

Title	Prediction of load-bearing capacity of FRP-steel composite tubed concrete columns: Using explainable machine learning model with limited data
Authors	Liu, Xiaoyang;Sun, Guozheng;Ju, Ruiqing;Li, Jing;Li, Zili;Jiang, Yali;Zhao, Kai;Zhang, Ye;Jing, Yucai;Yang, Guotao
Publication date	2024-12-03
Original Citation	Liu, X., Sun, G., Ju, R., Li, J., Li, Z., Jiang, Y., Zhao, K., Zhang, Y., Jing, Y. and Yang, G. (2024) 'Prediction of load-bearing capacity of FRP-steel composite tubed concrete columns: Using explainable machine learning model with limited data', Structures, 71, 107890 (12pp). https://doi.org/10.1016/j.istruc.2024.107890
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1016/j.istruc.2024.107890
Rights	© 2024, Institution of Structural Engineers. Published by Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies. This manuscript version is made available under the CC BY-NC-ND 4.0 license. - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-10-09 16:00:26
Item downloaded from	https://hdl.handle.net/10468/16734



UCC

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Prediction of load-bearing capacity of FRP-steel composite tubed concrete columns: using explainable machine learning model with limited data

Xiaoyang Liu^a, Guozheng Sun^a, Ruiqing Ju^a, Jing Li^a, Zili Li^b, Yali Jiang^c, Kai Zhao^{b,*}, Ye Zhang^a, Yucai Jing^d, Guotao Yang^a

^a *School of Civil Engineering, Qingdao University of Technology, Qingdao 266033, China*

^b *School of Engineering and Architecture, University College Cork, Cork T12 HW58, Ireland*

^c *College of Transportation Engineering, Dalian Maritime University, Dalian 116026, China.*

^d *Shandong Hi-Speed Road & Bridge Group Co., Ltd., Jinan 250014, China*

*Corresponding author. E-mail address: kzhao@ucc.ie (K. Zhao).

Abstract

The complex interaction mechanism between the FRP jacket, steel tube, and confined concrete in the FRP-steel composite tubed concrete (F-STC) column makes the prediction of its mechanical behaviour a challenging task. This paper specifically investigates the application of machine learning models to predict the load-bearing capacity of F-STC columns. Since relatively few experimental works are performed on F-STC columns, the database established based on the test results from previous research contains only 69 data samples. In order to circumvent the training difficulties caused by the relatively small database, the application of the Gaussian process regression (GPR) model is trialled in this study. To make the predictions of the GPR model explainable, the Shapley additive explanations (SHAP) approach is incorporated with the GPR model in this study. Besides, four contrasting prediction models based on artificial neural network (ANN), support vector regression (SVR), decision tree (DT), and random forest (RF), are also proposed. The k-fold validation and test results indicate that the GPR model provides strong potential in predicting the load-bearing capacity of F-STC columns with high prediction accuracy and generalisation capability. Besides, 95% and 99% confidence intervals obtained by the GPR model are provided to show the uncertainty of the prediction results. Furthermore, the effect of the database size on the prediction performance of GPR, ANN, and SVR models is further examined by gradually reducing the number of data samples, and the comparisons illustrate the superiority of the GPR model.

Keywords: F-STC columns, machine learning, gaussian process regression, Shapley additive explanation, small database.

1. Introduction

Steel tubed concrete (STC) columns have been experiencing a growing utilisation in civil engineering since their initial introduction in the 1980s [1]. As the name implies, STC columns possess an outer steel tube which supplies confinement to the internal concrete to improve the strength and ductility of the infilled concrete, making it attractive to large-span and high-rise structures. Different from concrete filled steel tube (CFST) columns [2-4], the outer steel tube in STC columns is discontinuous, as shown in Fig. 1, with two girth gaps at the ends to ease the construction process of the beam-column joints. More importantly, since no direct axial load is applied to the outer steel tube in STC columns, the steel tube could be thinner compared to the steel tube in CFST columns [5-7]. Several studies have proved that STC columns exhibit superior mechanical performance over the corresponding CFST columns [5, 8, 9]. Although the outward local buckling of the steel tube in STC columns can be efficiently suppressed, local buckling failures can still occur around the mid-height region [9, 10]. Besides, for structural members encased with steel jackets, such as STC and CFST columns, extensive precautions should be enforced to prevent corrosion issues.

Fibre reinforced polymer (FRP) is characterised as a promising engineering material, as it exhibits a high strength-to-weight ratio, strong resistance to impact, fatigue, and corrosion [11-14]. Therefore, numerous studies have been carried out to investigate the utilisation of FRP for confining concrete [15-23]. However, FRP exhibits a brittle post-peak response and usually fails by sudden rupture. Besides, the process of wrapping FRP sheets to concrete columns is complicated as the concrete surface should be well cleaned [24]. Nevertheless, the utilisation of a large amount of FRP could render the confined concrete column economically inefficient. To make use of the advantages of both steel and FRP jackets and overcome the disadvantages of each, the FRP-steel composite tubed concrete (F-STC) column was proposed, in which the FRP sheet is wrapped around the steel tube to form a composite confinement tube, contributing to a relatively ductile post-peak response. The addition of an FRP sheet also effectively relieves the buckling and corrosion concerns of the steel tube in F-STC columns, enhancing the load-bearing capacity and ductility of the composite columns [25-28].

Many efforts have been made to investigate the performance of FRP confined CFST (FCCFST) columns [16-20, 29, 30], based on which several prediction models that are usually categorised

into design-oriented models [21, 31] and analysis-oriented models [19, 32-34] have been developed. However, as the axial load applied on the FRP-steel composite tube in F-STC columns is transferred through interface friction between the core concrete and the composite tube, the performance of F-STC columns differs considerably from FCCFST columns [26, 28], and the models developed for FRP cannot be used for F-STC columns.

Several studies have been conducted on the performance of F-STC columns. In the work of Liu et al. [35], the performance of F-STC columns was initially studied, and the obtained results indicated that the confining effect was significantly improved by adding CFRP sheets, and the F-STC columns exhibited excellent load-bearing capacity and ductility. In their later work [36], an experimental work consisting of 25 specimens was further performed to study the effects of the number of CFRP layers, diameter-to-thickness ratio of the steel tube, as well as the concrete strength on the performance of F-STC columns. Based on the obtained test results, the performance of several well-known models for confined concrete in predicting the axial load-bearing capacity of F-STC columns was evaluated, and it was concluded that even the model with the highest accuracy provided conservative results for F-STC columns. Therefore, the development of models specifically for F-STC columns is essential. Xu et al. [37] developed a design-oriented model for F-STC columns based on existing experimental results, as F-STC columns with varied configurations exhibit different load-strain behaviour, which is basically categorised into three patterns (i.e. elastic-plastic, bi-linear, and linear-parabolic), different design formulas should be established for these three cases. Wang et al. [38] experimentally investigated the performance of F-STC columns, FRP concrete confined (FCC) columns, and STC columns, and the confinement mechanism of F-STC columns was investigated through the comparison between the FRP confinement and steel confinement. The obtained results revealed that the confinement mechanism of F-STC columns should be categorised into three phases, which were subsequently dominated by steel confinement, FRP-steel composite confinement, and FRP confinement during the loading process. Therefore, the confinement mechanism in the F-STC columns is much more complex than that in FCC columns and STC columns, inevitably causing more difficulties in the prediction of their mechanical behaviour.

Recently, machine learning methods have been successfully utilised for solving various civil engineering problems, especially the complex nonlinear problems [39-47]. Pham et al. [48]

developed an artificial neural network (ANN) model for obtaining the load-bearing capacity of FRP confined concrete columns, the obtained results revealed that the ANN model outperformed the conventional empirical models. ANN models have also been successfully utilised in the works of Cascardi et al. [49] and Ouni and Raza [50]. Although the proposed ANN models are able to provide results with high accuracy, the development of the ANN model usually requires a large database. In the work of Mozumder et al. [51], support vector regression (SVR) models were developed to predict the strength of FRP confined concrete based on a dataset containing 238 experimental data samples, and the comparison between SVR models and ANN models indicated that the developed SVR models were able to achieve better prediction results for this problem. Mansouri et al. [52] investigated the application of the ANN model, M5Tree model, multivariate adaptive regression splines model, and adaptive neuro-fuzzy inference system model to predict the ultimate conditions of FRP confined concrete, for which more than 1000 data samples were utilised for the training process. It was concluded that the four machine learning models exhibited superior performance than the majority of conventional models.

Although the machine learning models have outstanding prediction performance, the black-box nature of their prediction process makes them difficult to interpret, thereby limiting their application. To make the machine learning model explainable, several explainability approaches have been proposed, such as Shapley additive explanations (SHAP) [53] and interpretable mimic learning (IML) [54]. Wang et al. [55] proposed an extreme gradient boosting model for predicting bridge scour risk, and the SHAP approach was used to interpret the prediction results and aid in human understanding of how the predictions are made. Comparisons are made between the SHAP results and the results obtained from engineering perspectives, the obtained results indicated that the SHAP method could provide explanations that are neglected from human perspectives. In the work of Zhou et al. [56], a random forest (RF) model incorporated with SHAP was developed to predict the damage stage of bridge piers, providing clear explanations for the prediction results at both local and global levels. More importantly, the obtained results also demonstrated the superiority of the explainable machine learning model over conventional empirical models.

Despite the attractive prediction performance of machine learning models, they have not been applied to the prediction of F-STC columns due to the relatively limited number of existing experimental data samples, which causes great difficulties in model training. In civil engineering,

the experimental tests are usually costly and time-consuming, resulting in the scarcity of big data, for which the effective utilisation of limited data is of great importance. Gaussian process regression (GPR), which is commonly recognised as a promising machine learning model that can be well trained with a small database [57-59], provides a potential solution. In the present study, a GPR model incorporated with SHAP interpretation is proposed to predict the load-bearing capacity of F-STC columns based on a database containing 69 test samples obtained from existing literature. The use of the SHAP approach transforms this prediction problem into a white-box process. Besides, the application of ANN, SVR, RF, and decision tree (DT) to this prediction problem is also investigated, and the superiority of the GPR model is demonstrated through the comparison between these five models. In the following context, the experimental database utilised in this paper is introduced in Section 2. Section 3 describes the details of the machine learning models. The obtained results are presented and discussed in Section 4. A brief conclusion is provided in Section 5.

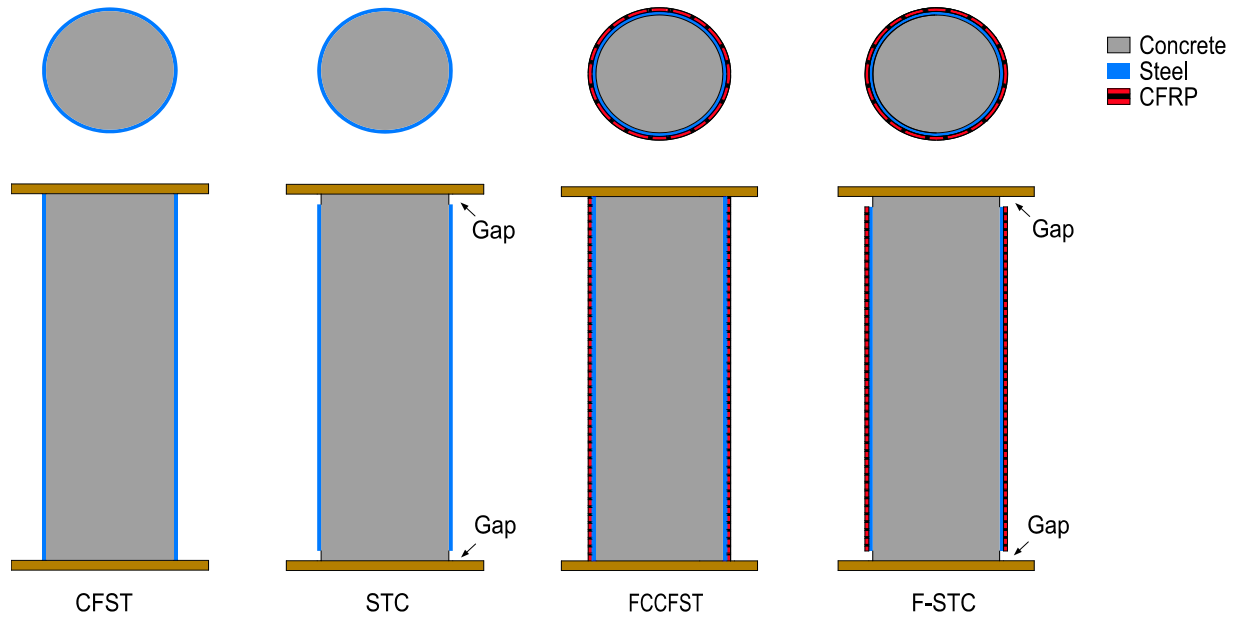


Fig. 1. Illustration of various composite columns.

2. Experimental database

In this study, 69 test results of the F-STC columns are collected from existing studies [35-38, 60-62], based on which an experimental database for the development of the machine learning models is established. The details of the test results are presented in Table 1, in which D and L represent the diameter and length of the columns; E_{frp} and ε_{frp} are the elastic modulus and rupture strain of the FRP sheet, respectively, and the values of ε_{frp} are taken from the measured rupture strain of the test. t_{frp} and t_s are the thicknesses of the FRP sheet and steel tube, respectively; f_y is the yield strength of the steel tube, and f_{cu} is the cubic compressive strength of the concrete; N_p represents the bearing capacity of the columns.

Since most F-STC columns are not reinforced with steel bars, the specimens containing steel bars are excluded from the database. All the specimens in the database meet the requirement that the ratio L/D is less than 3, and the values of the ratio D/t_s range between 58.7 to 172. In this study, the load-bearing capacity N_p is taken as the output of the machine learning models, whereas the other parameters in Table 1 are taken as the inputs. In order to improve the prediction accuracy, the inputs and outputs are normalised within the range of 0 to 1.

Table 1 Experimental results of existing studies.

Refs	D (mm)	L (mm)	f_{cu} (MPa)	f_y (MPa)	t_s (mm)	E_{frp} (GPa)	ε_{frp} ($\mu\varepsilon$)	t_{frp} (mm)	N_p (kN)
[38]	172	510	68.7	189	1	245	5120	0.167	1820
	172	510	68.7	189	1	245	5580	0.334	2294
	172	510	68.7	189	1	245	6710	0.501	2834
	174	510	68.7	190	2	245	6350	0.167	2148
	174	510	68.7	190	2	245	6980	0.334	2607
	174	510	68.7	190	2	245	7520	0.501	3081
	176	510	68.7	199	3	245	7270	0.167	2323
	176	510	68.7	199	3	245	6300	0.334	2659
	176	510	68.7	199	3	245	6250	0.501	3183
[36]	260	780	72.5	299	2	246	10862	0.334	4494
	260	780	72.5	299	2	246	13257	0.334	4747
	260	780	72.5	299	2	246	11863	0.334	4581
	260	780	72.5	299	2	246	7100	0.668	5567
	260	780	72.5	299	2	246	8718	0.668	5419
	260	780	72.5	299	2	246	15355	0.668	5558

[35]	200	600	105	299	2	246	6332	0.334	4127
	200	600	105	299	2	246	9392	0.334	4230
	200	600	105	299	2	246	9068	0.334	4212
	200	600	105	299	2	246	7442	0.668	4523
	200	600	105	299	2	246	11792	0.668	4939
	200	600	105	299	2	246	10566	0.668	4642
	260	780	105	299	2	246	7101	0.334	6567
	260	780	105	299	2	246	7453	0.334	6568
	260	780	105	299	2	246	7161	0.334	6514
	260	780	105	299	2	246	5778	0.668	6269
	260	780	105	299	2	246	6401	0.668	6332
	260	780	105	299	2	246	9408	0.668	6374
	260	780	105	299	2	246	8377	0.835	6927
	200	600	57.1	264.3	2	235	9315	0.334	2607
	200	600	57.1	264.3	2	235	10975	0.668	3456
	200	600	57.1	264.3	2	235	11338	0.668	3083
	200	600	57.1	264.3	2	235	9723	0.668	3327
	200	600	66.8	264.3	2	235	9182	0.668	3846
	200	600	66.8	264.3	2	235	8957	0.668	3469
	200	600	66.8	264.3	2	235	10513	0.668	3776
	260	780	57.1	264.3	2	235	8393	0.334	3360
	260	780	57.1	264.3	2	235	6799	0.668	4780
	260	780	57.1	264.3	2	235	4800	0.668	4050
	260	780	57.1	264.3	2	235	10157	0.668	4757
	260	780	66.8	264.3	2	235	3546	0.334	4321
	260	780	66.8	264.3	2	235	7278	0.668	4711
	260	780	66.8	264.3	2	235	9668	0.668	5374
	260	780	66.8	264.3	2	235	9869	0.668	5131
[62]	172	510	68.7	188	1	245	8200	0.167	1787
	172	510	68.7	188	1	245	6300	0.334	2289
	172	510	68.7	188	1	245	5200	0.501	2574
	174	510	68.7	192	2	245	6800	0.167	2090
	174	510	68.7	192	2	245	7000	0.334	2602
	174	510	68.7	192	2	245	7400	0.501	3076
	174	510	81.1	192	2	245	6300	0.334	2846
	174	510	81.1	192	2	245	7000	0.501	3316
	172	510	68.7	188	1	80	6900	0.23	1767
	172	510	68.7	188	1	80	6900	0.46	1714
	174	510	68.7	192	2	80	8300	0.23	1921
	174	510	68.7	192	2	80	9900	0.46	2182
[61]	200	600	57.1	264.3	2	235	9278	0.334	2633
	200	600	57.1	264.3	2	235	9004	0.334	2755
	200	600	66.8	264.3	2	235	9169	0.334	3048
	200	600	66.8	264.3	2	235	10111	0.334	3017
	260	780	57.1	264.3	2	235	8629	0.334	3736

	260	780	57.1	264.3	2	235	8463	0.334	3610
	260	780	66.8	264.3	2	235	9857	0.334	4345
	260	780	66.8	264.3	2	235	2245	0.334	3847
[60]	260	780	76.5	299	2	246	10169	0.334	4788
	260	780	76.5	299	2	246	10664	0.334	4779
	260	780	76.5	299	2	246	9212	0.334	4677
	260	780	76.5	299	2	246	10828	0.668	5890
	260	780	76.5	299	2	246	8397	0.668	5596
	260	780	76.5	299	2	246	8602	0.668	5340

3. Model development

In this section, based on the database discussed above, five machine learning models, including GPR, ANN, SVR, DT and RF, are developed for predicting the load-bearing capacity of F-STC columns.

3.1 Gaussian process regression model

The GPR model has the advantages of providing probabilistic output, adaptively acquiring hyperparameters, as well as being easily implemented, making it a powerful tool in regression analysis [63]. Instead of predicting one single target value, the GPR model is able to provide a probability distribution of the target based on the mean and variance of a Gaussian process, as below:

$$f(\mathbf{x}) \square G(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \quad (1)$$

where $\mathbf{x}$ and $\mathbf{x}'$ are input vectors, $m(\mathbf{x})$ is the mean function which is usually set as zero mean, and $k(\mathbf{x}, \mathbf{x}')$ is the covariance function.

Based on Eq. (1), the regression problem can be described as:

$$\mathbf{y} = f(\mathbf{x}) + \boldsymbol{\varepsilon} \quad (2)$$

where $\mathbf{y}$ represents the vector containing the observed values, and $\boldsymbol{\varepsilon}$ represents the noise terms, assumed to follow a Gaussian distribution, defined as:

$$\boldsymbol{\varepsilon} \square N(\mathbf{0}, \boldsymbol{\sigma}_n^2) \quad (3)$$

where $\boldsymbol{\sigma}_n^2$ represents the variance. Therefore, the prior distribution of $\mathbf{y}$ can be expressed as:

$$\mathbf{y} \sim N(\mathbf{0}, \mathbf{K}(X, X) + \sigma_n^2 \mathbf{I}_n) \quad (4)$$

where $\mathbf{I}_n$ is an identity matrix of dimension n . The joint distributions of the observed and predicted values can be described as:

$$\begin{bmatrix} \mathbf{y} \\ f^* \end{bmatrix} \sim N\left(0, \begin{bmatrix} \mathbf{K}(X, X) + \sigma_n^2 \mathbf{I}_n & \mathbf{K}(X, x^*) \\ \mathbf{K}(x^*, X) & k(x^*, x^*) \end{bmatrix}\right) \quad (5)$$

where f^* is the predicted value, $\mathbf{K}(X, X)$ is a $n \times n$ order covariance matrix, in which the element $k_{ij} = k(x_i, x_j)$ quantifies the correlation between x_i and x_j , $\mathbf{K}(X, x^*)$ is a $n \times 1$ order covariance matrix between the testing sample x^* and all training samples set X , and $k(x^*, x^*)$ represents the covariance matrix of the testing samples.

Therefore, the posterior distribution of the predicted value f^* is obtained as below:

$$f^* | X, \mathbf{y}, x^* \sim N(\bar{f}^*, \text{cov}(f^*)) \quad (6)$$

where the predicted value $\bar{f}^*$ is expressed as:

$$\bar{f}^* = \mathbf{K}(x^*, X) [\mathbf{K}(X, X) + \sigma_n^2 \mathbf{I}_n]^{-1} \mathbf{y} \quad (7)$$

and the corresponding variance is obtained as:

$$\text{cov}(f^*) = k(x^*, x^*) - \mathbf{K}(x^*, X) [\mathbf{K}(X, X) + \sigma_n^2 \mathbf{I}_n]^{-1} \mathbf{K}(X, x^*) \quad (8)$$

3.2 Artificial neural network model

In this section, the ANN model, which is commonly recognised as one of the most effective prediction models for highly non-linear problems, is proposed based on the biological neuron system. An ANN model usually contains an input layer, an output layer, and several hidden layers, each layer is given a specific number of neurons, and the neurons in the preceding and succeeding layers are interconnected to form a neural network. Fig. 2 shows a typical ANN, as can be seen that, once the inputs are received in the input layer, they are processed in the neurons and fed forward to the output layer. During the process, once the inputs of a neuron are received, they are weighted and summed, as below:

$$s = \sum_{j=1}^n w_{ij} p_j + b_i \quad (9)$$

where p_j is the input received from the j^{th} neuron of the preceding layer, w_{ij} and b_i are the weight and bias for the i^{th} neuron of the processing layer.

The weighted sum of the inputs is then applied to an activation function to obtain the output of the neuron. The commonly used activation functions include the rectified linear unit (ReLU), hyperbolic tangent (tanh), and sigmoid function, the expressions of the three activation functions are shown as:

$$\text{ReLU} = \max(0, s) \quad (10)$$

$$\tanh = \frac{e^s - e^{-s}}{e^s + e^{-s}} \quad (11)$$

$$\text{sigmoid} = \frac{1}{1 + e^{-s}} \quad (12)$$

Once the output result is derived at the final layer, the difference between the target and predicted outputs is turned into the loss function. For the regression problem, the mean squared error (MSE) is typically chosen as the loss function, the expression of which is shown below:

$$\text{MSE} = \frac{1}{m} \sum_{i=1}^m (t_i - p_i)^2 \quad (13)$$

where t_i is the target value from the experimental database, p_i is the predicted value of the model, and m is the number of data samples for training.

Once the value of the loss function is obtained, the back propagation process is performed to minimise the value of the loss function through the optimisation of the weights and biases for each neuron. This process repeats until the maximum number of iterations is reached. For the ANN model, the prediction performance is significantly affected by its hyperparameters, such as the learning rate that determines the step size of the optimisation, and the regularisation parameter that controls the complexity of the model. Therefore, a grid search is implemented in this study to obtain the optimal hyperparameters that lead to the best performance of the ANN model. In this

paper, the hyperparameters include the number of hidden layers, the number of neurons in each layer, the learning rate, activation functions, batch size, and optimisation solvers, the ranges of which are outlined in Table 2. It should be noted that the hidden layer structure shown in Table 2 represents both the number of hidden layers and the number of neurons in each layer, for example, (10,) refers to an ANN model with one hidden layer containing 10 neurons, and (30, 30, 30) refers to an ANN model with three hidden layers, each containing 30 neurons.

Table 2 The ranges of hyperparameters.

Model	Hyperparameters	Range
ANN	learning rate	[0.0001, 0.001, 0.01, 0.1]
	activation function	['sigmoid', 'ReLU', 'tanh']
	solver	['L-BFGS', 'SGD', 'ADAM']
	batch size	[5, 10, 20, 30, 40]
	hidden layer structure	[(10,), (30,), (50,), (10,10), (30,30), (50,50), (10,10,10), (30,30,30), (50,50,50)]
DT	maximum features	['all', 'sqrt', 'log2']
	minimum sample leaf	[2, 3, 4, 5, 6, 7, 8, 9, 10]
	minimum sample split	[2, 3, 4, 5, 6, 7, 8, 9, 10]
	maximum depth	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15]
RF	number of trees	[5, 10, 15, 20, 25, 30]
SVR	regularisation parameter	[0.001, 0.01, 0.1, 1, 10, 100]
	kernel function	['linear', 'RBF', 'sigmoid']
	kernel function parameter	[0.001, 0.01, 0.1, 1, 10, 100]

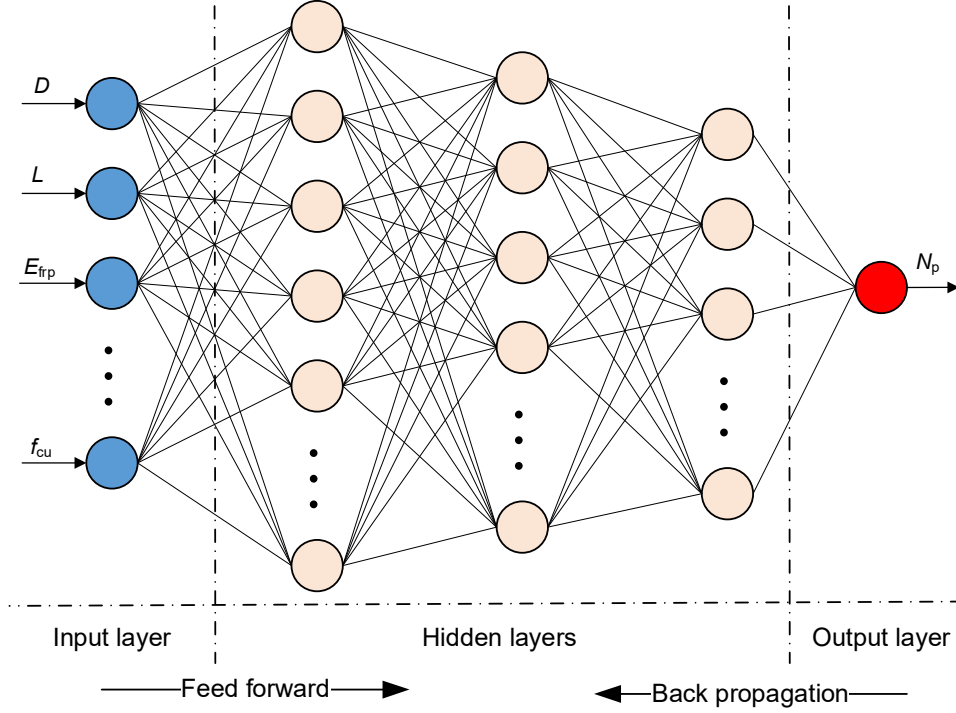


Fig. 2. Illustration of ANN model.

3.3 Decision tree and random forest models

The DT model is developed based on a tree-like structure which is composed of nodes and branches. The leaf nodes at the bottom of the tree structure represent the predicted values, while the internal nodes represent the attribute tests. Based on the result of an attribute test, two data subsets are generated and assigned to the lower internal nodes. The path from the upper nodes to the lower node is created through the branches.

In order to make a precise prediction, the attribute and split point of a node should be properly determined. In a DT regression model, each node is usually associated with three branches, two of which are assigned to extend below the node. The MSE criterion is usually used to assess the selection of attributes and split points, as below:

$$MSE = \frac{1}{n_L} \sum_{i=1}^{n_L} (y_i - \bar{y}_L)^2 + \frac{1}{n_R} \sum_{i=1}^{n_R} (y_i - \bar{y}_R)^2 \quad (14)$$

where n_L and n_M represent the number of data samples of the left subsets and right subsets for a node, respectively, $\bar{y}_L$ and $\bar{y}_R$ are the average target values of the left subsets and right subsets, respectively.

The optimal attribute a^* and split point s^* for a node d can be obtained by minimising the value of MSE:

$$(a^*, s^*) = \arg \min_{a,s} \text{MSE}(d) \quad (15)$$

The generation of new nodes proceeds until the stopping criterion is met. The hyperparameters of the DT model are shown in Table 2, and their values are obtained through a grid search.

As for the RF model, it is developed by combining multiple decision trees, each generated from a random subset of the training database and also a random subset of features. The decision trees in a random forest model are built independently, and they provide output results individually, as shown in Fig. 3. The prediction result of a RF model is achieved based on the combination of the output results of the component decision trees. Hence, compared to DT models, RF models have an additional hyperparameter which is the number of decision trees. The range of the hyperparameter for this study is shown in Table 2.

The output of a RF model is usually taken as the average of the outputs of the trees, as below:

$$f = \frac{1}{T} \sum_{t=1}^T f_t(\mathbf{x}) \quad (16)$$

where T is the number of decision trees in the RF model, f_t is the output of the t^{th} tree trained on a subset of the training dataset $\mathbf{x}$.

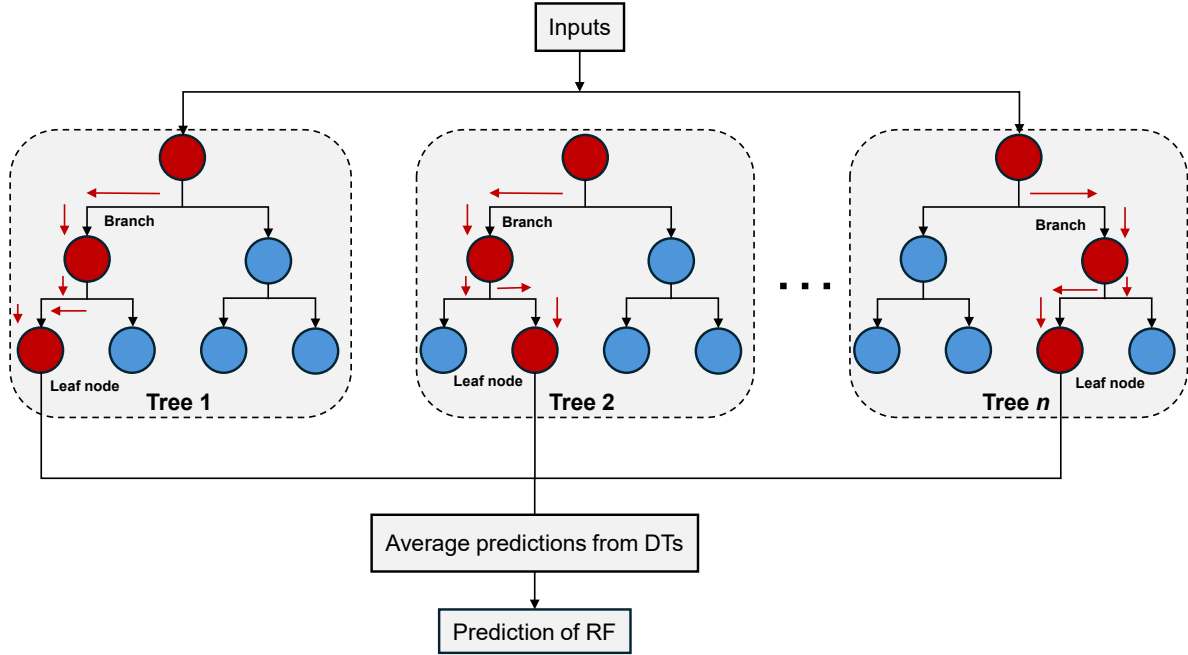


Fig. 3. Illustration of DT and RF models.

3.4 Support vector regression model

SVR solves complex non-linear prediction problems by mapping input features into higher-dimensional spaces, in which a hyperplane that aims to fit the targets inside the margin boundaries is established, as shown in Fig. 4. The expression of a hyperplane is shown as:

$$f(x) = w\phi(x) + b \quad (17)$$

where w is a weight term, b is a bias term, and $\phi(x)$ represents the non-linear mapping function.

In order to minimise the distance from all data points to the hyperplane, the values of w and b are determined by minimising the risk function R , as below:

$$\min : R = \frac{1}{2} \|w\|^2 + C \sum_i^m L_\varepsilon(y_i, f(x_i)) \quad (18)$$

where C is the regularisation parameter, L_ε is the ε insensitive loss function, the expression of which is shown as:

$$L_{\varepsilon}(y, f(x)) = \begin{cases} |y - f(x)| - \varepsilon, & \text{if } |y - f(x)| > \varepsilon \\ 0, & \text{if } |y - f(x)| \leq \varepsilon \end{cases} \quad (19)$$

where ε is the distance from the margin boundary (i.e. support vector) to the hyperplane.

As can be seen that the loss function is only activated for the outlier data points, hence, the risk function R can be transformed as:

$$\min : R = \frac{1}{2} \|w\|^2 + C \sum_i^m L_{\varepsilon}(\xi_i + \xi_i^*) \quad (20)$$

where ξ_i and ξ_i^* are two slack variables, which represent the distances from the outlier data points to the hyperplane, with ξ_i representing the distance for points above the margin boundary, and ξ_i^* representing the distance for points below the margin boundary. Besides, constraints are set for this optimisation problem, as shown below:

$$\text{subject to: } \begin{cases} y_i - f(x_i) \leq \varepsilon + \xi_i \\ f(x_i) - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, i = 1, \dots, m \end{cases} \quad (21)$$

The Lagrange method is employed to solve this optimisation problem, leading to the formulation of a dual problem, as shown below:

$$f(x) = \sum_{i=1}^{M_{sv}} (a_i - a_i^*) K(x_i, x) + b \quad (22)$$

where a_i and a_i^* are Lagrange multipliers, M_{sv} is the number of support vectors, and the kernel function $K(x_i, x)$ is expressed as:

$$K(x_i, x) = \varphi(x_i) \times \varphi(x) \quad (23)$$

In this paper, the hyperparameters of the SVR model are determined based on a grid search, and the ranges of which are shown in Table 2.

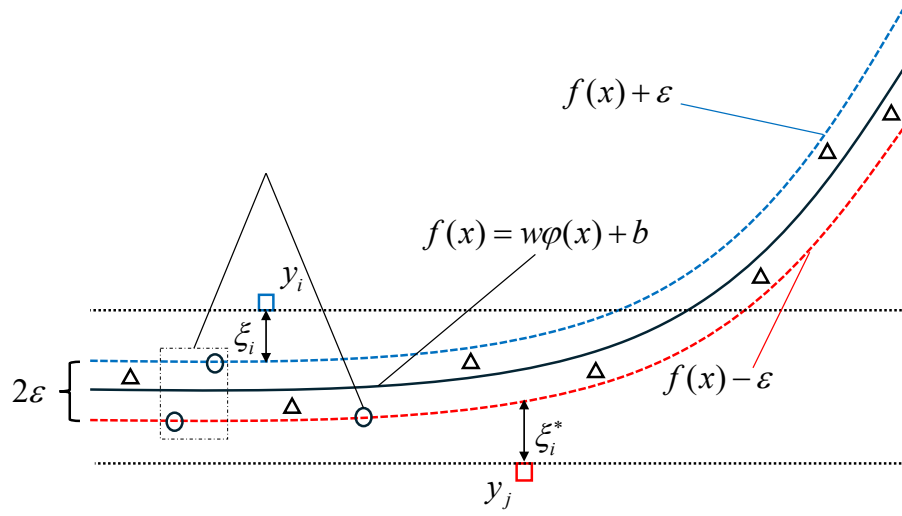


Fig. 4. Illustration of SVR model.

3.5 Model measures

In this study, the performance of the developed machine learning models is measured based on the mean square error (MSE), relative root mean square error (RRMSE), and the coefficient of determination (R^2), and the expressions of the RRMSE and R^2 are shown below:

$$\text{RRMSE} = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (t_i - p_i)^2}}{\bar{t}} \quad (24)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (t_i - p_i)^2}{\sum_{i=1}^N (t_i - \bar{t})^2} \quad (25)$$

where N is the number of data samples, and $\bar{t}$ is the mean value of the targets.

Unlike the root mean square error (RMSE), the normalisation introduced in RRMSE makes the error measure independent of the scale of the target values. The machine learning model exhibiting superior prediction capability tends to achieve a RRMSE value close to 0 and a R^2 value close to 1. Besides, k-fold cross-validation is implemented in this study to evaluate the generalisation capability of the developed models. Since the database is small, the training dataset is split into

three folds, and the models are trained three times with each fold utilised once as a validation set while the left two folds are utilised for training.

3.6 Shapley additive explanation

In this study, the SHAP approach is employed as a post-hoc explainability approach for the GPR model. Based on game theory, the SHAP approach is developed to interpret the black-box prediction process of machine learning models by providing the contribution of each feature to the prediction using SHAP values [53]. The explanation can be made for the predictions of the whole database at the global level, as well as for each individual prediction at the local level. The explanation is performed based on an explanation model, which is expressed as:

$$E(\mathbf{s}') = \phi_0 + \sum_{i=1}^{N_f} \phi_i s'_i \quad (26)$$

where $\mathbf{s}'$ is a vector containing the simplified inputs, which links the original inputs using a mapping function $\mathbf{x} = h(\mathbf{s}')$, N_f is the number of features, ϕ_0 is the base value, and ϕ_i is the SHAP value of the i^{th} feature.

The SHAP value of the i^{th} feature ϕ_i is determined considering local accuracy, missingness, and consistency, the expression of which is shown as follows:

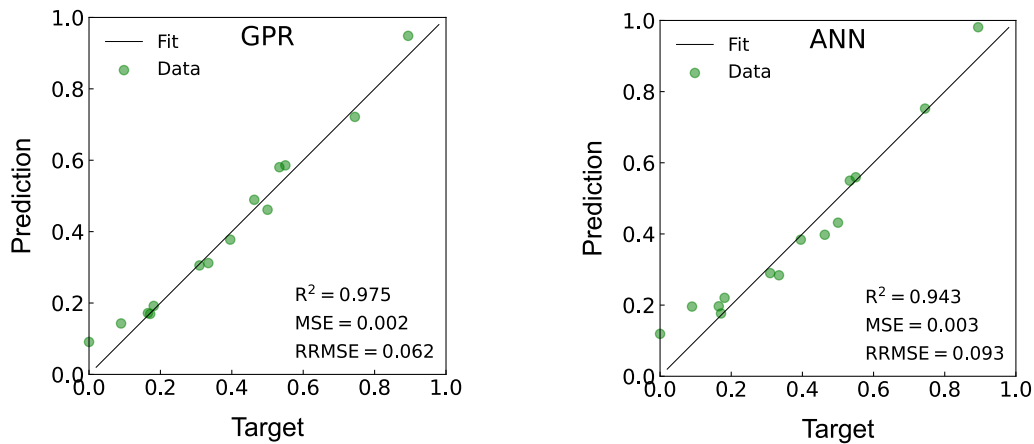
$$\phi_i(f, \mathbf{x}) = \sum_{\mathbf{z}' \subseteq \mathbf{s}'} \frac{|\mathbf{z}'|!(N_f - |\mathbf{z}'| - 1)!}{N_f!} [f(\mathbf{z}' \cup \{i\}) - f(\mathbf{z}')] \quad (27)$$

where f represent the GPR model for interpretation, $\mathbf{z}'$ is a subset of $\mathbf{s}'$, $|\mathbf{z}'|$ are the nonzero elements in $\mathbf{z}'$, $f(\mathbf{z}' \cup \{i\})$ and $f(\mathbf{z}')$ are the outputs of the GPR model with and without considering the i^{th} feature, respectively. A positive SHAP value for the i^{th} feature indicates a positive effect on the prediction result, with higher values indicating a greater effect. Conversely, a negative SHAP value indicates a negative effect, with lower values indicating a greater effect.

4. Results and discussion

4.1 Comparison between machine learning models

The machine learning models proposed in the last section are utilised herein for the prediction of the load-bearing capacity of F-STC columns based on the database shown in Table 1. Fig. 5 shows the prediction results of the machine learning models, where the abscissa represents the predicted results, while the ordinate denotes the experimental results. As expected, the GPR model achieves the highest value of R^2 , and the lowest values of MSE and RRMSE, indicating that it attains the highest goodness-of-fit and lowest prediction error, outperforming the other models. The SVR model exhibits a slightly better prediction performance than the ANN model, as the R^2 of the SVR model is 0.951, whereas the R^2 of the ANN model is 0.943. Although the values of MSE for the two models are the same, both being 0.003, the SVR model achieves a slightly lower value of RRMSE. As discussed in the last section, the values of hyperparameters are determined by a grid-search process. For the SVR model, the radial basis function (RBF) is chosen as the kernel function, with the kernel function parameter set to 0.01 and the regularisation parameter set to 100. The limited-memory broyden-fletcher-goldfarb-shanno algorithm, which is suitable for small databases, is chosen as the solver for the ANN model by the grid-search process. Since L-BFGS is a quasi-Newton optimisation method that utilises adaptive step size during the optimisation process, the learning rate is not required and, therefore, not set for the ANN model. Besides, one hidden layer containing ten neurons with the ReLU activation function is applied to the ANN model, and the batch size is set to 5.



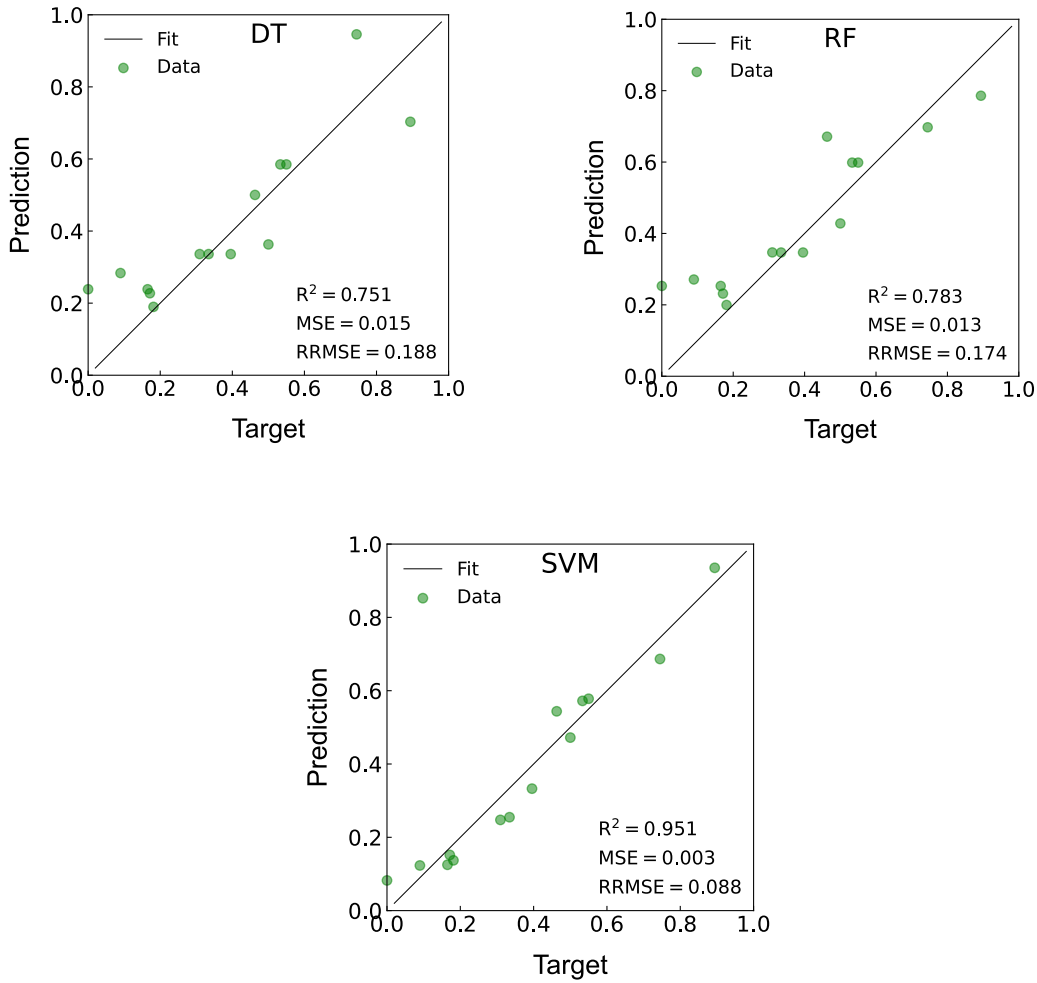


Fig. 5. Comparisons of R^2 , MSE, and RRMSE between machine learning models.

The values of R^2 of the DT model and RF model are 0.751 and 0.783, respectively, which are noticeably lower than those of the other three models. Correspondingly, both the DT model and RF model obtain higher values of MSE and RRMSE compared to the other models, demonstrating that more data is required for the development of these two models. It should be noted that the RF model exhibits slightly better prediction performance than the DT model, this should be attributed to the ensemble nature of the RF model, which combines multiple decision trees as explained in Section 3.3, hence, enhancing the prediction performance. Based on the grid-search process, the DT model with a maximum depth of 12 is chosen, where the minimum samples per leaf are set to 2, and the minimum samples for splitting are set to 4. The base-2 logarithm is selected to determine

the maximum number of features to consider for the best split at each node. Since eight features are considered in the database in this study, the maximum number of features is set to 3. The RF model is assigned the same number of minimum samples per leaf and minimum samples for splitting as the DT model. As the RF model consists of twenty decision trees, the maximum depth of each tree decreases to 4, and all eight features are considered for the split at each node. The obtained results demonstrate that the combination of predictions from multiple trees with limited depth and a larger number of splitting features leads to better prediction performance than the use of a single decision tree with deeper depth and fewer features.

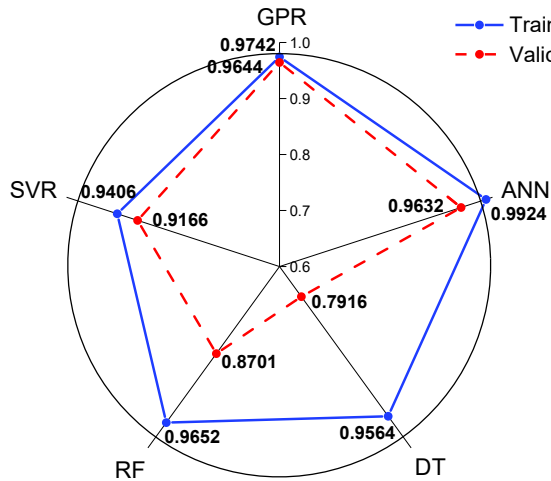
Table 3 presents the k-fold results of the machine learning models, and the average value and standard deviation of R^2 , MSE, and RRMSE are shown in Fig. 6. It can be observed that the GPR model performs consistently well across the three folds of the database, it achieves the smallest difference in R^2 between the training results and validation results, with average values of 0.974 and 0.964, respectively. Besides, the GPR model also achieves the smallest differences in terms of MSE and RRMSE between the training and validation results, indicating its robust performance and strong generalisation capability for this prediction problem. Compared to the GPR and ANN models, the SVR model achieves slightly lower values of R^2 and higher values of MSE and RRMSE in both the training and validation processes. As shown in Fig. 6(b), the standard deviation of R^2 for the validation of the SVR model is 0.0313, which is significantly higher than those of the GPR and ANN models. This is attributed to the overfitting of the SVR model in the second fold, where the value of R^2 equals to 0.875, as shown in Table 3.

Although the ANN model achieves the highest values of R^2 and lowest values of MSE and RRMSE during the training process, with average values of 0.9924, 0.0005, and 0.0542, respectively, the differences between the training and validation results are larger than those of the GPR model and SVR model, especially the MSE and RRMSE, as shown in Fig. 6 (c) and (e). This is because the ANN model is slightly overfitting due to the small amount of the training data. Besides, it can be observed from Fig. 6(f) that, compared to the GPR model, the ANN and SVR models obtain larger values of standard deviation for the validation set, indicating that the GPR model has superior generalisation capability. Furthermore, it can be seen that the DT and RF models achieve acceptable prediction results during the training process but poor results on the validation sets, indicating that the two models are heavily overfitted. Moreover, the RF model achieves slightly

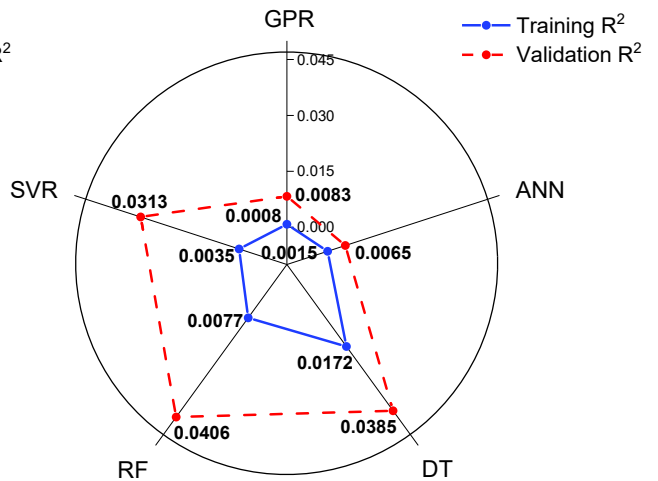
better generalisation performance compared to the DT model, this should be attributed to the ensemble nature of the RF model, which combines predictions from multiple trees, reducing the overfitting of a single DT model.

Table 3 k-fold results of the machine learning models.

Model	Fold	Training			Validation		
		R ²	MSE	RRMSE	R ²	MSE	RRMSE
GPR	1	0.975	0.002	0.099	0.965	0.002	0.119
	2	0.975	0.002	0.104	0.954	0.003	0.124
	3	0.973	0.002	0.099	0.974	0.002	0.108
ANN	1	0.990	0.001	0.056	0.954	0.004	0.156
	2	0.994	0.001	0.053	0.968	0.002	0.088
	3	0.992	0.001	0.054	0.966	0.002	0.118
DT	1	0.949	0.004	0.153	0.804	0.012	0.237
	2	0.980	0.001	0.079	0.832	0.015	0.307
	3	0.940	0.005	0.157	0.740	0.016	0.305
RF	1	0.967	0.003	0.116	0.894	0.007	0.198
	2	0.974	0.002	0.101	0.903	0.006	0.197
	3	0.955	0.003	0.132	0.813	0.016	0.270
SVR	1	0.936	0.004	0.164	0.951	0.0039	0.125
	2	0.945	0.004	0.147	0.875	0.0066	0.215
	3	0.941	0.004	0.146	0.924	0.0057	0.188



(a)



(b)

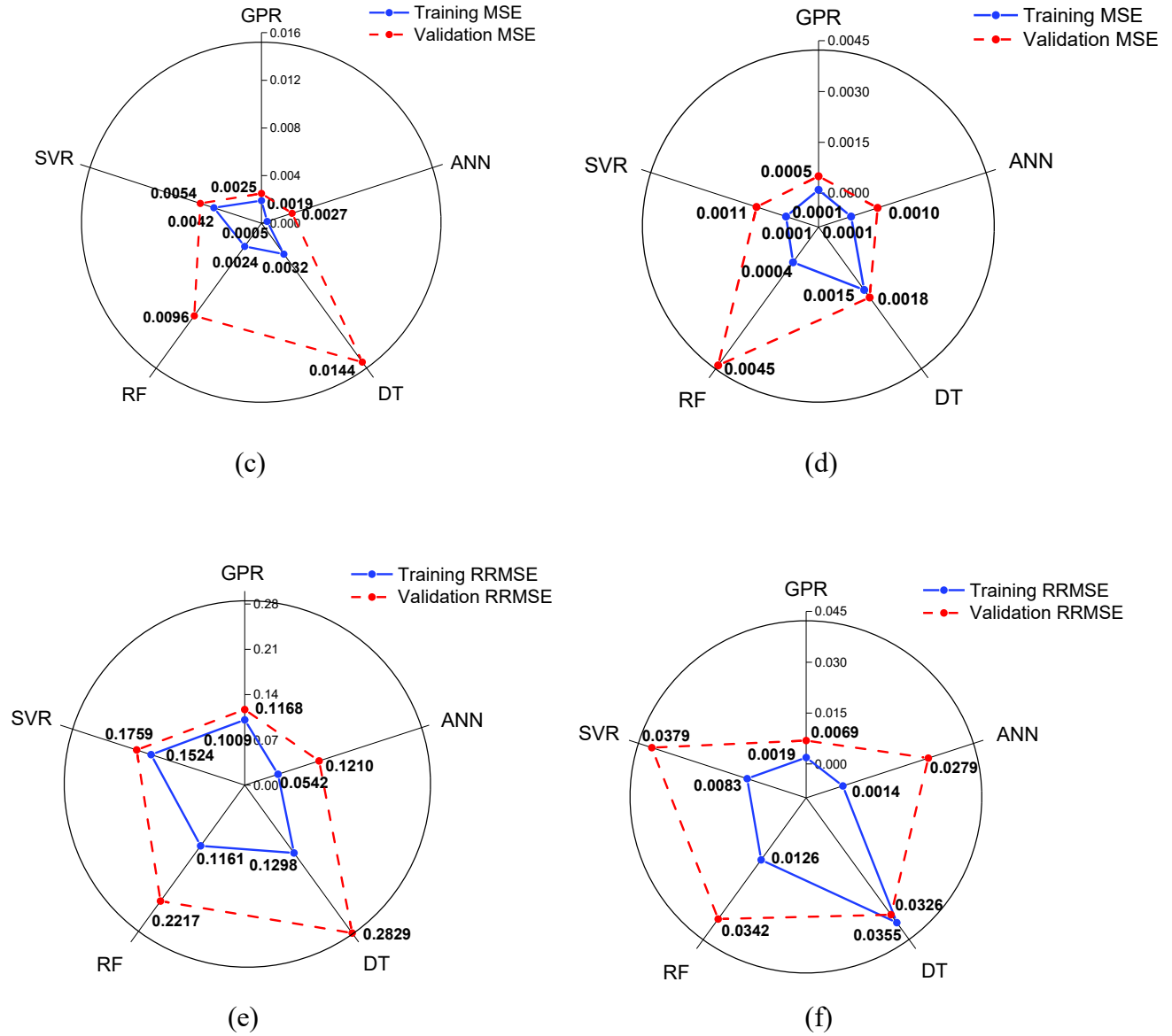
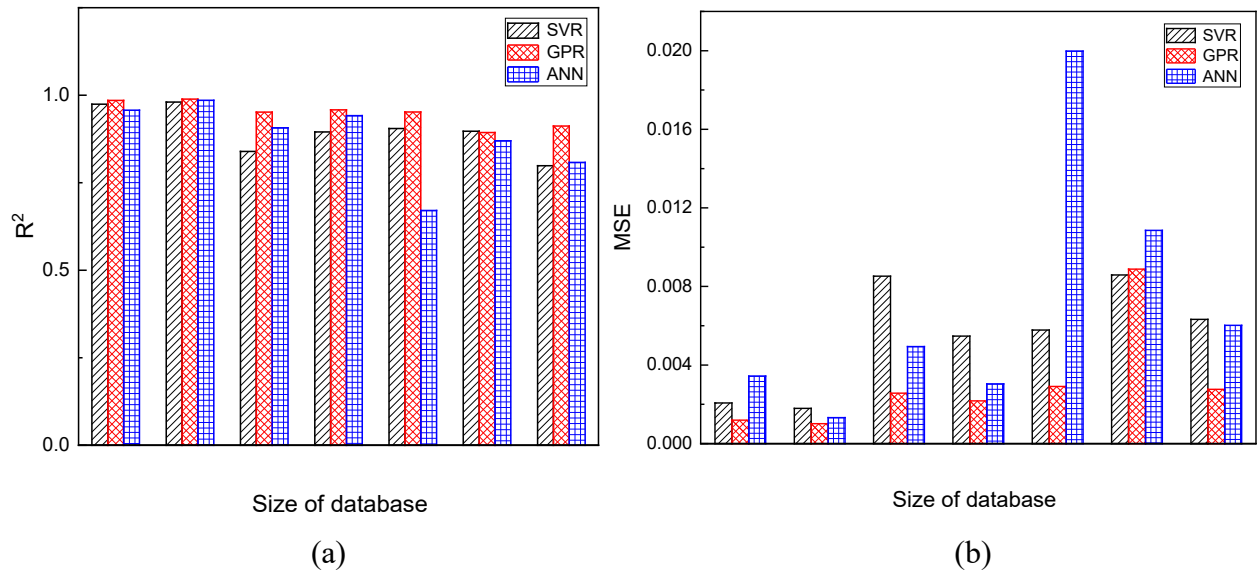


Fig. 6. Comparisons of training and validation results in terms of average value and standard deviation of R^2 , MSE, and RRMSE.

4.2 Effect of the size of database

In this section, the effect of database size on the prediction performance of the GPR, ANN, and SVR models is further investigated. As shown in Fig. 7, the data in the database is reduced by 10% each time until only 20 samples are left. For each database, the split is kept at 80% for training and 20% for testing. As expected, as the size of the database decreases, the prediction performance of

the three models generally exhibits a descending trend. It should be noted that the fluctuation in prediction performance should be attributed to the sampling variability for training and testing, and the effect of which is more pronounced with a smaller database. It can be observed from Fig. 7 that the GPR model outperforms the other two models across almost all the databases. For the last database with 20 data samples, the GPR model achieves prediction results with values of R^2 , MSE and RRMSE of 0.912, 0.003, and 0.166, respectively, while the values for the ANN model are 0.808, 0.006, and 0.245, respectively, and the values for the SVR model are 0.798, 0.006, and 0.251, respectively. The GPR model exhibits the lowest prediction performance with an R^2 value of 0.893 when the size of the database is 27, while the lowest R^2 value obtained by the ANN model is 0.671 when the database contains 34 data samples, and the corresponding MSE and RRMSE are 0.020 and 0.297, respectively, which are higher than the contrasting values observed across all the databases, as shown in Fig. 7. In addition to the advantages of the GPR model discussed above, the GPR model also exhibits less fluctuation in terms of R^2 , MSE, and RRMSE as the size of the database decreases compared to the SVR and ANN models, further indicating that the GPR model is less sensitive to the database size than the other two models.



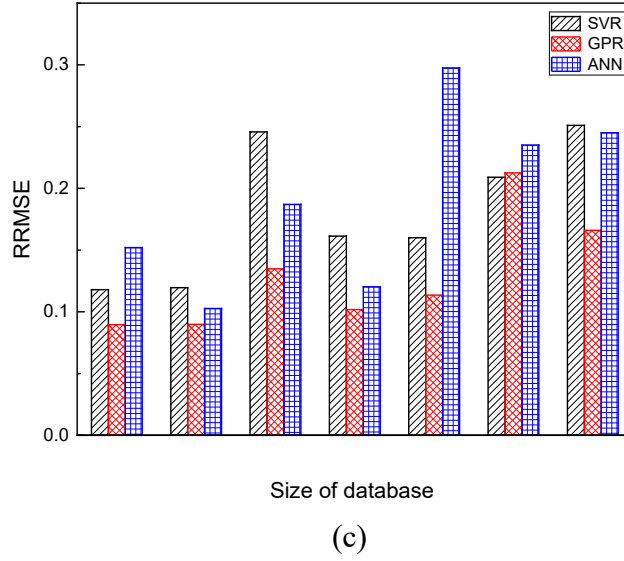


Fig. 7. Effect of the database size with regards to (a) R2, (b) MSE, and (c) RRMSE.

4.3 GPR results with confidence interval

Different from the contrasting machine learning models, the GPR model is able to provide results with confidence interval ranges, based on which the reliability of the prediction can be measured. Fig. 8 shows the prediction results of the GPR model with confidence intervals, where the green points at the middle of the range represent the predicted results, while the red points denote the experimental results. The prediction with higher uncertainty could exhibit a wider confidence interval, and the variations in the width of the confidence interval for different data points should be attributed to their proximity to the training data. It can be observed that only one point slightly exceeds the 95% confidence interval, two points fall within the narrow range between the 95% and 99% confidence intervals, and the 99% confidence interval covers all the remaining data points. The obtained results indicate that the predictions of the GPR model are generally within the expected range of certainty. The extensive coverage of the 99% confidence interval over the majority of data points demonstrates the robustness of the model in the prediction of the load-bearing capacity of the F-STC columns.

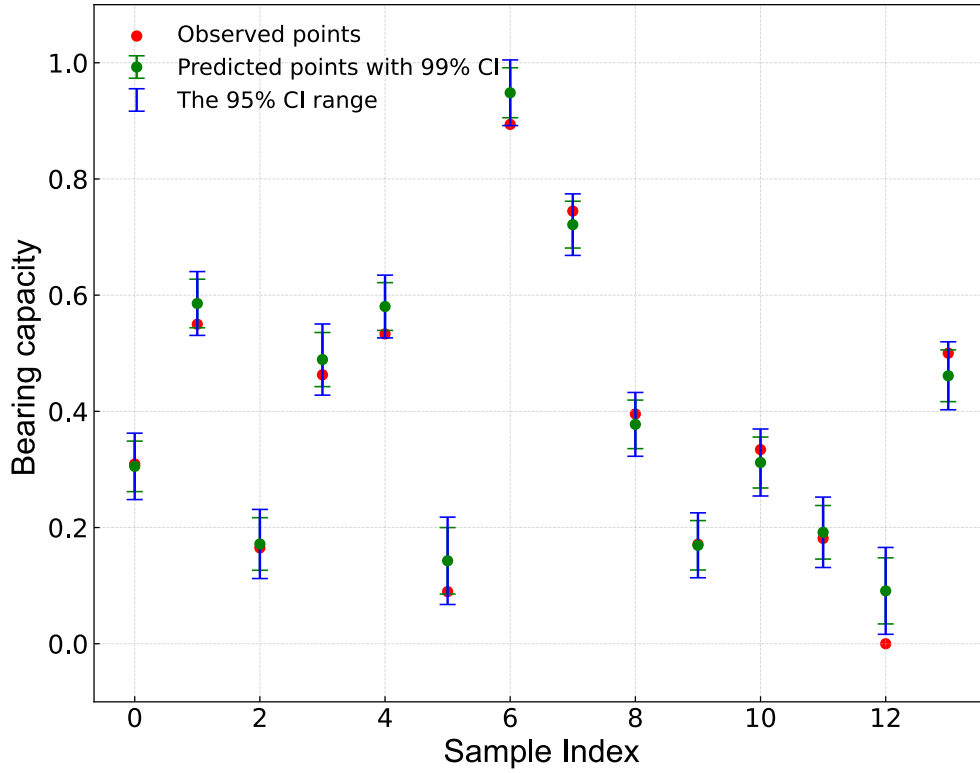


Fig. 8. Prediction results of the GPR model with confidence intervals.

4.4 Interpretations for the GPR model

Fig. 9 depicts a summary plot of the SHAP values for the whole database, providing a global interpretation of the prediction process. For each individual data point, the relationship between each feature value and its corresponding SHAP value is given using dots, with higher feature values marked in red and lower feature values in blue. The SHAP value is shown in the abscissa, and a positive SHAP value means a positive contribution to the load-bearing capacity of the F-STC column, while a negative SHAP value means a negative contribution. The mean absolute SHAP values are shown in Fig. 10, it can be observed that the concrete strength has the highest effect on the load-bearing capacity of the F-STC column, whereas the yield strength of the steel tube has the lowest effect, and the thickness of the steel tube is the second least importance.

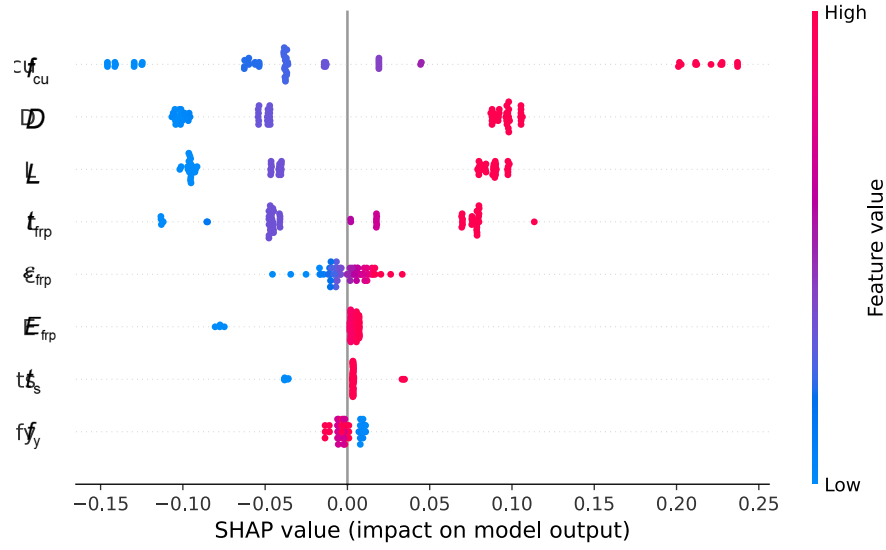


Fig. 9. Global interpretations of the GPR model for the whole dataset.

It can be observed that the effect of the thickness of the FRP sheet is considerably larger than that of the steel tube. Apart from the concrete strength, the diameter and length of the column also have a noticeable effect on its load-bearing capacity. It can be observed from Fig. 9, higher values of concrete strength, column diameter, column length, and thickness of the FRP sheet lead to higher load-bearing capacity. However, the positive effects of the elastic modulus of the FRP sheet, as well as the yield strength and thickness of the steel tube are limited. This could be attributed to the limited value ranges of these features in the database.

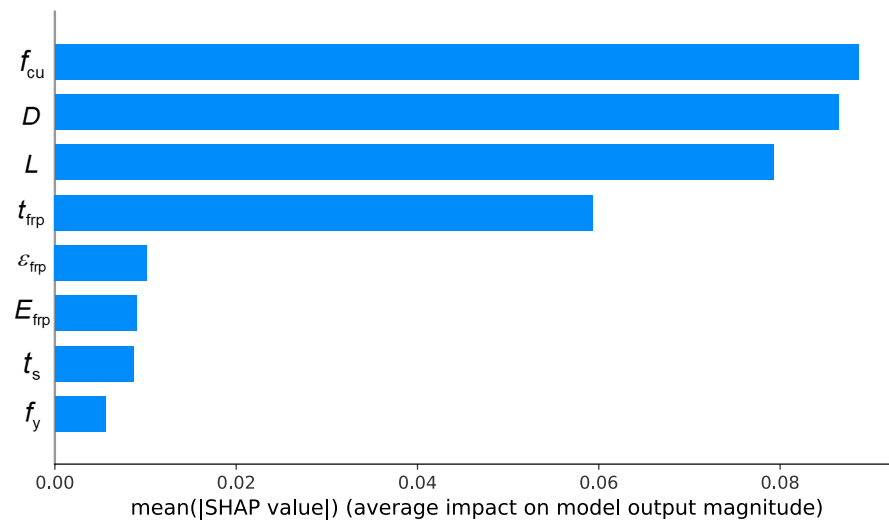
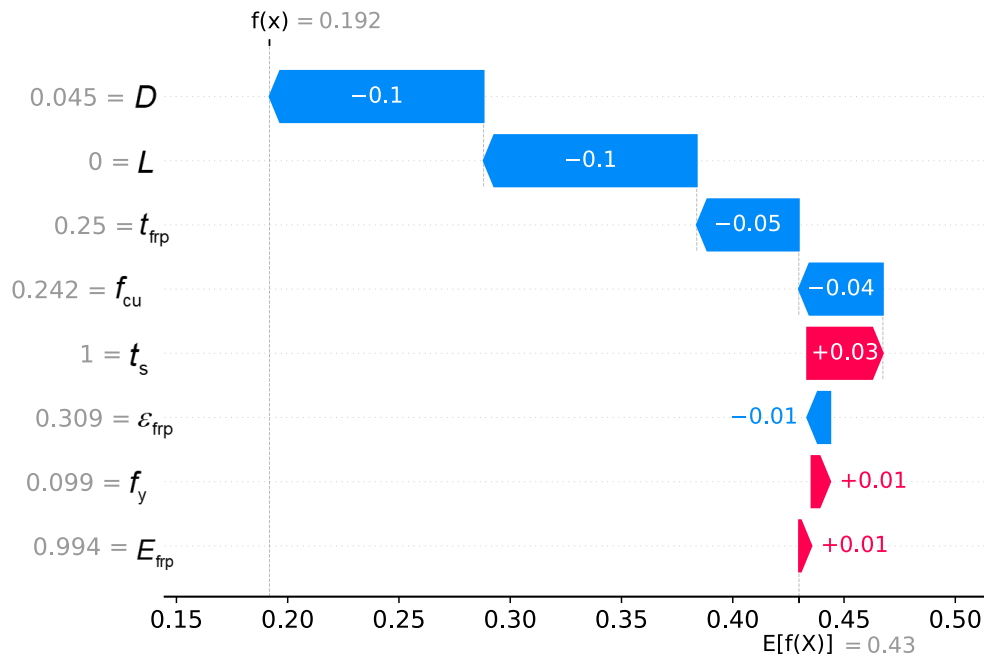


Fig. 10. Mean absolute SHAP values for inputs.

Fig. 11 shows the specific prediction interpretation for individual data samples. As discussed in Section 2, the feature values shown in Fig. 11 are normalised within the range of 0 to 1, as discussed in Section 2. For the example shown in Fig. 11(a), it can be observed that the small length and diameter of this column lead to a total decrease of 0.2 in the load-bearing capacity from the base value, which is the mean value of the predictions. The relatively lower values of concrete strength, thickness and rupture strain of FRP sheet also result in a decrease in the load-bearing capacity. Although the positive effects of the elastic modulus of the FRP sheet and the thickness and yield strength of the steel tube are limited, their high feature values lead to an increase in the load-bearing capacity. Unlike the example shown in Fig. 11(a), the high thickness of the FRP sheet in the example shown in Fig. 11(b) results in an increase in the load-bearing capacity, with a value of 0.8. Comparison between these two examples demonstrates that the higher load-bearing capacity of the second example mainly benefits from the greater thickness of the FRP sheet and the relatively larger column diameter and column length.



(a)

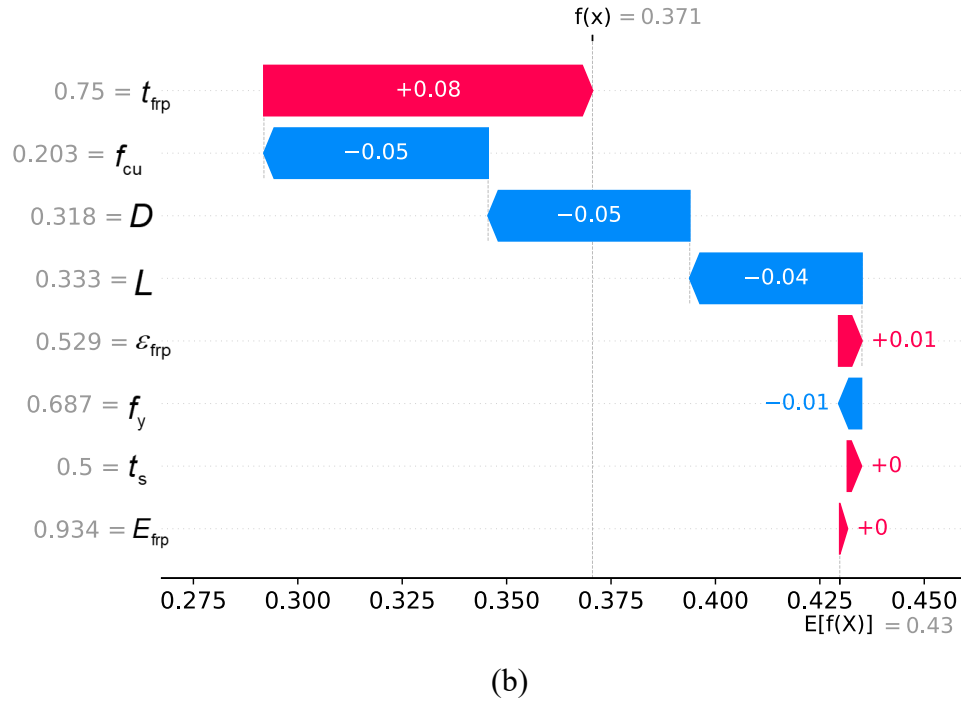


Fig. 11. Local interpretation of the GPR model for two data samples.

5. Conclusions

This paper investigates the application of machine learning models for predicting the load-bearing capacity of F-STC columns using limited data. Due to the limited number of experimental tests conducted on the F-STC columns, only 69 samples obtained from previous research were utilised in this study. To solve this problem, a GPR model is proposed, enabling reliable predictions for cases with limited data. More importantly, the predictions of the GPR model are interpretable by combining them with the SHAP approach. Based on the limited data, the contribution of each feature at both the global level and local level is provided. Besides, the application of four popular machine learning models, including ANN, SVR, DT, and RF, to this prediction problem is also studied, and comparisons between the five machine learning models are made. The obtained results demonstrate that the GPR model exhibits the lowest prediction deviation and highest goodness-of-fit, as well as the best generalisation capability for this prediction problem. Although the ensemble nature of the RF model improves its generalisation performance, both the DT and RF models exhibit significantly lower prediction accuracy compared to the other three models, hence, they are not appropriate for the use in this prediction problem. The ANN and SVR models show similar

prediction capabilities, and both models exhibit slight overfitting due to the limited size of the database. Besides, the further reductions in the database size significantly affect the prediction performance of the ANN and SVR models, whereas the effect on the GPR model is slight. Furthermore, the reliability of the GPR predictions is measured based on its confidence interval, the obtained results illustrate that the GPR model is able to achieve results with high confidence for this prediction problem. Nevertheless, in the present study, nine input features and one output are considered. For future study, to further enhance the interpretability of the predictions, it is recommended to consider more input features, such as the reinforcement ratio and fibre angle, as well as more outputs such as the failure mode and deformation of the F-STC column. This study primarily focuses on the performance of the machine learning models with limited data, future study could compare how performance changes when extensive databases are utilised.

Acknowledgements

This research is sponsored by the Natural Science Foundation of Shandong Province, China (No. ZR2021QE235). This support is gratefully acknowledged.

References

- [1] Tomii M, Sakino K, Xiao Y, Watanabe K. Earthquake resisting hysteretic behavior of reinforced concrete short columns confined by steel tube. Proceedings of the international speciality conference on concrete filled steel tubular structures, Harbin, China 1985. p. 119-25.
- [2] Kwan A, Dong C, Ho J. Axial and lateral stress-strain model for circular concrete-filled steel tubes with external steel confinement. *Engineering Structures*. 2016;117:528-41.
- [3] Wei Y, Jiang C, Wu Y-F. Confinement effectiveness of circular concrete-filled steel tubular columns under axial compression. *Journal of Constructional Steel Research*. 2019;158:15-27.
- [4] Ouyang Y, Kwan A. Finite element analysis of square concrete-filled steel tube (CFST) columns under axial compressive load. *Engineering Structures*. 2018;156:443-59.
- [5] Le Hoang A, Fehling E. Numerical study of circular steel tube confined concrete (STCC) stub columns. *Journal of Constructional Steel Research*. 2017;136:238-55.
- [6] Liu J, Li X, Zang X, Wang X, Chen YF. Seismic behavior of shear-critical circular TSRC columns with a shear span-to-depth ratio of 1.3. *Thin-Walled Structures*. 2019;134:373-83.
- [7] Du P, Gan D, Yang Y, Tan KH. Simplified method for analysis of circular steel tubed-reinforced-concrete columns under eccentric compression. *Engineering Structures*. 2019;198:109524.
- [8] Wang X, Liu J, Zhang S. Behavior of short circular tubed-reinforced-concrete columns subjected to eccentric compression. *Engineering Structures*. 2015;105:77-86.

- [9] Liu J, Teng Y, Zhang Y, Wang X, Chen YF. Axial stress-strain behavior of high-strength concrete confined by circular thin-walled steel tubes. *Construction and Building Materials*. 2018;177:366-77.
- [10] Bai J, Chen H, Zhao J, Liu M, Jin S. Seismic design and subassembly tests of buckling-restrained braced RC frames with shear connector gusset connections. *Engineering Structures*. 2021;234:112018.
- [11] Xie J, Li Y, Lu Z, Fan Z, Li J, Li S. Effects of immersion in water, alkaline solution, and seawater on the shear performance of basalt FRP bars in seawater-sea sand concrete. *Journal of Composites for Construction*. 2022;26:04021071.
- [12] Liao J, Zeng J-J, Jiang C, Li J-X, Yuan J-S. Stress-strain behavior and design-oriented model for FRP spiral strip-confined concrete. *Composite Structures*. 2022;293:115747.
- [13] Wang J, Lu S, Yang J. Behavior of eccentrically loaded rectangular RC columns wrapped with CFRP jackets under different preloading levels. *Journal of Building Engineering*. 2021;34:101943.
- [14] Yang J, Lu S, Wang J, Wang Z. Behavior of CFRP partially wrapped square seawater sea-sand concrete columns under axial compression. *Engineering Structures*. 2020;222:111119.
- [15] Lu Y, Li N, Li SJP. Behavior of FRP-confined concrete-filled steel tube columns. *Polymers*. 2014;6:1333-49.
- [16] Che Y, Wang Q, Shao Y. Compressive performances of the concrete filled circular CFRP-steel tube (C-CFRP-CFST). *Advanced Steel Construction*. 2012;8:331-58.
- [17] Wei Y, Wu G, Li G. Performance of circular concrete-filled fiber-reinforced polymer-steel composite tube columns under axial compression. *Journal of Reinforced Plastics Composites*. 2014;33:1911-28.
- [18] Na L, Yiyan L, Shan L, Lan L. Slenderness effects on concrete-filled steel tube columns confined with CFRP. *Journal of Constructional Steel Research*. 2018;143:110-8.
- [19] Ding F-x, Lu D-r, Bai Y, Gong Y-z, Yu Z-w, Ni M et al. Behaviour of CFRP-confined concrete-filled circular steel tube stub columns under axial loading. *Thin-Walled Structures*. 2018;125:107-18.
- [20] Wang Q-L, Zhao Z, Shao Y-B, Li Q-L. Static behavior of axially compressed square concrete filled CFRP-steel tubular (S-CF-CFRP-ST) columns with moderate slenderness. *Thin-Walled Structures*. 2017;110:106-22.
- [21] Zhang Y, Wei Y, Bai J, Zhang Y. Stress-strain model of an FRP-confined concrete filled steel tube under axial compression. *Thin-Walled Structures*. 2019;142:149-59.
- [22] Park J, Hong Y, Hong G, Kim J, Choi S. Design formulas of concrete filled circular steel tubes reinforced by carbon fiber reinforced plastic sheets. *Procedia engineering*. 2011;14:2916-22.
- [23] Tao Z, Han L-H, Zhuang J-P. Axial loading behavior of CFRP strengthened concrete-filled steel tubular stub columns. *Advances in Structural Engineering*. 2007;10:37-46.
- [24] Xu T, Liu J, Wang X, Guo Y, Chen YF. Behaviour of short CFRP-steel composite tubed reinforced normal and high strength concrete columns under eccentric compression. *Engineering Structures*. 2020;205:110096.
- [25] Xu T, Liu J, Wang X, Guo Y, Chen YF. Behaviour of short CFRP-steel composite tubed reinforced normal and high strength concrete columns under eccentric compression. *Engineering Structures*. 2020;205.
- [26] Lu S, Wang J, Yang J, Wang L. Compressive behavior of circular steel tubed high-strength concrete columns with embedded steel tubes. *Journal of Constructional Steel Research*. 2022;193.
- [27] Xu T, Liu J, Guo Y. Design method of short circular FRP-steel composite tubed RC columns under eccentric compression. *Composite Structures*. 2021;262.

- [28] Chen G, Wang Y, Cai G, Si Larbi A, Wan B, Hao Q. Performance and modeling of FRP-steel dually confined reinforced concrete under cyclic axial loading. *Composite Structures*. 2022;300.
- [29] Lu Y, Li N, Li S. Behavior of FRP-confined concrete-filled steel tube columns. *Polymers*. 2014;6:1333-49.
- [30] Hu Y, Yu T, Teng J. FRP-confined circular concrete-filled thin steel tubes under axial compression. *Journal of Composites for Construction*. 2011;15:850-60.
- [31] Yu F, Wu P. Study on stress-strain relationship of FRP-confined concrete filled steel tubes. *Advanced Materials Research*. 2011;163:3826-9.
- [32] Choi K-K, Xiao Y. Analytical model of circular CFRP confined concrete-filled steel tubular columns under axial compression. *Journal of Composites for Construction*. 2010;14:125-33.
- [33] Teng J, Hu Y, Yu T. Stress-strain model for concrete in FRP-confined steel tubular columns. *Engineering Structures*. 2013;49:156-67.
- [34] Dong C, Kwan A, Ho J. Axial and lateral stress-strain model for concrete-filled steel tubes with FRP jackets. *Engineering Structures*. 2016;126:365-78.
- [35] Liu J-P, Xu T-X, Wang Y-H, Guo Y. Axial behaviour of circular steel tubed concrete stub columns confined by CFRP materials. *Construction and Building Materials*. 2018;168:221-31.
- [36] Liu J, Xu T, Guo Y, Wang X, Frank Chen Y. Behavior of circular CFRP-steel composite tubed high-strength concrete columns under axial compression. *Composite Structures*. 2019;211:596-609.
- [37] Xu T, Liu J, Yu T, Guo Y. Design-oriented model for circular F-STC stub columns subjected to axial compression. *Journal of Constructional Steel Research*. 2020;170.
- [38] Wang Y, Chen G, Wan B, Han B, Ran J. Axial compressive behavior and confinement mechanism of circular FRP-steel tubed concrete stub columns. *Composite Structures*. 2021;256.
- [39] Feng J, Gao K, Gao W, Liao Y, Wu G. Machine learning-based bridge cable damage detection under stochastic effects of corrosion and fire. *Engineering Structures*. 2022;264.
- [40] Lim S, Chi S. Xgboost application on bridge management systems for proactive damage estimation. *Advanced Engineering Informatics*. 2019;41.
- [41] Chou J-S, Tsai C-F, Pham A-D, Lu Y-H. Machine learning in concrete strength simulations: Multi-nation data analytics. *Construction and Building Materials*. 2014;73:771-80.
- [42] Safarzaghegan Gilan S, Bahrami Jovein H, Ramezani-pour AA. Hybrid support vector regression – Particle swarm optimization for prediction of compressive strength and RCPT of concretes containing metakaolin. *Construction and Building Materials*. 2012;34:321-9.
- [43] Cheng M-Y, Chou J-S, Roy AFV, Wu Y-W. High-performance Concrete Compressive Strength Prediction using Time-Weighted Evolutionary Fuzzy Support Vector Machines Inference Model. *Automation in Construction*. 2012;28:106-15.
- [44] Perera R, Arteaga A, Diego AD. Artificial intelligence techniques for prediction of the capacity of RC beams strengthened in shear with external FRP reinforcement. *Composite Structures*. 2010;92:1169-75.
- [45] Tixier AJP, Hallowell MR, Rajagopalan B, Bowman D. Application of machine learning to construction injury prediction. *Automation in Construction*. 2016;69:102-14.
- [46] Perera R, Tarazona D, Ruiz A, Martín A. Application of artificial intelligence techniques to predict the performance of RC beams shear strengthened with NSM FRP rods. Formulation of design equations. *Composites Part B: Engineering*. 2014;66:162-73.

- [47] Duan J, Asteris PG, Nguyen H, Bui X-N, Moayedi H. A novel artificial intelligence technique to predict compressive strength of recycled aggregate concrete using ICA-XGBoost model. *Engineering with Computers*. 2020;37:3329-46.
- [48] Pham TM, Hadi MNS. Predicting Stress and Strain of FRP-Confined Square/Rectangular Columns Using Artificial Neural Networks. *Journal of Composites for Construction*. 2014;18.
- [49] Cascardi A, Micelli F, Aiello MA. An Artificial Neural Networks model for the prediction of the compressive strength of FRP-confined concrete circular columns. *Engineering Structures*. 2017;140:199-208.
- [50] El Ouni MH, Raza A. Data-driven analysis of concrete-filled steel-tube CFRP-confined NSC columns. *Mechanics of Advanced Materials and Structures*. 2021;29:5667-88.
- [51] Mozumder RA, Roy B, Laskar AI. Support Vector Regression Approach to Predict the Strength of FRP Confined Concrete. *Arabian Journal for Science and Engineering*. 2016;42:1129-46.
- [52] Mansouri I, Ozbakkaloglu T, Kisi O, Xie T. Predicting behavior of FRP-confined concrete using neuro fuzzy, neural network, multivariate adaptive regression splines and M5 model tree techniques. *Materials and Structures*. 2016;49:4319-34.
- [53] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. 31st Conference on Neural Information Processing Systems (NIPS 2017). Long Beach, CA, USA. 2017.
- [54] Che Z, Purushotham S, Khemani R, Liu Y. Interpretable deep models for ICU outcome prediction. *AMIA annual symposium proceedings: American Medical Informatics Association*; 2016. p. 371.
- [55] Wang T, Reiffsteck P, Chevalier C, Chen C-W, Schmidt F. An interpretable model for bridge scour risk assessment using explainable artificial intelligence and engineers' expertise. *Structure and Infrastructure Engineering*. 2023;1-13.
- [56] Zhou C, Xie Y, Wang W, Zheng Y. Machine learning driven post-impact damage state prediction for performance-based crashworthiness design of bridge piers. *Engineering Structures*. 2023;292.
- [57] Hoang N-D, Pham A-D, Nguyen Q-L, Pham Q-N. Estimating compressive strength of high performance concrete with Gaussian process regression model. *Advances in civil engineering*. 2016;2016.
- [58] Le T-T, Le MV. Development of user-friendly kernel-based Gaussian process regression model for prediction of load-bearing capacity of square concrete-filled steel tubular members. *Materials and Structures*. 2021;54:59.
- [59] Yu Z, Xie W, Yu B, Cheng H. Probabilistic prediction of joint shear strength using Gaussian process regression with anisotropic compound kernel. *Engineering Structures*. 2023;277:115413.
- [60] Guo Y, Xu TX, Liu JP. Experimental study and analysis on axial behavior of circular CFRP-steel composite tubed high-strength concrete stub columns (In Chinese). *Journal of building Structures*. 2019;40:124-31.
- [61] Guo Y, Liu JP, Miao YJ, Wang YH, Xu TX. Experimental study on axial behavior of circular CFRP-steel composite tube confined concrete stub columns (In Chinese). *Engineering Mechanics*. 2017;34:41-50.
- [62] Ran JH. Axial compression mechanical behavior of FRP-steel composite tube confined concrete stub columns (In Chinese) [Master's Thesis]: Dalian University of Technology; 2014.
- [63] Zhao W, Chen P, Liu X, Wang L. Impact response prediction and optimization of SC walls using machine learning algorithms. *Structures*. 2022;45:390-9.

Highlights

- The development of five machine learning models using a small database is studied.
- To make the machine learning predictions explainable, a SHAP approach is employed.
- The GPR model outperforms the others with superior accuracy and generalisation.
- The effect of database size is further studied by gradually reducing the samples.