surrogate models, providing quantitative measures to evaluate how well a model performs. Four metrics are used in this context: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and Coefficient of Determination (R^2). Specifically, MAE calculates the average of the absolute

Table 1
Statistic properties of input features and axial compressive strength of FRP confined CAC.

Symbol	Description	Unit	Min	Median	Max	Mean	STD
D	Diameter of CAC column	mm	100	150	150	130.26	22.67
H	Height of CAC column	mm	200	300	400	295.65	69.02
<i>Type</i>	FRP Type	/			BFRP:1, CFRP: 2, GFRP:3		
E_{frp}	Elastic modulus of FRP	GPa	17	90	257	126.86	85.76
f_{frp}	Ultimate strength of FRP	MPa	191	2380	3900	2457.75	1222.42
$\varepsilon_{cu,frp}$	Ultimate strain of FRP	%	0.93	2.2	2.2	1.81	0.44
T_{frp}	Thickness of FRP layer	mm	0	0.26	4	0.66	1.06
f_{co}	Compressive strength of unconfined CAC	MPa	19.4	26.17	44.1	28.35	7.18
f_{cc}	Compressive strength of FRP confined CAC	MPa	17.69	45.26	88	46.94	16.65

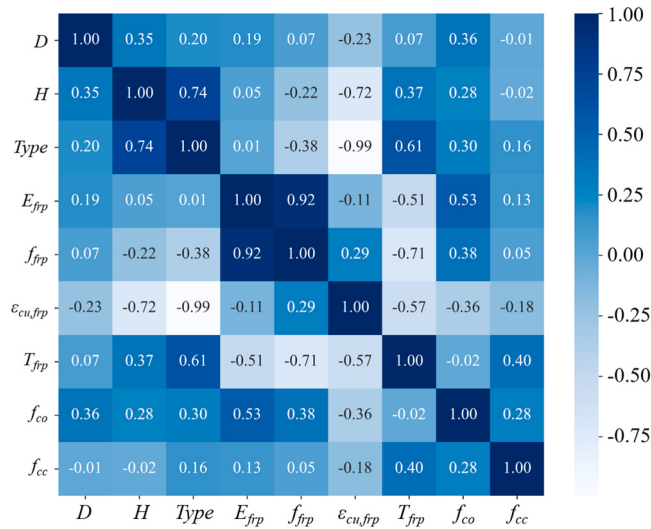


Fig. 2. Heatmap of Pearson correlation coefficient.

Table 2

Value of coefficients and prediction performances of existing empirical methods.

Reference	Coefficients			Performance evaluation matrix			
	α	β	γ	R^2	MAE	MAPE	RMSE
Ma et al. [26]	1	4.68	1	0.22	10.08	19.28 %	14.68
Wu et al. [27]	0.75	2.5	1	0.44	10.85	22.93 %	12.49
Teng et al. [28]	1	3.5	1	0.79	5.02	10.12 %	7.66
Khan et al. [29]	1	3.3	1	0.82	4.73	9.94 %	7.01
Zhou et al. [30]	1	1.4	0.65	0.54	7.60	13.41 %	11.31
Zhang et al. [10]	1	3.05	1	0.84	4.60	8.98 %	6.65

differences between predicted and actual values. RMSE is the square root of the average of squared differences between predicted and actual values. MAPE measures the average of the absolute percentage differences between predicted and actual values. R^2 indicates the proportion of variance in the target variable that can be explained by the independent variables, ranging from 0 to 1, with higher values suggesting a better fit. The computational formulae for MAE, MAPE, RMSE, and R^2 can refer to [16].

The precision of selected empirical formulae are illustrated in Fig. 3 and tabulated in Table 2. In general, formulae proposed by Teng et al. [28], Khan et al. [29] and Zhang et al. [10] achieve high prediction precision for their high value of R^2 and lower values of MAE, MAPE, and RMSE. In contrast, the performance of formulae proposed by Ma et al. [26] and Wu et al. [27] are inferior as the former overestimate the axial compressive strength while the latter achieves a conservative prediction. Among these models, the formula proposed by Zhang et al. [10] achieves highest precision, which might be due to the fact that it is proposed for FRP-confined geopolymer-based coral aggregate concrete which others for FRP-confined normal concrete.

4. Performance of machine learning algorithms

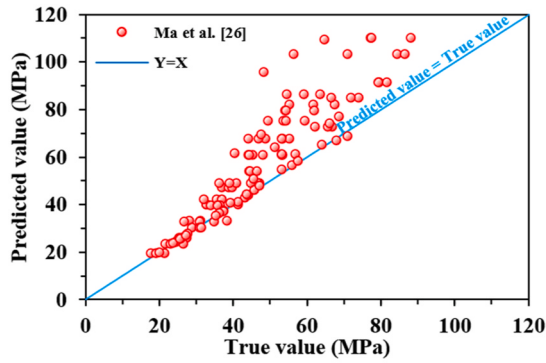
4.1. Data preprocessing

Normalization techniques are essential tools in data preprocessing, used to transform data into a consistent and comparable format. This ensures equal contribution of all features to model training regardless of their original scales [31]. Common normalization techniques include Min-Max Scaling, Log Normalization, Robust Normalization, and Batch Normalization [32]. In this study, Min-Max Scaling is used to normalize the dataset with nine input features to the range of [0, 1]. This method applies a linear transformation to the original data, and the formula for Min-Max Scaling is presented as follows:

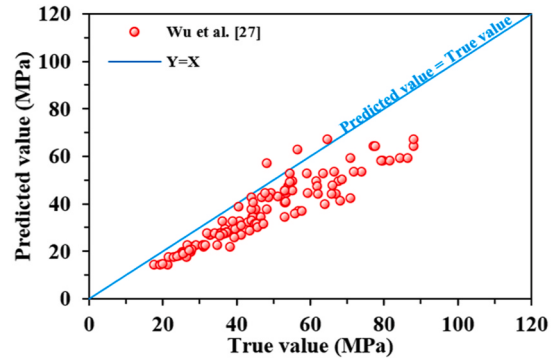
$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (3)$$

where x^* donates the normalized value of the input feature; x represents the value of the input feature; $x_{\max}$ and $x_{\min}$ are the maximum and minimum value of the input feature in the complete dataset, respectively.

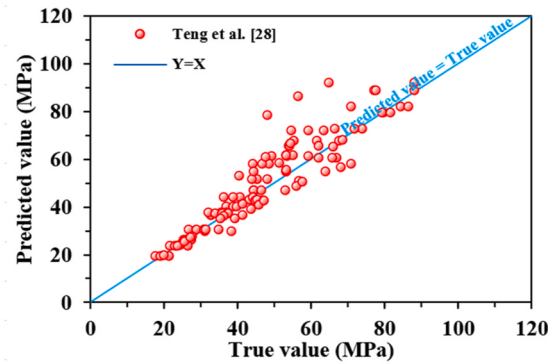
After normalization, the dataset is shuffled to prevent the model from learning order-specific patterns, ensuring the training set represents the overall data distribution [33]. The dataset is subsequently split into 80 % for training and 20 % for testing to evaluate



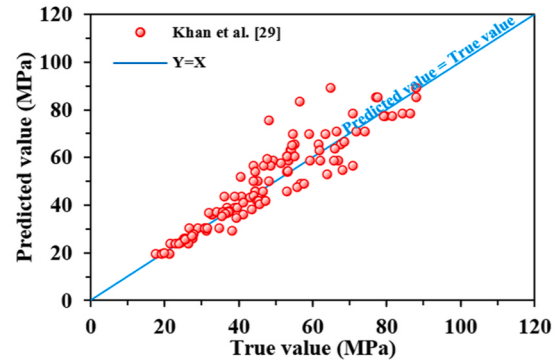
(a) Ma et al. [26]



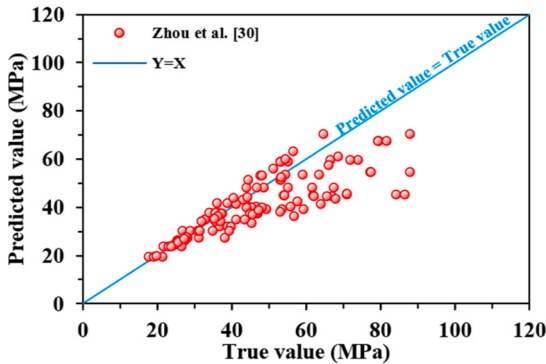
(b) Wu et al. [27]



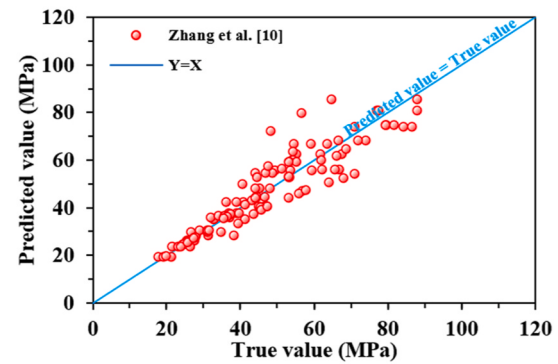
(c) Teng et al. [28]



(d) Khan et al. [29]



(e) Zhou et al. [30]



(f) Zhang et al. [10]

Fig. 3. Performance of existing empirical methods.

model performance.

4.2. Algorithms for machine learning model training

This study employs six ML algorithms for axial compressive strength prediction. Among these algorithms, Support Vector Machine (SVM) works by using kernel functions to transform data and capture complex relationships to minimizing error within a certain margin [34]. K-Nearest Neighbors (KNN) regression is a lazy learning approach that predicts values based on the nearest neighbors in the training set [35]. Decision Tree (DT) creates a tree-like model by recursively partitioning the data based on input features [36]. Random Forest (RF) regression is a supervised learning algorithm that integrates multiple decision trees to improve precision and robustness through bootstrapping and feature bagging [37]. Gradient Boosting Decision Trees (GBDT) builds an ensemble of trees sequentially to correct errors made by previous trees [38]. Extreme Gradient Boosting Trees (XGBoost) further enhances gradient

boosting by optimizing a differentiable loss function and incorporating regularization techniques to prevent overfitting [39].

4.3. Hyperparameter optimization

Hyperparameters, which are essential settings that dictate algorithm's behavior, must be specified before training begins. Several hyperparameter tuning technique, including Grid-Search, Random Search and Bayesian Optimization, are widely utilized to identify the best hyperparameter combination. Grid Search is a brute-force approach which defines a grid of hyperparameters and evaluates the performance of models for every combination of these hyperparameters [40]. This method is exhaustive and ensures to find the best combination within the specified grid. Random Search randomly samples hyperparameters from a predefined distribution. Unlike Grid Search, it does not evaluate all possible combinations, making it more efficient [41]. Bayesian Optimization uses a probabilistic model to predict which hyperparameters are likely to yield better performance. It iteratively selects the most promising hyperparameters based on the predictions outputted by model and updates the model with the new results [42].

To enhance the precision of ML models, the Grid-Search Cross-Validation algorithm [43] is employed herein. In this process, the dataset is split into ten subsets. In each iteration, 9 folds serve as the training set, while the remaining 1 fold is utilized as the validation set to evaluate performance. Following 10 iterations, the overall model performance is assessed by calculating the average performance across all folds, providing a reliable estimate. The algorithm records the performance metrics for each hyperparameter configuration. Ultimately, the combination that achieves the highest average prediction accuracy is selected as the optimal configuration, and the details are summarized in Table 3.

4.4. Precision of machine learning models

Fig. 4 and Fig. 5 presents the performance of six ML models on the training and testing dataset, revealing distinct patterns in their predictive capabilities. These models exhibited superior prediction performance on the training subset, with R^2 values exceeding 0.95, indicating strong correlations between predicted and actual values. In addition, all models met the acceptable criterion for ML models, where the MAE was smaller than the RMSE [44]. The MAPE values for all models were below 6 %, further highlighting their excellent performance on the training data. Among the models, KNN model stood out as the most reliable for it achieved the lowest RMSE, MAPE, and MAE values, while also attaining an R^2 value of 0.98, indicating high prediction accuracy on the training dataset. In contrast, SVM, Random Forest, GBDT and XGBoost models showed slightly lower precision but still acceptable performance, with R^2 values of 0.97 or 0.98. However, the Decision Tree model exhibited the highest RMSE, MAE, and MAPE values among all models, suggesting lower efficiency in capturing complex patterns within the training data.

Table 3
Hyperparameter values for each ML algorithms.

Algorithms	Hyperparameters	Explanations	Search Space	Optimized value
SVM	Kernel	Type of kernel function	{linear, rbf}	rbf
	γ	RBF kernel coefficient	{0.0001, 0.001, 0.01, 0.1, 1, 10}	1
	C	Regularization parameter	{1, 10, 100, 1000, 10000}	10000
KNN	K	Number of nearest neighbors	{1, 2, 3, 4, 5, 6, 7, 8, 9, 10}	4
	P	power parameter in the distance metric	{1, 2, 3, 4, 5, 6, 7, 8, 9, 10}	3
DT	weight	Neighbors closer in distance have greater influence	{distance, uniform}	distance
	d_{max}	Maximum depth of the tree	{4, 6, 8, 10}	8
	n_{max}	Maximum number of leaf nodes	{10, 15, 20, 25, 30}	25
	n_f	Number of features to consider for the best split	{1, 2, 3, 4, 5, 6, 7}	6
RF	criterion	The function to measure the quality of a split	{poisson, absolute_error, squared_error, friedman_mse}	absolute_error
	n_e	Number of estimators	{20, 40, 60, 80, 100}	40
	n_f	Number of features to consider for the best split	{2, 4, 6, 8}	8
	d_{max}	Maximum depth of the tree	{4, 6, 8, 10}	10
	n_{max}	Maximum number of leaf nodes	{10, 15, 20, 25, 30}	30
GBDT	criterion	The function to measure the quality of a split	{poisson, absolute_error, squared_error, friedman_mse}	poisson
	n_e	Number of estimators	{20, 40, 60, 80, 100}	40
	l_r	Learning rate	{0.01, 0.1, 1.0}	0.1
	n_f	Number of features to consider for the best split	{2, 4, 6, 8}	4
	d_{max}	Maximum depth of each tree	{4, 6, 8, 10}	10
XGBoost	n_{max}	Maximum number of leaf nodes	{10, 15, 20, 25, 30}	10
	n_e	Number of estimators	{20, 40, 60, 80, 100}	100
	$l_r=1.0$	Learning rate	{0.01, 0.1, 0.3, 0.5, 1.0}	0.5
	$\gamma_0=1$	minimum reduction in the loss function	{0.001, 0.01, 0.1, 1.0}	0.01
	$d_{max}=4$	Maximum depth of each tree	{2, 4, 6, 8}	2
	$\alpha=0$	Coefficient of Lasso regularization	{0, 0.5, 1.0}	0.5
	$\lambda=1$	Coefficient of Ridge regularization	{0, 0.5, 1.0}	1.0
	Objective	Objective function that defines specific goal that the model seeks to achieve	{reg:squarederror}	reg:squarederror

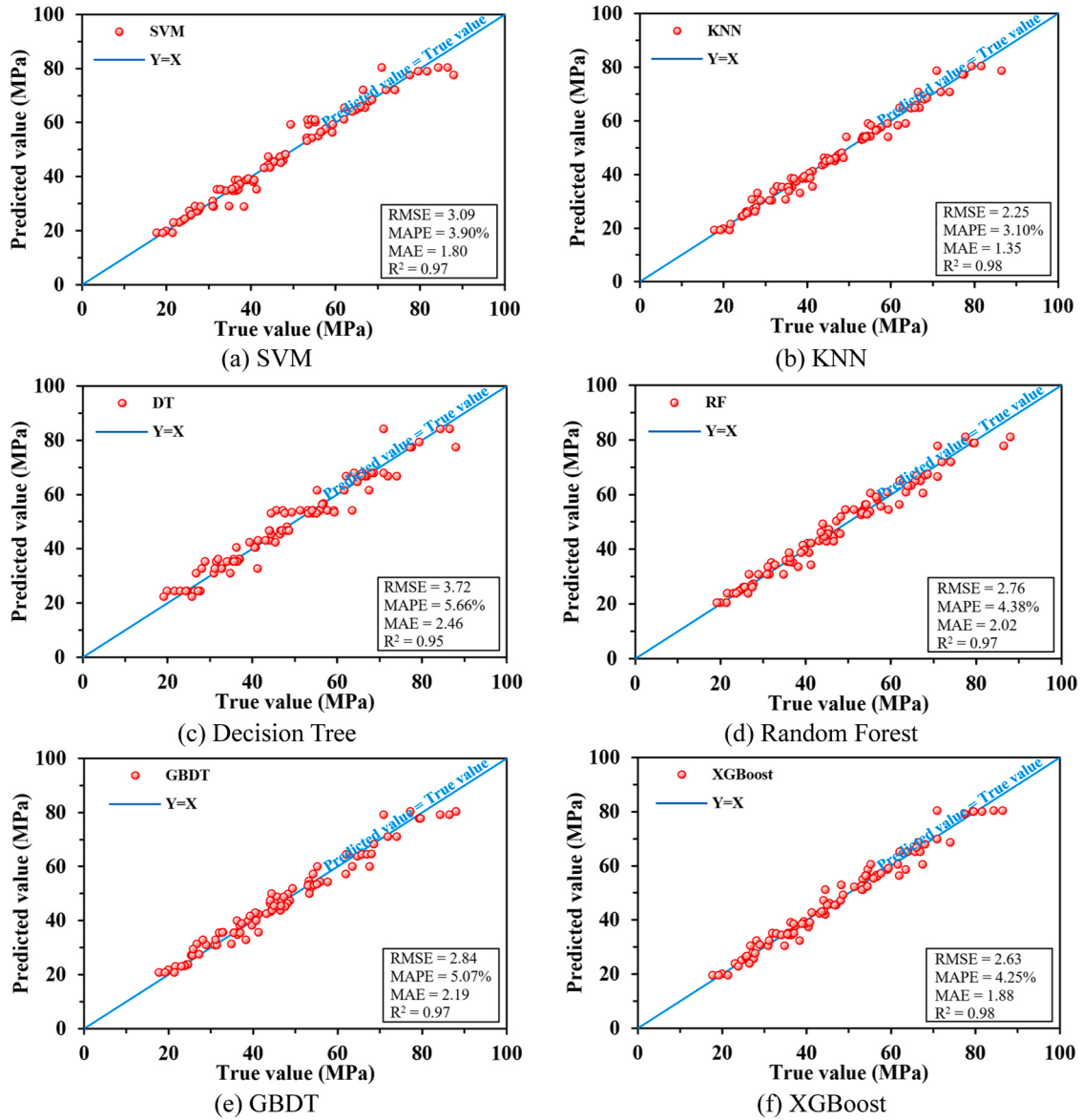


Fig. 4. Performance of ML models on training dataset.

On the testing dataset, the performance of ML models shifted slightly. All models maintained high accuracy, with R^2 values achieving or exceeding 0.88, indicating their ability to generalize well to unseen data. However, although the KNN model showed the highest precision on the training dataset, it exhibited the poorest performance on the testing dataset, with the lowest R^2 value and an MAPE value higher than 10 %. This discrepancy between training and testing performance suggests a potential overfitting issue. Meanwhile, Decision Tree, Random Forest and GBDT models demonstrated exceptional generalization and robustness, with R^2 values surpassing 0.95 on both training and testing datasets, highlighting their ability to maintain high precision across different data sets. The SVM and XGBoost models also performed well on the testing dataset, despite having a slightly higher MAPE and lower R^2 value compared to Decision Tree, Random Forest and GBDT models. Overall, based on the prediction precision and generalization ability, Random Forest is emerged as the most precise model across both training and testing datasets.

5. Explanation of ML models based on SHAP analysis

Although ML models often achieve high precision, they are frequently viewed as black boxes because it is challenging to know how they generate their predictions. To tackle this problem, several effective methods, including Local Interpretable Model-agnostic Explanations (LIME) [45], Counterfactual Explanations [46], and SHAP [47], are proposed. Among these algorithms, LIME operates under the assumption that while global interpretability of complex models might be challenging, local interpretability is more feasible.

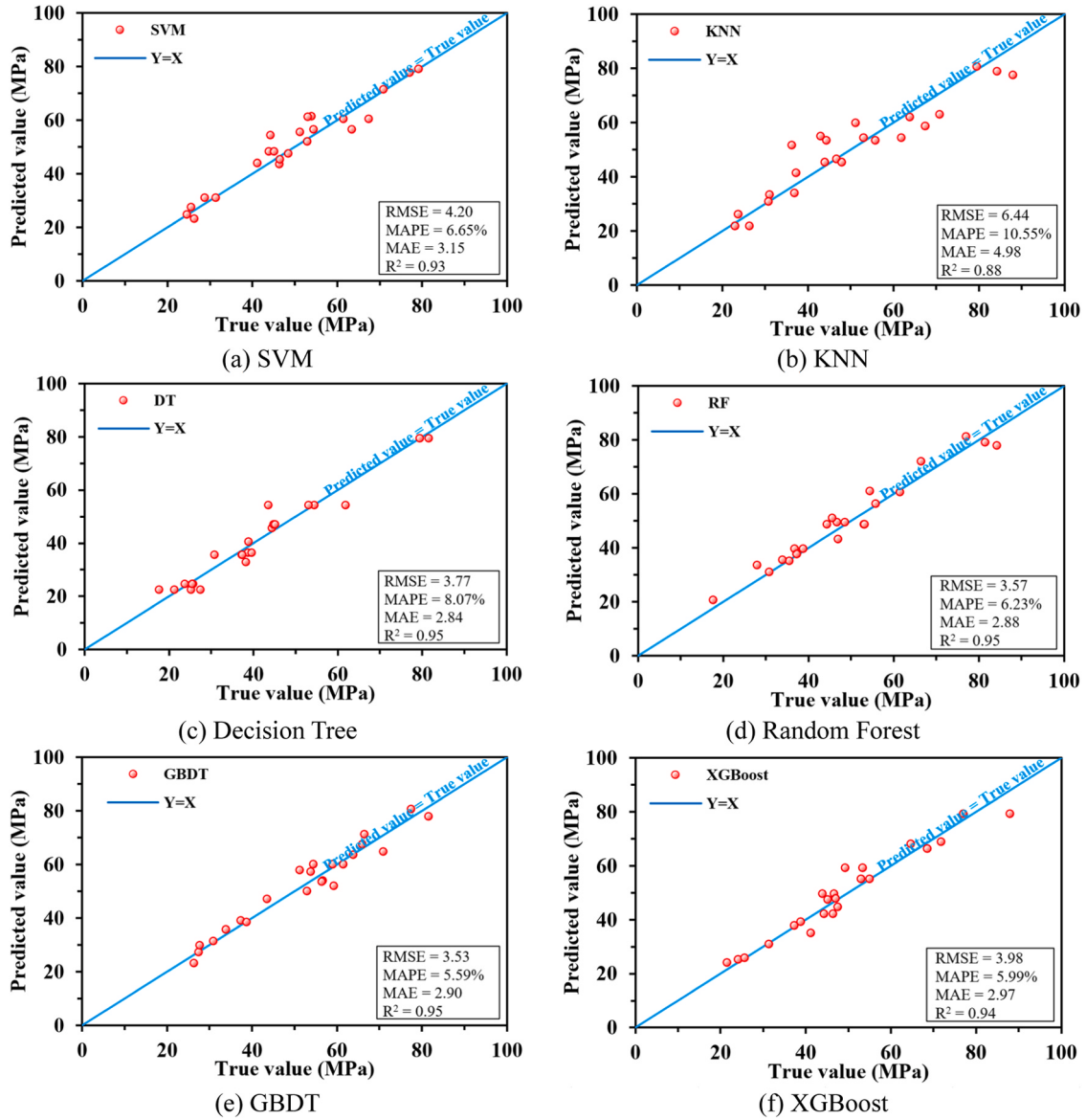


Fig. 5. Performance of ML models on testing dataset.

It identifies the most important features for a particular prediction by generating perturbations in the vicinity of the instance of interest [45]. Counterfactual Explanations offer hypothetical scenarios by clarifying how minor alterations in the input features can result in a different prediction [46]. Meanwhile, SHAP ensures that the attribution of predictions to input features is fair and consistent. In this study, the SHAP algorithm is selected because it can determine which features are most crucial for making a prediction, the extent to which each feature contributes to the prediction, and whether the contribution is positive or negative [47].

5.1. Prediction process of machine learning models

Table 4 identifies a selected sample from the database, and Fig. 6 details the prediction process of this sample by four models:

Table 4
Input features of the investigated sample.

D (mm)	H (mm)	Type (/)	E_{frp} (GPa)	f_{frp} (MPa)	$\varepsilon_{cu,frp}$ (%)	T_{frp} (mm)	f_{co} (MPa)	f_{cc} (MPa)
100	200	BFRP	90	2380	2.2	0.52	19.4	54.07

Decision Tree, Random Forest, GBDT, and XGBoost. Within this figure, red columns denote each feature's positive contribution to the predicted axial compressive strength, whereas blue columns indicate negative contributions. The length of each column reflects the absolute magnitude of the corresponding feature's Shapley value. For example, in Fig. 6(a), in view of the thickness of FRP layer T_{frp} exhibits a positive correlation with the axial compressive strength of FRP-confined CAC, a large value of T_{frp} (i.e., 0.52 mm) exerts a positive influence on the prediction outcomes. Additionally, the Decision Tree model identifies the small compressive strength of unconfined CAC (i.e., 19.4 MPa) as a negative variable in enhancing axial compressive strength. Moreover, the small height (i.e., 200 mm) succeed in avoid the buckling of column, which contribute positively to the axial compressive strength. In addition, the tensile strength of FRP material (i.e., 2380 MPa) is viewed as a disadvantage because it lowers the predictive outcomes. Based on the four subplots in Fig. 6, the analysis reveals that for the studied sample, the Random Forest, GBDT, and XGBoost algorithms prioritize the FRP layer thickness as the most influential variable. This is evident from the longer column lengths associated with this parameter compared to other input features. Besides, Decision Tree model also consider the thickness of FRP layer is also be recognized as the important features. These models integrate the base value along with both positive and negative contributions to produce their final predictions. The Decision Tree, Random Forest, GBDT, and XGBoost models yield compressive strength estimates of 53.53 MPa, 54.49 MPa, 51.88 MPa, and 59.20 MPa, respectively. In comparison, the actual values of axial compressive strength of FRP-confined CAC measured using three identical specimens are 49.32 MPa, 53.53 MPa and 59.36 MPa, respectively, leading to an average value of 54.07 MPa. These difference between the average value of measured results and predictions outputted by Decision Tree, Random Forest, GBDT and XGBoost models results in relative errors of -1.00% , -0.78% , -4.05% , and 9.49% , respectively.

5.2. Impact of each feature on axial compressive strength of FRP-confined CAC

Fig. 7 presents the Shapley values of features for individual samples across four models: Decision Tree, Random Forest, GBDT, and XGBoost. Within these figures, each sample is denoted by a dot—red dots correspond to samples with large feature values, while blue dots represent those with small ones. The features are ordered by their average absolute contribution to axial compressive strength across all samples, with those positioned higher in the ranking exerting a greater impact than their lower-ranked counterparts. Fig. 7 clarifies that all these selected models consider the thickness of FRP layer (T_{frp}) as the most influential factor, which might attribute to the fact that the FRP layer provides confinement to CAC and enhances its ability to withstand compression. A thicker FRP layer can provide more effective confinement stress, leading to higher compressive strength. Furthermore, all models acknowledge that the compressive strength of unconfined CAC (f_{co}), height of specimen (H) and elastic modulus of FRP material (E_{frp}) are influential variables. Specially, the Decision Tree, Random Forest and GBDT models believe that the ultimate strength of FRP (f_{frp}) critically affect the strength. Furthermore, the XGBoost model indicates that f_{frp} is less critical. One the other hand, all these models show the type and

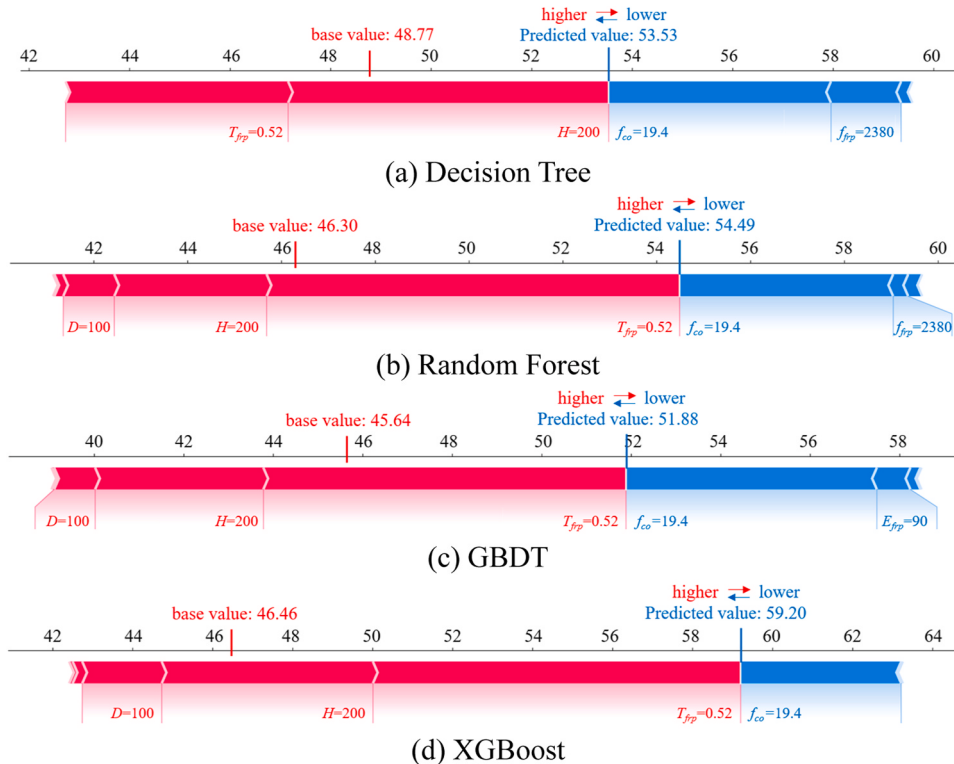


Fig. 6. Prediction process of various ML model.

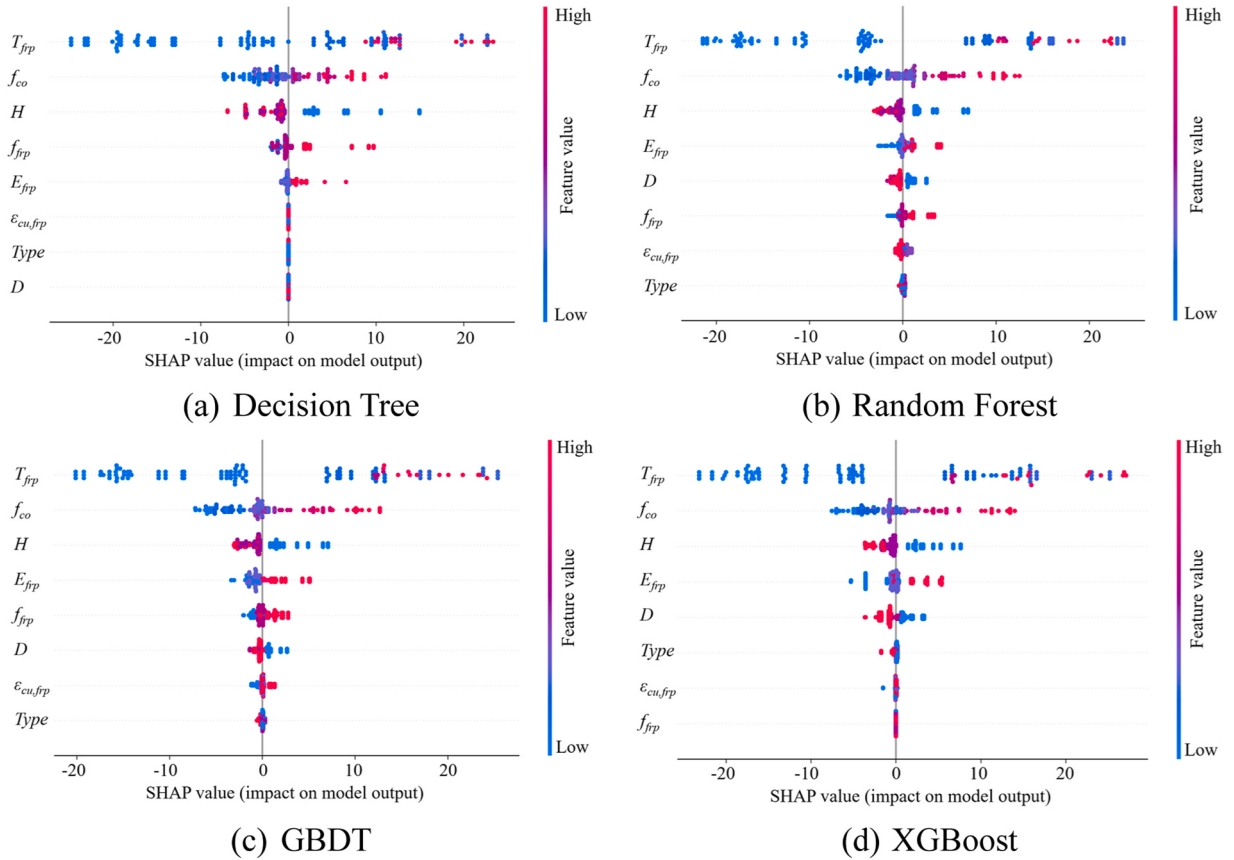


Fig. 7. Summary plot of ML models on entire dataset.

ultimate strain of FRP (*Type* and $\varepsilon_{cu,frp}$) and has a smaller influence as the primary mechanism through which FRP jackets enhance the compressive strength of CAC involves lateral confinement, which is more directly influenced by the thickness and elastic modulus of the FRP layer rather than its material type or ultimate strain.

5.3. Parametric analysis

Given its high precision on both the training and testing datasets, Random Forest model is selected to perform parametric analysis based on dependence plot of SHAP. In Fig. 8, the dot with blue or red color indicates the corresponding sample has a small or large value of the input feature, and its contribution to the axial compressive strength of FRP-confined CAC is presented as the value of vertical coordinate. As shown in Fig. 8(g), the contribution of thickness of FRP layer (T_{frp}) increase rapidly with its increasement within the range from 0 to 1. Subsequently, the positive contribution made by T_{frp} fluctuates within a small range, indicating that over confinement by FRP will not further enhance the axial compressive strength of confined concrete. A proper explanation for this phenomenon is that when the thickness of the FRP layer is within the range from 0 to 1 mm, the FRP is able to effectively transfer the compressive stress to the concrete core. The FRP acts like a confining jacket that restricts the lateral expansion of the CAC under axial compression. As the thickness increases in this initial range, the confining pressure exerted by the CAC on the concrete increases significantly. When the FRP layer becomes too thick, other failure modes may start to dominate. For example, the FRP jacket itself may experience buckling under excessive confinement pressure. In addition, CAC may experience different failure modes such as splitting due to excessive lateral pressure, which can also limit the further enhancement of the axial compressive strength. Apart from this, Fig. 8 (h) shows that the unconfined compressive strength exhibits a significant positive correlation with the strength of FRP-confined CAC, and this correlation can be approximated by a linear function. Besides, Fig. 8 also indicates that the elastic modulus and ultimate strength of FRP material also positively correlated with the compressive strength of confined CAC, while the diameter and height of column shows a negation correlation. The reason for this finding might attribute to the fact that FRP with a higher elastic modulus and ultimate strength is able to provide a higher lateral confinement on CAC. However, as indicated by Eq. (2), the confining pressure f_l is inversely proportional to the diameter of column D . Consequently, to sustain the same confining pressure, a larger diameter necessitates either a thicker FRP layer or a higher FRP strength. Besides, as the height of the column increases, the axial compressive strength decreases due to the increased slenderness ratio and the potential for buckling. Finally, compared to ultimate strength of FRP, the remaining two factors, including FRP type and ultimate strain of FRP show slighter influence on the compressive strength, as plotted in

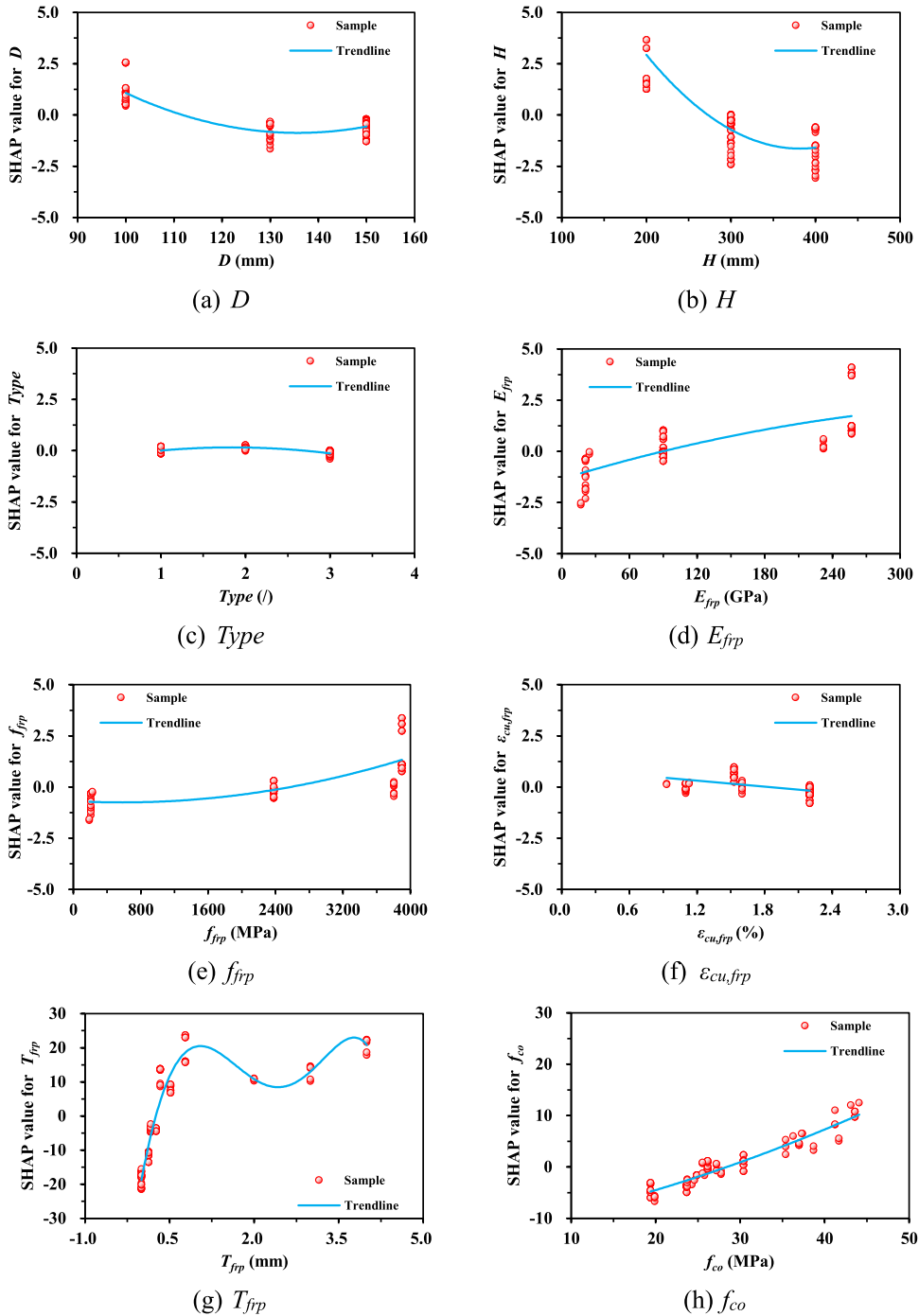


Fig. 8. SHAP value distribution of different input variables.

Fig. 8(c) and Fig. 8(f).

6. Conclusion

This research utilized ML algorithms to predict the axial compressive strength of FRP-confined CAC. A dataset comprising 115 samples and eight input features were established for training and testing the ML models. The accuracy of six empirical formulas was compared with that of the ML models using four evaluation metrics: MAE, MAPE, RMSE, and R^2 . The prediction process of the ML models was elucidated, and a parametric analysis was conducted using the SHAP algorithm to enhance the interpretability and

reliability of the models. Based on the aforementioned analysis, the following conclusions were drawn:

- 1) Three out of six empirical formulae achieve relative high prediction precision for their value of R^2 exceeding 0.79 and lower values of MAE, MAPE, and RMSE. The empirical model, which is proposed for compressive strength of FRP-confined geopolymer-based coral aggregate concrete prediction, achieves superior precision among six empirical formulae, with the highest R^2 value of 0.84 and the smallest values for MAE (4.60), MAPE (8.98 %) and RMSE (6.65).
- 2) With the exception of the KNN model, the remaining five data-driven ML models exhibit high precision in predicting the axial compressive strength of FRP-confined coral aggregate concrete, with metric R^2 values exceeding 0.93 on both the training and testing dataset. Besides, the MAPE values of these five models are all smaller than 9 % on the testing dataset. Specially, compared with other ML models, the Random Forest model is regarded as the most precise model across the entire dataset.
- 3) The axial compressive strength of confined CAC is primarily influenced by thickness of FRP layer and unconfined compressive strength of CAC. In addition, the elastic modulus and ultimate strength of FRP are also critical factors. Besides, compared to ultimate strength of FRP, type and ultimate strain of FRP show slighter influence on the compressive strength of confined CAC.
- 4) Although thickness of FRP layer has a significant positive correlation with the compressive strength, excessive confinement by FRP will not further enhance the axial compressive strength of confined CAC. Besides, to sustain the same confining pressure provide by FRP, CAC column with a larger diameter necessitates either a thicker FRP layer or a higher FRP strength.

CRediT authorship contribution statement

Chang Zhou: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Kai-Di Peng:** Writing – review & editing, Validation, Project administration. **Yu-Lei Bai:** Writing – review & editing, Validation, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are grateful for the financial support received from the Non-Metallic Excellence and Innovation Center for Building Materials (No. 24SFP-2).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.cscm.2025.e05065](https://doi.org/10.1016/j.cscm.2025.e05065).

Data availability

Data will be made available on request.

References

- [1] H. Li, Y. Liu, B. Liang, F. Liu, G. Wu, J. Du, H. Hou, A. Li, L. Shi, Demands and challenges for construction of marine infrastructures in China, *Front. Struct. Civ. Eng.* 16 (5) (2022) 551–563.
- [2] Y. Cao, J. Bao, P. Zhang, Y. Sun, Y. Cui, A state-of-the-art review on the durability of seawater coral aggregate concrete exposed to marine environment, *J. Build. Eng.* 60 (2022) 105199.
- [3] S. He, C. Jiao, Y. Niu, S. Li, Utilizing of coral/sea sand as aggregates in environment-friendly marine mortar: physical properties, carbonation resistance and microstructure, *Case Stud. Constr. Mater.* 16 (2022) e00981.
- [4] Q. Qin, Q. Meng, H. Yang, W. Wu, Study of the anti-abrasion performance and mechanism of coral reef sand concrete, *Constr. Build. Mater.* 291 (2021) 123263.
- [5] J. Zhu, K. Weng, W. Liu, B. Huang, K. Peng, J. Zhu, J. Dai, Thin-layer ultra-high-strength engineered cementitious composites (UHS-ECC) reinforced with small-diameter FRP bars for structural strengthening, *ThinWalled Struct.* 205 (2024) 112592.
- [6] J. Zhu, K. Weng, B. Huang, L. Xu, J. Dai, Ultra-high-strength engineered cementitious composites (UHS-ECC) panel reinforced with FRP bar/grid: Development and flexural performance, *Eng. Struct.* 302 (2024) 117193.
- [7] W. Zhou, P. Feng, J.-Q. Yang, Advances in coral aggregate concrete and its combination with FRP: a state-of-the-art review, *Adv. Struct. Eng.* 24 (6) (2021) 1161–1181.
- [8] B. Zhang, H. Zhu, Z. Dong, Z. Yang, Mechanical properties and durability of FRP-reinforced coral aggregate concrete structures: a critical review, *Mater. Today Commun.* 35 (2023) 105656.
- [9] T. Zhang, D. Niu, C. Rong, GFRP-confined coral aggregate concrete cylinders: the experimental and theoretical analysis, *Constr. Build. Mater.* 218 (2019) 206–213.
- [10] B. Zhang, Q. Wang, H. Zhu, Z. Yang, H. Peng, Behavior of BFRP-confined geopolymer-based coral aggregate concrete columns under axial compression: effects of specimen sizes, *J. Build. Eng.* 98 (2024) 111106.
- [11] J. Wang, P. Feng, T. Hao, Q. Yue, Axial compressive behavior of seawater coral aggregate concrete-filled FRP tubes, *Constr. Build. Mater.* 147 (2017) 272–285.

- [12] P. Li, D. Huang, R. Li, R. Li, F. Yuan, Effect of aggregate size on the axial compressive behavior of FRP-confined coral aggregate concrete, *Polymers* 14 (18) (2022) 3877.
- [13] Y. Tang, W. Wang, Q. Huang, Y. Cheng, C. Zhou, Shear performance prediction of RC beams shear-strengthened with FRP sheet: a machine learning driven design-oriented method, *Eng. Struct.* 334 (2025) 120240.
- [14] C. Zhou, Y. Xie, W. Wang, Y. Zheng, Machine learning driven post-impact damage state prediction for performance-based crashworthiness design of bridge piers, *Eng. Struct.* 292 (2023) 116539.
- [15] W. Zhang, W. Wang, Y. Tang, K. Sun, C. Zhou, Explainable machine learning driven strength degradation investigation of BFRP bar in seawater and sea sand concrete environment, *Structures* 71 (2025) 108205.
- [16] C. Zhou, W. Wang, Y. Zheng, Data-driven shear capacity analysis of headed stud in steel-UHPC composite structures, *Eng. Struct.* 321 (2024) 118946.
- [17] S. Zeng, X. Wang, L. Hua, M. Altayeb, Z. Wu, F. Niu, Prediction of compressive strength of FRP-confined concrete using machine learning: a novel synthetic data driven framework, *J. Build. Eng.* 94 (2024) 109918.
- [18] T. Hu, H. Zhang, C. Cheng, H. Li, J. Zhou, Explainable machine learning: compressive strength prediction of FRP-confined concrete column, *Mater. Today Commun.* 39 (2024) 108883.
- [19] B. Keshtegar, A. Gholampour, D.-K. Thai, O. Taylan, N.-T. Trung, Hybrid regression and machine learning model for predicting ultimate condition of FRP-confined concrete, *Compos. Struct.* 262 (2021) 113644.
- [20] A. Benzaamia, M. Ghrici, R. Rebouh, K. Pilakoutas, P.G. Asteris, Predicting the compressive strength of CFRP-confined concrete using deep learning, *Eng. Struct.* 319 (2024) 118801.
- [21] H. Tao, Z.H. Ali, F. Mukhtar, A.W. Al Zand, H.A. Marhoon, L. Goliatt, Z.M. Yaseen, Coupled extreme gradient boosting algorithm with artificial intelligence models for predicting compressive strength of fiber reinforced polymer-confined concrete, *Eng. Appl. Artif. Intell.* 134 (2024) 108674.
- [22] Z. Deng, W. Chai, B. Liu, Y. Hu, Compression behavior of CFRP-confined coral aggregate concrete (CCAC) circular stub columns, *Case Stud. Constr. Mater.* 16 (2022) e00863.
- [23] J. Sun, Z. Ding, Y. Wei, X. Li, Z. Wang, Compressive behavior of GFRP tube filled with structural polypropylene fiber reinforced seawater coral aggregates concrete, *Constr. Build. Mater.* 312 (2021) 125372.
- [24] Z. Dong, G. Wu, X.-L. Zhao, H. Zhu, Y. Wei, Z. Yan, Mechanical properties of discrete BFRP needles reinforced seawater sea-sand concrete-filled GFRP tubular stub columns, *Constr. Build. Mater.* 244 (2020) 118330.
- [25] Z.-F. Zhu, W.-W. Wang, Y.-X. Hui, S.-W. Hu, G.-Y. Men, J. Tian, H. Huang, Mechanical behavior of concrete columns confined with CFRP grid-reinforced engineered cementitious composites, *J. Compos. Constr.* 26 (1) (2022) 04021060.
- [26] G. Ma, X. Chen, L. Yan, H.-J. Hwang, Monotonic and cyclic axial compressive properties and modeling of basalt FRP-retrofitted predamaged short columns, *J. Compos. Constr.* 24 (4) (2020) 04020023.
- [27] G. Wu, Z. Lü, Z. Wu, Strength and ductility of concrete cylinders confined with FRP composites, *Constr. Build. Mater.* 20 (3) (2006) 134–148.
- [28] J. Teng, Y. Huang, L. Lam, L. Ye, Theoretical model for fiber-reinforced polymer-confined concrete, *J. Compos. Constr.* 11 (2) (2007) 201–210.
- [29] Q.S. Khan, M.N. Sheikh, M.N. Hadi, Experimental results of circular FRP tube confined concrete (CFFT) and comparison with analytical models, *J. Build. Eng.* 29 (2020) 101157.
- [30] Y. Zhou, X. Liu, F. Xing, H. Cui, L. Sui, Axial compressive behavior of FRP-confined lightweight aggregate concrete: An experimental study and stress-strain relation model, *Constr. Build. Mater.* 119 (2016) 1–15.
- [31] D. Singh, B. Singh, Investigating the impact of data normalization on classification performance, *Appl. Soft Comput.* 97 (2020) 105524.
- [32] A. Jain, K. Nandakumar, A. Ross, Score normalization in multimodal biometric systems, *Pattern Recognit.* 38 (12) (2005) 2270–2285.
- [33] C. Zhou, X. Tan, Y. Zheng, Y. Wang, S. Mahjoubi, Data-driven axial bearing capacity analysis of steel tubes infilled with rubberized alkali-activated concrete, *Adv. Struct. Eng.* 28 (1) (2025) 145–158.
- [34] C. Cortes, V. Vapnik, Support-vector networks, *Mach. Learn.* 20 (1995) 273–297.
- [35] T. Cover, P. Hart, Nearest neighbor pattern classification, *IEEE Trans. Inf. Theory* 13 (1) (1967) 21–27.
- [36] C. Zhou, Y. Zheng, Data-driven compressive strength investigation and design suggestions for rubberized concrete, *Mater. Today Commun.* 46 (2025) 112477.
- [37] L. Breiman, Random forests, *Mach. Learn.* 45 (2001) 5–32.
- [38] J.H. Friedman, Stochastic gradient boosting, *Comput. Stat. data Anal.* 38 (4) (2002) 367–378.
- [39] T. Chen, C. Guestrin, Xgboost: a scalable tree boosting system, *Proc. 22nd acm sigkdd Int. Conf. Knowl. Discov. data Min.* (2016) 785–794.
- [40] C. Zhou, H.-Y. Guo, J.-G. Dai, Data-driven investigation and multi-objective optimization design of alkali-activated concrete, *J. Asian Concr. Fed.* 11 (1) (2025) 53–70.
- [41] J. Bergstra, Y. Bengio, Random search for hyper-parameter optimization, *J. Mach. Learn. Res.* 13 (1) (2012) 281–305.
- [42] Z. Ghahramani, Probabilistic machine learning and artificial intelligence, *Nature* 521 (7553) (2015) 452–459.
- [43] Y. Sun, S. Ding, Z. Zhang, W. Jia, An improved grid search algorithm to optimize SVR for prediction, *Soft Comput.* 25 (2021) 5633–5644.
- [44] M. Khan, C. McNally, A holistic review on the contribution of civil engineers for driving sustainable concrete construction in the built environment, *Dev. Built Environ.* 16 (2023) 100273.
- [45] M.T. Ribeiro, S. Singh, C. Guestrin, Why should i trust you?" Explaining the predictions of any classifier, *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. data Min.* (2016) 1135–1144.
- [46] R. Guidotti, Counterfactual explanations and how to find them: literature review and benchmarking, *Data Min. Knowl. Discov.* 38 (5) (2024) 2770–2824.
- [47] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, *Advances in neural information processing systems*, Long Beach, CA, USA, 2017.