Explainable Machine Learning Model for Load-Carrying Capacity Prediction of FRP-Confined Corroded RC Columns

Huihui Li[1,*]; Haoran Li[2]; Shuwen Deng[3,*]; and Qian Chen[4]

Abstract: This paper proposes a novel explainable machine learning (ML) model to predict the axial load-carrying capacity of FRP-confined corroded RC columns utilizing the eXtreme Gradient Boosting (XGBoost) algorithm and Shapley Additive exPlanations (SHAP) technique. The XGBoost predictive model is developed based on a thorough database of experimental tests for 285 FRP-confined corroded RC columns collected from the existing studies (231 specimens) and those performed by the authors (54 specimens). Twenty parameters are considered as the critical input variables for developing the predictive model. SHAP technique is employed for evaluating the importance and interpreting the XGBoost model's prediction performance. In addition, the feasibility and effectiveness of the developed XGBoost predictive model are assessed using several empirical design models and other ensemble ML algorithms. The results indicate that, (1) the suggested XGBoost model is feasible to predict the axial load-carrying capacity of FRP-confined corroded RC columns; (2) the SHAP technique provides good explainability and interpretability to the XGBoost predictive model; (3) the input variables can be comprehensively studied concerning the feature importance through the SHAP technique, and the most important ones affecting the axial load-carrying capacity of FRP-confined corroded RC columns are the gross sectional area of column, FRP thickness, elastic modulus of FRP, eccentricity ratio, corrosion rate, and concrete compressive strength; (4) the prediction effectiveness and feasibility of the developed XGBoost model outperform that of the existing empirical models and other ML algorithms, and the mean values of R^2, RMSE, MAE, and MAPE of the XGBoost model are 0.978, 122 kN, 703.6 kN, and 7.7%, respectively; and (5) the XGBoost model can offer the alternative approach to determine the axial load-carrying capacity of FRP-confined corroded RC columns for design practice, in addition to current mechanics-based design models.

Author keywords: FRP-confined corroded RC columns; deteriorating effect; machine learning; XGBoost algorithm; SHAP technique; empirical models

Introduction

Due to their superior structural resistance, RC structures are widely applied for the protective design of civil infrastructures nowadays.[1,2] However, they are prone to numerous deterioration mechanisms because of environmental effects and climate change, such as erosion, carbonation, freeze-thaw cycles, fatigue, and chloride-induced corrosion (CIC).[2–8] Among these deterioration effects, CIC could lead to significant corrosion of steel bars and has been recognized as one of the primary causes impairing the mechanical properties and durability of aging RC structures. Numerous studies have focused on the deterioration impacts of CIC on the degraded structural response and load-carrying capacity of aging RC structures.[2,9–11] RC columns are critical structural members of many highway bridges and buildings, and the tragic damage caused by CIC could trigger progressive collapse.[1,2,12,13] In addition, the structural redundancy of RC columns is generally weaker than their beams and slab counterparts.[2,12] Thus, it is significant to improve the deteriorated structural resistance and structural performance of corroded RC columns, reduce tremendous social and economic losses, and, more importantly, mitigate human casualties.[1,2,13,14] This also necessitates investigations regarding how to improve the deteriorated resistance and residual strength of corroded RC columns, which is one of the primary research focuses of this study.

Fiber-reinforced polymer (FRP) has been widely employed to strengthen and retrofit corroded RC structures because of its inherited advantages of high strength, lightweight, superior corrosion resistance, simple on-site construction, and lower maintenance expense.[15–19] The

*Corresponding Authors: Huihui Li. Email: huihui.li@polyu.edu.hk; Shuwen Deng. Email: dengsw@hunau.edu.cn
[1]Department of Civil & Environmental Engineering, Hong Kong Polytechnic University, Hong Kong, China
[2]College of Civil Engineering, Chongqing University, Chongqing, China
[3]College of Water Resources & Civil Engineering, Hunan Agricultural University, Changsha, China
[4]Thornton Tomasetti Inc., New York 10031, NY, USA

advantages of FRP in strengthening the corroded RC structures mainly depend on the following aspects.[20,21] Firstly, FRP-wrapped structures could apply the confining pressure to offset the expansive forces generated by corrosion products. Secondly, FRP composites could act as the physical diffusion barrier to prevent the ingress of chloride ions and oxygen into RC structures, delaying corrosion of steel bars and thus protecting them from CIC.[20,21] Thus, strengthening or rehabilitation of corroded RC columns by wrapping FRP composites has been extensively investigated both experimentally[16,17,22,23] and theoretically.[24,25] Moreover, significant efforts were dedicated to investigating the structural response and mechanical properties of FRP-strengthened RC columns, i.e., stress–strain behavior,[26–28] seismic performance,[11,29–33] and axial and eccentric compression behavior.[16,34–37] These studies indicate that additional confinement provided by the wrapped FRP composites could significantly enhance the structural resistance of corroded RC columns.

However, numerous studies mainly focused on the mechanical performance of uncorroded RC columns. In contrast, limited ones were performed to examine the strength prediction of corroded RC columns confined by FRP composites (FRP-confined corroded RC columns). Zhou et al.[11] experimentally studied the seismic behavior of several corroded RC columns strengthened by FRP. They found that corrosion of steel bars could significantly deteriorate the strength and ductility of columns. Bae and Belarbi[16] experimentally studied the corrosion of steel bars on the bearing capacity of CFRP-wrapped corroded RC columns. They found that CFRP wrapping was helpful in decreasing the steel corrosion rate and reducing the degradation of stiffness and bearing ability of columns. Dai et al.[33] studied the deformation ability of FRP-retrofitted corroded RC columns and suggested an improved prediction model for the evaluation of the yield rotation of columns. In addition, Chotickai et al.[34] experimentally examined the influence of corrosion damage and volumetric CFRP ratio on the eccentric compressive behavior of CFRP-strengthened corroded RC columns. They suggested that the effectiveness of CFRP jacketing in enhancing the ultimate compressive strength of the corrosion-damaged columns depended on the volumetric CFRP ratio, and CFRP jacketing with a higher volumetric CFRP ratio could achieve a more effective confinement contribution and restore a more effective cross-sectional area of the cracked concrete. Li et al.[37] studied the effects of corrosion-induced damages under different corrosion rates of steel rebar on the structural behavior of several LRS-FRP-confined corroded RC columns. They observed that steel rebar corrosion could accelerate the steel rebar's bucking and concrete deterioration, thus reducing the ultimate compressive strength of columns. Also, compared with the unconfined corroded ones, the load-carrying capacity (P_{max}), ductility, and energy-absorbing capacity of LRS-FRP-confined corroded RC columns were much superior, indicating the effective confinement provided by LRS-FRP composites.

Based on the above-mentioned literature review, most of the previous studies mainly considered corrosion of

steel bars through the degradation of the steel rebar's cross-sectional area. However, in practical situations, the deterioration effects of CIC are more complicated. In addition to the degradation of the rebar's area, the deterioration effects of CIC should be non-uniform. Non-uniform CIC could lead to many other secondary effects, such as (i) degradation of the yield and ultimate strengths of rebar and (ii) degradation of the compressive strengths of cover and confined concrete.[2,4,8,10] The accumulation of the corrosion products could also lead to the cover concrete being cracked and spalled off, which would further impair the bond strength between the steel rebar and concrete.[11,38,39] Besides, apart from its deterioration impacts on the degradation of material properties, non-uniform CIC could also result in the degradation of stiffness, ductility, and P_{max} of corroded RC columns, particularly those under compression.[34,36,40–42] Moreover, corrosion of steel bars induced by CIC could affect the strain distributions of FRP composites and, thus, further impairing the confinement efficiency for FRP-confined RC columns.[29,31,32,39] Thus, because of the combined action of FRP confinement and corrosion-induced damages, it is difficult to predict the P_{max} of FRP-confined corroded RC columns accurately. Although several existing available empirical models suggested by some scholars[43–45] could be employed to predict P_{max} of FRP-confined RC columns, the feasibility should be further validated. Additionally, since these empirical models were developed based on predefined formulas, and a limited number of test results, there should exist significant discrepancies in predicting the P_{max} of columns.[18,46] Therefore, it is necessary to develop an accurate model for predicting the P_{max} of FRP-confined corroded RC columns for the safe design and retrofitting purposes.

Recently, with the rapid development of computing technology, data-driven and machine learning (ML) algorithms have emerged as robust and powerful techniques to address many complicated civil engineering problems.[18,19,46–56] Compared to conventional empirical models, the featured merits of ML algorithms are primarily attributed to their capability to assess the relationship between the input critical variables and output parameters without the requirements of the prior setting of assumptions and the predefined mathematical or physical models.[46] Hence, many scholars have applied ML algorithms as one of the primary alternative techniques to determine the compressive strength,[47–49,57] stress–strain model,[52,54] and load-carrying capacity or failure modes of RC members with superior accuracy.[18,46,51,53,56]

Owing to the superior computation efficiency and strong capability in modeling datasets, eXtreme Gradient Boosting (XGBoost) is known as one of the most advanced ML algorithms.[18] Thus, the XGBoost algorithm has been widely applied in civil engineering.[18,19,46,58–62] For instance, to predict the P_{max} of FRP-RC columns, Bakouregui et al.[18] developed the XGBoost model based on 283 experimental results for FRP-RC columns, and the effectiveness and feasibility of the model were evaluated through several code-based design models and empirical equations. They suggested that the XGBoost predictive model outperformed the numerical equations and code-based design models.

Liu et al.[19] developed an XGBoost model to predict the life-cycle mechanical performance of the pultruded FRP composites. They suggested that the XGBoost model could provide a good prediction interpretability, and its prediction results agreed well with the test data. Similarly, to develop the predictive model in determining the flexural capacity of FRCM-strengthened RC beams, Wakjira et al.[46] assessed the prediction performance of the XGBoost algorithm and the other six ML models. They suggested that the XGBoost model outperformed other ML algorithms and exhibited optimal accuracy. Likewise, based on a comprehensive experimental database, Ma et al.[62] proposed a novel XGBoost algorithm for predicting the P_{max} of CFRP-confined CFST columns with superior efficiency and accuracy. Thus, the aforementioned studies have confirmed that the XGBoost model has high computational efficiency, and a well-trained XGBoost predictive model could achieve reasonable prediction results with excellent accuracy. Therefore, this paper proposes to employ the XGBoost algorithm to predict the P_{max} of FRP-confined corroded RC columns.

However, the XGBoost algorithm also has several inevitable limitations. For example, similar to other typical ML models, the XGBoost algorithm is considered as "black boxes" owing to it is usually impossible to explain the involved mechanisms.[18,46] Thus, the explainability of ML models should be an imperative step to support a desirable prediction. In this regard, to achieve the interpretable and explainable XGBoost model, the Shapley Additive exPlanations (SHAP) technique[63] could be utilized. However, to date, very limited research has focused on the interpretability and explainability of ML algorithms using the SHAP technique.[18,19,46,55,64,65]

Therefore, this study aims to propose a novel, explainable predictive model to achieve an alternative and robust prediction of P_{max} for FRP-confined corroded RC columns. Firstly, the XGBoost predictive model is constructed based on the thorough test results of 285 FRP-confined corroded RC columns, including 231 experimental tests gathered from the existing studies reported in the literature and 54 from the authors. Then, through the correlation analysis, twenty parameters are selected as the critical input variables to construct the XGBoost model. Subsequently, the SHAP framework is applied to assess the feature significance of the input variables and interpret the XGBoost model. In addition, the capability and prediction performance of the model are compared and validated through several empirical design models reported in the literature and some widely used ML algorithms, such as the decision tree (DT), random forest (RF), and gradient boosting decision tree (GBDT). Finally, some major conclusions and possible future investigations are summarized.

Methodology

XGBoost algorithm

Fig. 1 shows the schematic information of the XGBoost algorithm. As seen from Fig. 1, the XGBoost framework mainly consists of several root nodes, a number of internal nodes, branches, and leaf nodes. Besides, the XGBoost algorithm is known as an advanced implementation, and it employs an additive strategy, which can be mathematically represented by Eq. (1) below.[18,19]

$$\hat{y}_i = \sum_{m=1}^{M} f_m(X_i) \tag{1}$$

where $\hat{y}_i$ is the predicted response with respect to the input X_i; M is the total number of classifications and regression trees (CARTs) (i.e., $m = 1, 2, \cdots, M$); and $f_m(X_i)$ is the predicted response of each CART. After the prediction results are attained, the objective function (L) is required to assess the performance and accuracy of the results. During the development of XGBoost model, L can be expressed by,[19]

$$L = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{2}$$
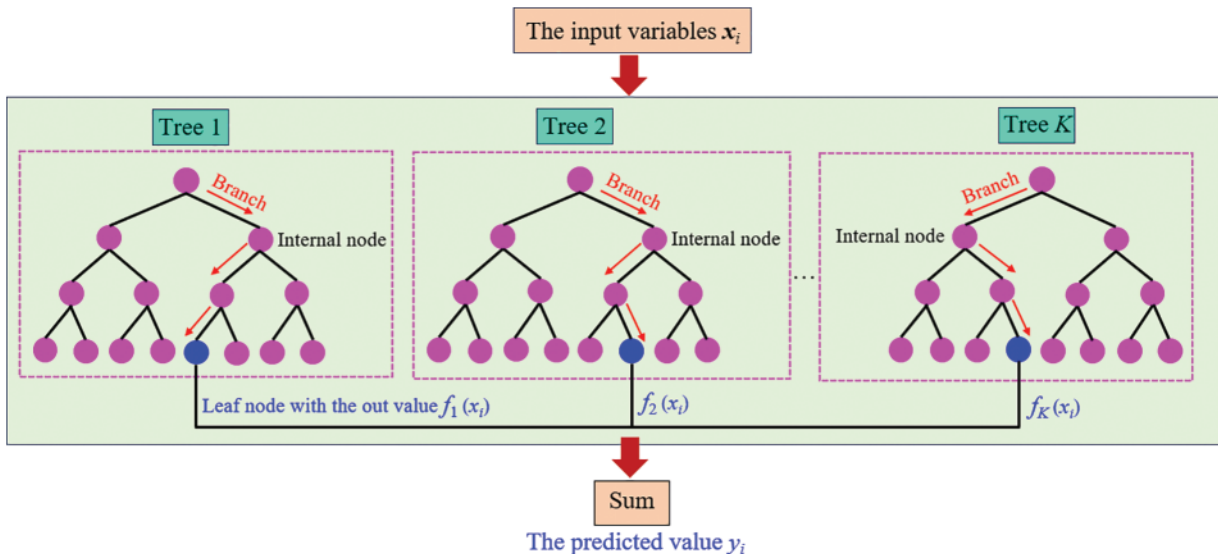

Figure 1. Schematic information of the XGBoost decision tree model

As given in Eq. (2), n is the total number of datasets (i.e., $i = 1, 2, \cdots, n$), K is the total number of trees (as illustrated in Fig. 1) (i.e., $k = 1, 2, \cdots, K$), and L contains two different parts, including (i) loss function $l\left(y_i, \hat{y}_i\right)$ and (ii) regularization item Ω, which can be represented by,

$$\Omega\left(f\right) = \gamma T + \frac{1}{2}\lambda \cdot \sum_{j=1}^{T} \omega_j^2 \qquad (3)$$

where T is the number of leaf nodes of a CART (i.e., $j = 1, 2, \cdots, T$); ω_j is the predicted value of the jth leaf node; γ and λ are the hyperparameters of the model. To minimize L and attain the optimized predictions, the XGBoost model training is generally required. Such a training process is an optimization problem, which should be performed in a step-by-step manner. During each step, a new CART is developed based on the existing CARTs, so L can be further minimized. Thus, the objective function of the tth step can be determined by,

$$L^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t)}\right) + \sum_{i=1}^{t} \Omega\left(f_i\right) = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t\left(x_i\right)\right)$$
$$+ \sum_{i=1}^{t-1} \Omega\left(f_i\right) + \Omega\left(f_t\right) \qquad (4)$$

During the tth step, the existing $(t - 1)$ CART is usually known and it can be considered as a constant. Thus, the objective function $L^{(t)}$ can be further simplified as,

$$L^{(t)} = \sum_{i=1}^{n} l\left[y_i, \hat{y}_i^{(t-1)} + f_t\left(x_i\right)\right] + \Omega\left(f_t\right) + c \qquad (5)$$

In addition, the second-order Taylor approximation can be employed to optimize the $L^{(t)}$, so Eq. (5) can be further transformed into Eq. (6).

$$L^{(t)} = \sum_{i=1}^{n} l\left[y_i, \hat{y}_i^{(t-1)} + g_i f_t\left(x_i\right) + \frac{1}{2}h_i f_t^2\left(x_i\right)\right]$$
$$+ \Omega\left(f_t\right) + c \qquad (6)$$

In which,

$$g_i = \frac{\partial l\left[y_i, \hat{y}_i^{(t-1)}\right]}{\partial \hat{y}_i^{(t-1)}} \qquad (7)$$

$$h_i = \frac{\partial^2 l\left[y_i, \hat{y}_i^{(t-1)}\right]}{\partial \left[\hat{y}_i^{(t-1)}\right]^2} \qquad (8)$$

Moreover, for the loss function $l\left(\cdot\right)$, the only requirement is that it should permit the second-order derivative.[19] Additionally, because the input variables X_i should be projected to the leaf nodes of the CARTs, $f_k\left(X_i\right)$ can be represented by,

$$f_k\left(X_i\right) = \omega_{q(X_i)}, \omega \in R^T, q: R^d \rightarrow \{1, 2, \cdots, T\} \qquad (9)$$

where $q(X_i)$ is a map function; ω is the leaf node value; d is the attribute number of the input X_i; and R^T and R^d are the T-dimensional and d-dimensional vectors, respectively. Submitting Eqs. (3), (7)–(9) into Eq. (6), $L^{(t)}$ can be determined by,

$$L^{(t)} \approx \sum_{i=1}^{n} \left[g_i \omega_{q(X_i)} + \frac{1}{2}h_i \omega_{q(X_i)}^2\right] + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 + c$$
$$= \sum_{j=1}^{T} \left[\left(\sum_{i \in I_j} g_i\right) \cdot \omega_j + \frac{1}{2} \cdot \left(\sum_{i \in I_j} (h_i + \lambda) \omega_j^2\right)\right]$$
$$+ \gamma T + c \qquad (10)$$

Letting $G_j = \sum\limits_{i \in I_j} g_i$ and $H_j = \sum\limits_{i \in I_j} h_i$, Eq. (10) can be further simplified as,

$$L^{(t)} = \sum_{j=1}^{T} \left[G_j \omega_j + \frac{1}{2} \cdot \left(H_j + \lambda\right) \omega_j^2\right] + \gamma T + c \qquad (11)$$

To obtain $L_{\min}$, the first derivative of Eq. (11) can be acquired, and hence $L_{\min}$ can be determined by using Eq. (12).

$$L_{\min} = \frac{1}{2} \cdot \sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} + \gamma T + c \qquad (12)$$

Additionally, $L_{\min}$ can be achieved when ω_j is represented by,

$$\omega_j = -\frac{G_j}{H_j + \lambda} \qquad (13)$$

Explaining the XGBoost model using the SHAP technique

Owing to the difficulty within the interpretation and explanation of the involved mechanisms of ML models, they are usually considered as "blacked boxes." Both the interpretability and explainability of the models are important in understanding the complicated nonlinear relationships between the input and output variables of ML algorithms.[18,51] In which, interpretability is usually defined as the ability to explain or to provide meaning in understandable terms to a human.[18] Besides, explainability is associated with the notion of explanation as an interface between humans and a decision-maker, that is, at the same time, both an accurate proxy of the decision-maker and comprehensible to humans.[18] Explanations supporting the output of an ML model are crucial, especially in civil engineering. The Shapley Additive exPlanations (SHAP) technique proposed by Lundberg and Lee[63] is one of the explainable artificial intelligence (XAI) tools that can be used to explain these complex models. The SHAP technique is a unified approach to explain the output of any ML model. The SHAP technique aims to provide local explainability by building surrogate models based on the ML models. The SHAP technique has a fast implementation for tree-based models, and it is very popular in interpreting ML models.[18] Thus, in this study, the SHAP technique is employed to interpret and explain the developed XGBoost predictive model.

The SHAP algorithm calculates the contribution of each input variable to the prediction for each observation. This contribution is calculated by using the input variables and the prediction. SHAP values are based on conditional expectation and Shapley game theory, whose aims are to investigate how each feature affects the prediction. The Shapley game theory aims at distributing the total gain or payoff among players, depending on the relative importance of their contributions to the final outcome of a game.[18] In order to generate an interpretable and explainable predictive model, the SHAP technique employs an additive feature attribution, e.g., an output model is defined as a linear addition of the input variables. Assuming a model with input variables $x = (x_1, x_2, \ldots, x_n)$, where n is the number of input variables, the explanation model $g(x')$ with simplified input x' for an original model $f(x)$ can be expressed as[18,51,56]:

$$f(x) = g(x') = \varphi_0 + \sum_{j=1}^{N} \varphi_j x'_j \qquad (14)$$

where N is the number of all input features; φ_0 is a constant when all input variables are missing; and φ_j is the contribution of the jth feature to the model output, which is the core computed SHAP value. The input variables x and x' are correlated through a mapping function, $x = h_x(x')$. Generally, Eq. (14) can be illustrated by Fig. 2, in which φ_0, φ_1, φ_2, and φ_3 increase the predicted value of $g(x')$, while φ_4 decreases this value. According to Lundberg and Lee,[63] a unique solution should exist for Eq. (14), which has three desirable features, i.e., (i) local accuracy, (ii) missingness, and (iii) consistency.[51,56] In specific, local accuracy ensures that the output of the function is the sum of the feature attributions and requires the model to match the output of $f(\cdot)$ for the simplified input x'. The local accuracy happens when $x = h_x(x')$. Missingness ensures that no importance is assigned to missing features. As $x_i' = 0$ implies $\phi_i = 0$ (i.e., ϕ_i is the Shapely value), missingness is satisfied. Through consistency, changing a larger impact feature will not decrease the attribution assigned to that feature. For a setting $z'\backslash i$ when $z_i' = 0$, $f_x'(z') - f_x'(z'\backslash i) \geq f_x(z') - f_x(z'\backslash i)$ implies $\phi_i(f', x) \geq \phi_i(f, x)$. Thus, the only possible model that satisfies these properties can be determined by,[51]

$$\phi_j(f, x) = \sum_{z' \subseteq x'} \frac{|z'|! \cdot (N - |z'| - 1)!}{N!} \cdot [f_x(z') - f_x(z'\backslash j)] \qquad (15)$$
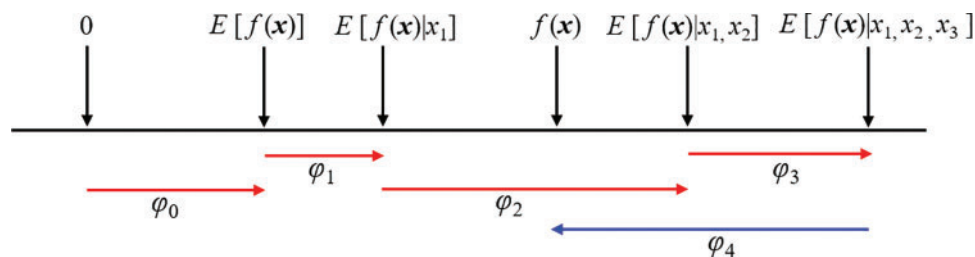
where $|z'|$ is the number of non-zero entries in z'; Lundberg and Lee[63] suggested a solution to Eq. (15) where $f_x(z') = f(h_x(z')) = E[f(z)|z_S]$; and S is the set of non-zero indices z', which is known as SHAP values.

Based on the aforementioned introductions, the SHAP technique can provide good explanations for local and global models. SHAP values can be approximated by various methods, such as Kernel SHAP, Deep SHAP, and Tree SHAP.[18] Among these methods, Tree SHAP, a version of SHAP for tree-based ML models (e.g., decision trees, random forest (RF), and gradient-boosted trees (i.e., XGBoost and Cat-Boost)), is used in this study. Tree SHAP considers tree-based models alongside an input dataset X of size $N \times M$ and produces an $N \times M$ matrix with the SHAP values. The SHAP interaction values guarantee consistency in explaining the effects of interaction on individual predictions. The two unique advantages of SHAP values are its global and local interpretability. Contrary to the existing important features in ML models, the SHAP technique can identify whether the contribution of each input feature is positive or negative. Also, each observation can get its SHAP value. Thus, the SHAP can help interpret the model globally as well as locally. A more detailed description and application of the SHAP technique in civil engineering practice could be referred to several previous researches.[18,51,56]

Determination of the XGBoost Predictive Model

Experimental database

To establish the XGBoost predictive model, a comprehensive database of experimental tests for 285 FRP-confined corroded RC columns was collected from 16 previous studies (231 specimens) and those conducted by the authors (54 specimens), as summarized in Tables 1 and 2, respectively. As per the collected columns, 202 were circular and 83 were square or rectangular specimens. Additionally, the collected specimens consisted of 225 and 60 columns under concentric and eccentric compression, respectively. As summarized in Table 3, the experimental database included 20 critical parameters. In addition, Table 4 summarizes the statistical information. As seen from Table 4, the tensile strength of FRP (F_{frp}), load-carrying capacity (P_{max}), and gross cross-sectional area of the collected columns tended to exhibit the largest variations.

Figure 2. SHAP attributes[51]

Table 1. Summary of the existing studies used to develop the experimental database

No.	References	Number of specimens and type of loading	
		Concentric compression	Eccentric compression
1	Bae and Belarbi[16]	7	—
2	Li et al.[37]	16	—
3	Tastani and Pantazopoulou[66]	11	—
4	Jayaprakash et al.[67]	—	15
5	Chotickai et al.[68]	—	12
6	Maaddawy[69]	—	12
7	Radhi et al.[70]	8	—
8	Nematzadeh et al.[71]	—	9
9	Shaikh and Alishahi[72]	4	12
10	Bai et al.[73]	6	—
11	Shan[74]	34	—
12	Yu[75]	16	—
13	Li et al.[76]	3	—
14	Wen[77]	10	—
15	Chen[78]	28	—
16	Gao[79]	28	—

Determination of the input variables

Reasonable determination of the input variables is significant to accurately predict the P_{max} of FRP-confined corroded RC columns. Thus, a comprehensive investigation of the constructed experimental database was conducted by determining the correlation coefficient (φ_k) and corresponding statistical importance.[18] The primary aim of correlation analysis is to investigate the potential association relationship between the independent input parameters and output response. The concept of φ_k was proposed by Baak et al.,[80] and it has several advantages. The statistical importance is usually utilized to determine the accuracy and relevance of φ_k. Indeed, a high coefficient of correlation might be statistically significant or insignificant. On the other hand, a small correlation might be very significant. The statistical significance of each correlation was based on a hybrid method of Monte Carlo simulations (MCS) and adjustments of Pearson's χ^2 test.[18,80] The significance is obtained by converting the p-value of the hypothesis test to a normal Z-score. The significance is defined as follows:

$$Z = \Phi^{-1}(1-p) \tag{16}$$

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{z} e^{-t^2/2} dt \tag{17}$$

where Z is the significance in 1-sided Gaussian standard deviations and Φ^{-1} is the quantile of the standard Gaussian. It should be noted that the input variables were simplified before the correlation analysis. For example, D, H, b, and h were integrated by using A_g. r was simplified by ρ, which is more suitable in practical situations. For the input material properties of FRP composites, FRP_{type} could be represented by E_{frp} and F_{frp}, as well as N_{frp} could be represented by

t_{frp}. Similarly, for the input material properties of steel bars, L_{type} could be represented by ρ_s. Fig. 3 shows the φ_k and statistical significance matrixes of the input parameters. φ_k varies between 0 and 1, where 0 means no association and 1 means complete association, respectively.

As illustrated in Fig. 3, a darker color means a more pronounced correlation. For the selected variables, P_{max} of the specimens exhibited a strong correlation with f_c ($\varphi_k = 0.81$, significance = 8.44), A_g ($\varphi_k = 0.77$, significance = 8.45), t_{frp} ($\varphi_k = 0.76$, significance = 5.87), E_{bar} ($\varphi_k = 0.76$, significance = 6.86), F_{frp} ($\varphi_k = 0.75$, significance = 6.93), and F_{bar} ($\varphi_k = 0.75$, significance = 9.38), respectively. Likewise, P_{max} also correlated well with ρ_s ($\varphi_k = 0.71$, significance = 7.38), e ($\varphi_k = 0.71$, significance = 7.17), E_{frp} ($\varphi_k = 0.67$, significance = 7.13), and ρ ($\varphi_k = 0.63$, significance = 4.8), respectively. Fig. 4 shows the linear regression analysis results of P_{max} of the columns with different input variables. Obviously, as seen from Fig. 4, the P_{max} of the specimens exhibited an increasing trend with the increase of A_g, ρ, f_c, t_{frp}, E_{frp}, F_{frp}, E_{bar}, and F_{bar}, but it decreased with the increase of ρ_s, e, and η. Based on the above-mentioned preliminary correlation analyses, the following function was considered as the XGBoost predictive model for predicting the P_{max} of FRP-confined corroded RC columns.

$$P_{max} = f\left(A_g, \rho, f_c, t_{frp}, E_{frp}, F_{frp}, \rho_s, E_{bar}, F_{bar}, e, \eta\right) \tag{18}$$

Model training and performance evaluations

In this study, the experimental database was randomly categorized into two different parts, including (1) the training datasets and (2) the testing datasets. In specific, 80% and 20% of specimens were used to construct the training and testing datasets, respectively. The former was employed to

Table 2. Detailed information and experimental results of the specimens tested in this study

No.	Specimen ID	D (mm)	H (mm)	b (mm)	h (mm)	A_g (mm²)	Circular	R (mm)	ρ (%)	f_c (MPa)	FRP_{type}	N_{frp}	t_{frp} (mm)	E_{frp} (GPa)	F_{frp} (MPa)	L_{type}	ρ_s (%)	E_{bar} (GPa)	F_{bar} (MPa)	e (%)	η (%)	P_{max} (kN)
1	A0-1	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	0	0.167	265	4525	4NO.12	2.56	201	445	0	0	513.1
2	A0-2	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	0	0.167	265	4525	4NO.12	2.56	201	445	0	0	525.1
3	A0-3	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	0	0.167	265	4525	4NO.12	2.56	201	445	0	0	482.6
4	A12.5-1	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.12	2.56	201	445	0	16.32	319.5
5	A12.5-2	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.12	2.56	201	445	0	14.87	388.2
6	A12.5-3	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.12	2.56	201	445	0	14.22	426.5
7	AF0-1	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.12	2.56	201	445	0	0	1094.7
8	AF0-2	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.12	2.56	201	445	0	0	1044.5
9	AF0-3	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.12	2.56	201	445	0	0	1023.4
10	AF5-1	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.12	2.56	201	445	0	8.38	1068.9
11	AF5-2	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.12	2.56	201	445	0	7.11	1064.8
12	AF5-3	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.12	2.56	201	445	0	7.34	1062.5
13	AF12.5-1	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.12	2.56	201	445	0	14.74	972.5
14	AF12.5-2	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.12	2.56	201	445	0	15.67	946.2
15	AF12.5-3	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.12	2.56	201	445	0	15.01	944.8
16	AF20-1	150	300	—	—	17662.5	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.12	2.56	201	445	0	23.24	802.5
17	AF20-2	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	445	0	24.87	892.9
18	AF20-3	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	445	0	23.05	906.5
19	B0-1	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	0	903.4
20	B0-2	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	0	929.4
21	B0-3	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	0	808.7
22	B12.5-1	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	15.44	746.5
23	B12.5-2	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	15.13	764.3
24	B12.5-3	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	16.07	662.7
25	BF0-1	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	0	1801.1
26	BF0-2	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	0	1688.4
27	BF0-3	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	0	1724.8
28	BF5-1	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	7.12	1685.5
29	BF5-2	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	7.42	1674.5
30	BF5-3	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	8.14	1443.3
31	BF12.5-1	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	14.19	1371.2
32	BF12.5-2	200	400	—	—	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	15.20	1300.5

Table 2. Continued

No.	Specimen ID	D (mm)	H (mm)	b (mm)	h (mm)	A_g (mm²)	Circular	R (mm)	ρ (%)	f_c (MPa)	FRP_{type}	N_{frp}	t_{frp} (mm)	E_{frp} (GPa)	F_{frp} (MPa)	L_{type}	ρ_s (%)	E_{bar} (GPa)	F_{bar} (MPa)	e (%)	η (%)	P_{max} (kN)
33	BF12.5-3	200	400	–	–	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	15.56	1306.7
34	BF20-1	200	400	–	–	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	23.61	1179.0
35	BF20-2	200	400	–	–	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	24.22	1117.6
36	BF20-3	200	400	–	–	31400.0	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.16	2.56	230	460	0	23.06	1212.8
37	C0-1	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	0	1538.9
38	C0-2	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	0	1485.6
39	C0-3	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	0	1519.9
40	C12.5-1	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	14.57	1081.7
41	C12.5-2	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	13.88	1124.9
42	C12.5-3	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	13.61	1112.4
43	CF0-1	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	0	2350.1
44	CF0-2	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	0	2202.2
45	CF0-3	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	0	2274.1
46	CF5-1	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	7.24	2261.2
47	CF5-2	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	7.33	2256.4
48	CF5-3	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	6.98	2290.9
49	CF12.5-1	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	14.25	1970.2
50	CF12.5-2	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	14.06	1999.9
51	CF12.5-3	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	15.11	1918.3
52	CF20-1	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	23.37	1719.9
53	CF20-2	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	22.49	1636.6
54	CF20-3	250	500	–	–	49087.4	Yes	0	1	25.3	CFRP	1	0.167	265	4525	4NO.20	2.56	220	455	0	22.82	1694.5

Table 3. Descriptions and representations of the input/output variables

Variables	Parameters (units)	Notation
Input	Diameter of circular cross-section (mm)	D
	Column height (mm)	H
	Width of rectangular cross-section (mm)	b
	Height of rectangular cross-section (mm)	h
	Column gross cross-sectional area (mm²)	A_g
	Column section type	$Section_{type}$
	Corner radius (mm)	r
	Corner radius ratio	ρ
	Compressive strength of concrete (MPa)	f_c
	Type of fiber-reinforced polymer	FRP_{type}
	Layer number of FRP	N_{frp}
	Thickness of FRP	t_{frp}
	Elastic modulus of FRP	E_{frp}
	Tensile strength of FRP	F_{frp}
	Type of longitudinal reinforcement	L_{type}
	Longitudinal reinforcement ratio (%)	ρ_s
	Elastic modulus of steel reinforcement (GPa)	E_{bar}
	Yield strength of the steel reinforcement (MPa)	F_{bar}
	Eccentricity ratio (%)	e_r
	Corrosion rate (%)	η
Output	Load-carrying capacity (kN)	P_{max}

Table 4. Summary of the statistical information of the input variables

Input and output variables	Minium	Mean	Standard deviation	Maximum
D (mm)	100	158.85	38.26	203
H (mm)	150	504.45	261.44	1375
b (mm)	120	150.96	27.17	200
h (mm)	120	150.96	27.17	200
A_g (mm²)	7850	21750.21	9132.59	40000
r (mm)	0	7.32	15.47	75
ρ	0	0.78	0.35	1
f_c (MPa)	17.7	33.06	7.35	47
t_{frp} (mm)	0	0.23	0.23	1.68
E_{frp} (GPa)	0	162.65	111.87	280
F_{frp} (MPa)	0	2826.66	1664.18	4900
ρ_s (%)	0.89	2.75	1.29	6.79
E_{bar} (GPa)	199.1	205.80	11.73	237
F_{bar} (MPa)	210	412.57	82.53	550
e_r (%)	0	0.19	0.52	3.44
η (%)	0	10.02	9.62	51
P_{max} (kN)	36.9	915.59	628.67	2536.11

train the model and parameter evaluation, whereas the latter was taken for model assessment. Thus, in the present developed XGBoost predictive model, the training and testing datasets had 228 and 57 FRP-confined corroded RC columns, respectively. As per the model training, the effectiveness and capability of the XGBoost model were assessed by using several crucial measures, including (i) the coefficient of determination (R^2); (ii) root mean square error (RMSE);
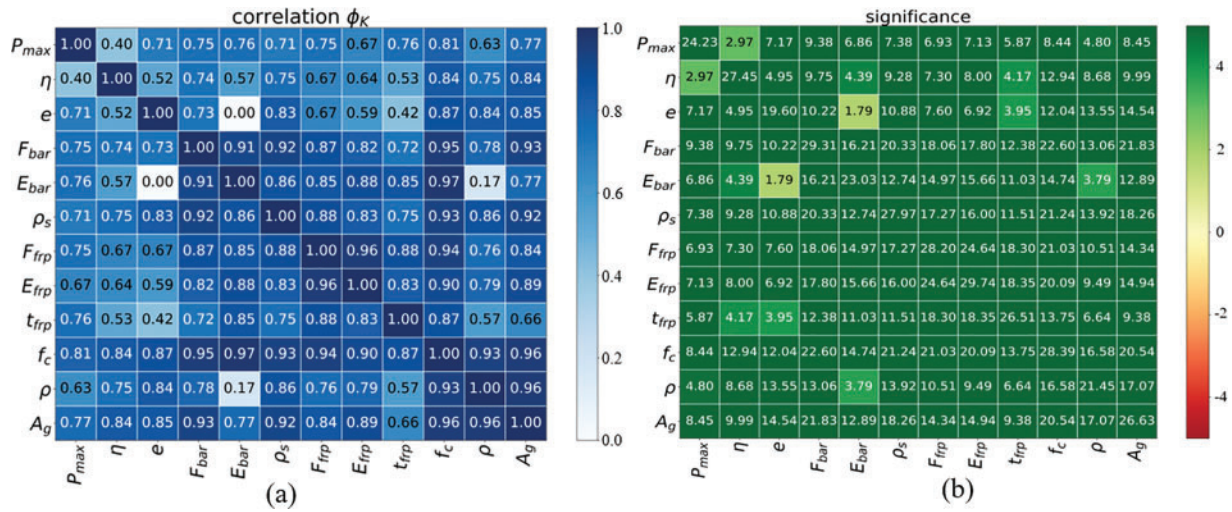
Figure 3. Correlation coefficients and statistical significance of each correlation: (a) correlation coefficient and (b) statistical significance

(iii) mean absolute error (MAE); and (iv) mean absolute percentage error (MAPE).[18,19] Their mathematical expressions are given in the following equations.

$$R^2\left(P_{max}, \hat{P}_{max}\right) = 1 - \frac{\sum_{i=1}^{m}\left(P_{max}(i) - \hat{P}_{max}(i)\right)^2}{\sum_{i=1}^{m}\left(P_{max}(i) - \overline{P}_{max}\right)^2} \quad (19)$$

$$RMSE\left(P_{max}, \hat{P}_{max}\right) = \sqrt{\frac{1}{m} \cdot \sum_{i=0}^{m-1}\left(P_{max}(i) - \hat{P}_{max}(i)\right)^2} \quad (20)$$

$$MAE\left(P_{max}, \hat{P}_{max}\right) = \frac{1}{m} \cdot \sum_{i=0}^{m-1}\left|P_{max}(i) - \hat{P}_{max}(i)\right| \quad (21)$$

$$MAPE\left(P_{max}, \hat{P}_{max}\right) = \frac{100}{m} \cdot \sum_{i=0}^{m-1}\left|\frac{P_{max}(i) - \hat{P}_{max}(i)}{P_{max}(i)}\right| \quad (22)$$

where m is the number of data points; P_{max} and $\hat{P}_{max}$ are the experimental and predicted ultimate strengths of columns, respectively; and $\overline{P}_{max}$ is the mean value of test results, which can be determined by Eq. (23). Among these statistical measures, a larger R^2 (i.e., close to 1.0) and the smaller values of RMSE, MAE, and MAPE indicate superior prediction accuracy of the model.

$$\overline{P}_{max} = \frac{1}{N} \cdot \sum_{i=1}^{N} P_{max}(i) \quad (23)$$

Model tuning and cross-validations

The performance and effectiveness of the XGBoost prediction model could be enhanced by determining the optimal combination of hyperparameter values. Grid search, random search, and Bayesian optimization methods are the most common techniques to tune machine-learning models.[18] In this study, the hyperparameters of the model were optimized through k-fold cross-validations combined with randomized and grid searches. The initial hyperparameters were determined by the randomized search, and then the acquired ones were further optimized using the grid search. Subsequently, the training dataset was randomly divided into k folds, in which $(k-1)$ folds were utilized for the model training, and 1-fold was used for performance assessment during the k-fold cross-validation process. Such a process would be repeated k times, where each of the k subsamples was employed once as the validation data. 10-fold cross-validation was used in this study. Hence, during each cross-validation process, 90% of the dataset was used as the training set, while the remaining 10% was used for performance assessment of the model. Results of the randomized and grid searches are summarized in Table 5.

Prediction Results and Discussions

Performance evaluation of the XGBoost model

As introduced in Sect. 2, the XGBoost algorithm builds the sequential trees. With this regard, a single XGBoost decision tree from the training model is presented in Fig. 5. As shown in Fig. 5, the root node was A_g, and the second layers were e and t_{frp}, respectively. These observations were consistent with the correlation analyses as presented in Sect. 3. In addition, the regression error and residual values of the predicted P_{max} using the XGBoost model could be obtained, and the predicted P_{max} was illustrated in Fig. 6. As seen from this figure, the XGBoost-predicted P_{max} was generally close to the experimental results, with the R^2 of 0.994. Thus, the developed XGBoost model could provide the acceptable P_{max} for FRP-confined corroded RC columns.

Table 6 summarizes the performance metrics of the XGBoost predictions of the training and testing datasets for different models, respectively. As seen from Table 6, for different models, the accuracy of the training dataset was generally superior to the testing one. For example, for the model of all data, values of R^2, RMSE, MAE, and MAPE
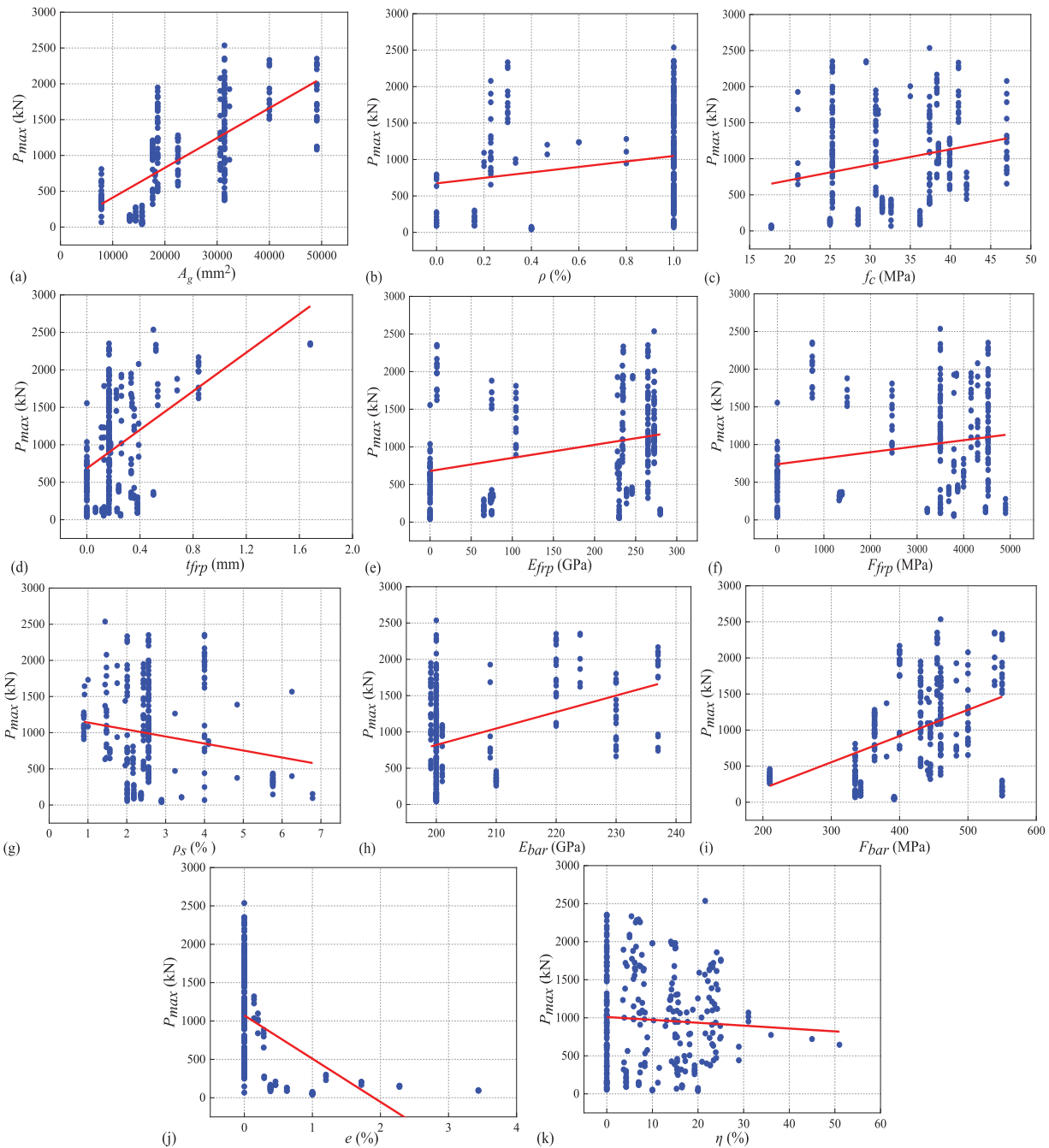
21425005-10

BER Open: Int. J. Bridge Eng., Manage. Res.

Figure 4. Linear regression analyses of the P_{max} versus the input variables

with the model for the training process were 0.993, 56 kN, 13.4 kN, and 2%, respectively, whereas those for the testing procedure were 0.978, 122 kN, 70.6 kN, and 7.7%, respectively. This suggests that the developed XGBoost model exhibited both good learning and predicting capacity. Additionally, values of MAPE for all the ML models were smaller than 10%, indicating the prediction accuracy of the developed XGBoost model was excellent.[18] Thus, this further demonstrated that the developed XGBoost model showed superior effectiveness and accuracy in determining P_{max} of FRP-confined corroded RC columns.

Fig. 7 shows the feature importance based on the developed XGBoost predictive model. This figure indicates how

each input variable affected the XGBoost model's predictions. The feature importance was automatically calculated by the XGBoost algorithm. F scores of the predictive model could be determined by three different evaluation criteria, including (i) weight, (ii) gain, and (iii) cover scores.[18] In specific, the F scores were obtained based on the number of times a feature appeared in a tree (XGBoost weight score), the average gain of splits using the feature (XGBoost gain score), or the average coverage of splits using the feature with coverage being defined as the number of samples affected by the split (XGBoost cover score).[18] There is a direct relationship between feature importance and the value of the F score. As observed from Fig. 7, the feature importance determined

Table 5. Randomized and grid search values investigated by the hyperparameter tuning and cross-validations

Hyperparameters	Description	Lower limit	Upper limit	Best hyperparameters		
				All data	Concentric	Eccentric
n_estimators	Number of gradient-boosted trees	0	200	150	150	150
max_depth	Maximum tree depth for base learners	1	13	8	7	6
learning_rate	Step size shrinkage used in the update to prevent overfitting	0.1	1	0.11	0.25	0.2
subsample	Subsample ratio of the training instances	0	1	1	1	1
colsample_bytree	Subsample ratio of columns when constructing each tree	0	1	1	0.95	0.75
alpha	L1 regularization term on weights	0	1	1	0.85	0.85

Figure 5. Single XGBoost decision trees from the trained model

by using different evaluation criteria was inconsistent. For example, by using the weight score as the evaluation criterion, the five most significant parameters influencing the predictions of P_{max} of the columns were η, t_{frp}, A_g, f_c, and ρ_s, whereas that were A_g, F_{bar}, e, E_{frp}, and t_{frp}; as well as η, E_{bar}, t_{frp}, f_c, and F_{bar}, respectively, by employing the gain and cover scores as the evaluation criteria, respectively. Such an inconsistence in the predicted feature importance from the XGBoost model based on different evaluation criteria could lead to the interpretation and explanations of the model's predictions being contradictory. However, this is inevitable because the traditional XGBoost models could have inconsistent assessments of feature importance; similar observations were also found in several previous studies.[18,51,56,81] Thus, an additional analysis of the significance of feature parameters was conducted and presented in the following subsection.

Explanation of the XGBoost model

Fig. 8 shows the SHAP summary plot and the relative feature importance of the input variables. As shown in Fig. 8, the SHAP plot illustrates the SHAP value for each variable, and the color represents the feature value from low (blue) to high (red). In addition, as shown in Fig. 8, the six most significant parameters influencing the prediction of P_{max} of the columns were A_g, t_{frp}, E_{frp}, e, η, and f_c, respectively. This observation agreed well with the correlation analysis in Sect. 3.2, indicating P_{max} of FRP-confined corroded RC columns mainly relied on these featured parameters (i.e., A_g, t_{frp}, e, E_{frp}, f_c, and η, respectively). In addition, as shown in

21425005-12   BER Open: Int. J. Bridge Eng., Manage. Res.

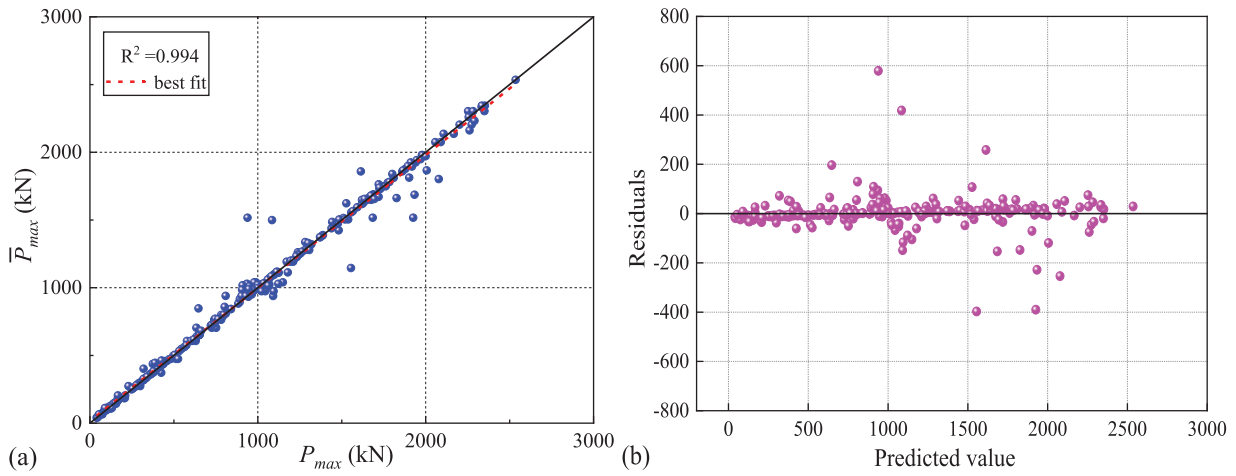BER Open: Int. J. Bridge Eng., Manage. Res., 2024, 2(1): 21425005

Figure 6. Regression error and residual values of XGBoost model: (a) prediction error; (b) residual values

Table 6. Performance metrics of XGBoost models

Models	Training dataset				Testing dataset			
	R^2	RMSE (kN)	MAE (kN)	MAPE (%)	R^2	RMSE (kN)	MAE (kN)	MAPE (%)
All data	0.993	56	13.4	2	0.978	122	70.6	7.7
Concentric	0.991	64.6	15	1.3	0.975	99.7	65.9	7.2
Eccentric	0.999	9.57	4.75	3.8	0.984	31.3	18.5	8
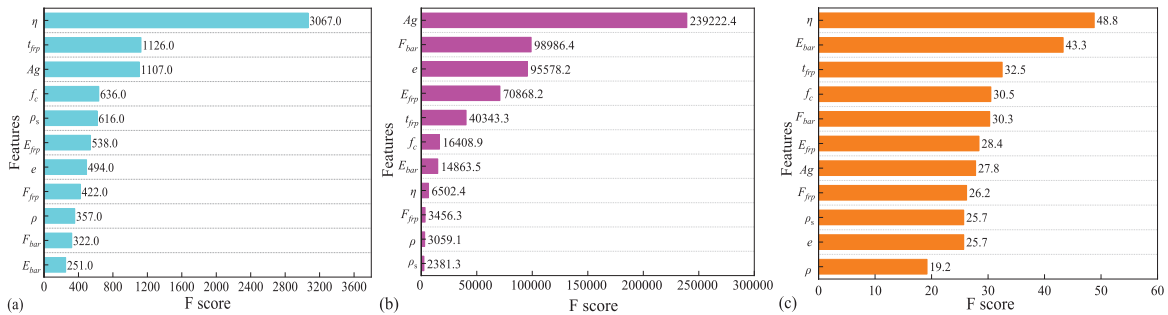
Figure 7. Feature importance based on XGBoost model: (a) weight; (b) gain; (c) cover
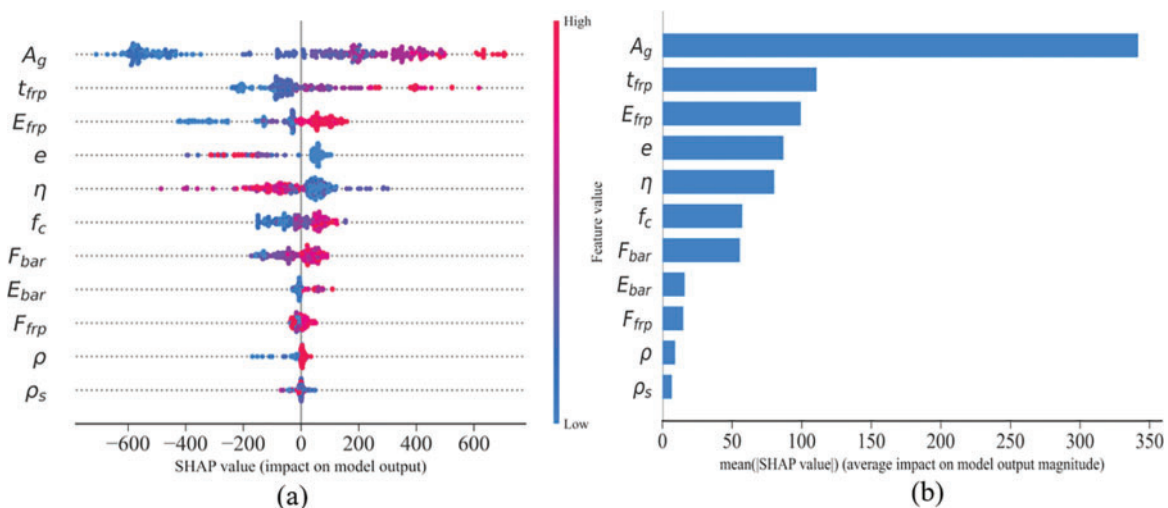
Figure 8. (a) SHAP summary plot and (b) the relative importance of each feature

21425005-13

BER Open: Int. J. Bridge Eng., Manage. Res.
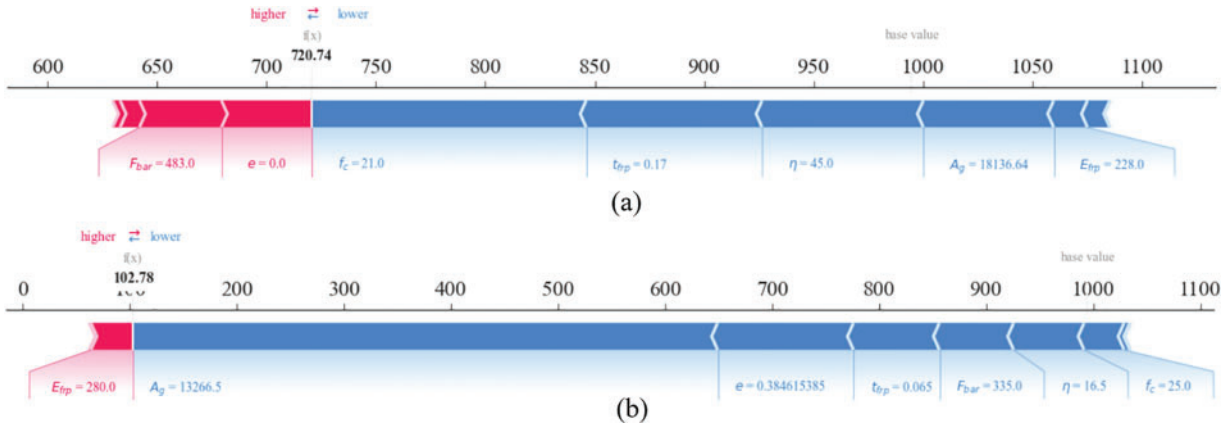
Figure 9. Explanation of typical individual prediction for (a) specimen No. 2 and (b) specimen 47

Table 7. Summary of existing models for predicting the axial ultimate strength of FRP-confined RC columns

Selected models	Cross-section	Model	Supplementary notation
Youssef et al.[44]	Circular	$\dfrac{f_{cu}}{f_{co}} = 1 + 2.25 \cdot \left(\dfrac{f_l}{f_{co}}\right)^{1.25}$	f_l is the lateral confining stress at the ultimate condition of the FRP jacket, which is represented by: $f_l = \dfrac{2E_{frp}t_{frp}\varepsilon_{fu}}{D \text{ or } b}$
	Rectangular	$\dfrac{f_{cu}}{f_{co}} = 0.5 + 1.225 \cdot \left(\dfrac{k_e f_l}{f_{co}}\right)^{0.6}$	k_e is the confinement effectiveness coefficient, which is represented by: $k_e = \dfrac{1 - \left[\dfrac{(b-2r_c)^2 + (h-2r_c)^2}{3bh}\right] - \dfrac{A_s}{bh}}{1 - \dfrac{A_s}{bh}}$
Wei and Wu[44]	Circular & rectangular	$\dfrac{f_{cu}}{f_{co}} = 1 + 2.2 \cdot \left(\dfrac{2r_c}{b}\right)^{0.72} \cdot \left(\dfrac{f_l}{f_{co}}\right)^{0.94}$ $\cdot \left(\dfrac{h}{b}\right)^{-1.9}$	$f_l = \dfrac{2E_{frp}t_{frp}\varepsilon_{fu}}{D \text{ or } b}$
Cao et al.[45]	Circular & rectangular	$\dfrac{f_{cu}}{f_{co}} = 1 + 8.34 \cdot \left(\dfrac{E_l}{E_c}\right)^{1.03}$ $\left(\dfrac{2r_c}{b}\right)^{0.81} \left(\dfrac{30}{f_{co}}\right)^{0.54} \left(\dfrac{h}{b}\right)^{-1.9} \left(\dfrac{\varepsilon_{fu}}{\varepsilon_{co}}\right)^{0.82}$	E_l is the confinement stiffness, which is represented by: $E_l = \dfrac{2E_{frp}t_{frp}}{D \text{ or } b}$

Note: D is the diameter of circular columns; h is the length of the longer side of cross-section of the rectangular columns; b is the length of the shorter side of cross-section the rectangular columns or the side of the square columns; r_c is the corner radius; E_c is the elastic modulus of concrete; A_s is the total area of longitudinal reinforcements in a column; f_{co} is the compressive strengths of the unconfined concrete; f_{cu} is the ultimate axial compressive stress of the FRP-confined concrete; ε_{co} is the peak strain of the unconfined concrete; and ε_{fu} is the ultimate tensile strain of FRP from the coupon tests.

Fig. 8(a), a high value of A_g, t_{frp}, E_{frp}, and f_c tended to boost the predictions of P_{max} of columns up, while low values could decrease the predictions. However, a high value of e and η tended to decrease the predictions, whereas a small value of e and η could increase the predictions.

Fig. 9 presents the explanation of predictions for specimens No. 2 and No. 47, respectively, which were experimentally tested under the concentric and eccentric loads, respectively. As illustrated in Fig. 9, the red arrows indicate the positive SHAP values and features that push up the model's predictions, whereas the blue arrows denote the negative SHAP values and features that push down the predictions. The base value was the average predicted P_{max} of the columns over the whole training dataset. As seen from Fig. 9, the XGBoost model's predicted P_{max} of specimens No. 2 and No. 47 were 720.74 and 102.78 kN, respectively.

The corresponding experimental test results were 720.60 and 101.65 kN, respectively. Hence, the XGBoost model's predicted P_{max} of these two specimens agreed well with the test results, indicating the superior prediction effectiveness of the XGBoost model. For specimen No. 2, F_{bar} and e were the most critical parameters that pushed up the base value, while f_c, t_{frp}, η, A_g, and E_{frp} decreased the base value. Similarly, for specimen No. 47, E_{frp} was the most crucial input variable, increasing the base value, whereas A_g, e, t_{frp}, F_{bar}, η, and f_c decreased the predictions.

Verification of the XGBoost predictive model

To further validate the effectiveness and feasibility of the XGBoost model, the predicted P_{max} of FRP-confined corroded RC columns was compared to those predicted by the

21425005-14

BER Open: Int. J. Bridge Eng., Manage. Res.

Table 8. Statistics performance metrics of the XGBoost model and existing empirical models

Models	Average	R^2	RMSE (kN)	MAE (kN)	MAPE (%)
XGBoost model	1.005	0.994	70.3	46.9	7.2
Youssef et al.[43]	1.148	0.897	338	250.1	23
Wei and Wu[44]	1.287	0.898	395.5	305	30.6
Cao et al.[45]	1.285	0.895	388.3	298.3	30.4
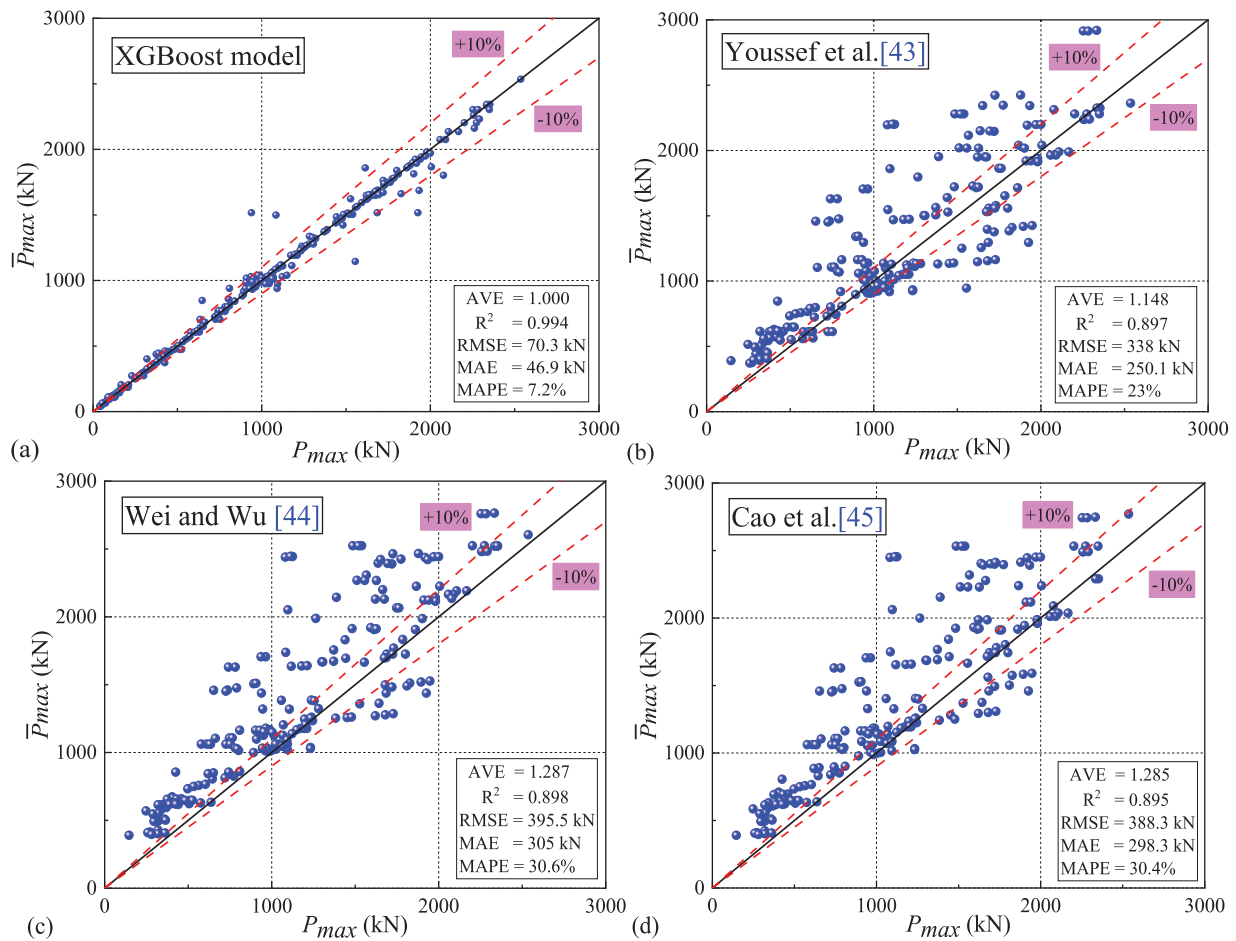
Figure 10. Comparative performance results of the XGBoost and the existing empirical models

empirical models available in several previous studies.[43–45] To date, there are many empirical models in predicting P_{max} of FRP-confined RC columns, but this paper only selected three representative ones[43–45] for analysis, which are summarized in Table 7. As seen from Table 7, the impacts of steel rebar corrosion on the mechanical performance of columns were not considered in these selected empirical models. Thus, to consider the corrosion effects on the degradation of mechanical properties of steel bars and FRP confining pressure, the design models in determining the P_{max} of FRP-confined corroded RC columns should be modified accordingly. According to several previous studies,[37,82] degradations of the mechanical performance of steel bars and FRP confining pressure of FRP-confined corroded RC

columns can be considered based on the corrosion rate (η):

$$A_s^* = (1 - \eta) \cdot A_{s0} \qquad (24)$$

$$\varepsilon_{rup}^* = (1 - 0.462\eta) \cdot \varepsilon_{rup} \qquad (25)$$

where A_{s0} and ε_{rup} are the initial cross-sectional area of steel bars and rupture strain of FRP before corrosion, whereas those after corrosion are represented by A_s^* and ε_{rup}^*, respectively. The statistical results of XGBoost and the empirical models[43–45] are presented in Table 8.

As seen from Table 8, values of R^2, RMSE, MAE, and MAPE of the XGBoost model were 0.994, 70.3 kN, 46.9 kN, and 7.2%, respectively, whereas that of the empirical model suggested by Wei and Wu[44] were 0.898, 395.5 kN, 305 kN, and 30.6%, respectively. Hence, the XGBoost model showed the best prediction results with the largest value of R^2 and the
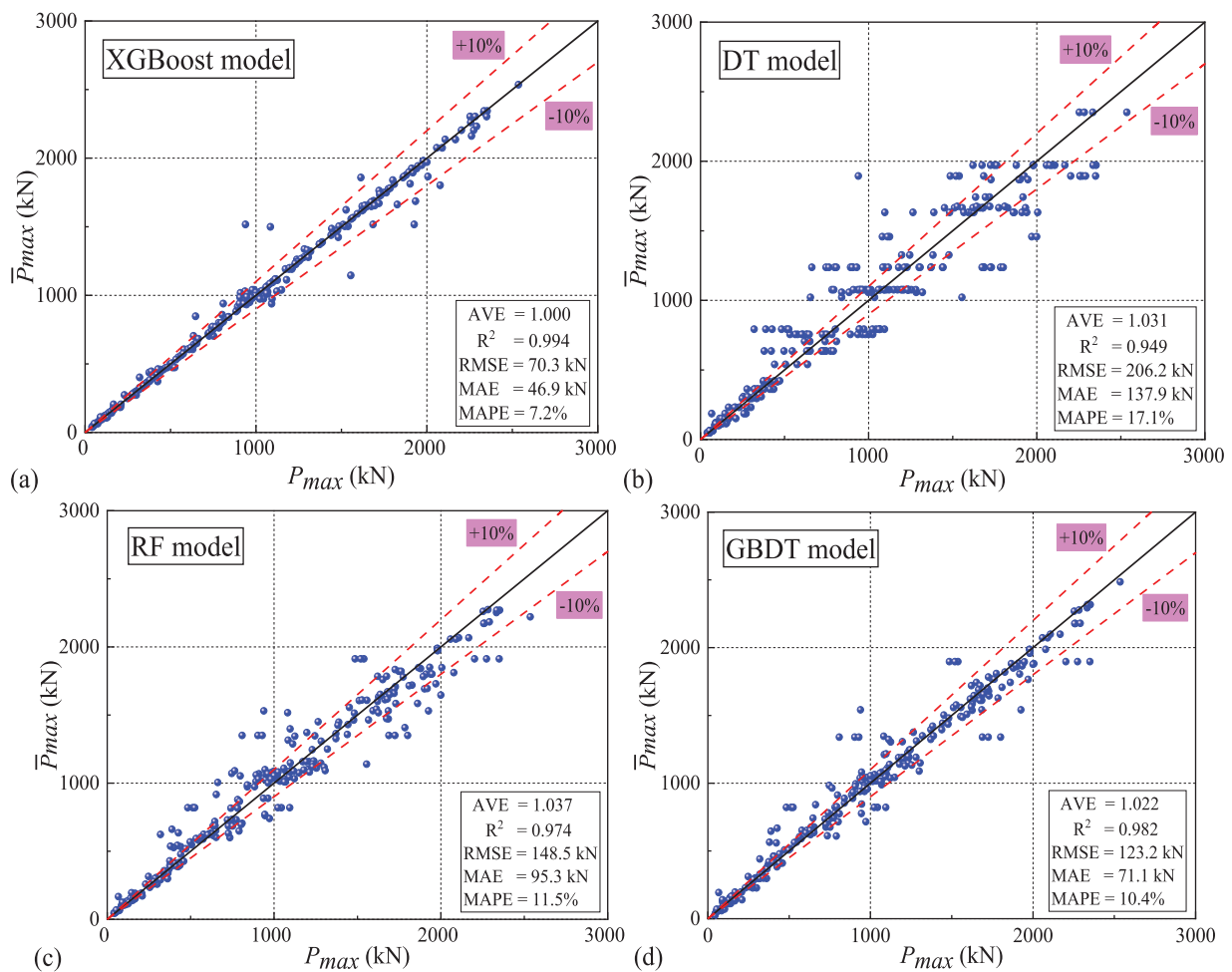
Figure 11. Comparative performance results of the XGBoost and other ML models

Table 9. Statistics performance metrics of XGBoost and other ML models

Model	Training dataset				Testing dataset				All data			
	R^2	RMSE (kN)	MAE (kN)	MAPE (%)	R^2	RMSE (kN)	MAE (kN)	MAPE (%)	R^2	RMSE (kN)	MAE (kN)	MAPE (%)
XGBoost	0.998	56	13.4	2	0.978	122	70	7.7	0.994	70.3	46.9	7.2
DT	0.901	196.8	138.9	17.5	0.878	240.2	152.5	15.4	0.949	206.2	137.9	17.1
RF	0.967	116.2	68.3	8.8	0.940	167.8	100.6	9.8	0.974	148.5	95.3	11.5
GBDT	0.969	112.6	66.3	10.4	0.946	158.6	90.5	10.2	0.982	123.2	71.1	10.4

smallest prediction errors (RMSE, MAE, and MAPE). This suggests that the feasibility and effectiveness of the XGBoost model in predicting P_{max} of FRP-confined corroded RC columns outperformed these empirical models.[43–45]

In addition, Fig. 10 displays the comparative performance results of the XGBoost and existing empirical models within a discreteness range of ±10%. As observed from Fig. 10, the XGBoost model exhibited the best prediction performance compared to existing empirical models.[43–45] Among these considered empirical models, the one suggested by Youssef et al.[43] tended to exhibit better predictions than Wei and Wu et al.[44] and Cao et al.,[44] but the most of prediction points generated by using these models were outside the desirable discreteness range (±10%), indicating the significant dispersions of the prediction results. Moreover,

it is worth noting that the P_{max} of FRP-confined corroded RC columns calculated by these empirical models was higher than the experimental ones. This is probably because the modeling of the corrosion effects on FRP-confined corroded RC columns is a complex problem. The simplified analysis of the corrosion effects through degradation of the cross-sectional area of steel bars (Eq. (24)) and reduction of the rupture strain of FRP composites (Eq. (25)) could not be effective and accurate enough.

Additionally, the XGBoost model was also compared with several other ML algorithms, such as the decision tree (DT), random forest (RF), and gradient boosting decision tree (GBDT). Fig. 11 compares the prediction results of the XGBoost model and the other three ML algorithms within a discreteness range of ±10%. It could be observed from

21425005-16

BER Open: Int. J. Bridge Eng., Manage. Res.

Table 9 and Fig. 11 that the RF and GBDT models exhibited good predictive performance in predicting the P_{max} of the columns. Still, the prediction results of the DT algorithm showed significant discreteness. Also, as seen from Fig. 11, the majority of the prediction points generated through the developed XGBoost model were inside the desirable discreteness range ($\pm 10\%$), whereas the prediction points of the other three ML models showed relatively pronounced dispersions. This further validated the superior effectiveness and capability of the XGBoost predictive model in predicting P_{max} of FRP-confined corroded RC columns, compared to the other three ML algorithms.

Conclusions

This study proposed a novel explainable machine learning (ML) model for the prediction of the axial load-carrying capacity (P_{max}) of FRP-confined corroded RC columns using the XGBoost algorithm and SHAP technique. The explainable XGBoost predictive model was established based on a thorough database of experimental tests for 285 FRP-confined corroded RC columns subjected to concentric and eccentric loadings. 20 parameters were selected as the critical input variables. Then, the SHAP technique was employed for the important evaluation and interpretation of the prediction performance of the model in predicting the P_{max} of the columns. Additionally, the effectiveness and accuracy of the developed XGBoost predictive model were validated through several empirical prediction models reported in the literature and some popularly used ML algorithms (DT, RF, and GBDT). Finally, the following conclusions are summarized:

1. A novel, explainable XGBoost decision tree-based ML method was proposed for quantitatively predicting P_{max} of FRP-confined corroded RC columns. The developed XGBoost predictive model was demonstrated to be capable and effective with good prediction performance and accuracy.
2. The proposed XGBoost predictive model could achieve good prediction interpretability using the SHAP technique. The feature importance of the selected critical variables could be quantitatively studied, and the most important ones influencing the prediction of P_{max} of FRP-confined corroded RC columns were A_g, t_{frp}, E_{frp}, e, η, and f_c, among the considered input variables.
3. The developed XGBoost model exhibited excellent prediction performance and accuracy in predicting the P_{max} of FRP-confined corroded RC columns. Values of R^2, RMSE, MAE, and MAPE of the XGBoost model were 0.978, 122 kN, 7036 kN, and 7.7%, respectively. The prediction effectiveness and capability of the model in predicting P_{max} of the columns significantly outperformed those of the existing empirical models. Also, the developed XGBoost predictive model was able to achieve superior predictions than the DT, RF, and GBDT algorithms.
4. The proposed XGBoost predictive model could provide new insights for addressing traditional engineering issues involving many critical influential parameters. In addition, if the database could be further enriched in the future, this developed XGBoost predictive model should be continuously updated and thereby making its prediction performance and accuracy superior and more reliable.

Although the developed XGBoost predictive model was suitable for predicting the load-carrying capacity of FRP-confined corroded RC columns, it also has several limitations. For instance, the database was constructed from 285 FRP-confined corroded RC columns collected from the existing studies (231 specimens) reported in the literature and those performed by the authors (54 specimens). The completeness of the experimental data, structural dimensions, environmental conditions, non-uniform corrosion effects, testing quality, and distributions of the input parameters play critical roles in the prediction accuracy and effectiveness of the developed XGBoost models. Thus, to further improve the prediction accuracy and effectiveness of the model, the experimental database, input feature variables, and the interactions among these considered variables should be updated and enriched with more test data. In addition, the generalizability of the SHAP explanations and XGBoost predictive results might be limited to the ranges of the input data tested, in addition to the cross-validation process, more advanced techniques, such as the Grid search, random search, and Bayesian optimization methods could be incorporated, to reduce the risk of overfitting. Moreover, the effectiveness and feasibility of the developed XGBoost predictive model were only validated against several empirical prediction models and some popularly used ML algorithms, such as the decision tree (DT), random forest (RF), and gradient boosting decision tree (GBDT). However, the developed XGBoost predictive model should be verified in the future through more advanced interpretable machine learning or deep learning models. Overall, the measured errors of the XGBoost predictions were very low from the perspective of engineering practice. XGBoost is an accurate tree-boosting system and it is designed as a regularized model formalized to control overfitting. Using the trained XGBoost predictive model, the user can theoretically predict the load-carrying capacity of FRP-confined corroded RC columns and other similar problems based on the assembled experimental database, which would have great application potential in practical engineering practice.

Acknowledgments

21425005-17   BER Open: Int. J. Bridge Eng., Manage. Res.

BER Open: Int. J. Bridge Eng., Manage. Res., 2024, 2(1): 21425005

Data Availability Statement

The code and the experimental database used in this study are published on the GitHub page (https://github.com/leeHR666/Load-carrying-capacity-prediction-of-the-FRP-confined-corroded-RC-columns).

Disclosure Statement

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.

References

[1] Li HH. *Alternate Load Paths and Retrofits for Long-Span Truss Bridges Under Sudden Member Loss and Blast Loads*. Ph.D. Thesis. Department of Civil Engineering, The City University of New York, New York, NY; 2021.

[2] Li HH, Li LF, Ye M, et al. Blast fragility assessment of aging coastal RC columns exposed to non-uniform CIC attacks using LBE function. *J Build Eng*. 2023;71(4):106510. doi:10.1016/j.jobe.2023.106510.

[3] Ghosh J, Sood P. Consideration of time-evolving capacity distributions and improved degradation models for seismic fragility assessment of aging highway bridges. *Reliab Eng Syst Saf*. 2016;154(1):197–218. doi:10.1016/j.ress.2016.06.001.

[4] Cui FK, Zhang HN, Ghosn M, Xu Y. Seismic fragility analysis of deteriorating RC bridge substructures subject to marine chloride-induced corrosion. *Eng Struct*. 2018;155(1):61–72. doi:10.1016/j.engstruct.2017.10.067.

[5] Li HH, Li LF. Bridge fragility analysis based on an improved uniform design-response surface methodology. *J Vib Shock*. 2018;37(22):245–254 (in Chinese).

[6] Li HH, Li LF. Bridge time-varying seismic fragility considering variables' correlation. *J Vib Shock*. 2019;38(9):173–183 (in Chinese).

[7] Cui FK, Li HH, Dong X, et al. Improved time-dependent seismic fragility estimates for deteriorating RC bridge substructures exposed to chloride attack. *Adv Struct Eng*. 2021;24(3):437–452. doi:10.1177/1369433220956812.

[8] Li HH, Li LF, Zhou GJ, et al. Time-dependent seismic fragility assessment for aging highway bridges subject to non-uniform chloride-induced corrosion. *J Earthq Eng*. 2022;26(7):3523–3553. doi:10.1080/13632469.2020.1809561.

[9] Li HH, Li LF, Wu WP, et al. Seismic fragility assessment framework for highway bridges based on an improved uniform design-response surface model methodology. *Bull Earthq Eng*. 2020;18(5):2329–2353. doi:10.1007/s10518-019-00783-1.

[10] Li HH, Li LF, Zhou GJ, et al. Effects of various modeling uncertainty parameters on the seismic response and seismic fragility estimates of the aging highway bridges. *Bull Earthq Eng*. 2020;18(14):6337–6373. doi:10.1007/s10518-020-00934-9.

[11] Zhou YW, Chen X, Wang XH, et al. Seismic performance of large rupture strain FRP retrofitted RC columns with corroded steel reinforcement. *Eng Struct*. 2020;216(6):110744. doi:10.1016/j.engstruct.2020.110744.

[12] Draganić H, Gazić G, Varevac D. Experimental investigation of design and retrofit methods for blast load mitigation-A state-of-the-art review. *Eng Struct*. 2019;190:189–209. doi:10.1016/j.engstruct.2019.03.088.

[13] Kishore KB, Gangolu J, Ramancha MK, et al. Performance-based probabilistic deflection capacity models and fragility estimation for reinforced concrete column and beam subjected to blast loading. *Reliab Eng Syst Saf*. 2022;227(7):108729. doi:10.1016/j.ress.2022.108729.

[14] Zhang XZ, Li ZX, Shi YC, et al. Fragility analysis for performance-based blast design of FRP-strengthened RC columns using artificial neural network. *J Build Eng*. 2022;52(6):104364. doi:10.1016/j.jobe.2022.104364.

[15] Al-Saadi NTK, Mohammed A, Al-Mahaidi R, et al. A state-of-the-art review: near-surface mounted FRP composites for reinforced concrete structures. *Constr Build Mater*. 2019;209(3):748–769. doi:10.1016/j.conbuildmat.2019.03.121.

[16] Bae SW, Belarbi A. Effects of corrosion of steel reinforcement on RC columns wrapped with FRP sheets. *J Perf Constr Facil*. 2009;23(1):20–31. doi:10.1061/(ASCE)0887-3828(2009)23:1(20).

[17] Neale KW, Demers M, Labossiere P. FRP protection and rehabilitation of corrosion-damaged reinforced concrete columns. *Int J Mater Prod Technol*. 2005;23(3/4):348–371. doi:10.1504/IJMPT.2005.007735.

[18] Bakouregui AS, Mohamed HM, Yahia A, et al. Explainable extreme gradient boosting tree-based prediction of load-carrying capacity of FRP-RC columns. *Eng Struct*. 2021;245(93):112836. doi:10.1016/j.engstruct.2021.112836.

[19] Liu X, Liu TQ, Feng P. Long-term performance prediction framework based on XGBoost decision tree for pultruded FRP composites exposed to water, humidity and alkaline solution. *Compos Struct*. 2022;284(5):115184. doi:10.1016/j.compstruct.2022.115184.

[20] Debaiky AS, Green MF, Hope BB. Long-term monitoring of carbon fiber-reinforced polymer-wrapped reinforced concrete columns under severe environment. *ACI Struct J*. 2006;103(6):865–873.

[21] Suh K, Mullins Sen GR, Winters D. Effectiveness of fiber-reinforced polymer in reducing corrosion in marine environment. *ACI Struct J*. 2007;104(1):76–83.

[22] Debaiky AS, Green MF, Hope BB. Carbon fiber-reinforced polymer wraps for corrosion control and rehabilitation of reinforced concrete columns. *ACI Mater J*. 2002;99(2):129–137.

[23] Harajli MH, Rteil AA. Effect of confinement using fiber-reinforced polymer or fiber-reinforced concrete on seismic performance of gravity load-designed columns. *ACI Struct J*. 2004;101(1):47–56.

[24] Bisby LA, Dent AJ, Green MF. Comparison of confinement models for FRP wrapped concrete. *ACI Struct J*. 2005;102(1):62–72.

[25] Chaallal O, Hassan M, LeBlanc M. Circular columns confined with FRP: Experimental versus predictions of models and guidelines. *J Composit Constr*. 2006;10(1):4–12. doi:10.1061/(ASCE)1090-0268(2006)10:1(4).

[26] Lam L, Teng JG. Design-oriented stress-strain model for FRP-confined concrete in rectangular columns. *J Reinforc Plast Compos*. 2003;22(13):1149–1186. doi:10.1177/0731684403035429.

21425005-18

BER Open: Int. J. Bridge Eng., Manage. Res.

BER Open: Int. J. Bridge Eng., Manage. Res., 2024, 2(1): 21425005

[27] Teng JG, Jiang T, Lam L, et al. Refinement of a design-oriented stress-strain model for FRP-confined concrete. *J Compos Constr*. 2009;13(4):269–278. doi:10.1061/(ASCE)CC.1943-5614.0000012.

[28] Wu Y-F, Wei Y. General stress-strain model for steel- and FRP-confined concrete. *J Compos Constr*. 2015;19(4):04014069. doi:10.1061/(ASCE)CC.1943-5614.0000511.

[29] Lee HS, Kage T, Noguchi T, et al. An experimental study on the retrofitting effects of reinforced concrete columns damaged by rebar corrosion strengthened with carbon fiber sheets. *Cement Concr Res*. 2003;33(4):563–570. doi:10.1016/S0008-8846(02)01004-9.

[30] Li J, Gong J, Wang L. Seismic behavior of corrosion-damaged reinforced concrete columns strengthened using combined carbon fiber-reinforced polymer and steel jacket. *Constr Build Mater*. 2009;23(7):2653–2663. doi:10.1016/j.conbuildmat.2009.01.003.

[31] Wang D, Huang L, Yu T, et al. Seismic performance of CFRP-retrofitted large-scale square RC columns with high axial compression ratios. *J Compos Constr*. 2017;21(5):04017031. doi:10.1061/(ASCE)CC.1943-5614.0000813.

[32] Wang D, Wang Z, Yu T, et al. Seismic performance of CFRP-retrofitted large-scale rectangular RC columns under lateral loading in different directions. *Compos Struct*. 2018;192(1):475–488. doi:10.1016/j.compstruct.2018.03.029.

[33] Dai KY, Yu YH, Qian K, et al. Deformation capacity of FRP retrofitted reinforced concrete columns with corroded reinforcing bars. *Eng Struct*. 2022;254(11):113834. doi:10.1016/j.engstruct.2021.113834.

[34] Chotickai P, Tongya P, Jantharaksa S. Performance of corroded rectangular RC columns strengthened with CFRP composite under eccentric loading. *Constr Building Mater*. 2021;268:121134. doi:10.1016/j.conbuildmat.2020.121134.

[35] Radhi MS, Gorgis IN, Hassan MS. Analytical analysis of design-oriented models for forecasting the performance of CFRP-confined corrosion-affected concrete columns. *Constr Build Mater*. 2021;313(6–7):125491. doi:10.1016/j.conbuildmat.2021.125491.

[36] Li N, Tang WS, Chen XY, et al. Compressive behavior degradation of FRP-confined RC columns exposed to a chlorine environment. *Mar Struct*. 2022;86(4):103277. doi:10.1016/j.marstruc.2022.103277.

[37] Li PD, Ren Y, Zhou YW, et al. Experimental study on the mechanical properties of corroded RC columns repaired with large rupture strain FRP. *J Build Eng*. 2022;54(8):104413. doi:10.1016/j.jobe.2022.104413.

[38] Zhou YW, Dang LJ, Sui LL, et al. Experimental study on the bond behavior between corroded rebar and concrete under dual action of FRP confinement and sustained loading. *Constr Build Mater*. 2017;155:605–616. doi:10.1016/j.conbuildmat.2017.08.049.

[39] Zhou YW, Zheng BW, Zhao DB, et al. Cyclic bond behaviors between corroded steel bar and concrete under the coupling effects of hoop FRP confinement and sustained loading. *Compos Struct*. 2019;224(6):110991. doi:10.1016/j.compstruct.2019.110991.

[40] Apostolopoulos CA, Papadakis VG. Consequences of steel corrosion on the ductility properties of reinforcement bar. *Constr Build. Mater*. 2008;22(12):2316–2324. doi:10.1016/j.conbuildmat.2007.10.006.

[41] Vu NS, Yu B, Li B. Predicting strength and drift capacities in corroded reinforced concrete columns. *Constr Build Mater*. 2016;115:304–318. doi:10.1016/j.conbuildmat.2016.04.048.

[42] Zhang DW, Zhao YX, Jin WL, et al. Shear strengthening of corroded reinforced concrete columns using pet fiber-based composites. *Eng Struct*. 2017;153(10):757–765. doi:10.1016/j.engstruct.2017.09.030.

[43] Youssef MN, Feng MQ, Mosallam AS. Stress-strain model for concrete confined by FRP composites. *Compos B Eng*. 2007;38(5–6):614–628. doi:10.1016/j.compositesb.2006.07.020.

[44] Wei YY, Wu Y-F. Unified stress-strain model of concrete for FRP-confined columns. *Constr Build Mater*. 2012;26(1):381–392. doi:10.1016/j.conbuildmat.2011.06.037.

[45] Cao YG, Jiang C, Wu Y-F. Cross-sectional unification on the stress-strain model of concrete subjected to high passive confinement by fiber-reinforced polymer. *Polymers*. 2016;8(5):186. doi:10.3390/polym8050186.

[46] Wakjira TG, Ibrahim M, Ebead U, et al. Explainable machine learning model and reliability analysis for flexural capacity prediction of RC beams strengthened in flexure with FRCM. *Eng Struct*. 2022;255(1):113903. doi:10.1016/j.engstruct.2022.113903.

[47] Naderpour H, Kheyroddin A, Amiri GG. Prediction of FRP-confined compressive strength of concrete using artificial neural networks. *Compos Struct*. 2010;92(12):2817–2829. doi:10.1016/j.compstruct.2010.04.008.

[48] Jalal M, Ramezanianpour AA. Strength enhancement modeling of concrete cylinders confined with CFRP composites using artificial neural networks. *Compos B Eng*. 2012;43(8):990–3000. doi:10.1016/j.compositesb.2012.05.044.

[49] Elsanadedy HM, Al-Salloum YA, Abbas H, et al. Prediction of strength parameters of FRP-confined concrete. *Compos B Eng*. 2012;43(2):228–239. doi:10.1016/j.compositesb.2011.08.043.

[50] Salehi H, Burgueno R. Emerging artificial intelligence methods in structural engineering. *Eng Struct*. 2018;171:170–189. doi:10.1016/j.engstruct.2018.05.084.

[51] Mangalathu S, Hwang SH, Jeon JS. Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. *Eng Struct*. 2020;219(6):110927. doi:10.1016/j.engstruct.2020.110927.

[52] Jiang K, Han Q, Bai Y, Du XL. Data-driven ultimate conditions prediction and stress strain model for FRP-confined concrete. *Compos Struct*. 2020;242(4):112094. doi:10.1016/j.compstruct.2020.112094.

[53] Fu B, Feng DC. A machine learning-based time-dependent shear strength model for corroded reinforced concrete beams. *J Build Eng*. 2021;36(4):102118. doi:10.1016/j.jobe.2020.102118.

[54] Munir MJ, Kazmi MS, Wu Y-F, et al. Development of novel design strength model for sustainable concrete columns: a new machine learning-based approach. *J Clean Prod*. 2022;357(8):131988. doi:10.1016/j.jclepro.2022.131988.

[55] Hu SL, Wang W, Lu YC. Explainable machine learning models for probabilistic buckling stress prediction of steel shear panel dampers. *Eng Struct*. 2023;288:116235. doi:10.1016/j.engstruct.2023.116235.

[56] Ding ZH, Zheng SX, Lei CH, et al. Machine learning-based prediction for residual bearing capacity and failure modes of rectangular corroded RC columns. *Ocean Eng*. 2023;281(1):114701. doi:10.1016/j.oceaneng.2023.114701.

[57] Cascardi A, Micelli F, Aiello MA. An artificial neural networks model for the prediction of the compressive strength of FRP-confined concrete circular columns. *Eng Struct*. 2017;140(6):199–208. doi:10.1016/j.engstruct.2017.02.047.

21425005-19   BER Open: Int. J. Bridge Eng., Manage. Res.

BER Open: Int. J. Bridge Eng., Manage. Res., 2024, 2(1): 21425005

[58] Dong W, Huang Y, Lehane B, Ma G. XGBoost algorithm-based prediction of concrete electrical resistivity for structural health monitoring. *Autom Constr*. 2020;114(8):103155. doi:10.1016/j.autcon.2020.103155.

[59] Duan J, Asteris PG, Nguyen H, et al. A novel artificial intelligence technique to predict compressive strength of recycled aggregate concrete using ICA-XGBoost model. *Eng Comput*. 2021;37(4):3329–3346. doi:10.1007/s00366-020-01003-0.

[60] Nguyen-Sy T, Wakim J, To QD, et al. Predicting the compressive strength of concrete from its compositions and age using the extreme gradient boosting method. *Constr Build Mater*. 2020;260(5):119757. doi:10.1016/j.conbuildmat.2020.119757.

[61] Zhang H, Zhou YH, Huang ZY, et al. Multiparameter identification of bridge cables using XGBoost algorithm. *J Bridge Eng*. 2023;28(5):04023016. doi:10.1061/JBENF2.BEENG-6021.

[62] Ma L, Zhou CG, Lee DK, et al. Prediction of axial compressive capacity of CFRP-confined concrete-filled steel tubular short columns based on XGBoost algorithm. *Eng Struct*. 2022;260(2):114239. doi:10.1016/j.engstruct.2022.114239.

[63] Lundberg S, Lee SI. A unified approach to interpreting model predictions. ArXiv: 170507874 [Cs, Stat]; 2017.

[64] Mangalathu S, Hwang SH, Jeon JS. Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. *Eng Struct*. 2020;219(6):110927. doi:10.1016/j.engstruct.2020.110927.

[65] Mangalathu S, Shin H, Choi E, et al. Explainable machine learning models for punching shear strength estimation of flat slabs without transverse reinforcement. *J Build Eng*. 2021;39(2):102300. doi:10.1016/j.jobe.2021.102300.

[66] Tastani SP, Pantazopoulou SJ. Experimental evaluation of FRP jackets in upgrading RC corroded columns with substandard detailing. *Eng Struct*. 2004;26(6):817–829. doi:10.1016/j.engstruct.2004.02.003.

[67] Jayaprakash J, Pournasiri E, De'nan F, et al. Effect of corrosion-damaged RC circular columns enveloped with hybrid and non-hybrid FRP under eccentric loading. *J Compos Mater*. 2015;49(18):2265–2283. doi:10.1177/0021998314545187.

[68] Chotickai P, Tongya P, Jantharaksa S. Performance of corroded rectangular RC columns strengthened with CFRP composite under eccentric loading. *Constr Build Mater*. 2021;268:121134. doi:10.1016/j.conbuildmat.2020.121134.

[69] Maaddawy TE. Post-repair performance of eccentrically loaded RC columns wrapped with CFRP composites. *Cement Concr Compos*. 2008;30(9):822–830. doi:10.1016/j.cemconcomp.2008.06.009.

[70] Radhi MS, Gorgis IN, Hassan MS. Analytical analysis of design-oriented models for forecasting the performance of CFRP-confined corrosion-affected concrete columns. *Constr Build Mater*. 2021;313(6–7):125491. doi:10.1016/j.conbuildmat.2021.125491.

[71] Nematzadeh M, Mousavimehr M, Shayanfar J, et al. Eccentric compressive behavior of steel fiber-reinforced RC columns strengthened with CFRP wraps: experimental investigation and analytical modeling. *Eng Struct*. 2021;226(2):111389. doi:10.1016/j.engstruct.2020.111389.

[72] Shaikh FUA, Alishahi R. Behaviour of CFRP wrapped RC square columns under eccentric compressive loading. *Structures*. 2019;20(4):309–323. doi:10.1016/j.istruc.2019.04.012.

[73] Bai YL, Dai JG, Teng JG. Buckling of steel reinforcing bars in FRP-confined RC columns: an experimental study. *Constr Build Mater*. 2017;140(2):403–415. doi:10.1016/j.conbuildmat.2017.02.149.

[74] Shan JL. *Studies on the Mechanical Properties of Corroded Reinforced Concrete Columns Confined with CFRP*. Dissertation. Harbin Institute of Technology; 2014 (in Chinese).

[75] Yu AM. *Calculation Method of Compressive Properties and Bearing Capacity of Damaged Reinforced Concrete Columns Reinforced by FRP Strip*. Dissertation. Zhenzhou University; 2018 (in Chinese).

[76] Li CC, Yu AM, Gao DY, et al. Experimental study on axial compression of corroded reinforced concrete columns strengthened with FRP strips under erosion environment. *Acta Mater Compos Sin*. 2020;37(8):2015–2028.

[77] Wen SC. Study on crushing resistance performance of corroded reinforced concrete columns confined by fiber reinforced polymer. *Hongshui River*. 2013;32(5):36–39 (in Chinese).

[78] Chen WX. *Experimental Study on the Mechanical Properties of FRP Reinforcement Corroded Reinforced Concrete Column and Corroded Steel Bar Buckling Characteristics*. Dissertation. Shenzhen University; 2015 (in Chinese).

[79] Gao J. *Performance Investigation of FRP Strengthened Reinforced Concrete Circular Columns Under Axial Compressive Loading*. Dissertation. Zhenzhou University; 2014 (in Chinese).

[80] Baak M, Koopman R, Snoek H, et al. A new correlation coefficient between categorical, ordinal and interval variables with Pearson characteristics. *Comput Stat Data Anal*. 2020;152(2):107043. doi:10.1016/j.csda.2020.107043.

[81] Lundberg SM, Erion GG, Lee SI, et al. Consistent individualized feature attribution for tree ensembles. ArXiv Preprint ArXiv:180203888; 2018.

[82] Du YG, Clark LA, Chan AHC. Residual capacity of corroded reinforcing bars. *Mag Concr Res*. 2005;57(3):135–147. doi:10.1680/macr.2005.57.3.135.

21425005-20   BER Open: Int. J. Bridge Eng., Manage. Res.

BER Open: Int. J. Bridge Eng., Manage. Res., 2024, 2(1): 21425005