



Ensemble XGBoost schemes for improved compressive strength prediction of UHPC

May Huu Nguyen^{a,b}, Thuy-Anh Nguyen^a, Hai-Bang Ly^{a,*}

^a University of Transport Technology, Hanoi 100000, Vietnam

^b Civil and Environmental Engineering Program, Graduate School of Advanced Science and Engineering, Hiroshima University, 1-4-1, Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8527, Japan

ARTICLE INFO

Keywords:

XGBoost
Ensemble methods
Compressive strength
Ultra-high performance concrete

ABSTRACT

XGBoost is a promising machine learning method capable of predicting essential concrete properties and enhancing advanced concrete design. However, its underlying version still requires further study and development. In this investigation, the effectiveness of advanced XGBoost versions, including Ada-XGBoost, Bagging-XGBoost, Stacking-XGBoost, and Voting-XGBoost, to predict the compressive strength (CS) of Ultra-high-performance concrete (UHPC) was accessed. A database covering 810 results from in the literature, including 15 inputs, such as 12 UHPC components, two curing conditions, and sample age, was utilized for training the models. The performance criteria for the five models, including RMSE, MAE, and R^2 , were evaluated using a combination of 10-Fold CV and Monte Carlo (MC) simulation. The results showed that the Stacking-XGBoost and XGBoost models outperformed other models in terms of prediction accuracy for the CS of UHPC. Based on SHAP values analysis, features such as age, fiber, slag, cement, sand, superplasticizer, water, relative humidity, and temperature were identified as the key parameters affecting UHPC's CS. Furthermore, a quantitative analysis of their combined impact on UHPC's CS was also provided.

1. Introduction

Ultra-high performance concrete (UHPC) is a new kind of concrete that has been studied, developed, and globally utilized for over 40 years. Several improvements in terms of strength and endurance over regular concrete [1,2] are offered by UHPC. Its ability to accelerate construction, improve aesthetics, and provide superior mechanical properties, including waterproofing, anti-corrosion, abrasion, and impact resistance, makes it advantageous for building applications [3]. Graybeal's study [4] highlights that UHPC is a cement-based conjugated concrete material, optimized for aggregate particles with a water-to-cement ratio of less than 0.25. This optimization results in reduced pores, a dense distribution of microstructures, and the use of dispersed fiber reinforcement to increase ductility under flex. On this basis, UHPC may obtain exceptional characteristics compared to standard concrete. The most essential mechanical characteristics of UHPC are its compressive strength (CS), which must be more than 150 MPa, and its direct tensile strength after breaking, which must be greater than 5 MPa.

Indeed, an extraordinarily high hardness and strength are exhibited by UHPC. Typically, the CS of UHPC falls between the range of 150 and

250 MPa. However, depending on the heat curing circumstances, it may exceed 250 MPa and even approach 800 MPa [5–8]. UHPC has an elastic modulus of between 45 and 55 GPa, but it may vary depending on the aggregate particle makeup [9]. Additionally, UHPC has a very high slump, with a spread higher than 600 mm, making it ideal for constructing vibration-resistant concrete structures [10]. When designing a structure with the same bearing capacity, UHPC allows for a smaller cross-section, reducing the quantity of material utilized and the stress on the structure itself. Notably, UHPC exhibits comparable capabilities to specific ordinary concrete and steel, commonly used for fabricating prestressed beams, thereby making it a potential competitor to steel materials. Combining UHPC with prestressing reinforcement can further maximize its characteristics, enhancing its flexural strength. In certain scenarios, the use of distributed fiber reinforcement in UHPC composites eliminates the need for flexural and shear reinforcing bars. Given these advantageous characteristics, UHPC finds successful application in constructing large-span buildings subjected to high loads or anomalous effects caused by external factors such as earthquakes, natural catastrophes, or terrorism. Moreover, UHPC enhances efficiency in projects vulnerable to significant erosion, such as marine, hydraulic, chemical, or

* Corresponding author.

E-mail address: bangh@utt.edu.vn (H.-B. Ly).

nuclear waste burial tanks [11,12].

The most significant drawback of UHPC is that the product price is raised due to the requirement of an enormous volume of cement, impacting both technical features and the environment in a significant manner. The relatively high quantity of cement in UHPC (i.e., 900–1200 kg/m³), is attributed to the scarcity of coarse material [3]. The negative environmental impact of concrete can be substantially mitigated by incorporating concrete additives that consume less energy, thereby partially substituting cement [2,13]. Consequently, greater attention has been directed towards eco-efficient supplementary cementitious materials (SCMs) such as fly ash (FA), blast furnace slag (GGBFS), and silica fume (SF), which can be employed to produce UHPC with reduced cement content [14–19]. Additionally, reinforcing strands, primarily steel fibers, are incorporated and manipulated to ensure UHPC's flexibility and impact resistance [20,21]. In comparison to regular concrete, the inclusion of fiber reinforcement significantly enhances the tensile and flexural strengths of UHPC. The key components of UHPC are composed of cement, SCMs, fine sand, quartz or glass powder, steel fiber, low water content, and high-range water reducer [22–24]. Additionally, UHPC's mechanical characteristics are influenced by the composition of the mixture and the curing conditions [2,11,25,26]. Recent years have witnessed several studies evaluating the mechanical properties of UHPC using diverse component types and mixing proportions [26–28]. In efforts by Zhang, experiments were carried out to investigate the microstructure and mechanical properties of UHPC containing FA and SF [3]. Compared to the reference sample with 30% FA replacement, combining FA with 10% and 20% SF incorporation reveals mechanical properties that are equal or superior. However, contrary to the aforementioned findings, Alsalmal et al. [29] observed that substituting 20% or 30% FA increased the CS of the concrete. A similar outcome can be achieved by simultaneously substituting 20% FA and 5% SF. Substitution with FA leads to a 40% reduction in the concrete's CS. According to Yazici et al. [18], UHPC is formulated by incorporating various slag replacement ratios (i.e., 20%, 40%, and 60%) into cement, with mixes containing 20% slag displaying the highest CS. Another study [19] highlights that UHPC mixes exhibit favorable compressive and flexural characteristics when utilizing 20–40% replacement of FA or slag. Moreover, the incorporation of a modest quantity of nanoparticles holds the potential to enhance UHPC's performance [30,31]. In supplementary investigations [13,32,33], it was determined that employing GGBFS to partially substitute cement leads to a reduction in the CS of UHPC.

Furthermore, the UHPC's CS was found to be influenced by the fiber through various laboratory tests [29,34,35]. Due to their inherent high strength and rigidity, additional steel fibers in the aggregate play a key role in significantly enhancing UHPC strength and ductility [21]. Despite numerous studies, the comprehensive examination of SCMs and fiber composition on UHPC's CS has not been carried out, as indicated by the findings of the previous study [36]. Conversely, laboratory experiments tend to be expensive, time-consuming, and labor-intensive [3]. Consequently, the creation of a model capable of quickly and reliably predicting UHPC's CS becomes crucial [3,36].

In recent decades, artificial intelligence (AI) has demonstrated its effectiveness in addressing complex engineering problems across various fields, including structural engineering [37–39], environmental engineering [40,41], and materials science [42–44]. For instance, the work by Araghani and Asteris [45] involved a comparison of ANN and ANFIS models for predicting the CS of cement-based mortar, providing insights into their respective performance. Asteris et al. [46] focused on predicting shear strength in reinforced concrete beams using ANNs, highlighting their suitability for modeling complex concrete properties. Apostolopoulou et al. [47] presented a comprehensive mapping and design approach for natural hydraulic lime mortars, optimizing compositions through AI techniques. Additionally, Asteris et al. [48] investigated metakaolin-based concrete materials, revealing essential mechanical characteristics using AI. The research by Barkhordari et al.

[49] demonstrated the utilization of ensemble learners for data-driven prediction of CS in FA concrete, resulting in improved prediction accuracy. Though not directly focused on CS, the potential of ensemble deep learning for structural damage identification was showcased by Barkhordari et al. [50]. Additionally, Cavalieri et al. [51] introduced convolution-based ensemble learning to estimate bond strength in corroded reinforced concrete. The application of AI techniques to predict concrete CS based on experimental data has been extensively implemented in recent years. Examples include rubber concrete [52], recycled aggregate concrete [53,54], fly ash concrete [55,56], self-compacting concrete [57], HPC and UHPC [58–60], silica fume concrete [61], geopolymers concrete [62]. These studies collectively demonstrate the viability of AI algorithms for predicting various mechanical characteristics of concrete. Moreover, AI algorithms have been employed by several researchers to predict the compressive strength of UHPC.

In Ghafari et al. [63] the prediction of UHPC's CS was accomplished using a backpropagation neural network (BPNN) in conjunction with a statistical combination of signals. This study involved two distinct curing techniques, namely steam and wet curing. For training the ANN model, a total of 53 unique UHPC mixtures were generated. The inputs for the BPNN model encompassed various mixture components. The findings revealed the BPNN's capability to accurately estimate UHPC's CS, attaining an R² value of 0.99 under steam curing and 0.98 under water storage conditions. In another investigation by Zhang et al. [3], an ANN model was developed using 78 groups of experimental data to estimate UHPC's CS containing complimentary binders. The model incorporated 11 input variables. The outcomes highlighted the high accuracy of the proposed ANN model in predicting UHPC's CS in the presence of binary additional materials. While the aforementioned studies underscore the excellence of ANNs as local search methods, it's important to acknowledge their inherent limitations. ANNs exhibit sensitivity to initial weights, resulting in challenges related to local minimum values and inconsistent training outcomes [64]. To address these drawbacks, Fan et al. [65] introduced a biology-based technology known as Genetic Algorithm (GA) to enhance the algorithmic structure of the ANN. This enhanced model was based on an examination of correlations among packing density, mechanical characteristics, and pore structure using a training dataset of 80 mixes. The performance of this model surpassed that of classical prediction models in terms of accuracy. Notably, prior research has successfully developed high-efficiency ANN models for UHPC's CS prediction. However, it's worth noting that these studies are limited by the scope of their employed datasets, encompassing a confined number of mixed components. Marani et al. [36], on the other hand, expanded the dataset significantly, incorporating 810 experimental data points in their analysis. They harnessed a synthetic adversarial network technique for UHPC's CS prediction, resulting in a more robust machine learning (ML) model with performance indicators including R² of 0.96, MAE of 6.72 MPa, and RMSE of 7.41 MPa. Nonetheless, a comprehensive examination of the factors influencing UHPC's CS remains a relatively unexplored area within the existing body of research.

This conducted study delves into the application of the Extreme Gradient Boosting (XGBoost) algorithm [66] within the realm of ML. This algorithm, renowned for its recent emergence, has found effective utility across diverse domains, often harnessed to surmount challenges tied to high accuracy in the realm of supervised learning. However, as of yet, no comprehensive exploration has addressed the efficiency of the XGBoost methodology in tackling the intricate conundrum of estimating CS within UHPC. To address this research gap, the authors extend the frontiers of XGBoost by proposing the incorporation of supplementary models. This augmentation involves the integration of two advanced models, specifically Adaptive Boosting (Ada-XGBoost) and Bagging-XGBoost. These combined models are subsequently amalgamated using ensemble techniques, specifically the Stacking and Voting approaches, giving rise to two distinct models: Stacking-XGBoost and

Table 1

Description of inputs and output of this study.

	No	Unit	Mean	Std	Min	Q _{25%}	Q _{50%}	Q _{75%}	Max
Cement	X ₁	kg/m ³	738	173	270	620	771	850	1251
Slag	X ₂	kg/m ³	25	74	0	0	0	0	375
Silica Fume	X ₃	kg/m ³	137	104	0	44	144	219	434
Limestone powder	X ₄	kg/m ³	42	133	0	0	0	0	1058
Quartz powder	X ₅	kg/m ³	33	80	0	0	0	0	397
Fly Ash	X ₆	kg/m ³	26	67	0	0	0	0	356
Nano silica	X ₇	kg/m ³	4	8	0	0	0	4	48
Water	X ₈	kg/m ³	180	26	90	163	177	193	273
Sand	X ₉	kg/m ³	995	283	0	786	1021	1231	1503
Gravel	X ₁₀	kg/m ³	155	358	0	0	0	0	1195
Fiber	X ₁₁	kg/m ³	56	75	0	0	0	156	234
Superplasticizer	X ₁₂	kg/m ³	30	14	1	18	30	44	57
Relative humidity	X ₁₃	%	98	8	50	100	100	100	100
Temperature	X ₁₄	°C	24	16	20	21	21	23	21.0
Age	X ₁₅	days	37	53	1	7	28	28	365
Compressive strength	Y	MPa	123	40	29	96	122	154	221

Voting-XGBoost. To substantiate the efficacy of these models, a comprehensive dataset encompassing 810 UHPC sample compression test outcomes, previously chronicled in academic literature, is employed for training purposes. The scope of the input parameters is broad, encompassing a spectrum spanning from constituents like cement, slag, silica fume, and fly ash to more nuanced components like quartz powder, limestone powder, nano-silica, water, sand, gravel, fiber, superplasticizer, and factors like temperature, relative humidity, and age. The robustness of these models is assessed through performance metrics, including the root mean square error (RMSE), coefficient of determination (R^2), and mean absolute error (MAE). As an overarching outcome, the analysis expounds on the pivotal role of various input parameters in shaping UHPC's CS. This exploration is made possible through the application of Shap values analysis, which unveils the intricate interplay of these factors and their varying degrees of influence on UHPC's CS.

2. Research significance

This study is marked as a novel contribution to the field of UHPC compressive strength prediction due to several key factors. In contrast to traditional ML algorithms relied upon by other papers, this research capitalizes on advanced ML techniques such as XGBoost and its upgraded counterparts (Ada-XGBoost, Bagging-XGBoost, Stacking-XGBoost, and Voting-XGBoost). These cutting-edge algorithms exhibit heightened precision and efficacy in forecasting UHPC's compressive strength. The dataset utilized in this study comprises 810 experimental outcomes, encompassing 15 input parameters that span diverse UHPC components, conditions, and sample ages, thereby ensuring the broad applicability and robustness of the predictive models. Furthermore, a comprehensive comparative and evaluative analysis of these ML models is undertaken beyond the scope of individual methodologies. The introduction of ensemble techniques, exemplified by Stacking-XGBoost and Voting-XGBoost, serves to enhance prediction accuracy and stability, thus differentiating this research from conventional single-model approaches. To delve into the factors that exert influence on UHPC's compressive strength, a SHAP values analysis is executed. This interpretive methodology sheds light on the relative significance of each input parameter, guiding engineers in data-informed decision-making during UHPC mixture design. The precise anticipation of UHPC's compressive strength equips civil engineers to optimize concrete mix formulations and construct more resilient and high-performance UHPC structures, culminating in outcomes of greater impact. In summary, the distinctive amalgamation of advanced ML techniques, ensemble methodologies, SHAP values analysis, and practical engineering applications distinguishes this study from other works in the realm of UHPC compressive strength prediction.

3. Database description and analysis

The experimental findings from 28 previous studies serve as the basis for the database utilized in this investigation, with 810 samples of UHPC using various mixtures [2,13,27,32,33,67-81], and summarized by Mirani et al. [36]. The final dataset consists of 15 input characteristics, including 12 components of UHPC (i.e., cement, SF, slag, FA, quartz powder, limestone powder, nano silica, water, sand and gravel, fiber, and superplasticizer), two curing conditions (i.e., relative humidity and temperature), and age of samples. The CS of UHPC is the model's output. Table 1 and Fig. 1, respectively, show the data analysis statistics and input-output relationships.

Before developing the predictive model, the correlations between the inputs and outputs were established to visualize the data, as shown in Fig. 2. Herein, the Pearson correlations are used to identify which parameters have the greatest influence on one another for this purpose. A basic correlation matrix is first built to eliminate redundant information and pinpoint the most important traits. The Pearson correlation coefficient (r_s) measures two random variables' linear relationship. Herein, the linear Pearson correlation coefficient is described as a statistic representing the linear relationship between two normally distributed continuous variables [82]:

$$r_s = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (1)$$

where, \bar{x}, \bar{y} are the mean of inputs (x_i) and output (y_i), respectively. The r_s in Eq. (1) ranges from -1 to 1, which remains unchanged when the two variables are transformed linearly. When $r_s = 1$, there is a perfect positive correlation between x and y . When r_s is zero, it is not possible to confirm that x and y are linearly related. Lastly, $r_s = -1$ means that x and y are strong negative correlated. Regarding the results, compressive strength (Y) of UHPC was moderately correlated with X₃, X₁₁, X₁₂, and X₁₅ ($0.4 \leq r_s \leq 0.6$) but weakly correlated with X₁, X₅, X₇, and X₉ ($0.2 \leq r_s \leq 0.39$). The CS was nearly independent of X₂, X₄, X₆, X₈, X₁₀, X₁₃, and X₁₄ ($0 \leq r_s \leq 0.19$). Input and output parameters are therefore regarded as independent variables. In other words, any input variables can be used to construct ML models and further evaluate the feature's relevance.

4. Methods

4.1. Machine learning methods

Five ML models are suggested to estimate the CS of UHPC concrete. In which XGBoost is the original approach, two methods, Ada-XGBoost

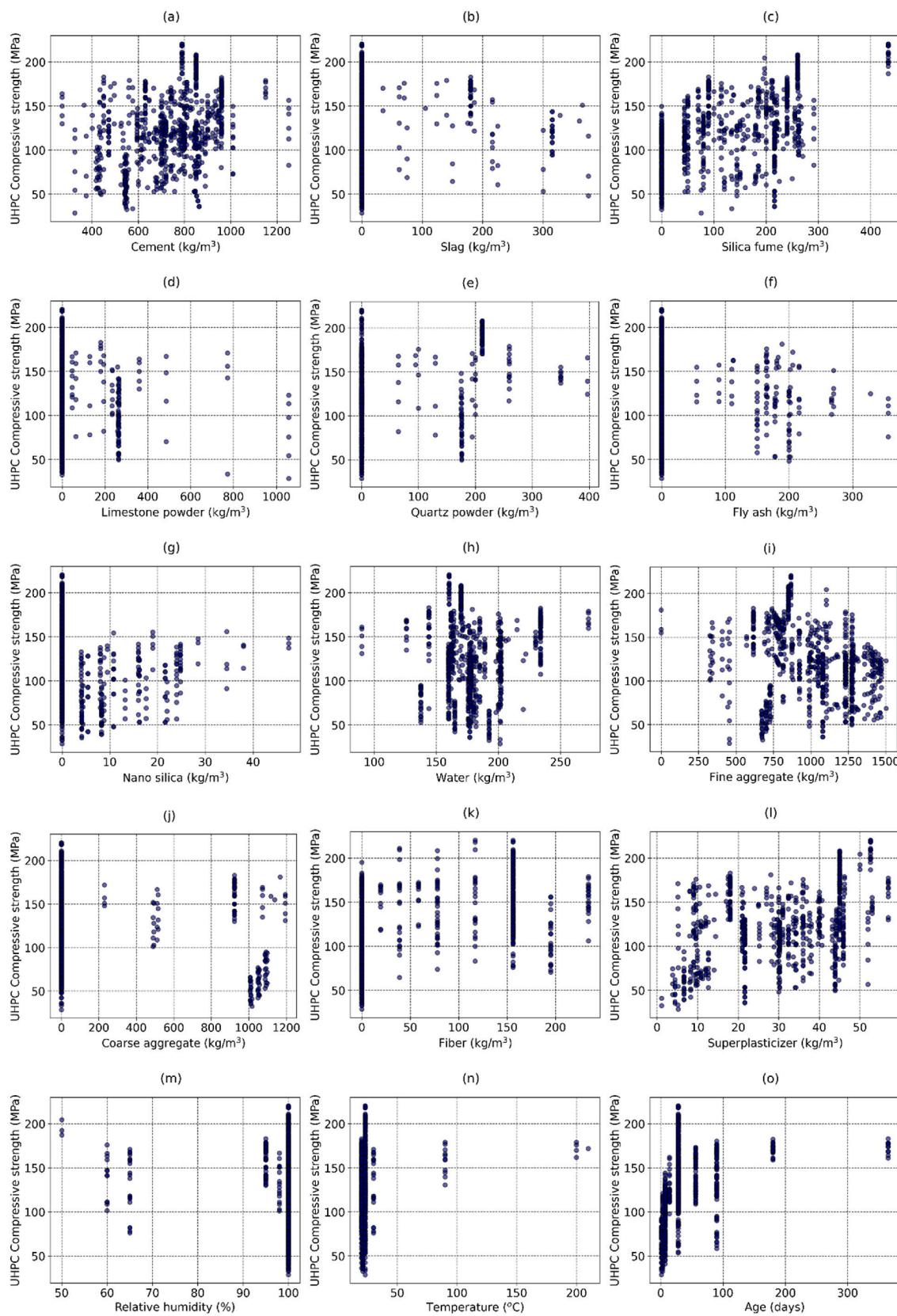


Fig. 1. Analysis of input–output relationships and distribution of variables.

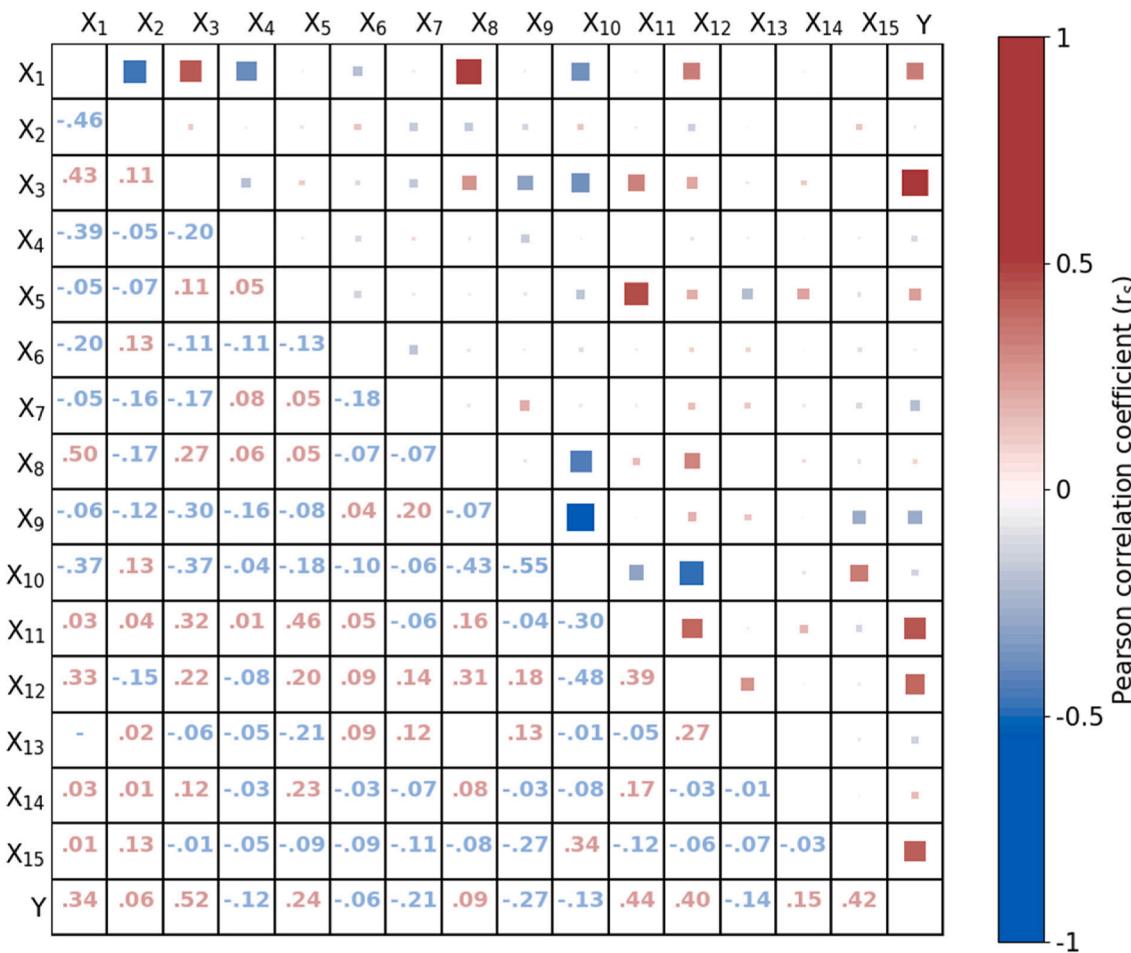
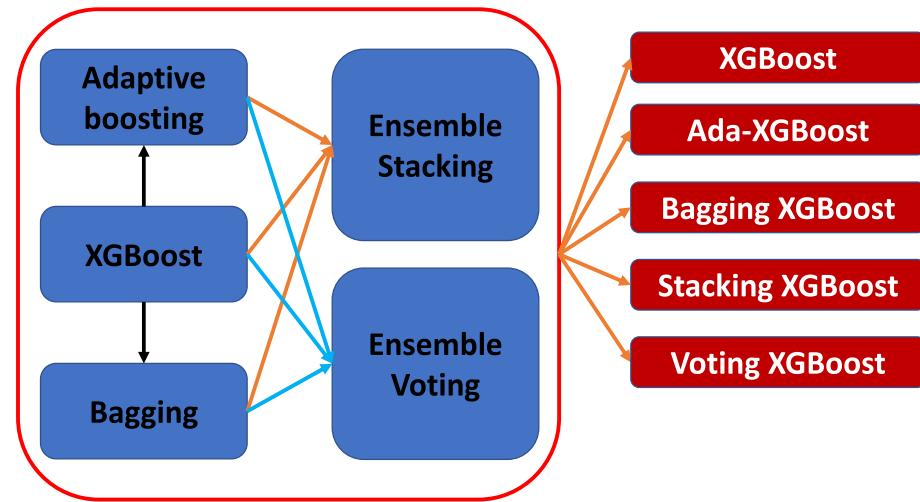
Fig. 2. Pearson correlation r of inputs and output of the current database.

Fig. 3. An overview of different modified XGBoost algorithms used in this study.

and Bagging-XGBoost, are enhanced from the original method, and two methods, Stacking-XGBoost and Voting-XGBoost, ensemble the three ML models mentioned above (Fig. 3). The following section summarizes the approaches' principles.

4.1.1. Xgboost algorithm

Chen and Guestin [66] presented XGBoost in 2016 based on the

fundamental premise of gradient tree boosting (GTB) algorithms, which combine several high-error decision trees (also known as base or weak learning trees) into a superior decision tree. XGBoost can create reinforcement trees quickly, operate in parallel, and solve regression and classification problems. This algorithm's core goal is to maximize the objective function's value. With parallel boost trees, XGBoost can tackle a wide range of data science issues rapidly and correctly. An XGBoost

Table 2
The selected hyper-parameters of ML models.

Model	Parameter settings
XGBoost	n_estimators = 100, max_depth = 6, learning_rate = 0.3, min_child_weight = 1, subsample = 1, sampling_method = uniform, grow_policy = depthwise, booster = gbtree,
Ada-XGBoost	base_estimator = XGBoost, n_estimators = 50, learning_rate = 1,
Bagging-XGBoost	base_estimator = modelXG, n_estimators = 10,
Voting-XGBoost	estimators = XGBoost + Ada-XGBoost + Bagging-XGBoost,
Stacking-XGBoost	estimators = XGBoost + Ada-XGBoost + Bagging-XGBoost, final_estimator = LinearRegression,

goal function is often divided into two pieces (training loss function and ruleset).

$$J(\Xi) = L(\Xi) + \Psi(\Xi) \quad (2)$$

where L represents the training process's loss function and Ψ is the decision tree's ruleset. The loss function, also called the error rate, is used to evaluate the model's progress while it is being trained. Rule sets are used to control the model's complexity and eliminate duplication or data gaps. The complexity of a model may be determined using a variety of ways. However, the following formula is commonly used to assess the complexity of each decision tree:

$$\Psi(\Xi) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (3)$$

where T denotes the leaves in decision tree, and w denotes the

decision tree leaf scores vector. The XGBoost enhancement method is based on the following objective function definition:

$$J(\Xi) = \sum_{i=1}^T \left[G_i w_j + \frac{1}{2} (H_j + \lambda) w_j^2 \right] + \gamma T \quad (4)$$

The objective of the XGBoost algorithm is to minimize the function $J(\Psi)$ so that the error rate is minimized.

4.1.2. Adaptive Boosting

Adaptive Boosting is a Boosting branch technique that generates many models of the same kind. This strategy's training procedure is sequential, with the straightforward concept of weighting data items through decision trees [83]. After the first tree is formed, the

Table 3

Summary of R^2 values for 25 MC simulations with 10-fold CV in training and validation datasets.

	Metrics	Dataset	Mean	Std	Min	Max
XG	R^2	Training	0.990	0.000	0.989	0.991
		Validation	0.953	0.003	0.945	0.958
Ada	R^2	Training	0.983	0.001	0.982	0.984
		Validation	0.945	0.002	0.940	0.951
Bagging	R^2	Training	0.983	0.001	0.982	0.984
		Validation	0.948	0.003	0.940	0.953
Stacking	R^2	Training	0.989	0.001	0.988	0.990
		Validation	0.953	0.002	0.946	0.958
Voting	R^2	Training	0.987	0.000	0.986	0.988
		Validation	0.952	0.002	0.946	0.956

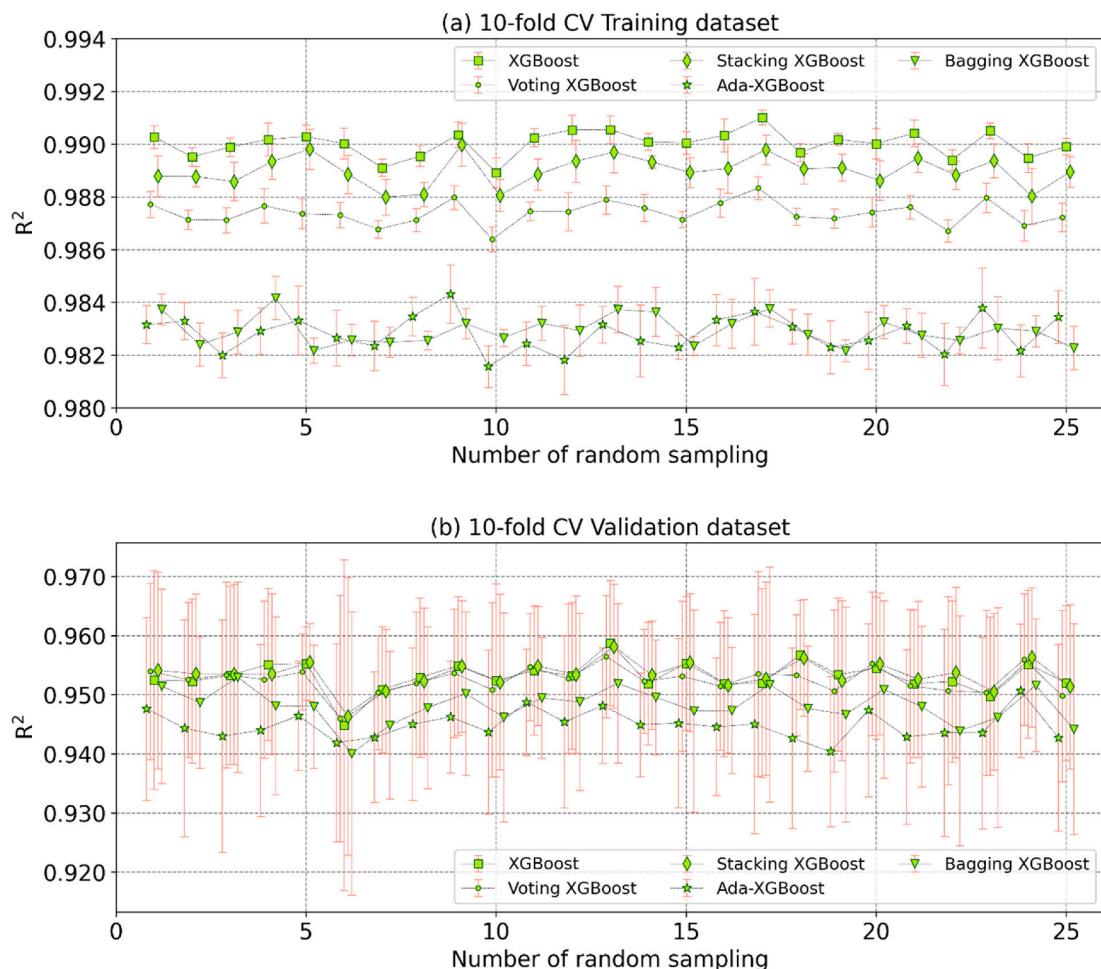


Fig. 4. Results of R^2 values for 25 MC simulations with 10-fold CV in training and validation datasets.

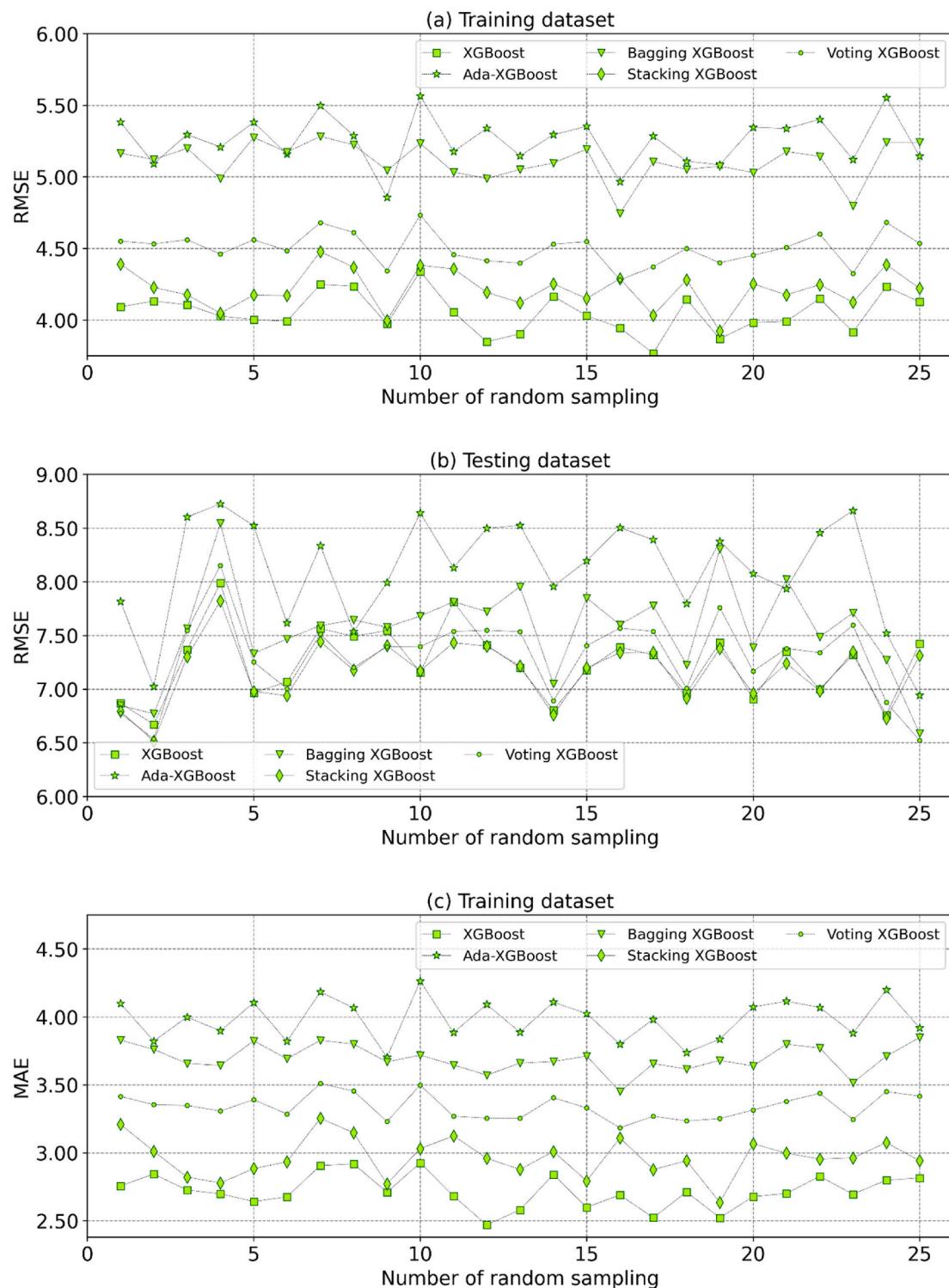


Fig. 5. Results of 25 MC results of training and validation parts (a, b) RMSE; (c, d) MAE; (e, f) R^2 .

performance of the tree per training sample is utilized to choose which training sample the subsequent tree should focus on. Training data that is harder to anticipate will be given a higher weight than more predictable data. Models are created sequentially, and their performance affects how they are constructed. Predictions are generated using the new data following all tree (models) construction. This time, the weighting of each tree is determined by its accuracy on the training data.

4.1.3. Bagging XGBoost

As a kind of ensemble learning, Bagging, or Bootstrap aggregation, improves the performance and accuracy of ML algorithms [84]. It is used to address trade-offs between bias and variance and decreases the variance of a prediction model. The Bagging technique's fundamental premise is to develop many models (often of the same kind) in parallel on distinct subsamples of the training data to create more accurate predictions. When new data is required for prediction, each model

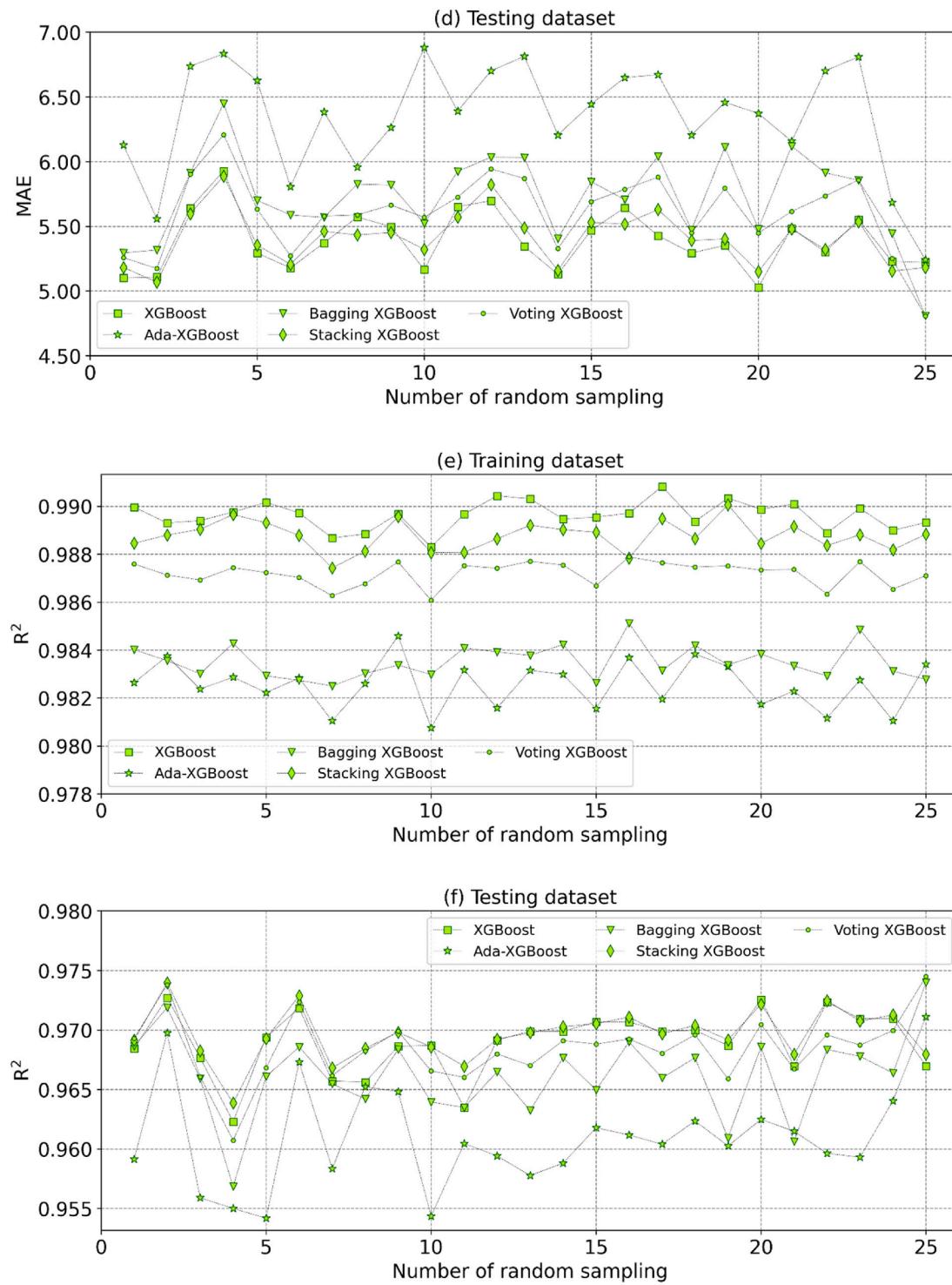


Fig. 5. (continued).

forecasts, averaging the predictions to provide a more accurate approximation of the actual output. Encapsulation is used to prevent data overfitting and applies to regression and classification models, most notably decision tree techniques.

4.1.4. Stacking XGBoost

Stacking is an example of an ensemble learning method that combines the results of many ML models trained on the same data. The strong point of stacking is that it may make use of the capacity of a diverse set of models to perform well in classification or regression tasks

and provide predictions that outperform all individual models within the ensemble [85]. When using a stacking technique, the best strategy to combine the predictions of many models is determined by a single supervisor model, the so-called *meta*-model. The *meta*-model is trained using the baseline models' out-of-sample predictions. Untrained data is provided to the base models, and predictions are generated. Then, the predictions and the predicted outputs comprise the input and output pairs for the training dataset required to construct the *meta*-model. As inputs to the *meta*-model, the outputs of the baseline models may be real values for regression or probabilistic values for classification, such as

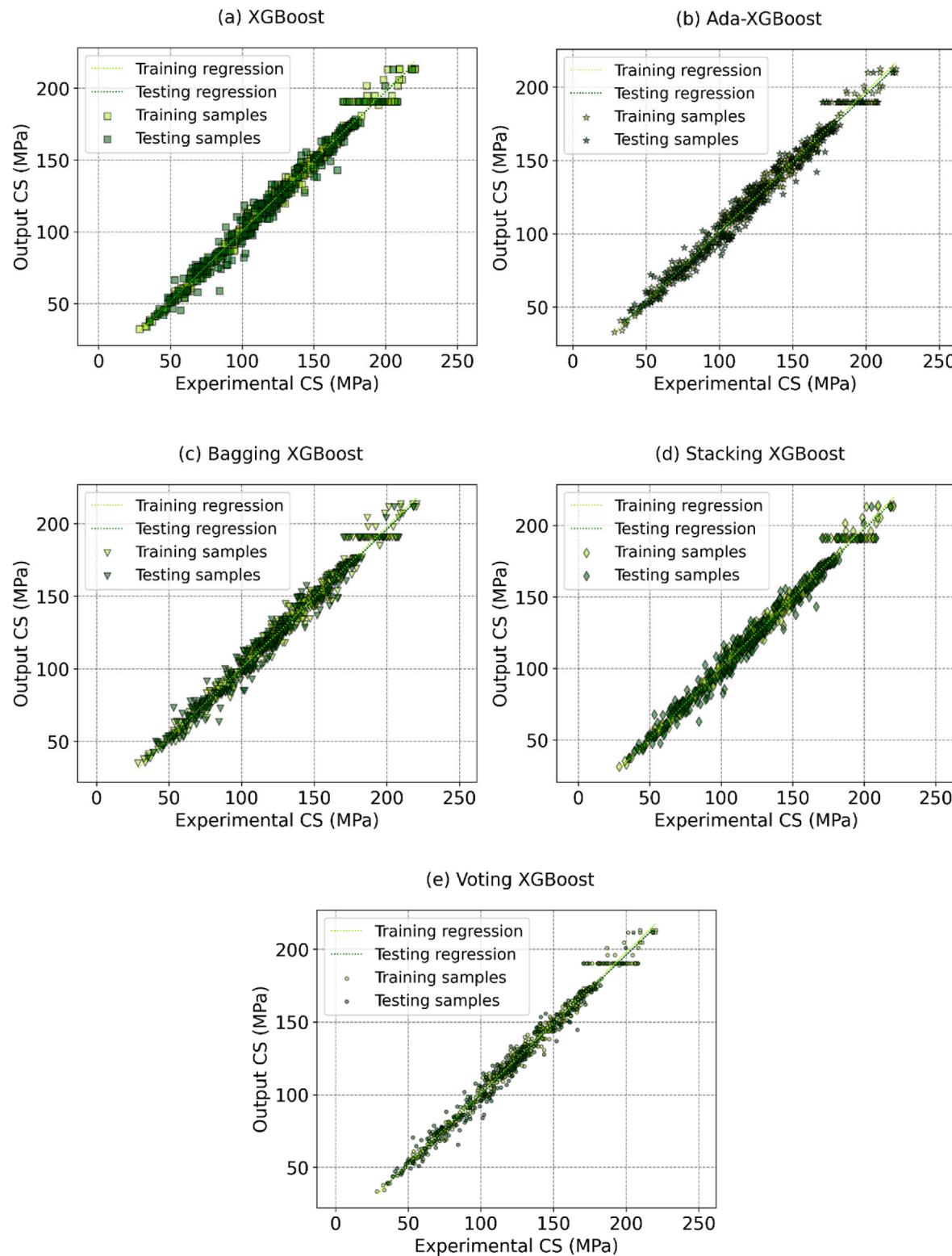


Fig. 6. Correlation between outputs and inputs CS of UHPC obtained from (a) XGBoost, (b) Ada XGBoost, (c) Bagging XGBoost, (d) Stacking XGBoost, and (e) Voting XGBoost.

actual values or class labels. The meta model's training data may also provide inputs for the basic model. This may offer context to the *meta*-model by indicating the optimal way to integrate the *meta*-model's predictions. Once the training set for the *meta*-model has been generated, the *meta*-model may be trained independently of the underlying models using the original training dataset in its entirety. Stacking is

suitable when a dataset has many distinct machine-learning models that are competent in various ways.

4.1.5. Voting XGBoost

Voting ensembles are the most straightforward technique of ensemble learning. Similar to Stacking, Voting is a collection of various

Table 4
Summary of prediction results on the training and testing sets.

Methods	Metrics	Mean	Std	Min	Max
XGBoost	Training dataset				
	RMSE	4.051	0.136	3.767	4.339
	MAE	2.716	0.121	2.468	2.925
	R ²	0.990	0.001	0.988	0.991
	Testing dataset				
	RMSE	7.238	0.320	6.671	7.988
Ada-XGBoost	MAE	5.387	0.220	5.028	5.928
	R ²	0.969	0.003	0.962	0.973
	Training dataset				
	RMSE	5.255	0.169	4.857	5.565
	MAE	3.982	0.148	3.702	4.263
	R ²	0.983	0.001	0.981	0.985
Bagging- XGBoost	Testing dataset				
	RMSE	8.111	0.486	6.941	8.725
	MAE	6.347	0.425	5.242	6.882
	R ²	0.961	0.004	0.954	0.971
	Training dataset				
	RMSE	5.107	0.132	4.744	5.282
Stacking- XGBoost	MAE	3.694	0.097	3.450	3.848
	R ²	0.984	0.001	0.983	0.985
	Testing dataset				
	RMSE	7.552	0.438	6.588	8.545
	MAE	5.728	0.338	4.808	6.449
	R ²	0.966	0.004	0.957	0.974
Voting- XGBoost	Training dataset				
	RMSE	4.216	0.132	3.924	4.477
	MAE	2.966	0.143	2.634	3.253
	R ²	0.989	0.001	0.987	0.990
	Testing dataset				
	RMSE	7.163	0.284	6.515	7.820
Voting- XGBoost	MAE	5.412	0.204	5.068	5.888
	R ²	0.970	0.002	0.964	0.974
	Training dataset				
	RMSE	4.500	0.112	4.279	4.733
	MAE	3.339	0.090	3.183	3.510
	R ²	0.987	0.000	0.986	0.988
Voting- XGBoost	Testing dataset				
	RMSE	7.298	0.373	6.523	8.152
	MAE	5.601	0.296	4.807	6.208
	R ²	0.968	0.003	0.961	0.974

models that make predictions on the same dataset, but rather than learning how to aggregate predictions through a *meta*-model, Voting employs a straightforward statistical technique [86]. Voting returns the mean or median of predictions from the regression problem-based models. In Voting, it is assumed that all base models are equal in relevance and perform equally well. Voting reduces the influence of irrelevant or duplicated features in the associative learning model. This is done by increasing the weights associated with individual classifiers performing well on test datasets. By minimizing the noise impact, Voting increases the efficiency of error prediction.

4.2. Cross-validation and random sampling

Cross-validation is a technique for determining the prediction performance of a model. Cross-validation is a word that relates to methods for segmenting a sample into several training and test datasets. K-folds are used to split the sample space. First, the original sample is divided into K equal (or nearly equal) subsamples at random. One subsample is kept as a validation set for error estimates, while the other K-1 subsamples are used for training. Then, K repetitions (folds) of cross-validation were performed, with each K subsample acting as the test set exactly once. The error estimates K for the folds may be averaged to give a single estimate. This technique benefits because all observations are used in the training and validation phases, whereas each observation is used only once for validation purposes. The fold (K) is often set to 5 or 10 to guarantee that the overall capacity for assessing the model's performance is not compromised. This study made use of the value K = 10, as suggested in [87]. Finally, the testing data from the original dataset is

separated into its part, and training (including model training and validation) makes no use of it.

Random subsampling isolates the data from the whole sample by dividing the data by K. For each data split, a fixed number of observations from the sample were kept as experimental data without replacement. Then, for each split K, the predictive model is trained from scratch to the training data, and the prediction error generated from each test set is calculated.

4.3. Shap values

A SHAP value is used in ML models to assign a tremendous influence to characteristics that may not have the same effect on predictions. The SHAP value is a notion from game theory used to quantify each participant's contribution in a coalition or cooperative game. In essence, the SHAP value shows the effect of having a particular value for a specific feature on the prediction we would make if the feature had acquired some underlying value. SHAP uses combinatorial calculus to evaluate the influence of each feature on the goal variable (dubbed the SHAP value) and retrains the model on all possible combinations of features [88]. The mean absolute value of a feature's influence on a target variable may be used to quantify its significance.

4.4. Model performance evaluation metrics

This study uses the following performance criteria to evaluate the performance and accuracy of ML models in predicting the CS of UHPC: R², RMSE, and MAE. The indicators can be calculated as the following Eqs. (5) to (7) [89].

$$R^2 = 1 - \frac{\sum_{i=1}^n (u_i - u_i)^2}{\sum_{i=1}^n (u_i - \bar{u})^2} \quad (5)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (u_i - u_i)^2} \quad (6)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |u_i - u_i| \quad (7)$$

where u_i and u_i are experimental and predicted CS of the ith UHPC samples, respectively; \bar{u} denotes the average of all experimental values; N is the total number of UHPC samples. Considering the proposed performance indicators, RMSE and MAE both attained their ideal value when equal to zero, whereas R^2 attained their best value when equal to one.

5. Results and discussion

5.1. Predictive capabilities study of ML models

The construction of a highly accurate and dependable model necessitates the meticulous selection of ML models to tackle the designated issues. Thus, every proposed model must be thoughtfully opted for, laying the groundwork for subsequent exploration. Various methodologies exist for the selection of models, each characterized by its own set of advantages and drawbacks. Among these approaches, K-fold cross-validation (K-fold CV) stands out as a widely adopted strategy within the realm of ML. In line with the guidance provided in reference [90], this study partitions the training dataset into 10 distinct folds. Moreover, to tackle challenges in civil engineering through the application of AI-ML, another frequently employed approach is the utilization of Monte Carlo (MC) cross-validation (CV) [91]. The main constraint associated with MC-CV lies in its inability to encompass an infinite array of dataset division permutations, as doing so would markedly escalate the number of iterations. In order to identify the optimal model, a synthesis of these two methodologies is implemented in this study. The amalgamation of the aforementioned strategies holds the potential to yield greater

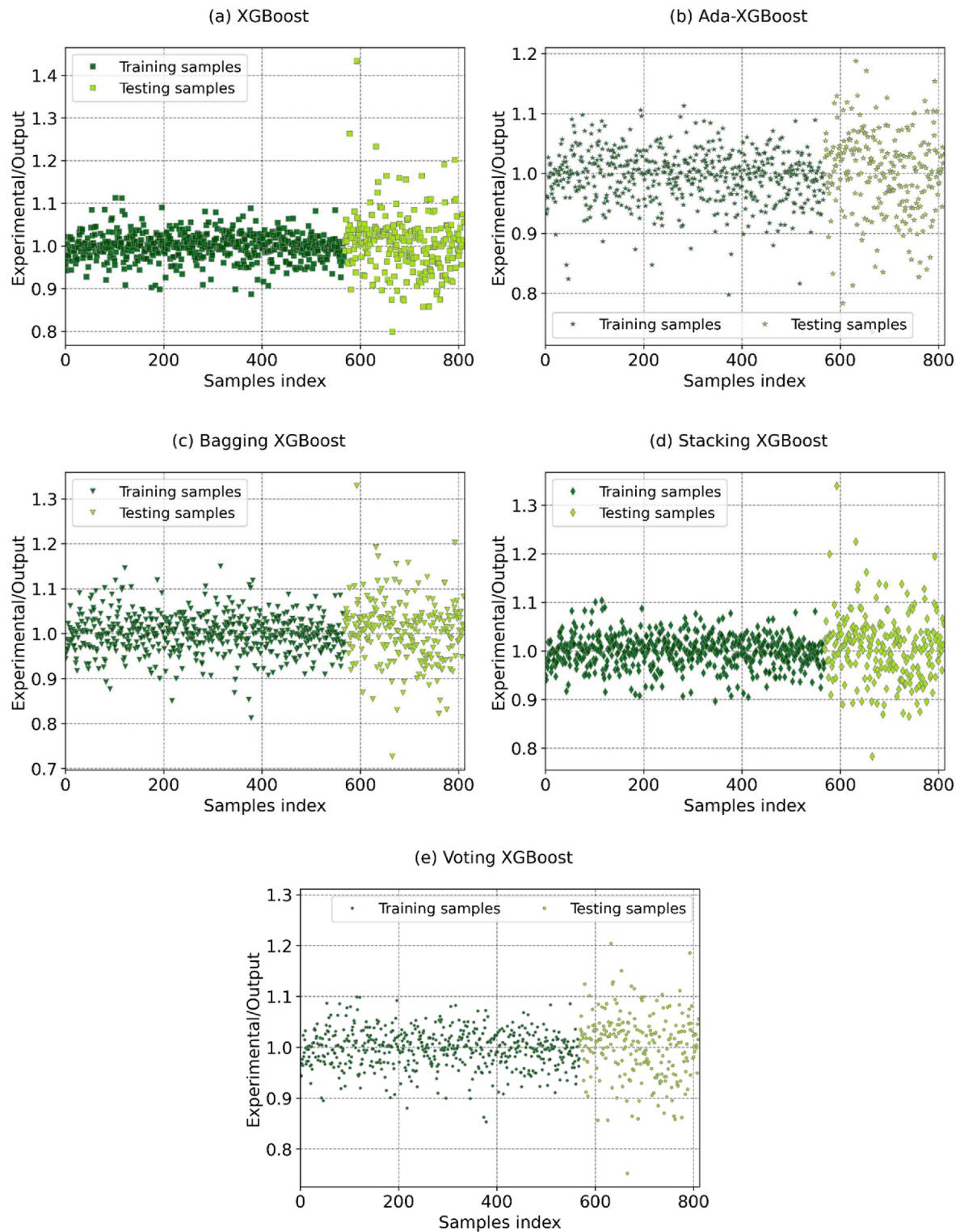


Fig. 7. Error prediction between outputs and inputs CS of UHPC obtained from (a) XGBoost, (b) Ada XGBoost, (c) Bagging XGBoost, (d) Stacking XGBoost, and (e) Voting XGBoost.

advantages and mitigate the individual drawbacks inherent to each. In this investigation, a total of 25 MC iterations are employed in conjunction with a 10-fold CV. For the sake of comparative analysis, evaluation metrics including R^2 , RMSE, and MAE are employed.

It is essential to take into account that although thorough trial and error tests were carried out, the hyper-parameters optimization procedure of the models is not shown in this paper. To increase the

prediction accuracy, the strategy was to start with the hyper-parameters of the model's default values and then fine-tune each parameter individually. The final selection of the model's hyper-parameters is detailed in Table 2. The training set (training and validation parts) and the testing part are assessed to demonstrate the dependability of the ML models. It should be highlighted that the testing part used to test the trained model is unknown during the model training process.

Table 5

The evaluation criteria for ML models' representative prediction results for the training and testing sets.

	Dataset	RMSE	MAE	R ²
XGBoost	Training	4.132	2.843	0.989
	Testing	6.671	5.109	0.973
Ada-XGBoost	Training	5.092	3.821	0.984
	Testing	7.023	5.557	0.970
Bagging XGBoost	Training	5.121	3.760	0.984
	Testing	6.771	5.317	0.972
Stacking XGBoost	Training	4.227	3.009	0.989
	Testing	6.515	5.068	0.974
Voting XGBoost	Training	4.531	3.353	0.987
	Testing	6.535	5.174	0.974

The mean and standard deviation values of R² of 10-fold CV are shown in Fig. 4, respectively, for the training and validation phases. They are also often called the training and validation scores corresponding to the training and validation phases. When considering the training scores, the prediction performance of five ML algorithms is very close to 1, suggesting that they work very well. Based on the R² training scores, the performance of ML models is ordered as XGBoost > Stacking-XGBoost > Voting-XGBoost > Bagging-XGBoost > Ada-XGBoost (see Fig. 4a).

It was noticed that the prediction performance of five ML models in the validation dataset, also known as validation scores, changed. Accordingly, the validation scores of the models are Stacking-XGBoost, XGBoost, Voting-XGBoost, Bagging-XGBoost, and Ada-XGBoost ranked from the best to the poorest predictor. The prediction results of R² for the training and validation scores over 25 runs of all models are also summarized (Table 3). With an average R² value of more than 0.95 for both training and validation scores, the prediction accuracy of the five ML models was determined to be relatively high.

Remarkably, after 25 trials, there is minimal variation in the maximum and lowest values of each assessment criteria, indicating that the Stacking-XGBoost model has highly stable predicting performance: R_{max}² = 0.958, R_{min}² = 0.946, R_{mean}² = 0.953, and R_{Std}² = 0.002. The obtained results illustrate the practicability and dependability of the Stacking-XGBoost model, validated by excellent scores and slight negligible variation.

The RMSE and MAE assessment criteria are also estimated to evaluate ML's performance more thoroughly. In addition, since the predictive performance of ML models is frequently determined by their ability to predict unknown data outside of the training phase, the next step is to evaluate how well the models perform on the testing dataset. It should be mentioned that the 10-fold CV was not applied in this evaluation round.

Fig. 6 illustrates the changes in five ML models' RMSE, MAE, and R²

values. In the training dataset, as can be seen, there is a high degree of concordance between the three evaluation criteria. The results of three indices validate the performance of five models, increasing from Ada-XGBoost, Bagging-XGBoost, Voting-XGBoost, Stacking-XGBoost, and XGBoost, as evidenced by low RMSE, MAE, and high R² values in Fig. 5a, c, and e, respectively. Regarding the testing set, the models' accuracy remains the same, and only the order changes between XGBoost and Stacking-XGBoost. As a result, the Stacking-XGBoost and XGBoost models' prediction abilities are consistent and close to one another. In other words, the results show that both the Stacking-XGBoost and XGBoost models are reliable and practical, as seen by the high scores and minimal variance. According to the significant criterion, which was confirmed after 25 simulations, the best model for predicting the CS of UHPC is Stacking-XGBoost, with RMSE_{mean} = 7.163 (MPa), MAE_{mean} = 5.412 (MPa), and R_{mean}² = 0.970. A summary of the prediction results of the training and testing sets of five models is presented in Table 4.

5.2. Representative prediction results

This section discusses the representative results obtained at the 16th iteration of each of the five models. Here, the results that were obtained from the models for the 16th run were based on its outstanding performance (i.e., high R² and low RMSE, and MAE values) observed in Fig. 5. The usual regression charts for the XGBoost, Ada-XGBoost, Bagging-XGBoost, Stacking-XGBoost, and Voting-XGBoost models are shown in Fig. 6a, b, c, d, and e, respectively.

In detail, the regression model illustrates the correlation between the target value (obtained from the experiment) and the value that was outputted from the ML model. The best line to draw is the one that passes across the spots where the predictions and experiment results coincide. The model's accuracy in making predictions improves, corresponding to the distance that separates the data from the fit line. As can be demonstrated, the value produced from any of the five ML models is rather close to the values intended to be achieved. In a closer view, the Ada-XGBoost model produces very few results that deviate significantly from the fit line (Fig. 6b). In other methods, in particular, XGBoost and Stacking-XGBoost, the outputs and targets are almost similar, and closely approach fit lines. These findings illustrate the capacity of XGBoost and Stacking-XGBoost to provide accurate prediction outputs by establishing a relationship between input and output parameters.

Fig. 7 shows the changes in relative errors, i.e., a ratio of Targets-to-Outputs of both training and validation datasets. All models perform excellently because most relative errors fall between 0.9 and 1.1. (i.e., -10%). Here, most mistakes in the training data set for each of the five models are small, proving the high training accuracy mentioned before. Compared to the other models, the Ada-XGBoost model predicts outcomes the least well, and its training and testing phases have a more

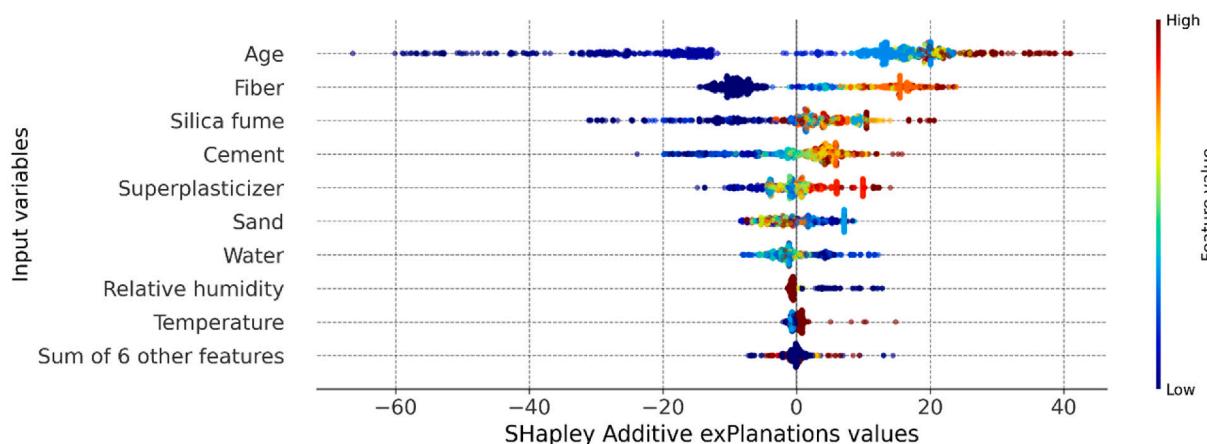


Fig. 8. Feature important analysis utilizing Beeswarm-SHAP value.

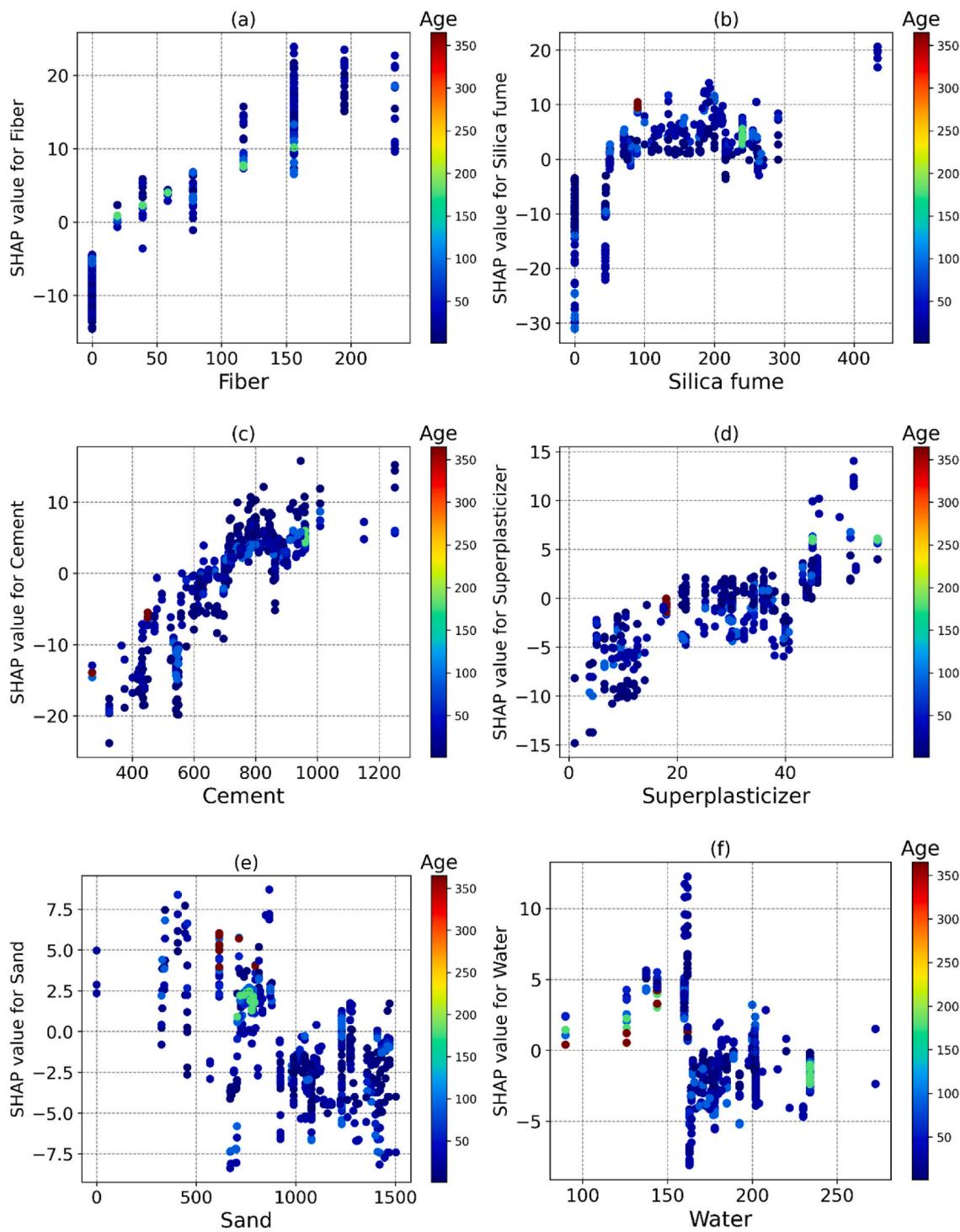


Fig. 9. SHAP dependence plots for Age of sample (X_{15}) and UHPC composition: (a) Fiber; (b) Silica fume; (c) Cement; (d) Superplasticizer; (e) Sand; and (f) Water.

comprehensive relative error range. The error ranges of the remaining four models are decreased in the following order: Stacking-XGBoost, XGBoost, Voting-XGBoost, and Bagging-XGBoost.

The XGBoost and Stacking-XGBoost models consistently disclose minor errors compared to the other models, except for a few atypical examples. The results demonstrate that the XGBoost and Stacking-XGBoost models beat the others in this testing dataset.

Finally, Table 5 calculates and shows the prediction metrics for five models. The fact that only one typical prediction result is displayed in this section should be noted. The training and testing datasets revealed that the Stacking-XGBoost and XGBoost models have outstanding

performance, as shown by the high R^2 value and the low RMSE and MAE values. In more detail, the calculated RMSE, MAE, and R^2 values for the Stacking-XGBoost model in the training dataset were 6.515, 5.068, and 0.974, while they were 6.671, 5.109, and 0.973 for the XGBoost model, respectively. It follows that Stacking-XGBoost and XGBoost are reliable techniques for correctly predicting the CS of UHPC.

5.3. Sensitivity analysis

According to the findings, Stacking-XGBoost is the most accurate model for determining the CS of UHPC, as was already mentioned in the

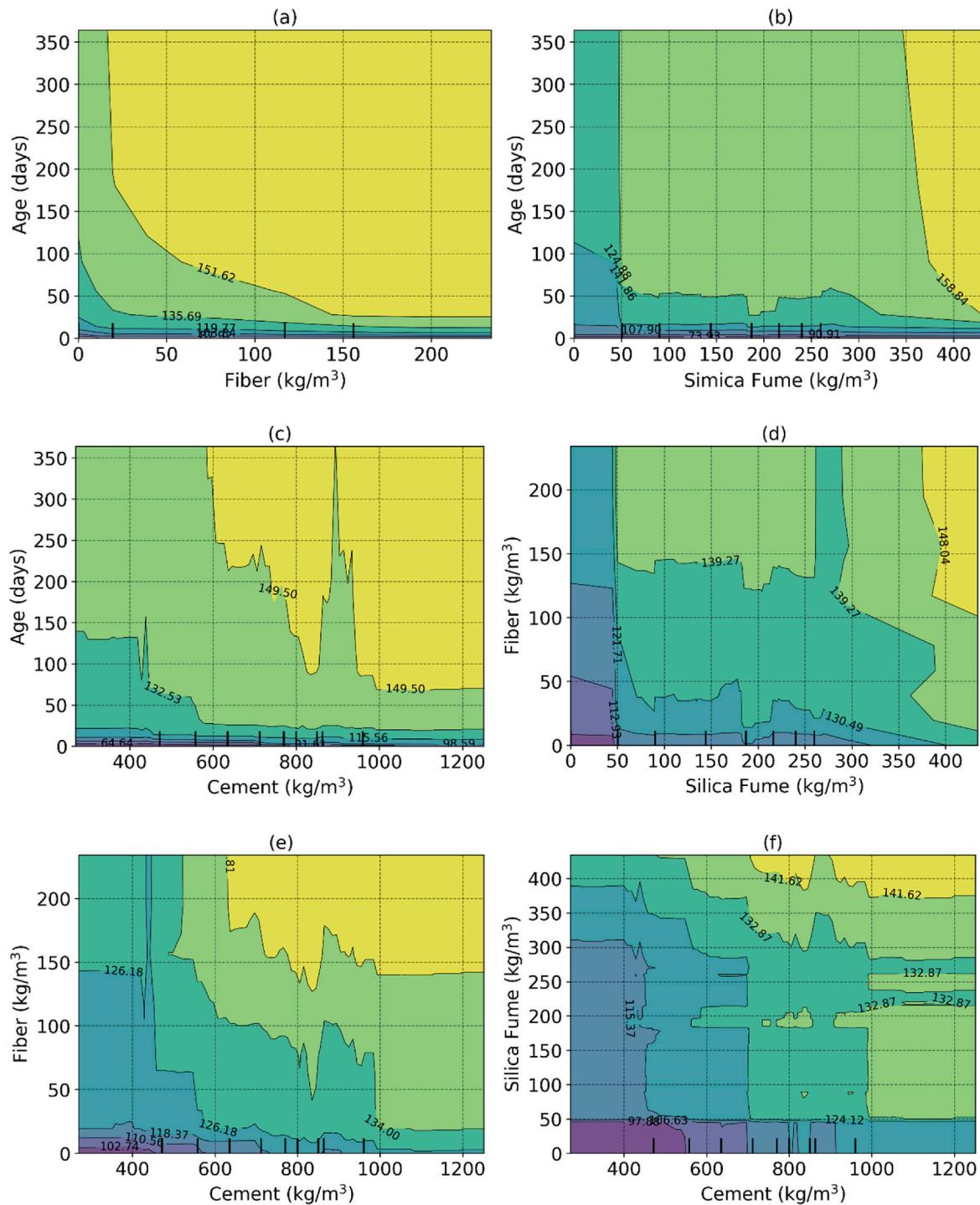


Fig. 10. Maps of UHPC's CS between different important components: (a) Age and fiber content; (b) Age and silica fume; (c) Age and cement; (d) Fiber and silica fume; (e) Fiber and cement; and (f) Silica fume and cement.

preceding discussion. In order to highlight the relevance of the roles that input variables play, Beeswarms-SHAP studies are used to determine the influence that the variables have on the Stacking-XGBoost model.

In Fig. 8, the influence of nine different input factors can be seen as having an effect on the CS of the UHPC. The SHAP value, in general, reflects two crucial pieces of information: the influence of the input variable on the CS of UHPC prediction accuracy via the y-axis and the impact of the input variable on the CS of UHPC via the x-axis. Both of these pieces of information are reflected by the scatter color.

Based on y-axis, the SHAP value for the first effect shows that the

most significant factor that affects the CS of UHPC is the age of the specimens. Precisely, Age > Fiber > Silica fume > Cement > Superplasticizer > Sand > Water > Relative humidity > Temperature are the input variables in descending order of their impact on the Stacking-XGBoost model's prediction accuracy. Six other variables have a negligible impact on UHPC's CS.

In addition, the SHAP value in the second information demonstrates the effect of each input variable on the magnitude of CS of UHPC. This can be seen based on the x-axis of the information. It is important to keep in mind that the zero values on the y-axis of the SHAP value represent

the mean value of the CS of UHPC, as shown in Table 1. Accordingly, the Age of the sample varied in the most extensive range, as shown by SHAP value ranging from -70 to 40, followed by Fiber, Silica fume, and Cement contents. The cement hydration products may affect the growth of CS in UHPC as it ages. Since the number of hydration products steadily increases throughout the hydration phases, resulting in a more compact microstructure and, therefore, improved CS [13,35]. Based on the color of scatters, SHAP values demonstrate that Age, Fiber, Silica fume, Cement, and Superplasticizer positively impact the CS of UHPC. When these inputs tend to rise, UHPC's CS increases. In contrast, relative humidity has a negative effect on the CS of UHPC. For the other parameters, the CS of UHPC is affected in a mixed manner (both positively and negatively).

Based on Fig. 8, the SHAP analysis results show that the sample's Age (X_{15}) significantly influences the Stacking-XGBoost model's ability. The role impact of the six following variables reduces from Fiber (X_{11}), Silica fume (X_3), Cement (X_1), Sand (X_9), Superplasticizer (X_{12}), and Water (X_8). The SHAP dependency plots method is used to understand better the influence that X_1 has on the prediction performance of the model as well as its connection with the six variables to follow. The SHAP dependence charts for Age (X_{15}) and other input parameters are shown in Fig. 9.

Based on the obtained results, it can be quantitatively discussed the combined effect of each two variables in the changes in CS of UHPC. For instance, when the amount of fiber ranges from 20 to 234 kg/m³, it discloses a positive effect on CS of UHPC, regardless of the sample age (Fig. 9a). At the age prior to 120 days, silica fume negatively impacts the CS of UHPC when its content is smaller than 50 kg/m³, but positively in the range from 50 to 434 kg/m³. The positive effect tends to be reduced when the silica fume content is within the range of 200–300 kg/m³ with small sample age (Fig. 9b). For cement, the obtained results show that CS of UHPC gradually increases when Cement rises from 400 to 1251 kg/m³ (see Fig. 9c). The similar tendency was also observed in Fig. 9d that proved the positive effect of Superplasticizer on the development of CS of UHPC with time. In contrast, the effects of sand and water on the CS of UHPC are negative, as shown by mix patterns in Fig. 9e and f, respectively. As can be seen, a content of more than 900 kg/m³ of sand, or 160 kg/m³ of water, negatively impacts the CS of UHPC.

5.4. Mapping of UHPC through different components

The crucial concern of addressing overfitting in soft computing models necessitates the implementation of rigorous measures to ensure the generalizability of the model. This is achieved by meticulously tuning hyperparameters and applying regularization techniques. In the present study, a series of graphs has been presented (Fig. 10), illustrating the relationships between pairs of variables. These graphs manifest smooth and consistent curves, indicative of the model's capacity to capture underlying trends without succumbing to noise or outliers, thereby averting overfitting [45,48,92–94]. Additionally, these graphs unveil the pronouncedly nonlinear nature of the underlying problem. Material engineers can leverage this insight to enhance the design of UHPC with the desired strength. As an illustration, the graph illustrating the interplay between age and fiber content demonstrates the substantial impact of fiber on concrete strength after 28 days, particularly when fiber content exceeds 145 kg/m³. Furthermore, the content of silica fume exhibits two discernible thresholds around 50 kg/m³ and 350 kg/m³, both exerting considerable influence on UHPC strength. These observations underscore the potential utility of ML models in estimating UHPC's compressive strength during the pre-design stage, affirming the robustness of this approach.

Nonetheless, it is pertinent to acknowledge that these figures solely underscore the potential and viability of ML models in predicting UHPC's CS. To further validate and augment these findings, a more expansive dataset should be amassed and subjected to analysis. The inclusion of supplementary data points will not only complement the

results obtained but will also contribute to the dependability and precision of the predictive models.

6. Conclusions

Despite the availability of several ML strategies for predicting the CS of UHPC, further research and development are still required to enhance the accuracy and robustness of ML algorithms. This study investigated the performances of XGBoost and its upgraded versions, namely, Ada-XGBoost, Bagging-XGBoost, Stacking-XGBoost, and Voting-XGBoost, in predicting the CS of UHPC. The dataset used for the models consisted of 810 experimental results encompassing 15 inputs, including 12 UHPC components, two curing conditions, and sample age. The CS of UHPC served as the output parameter. The performance criteria, including RMSE, MAE, and R², were assessed using 10-Fold CV and 25 MC simulations.

The obtained findings indicated that all XGBoost-modified models demonstrated good performance and stability, as evidenced by high R² and low RMSE and MAE values. Among the models, Stacking-XGBoost exhibited the highest predictive performance, followed by XGBoost, Voting-XGBoost, and Bagging-XGBoost. Particularly, the Stacking-XGBoost and XGBoost models showcased outstanding performance in both the training and testing datasets. The calculated RMSE, MAE, and R² values for Stacking-XGBoost in the training dataset were 6.515, 5.068, and 0.974, respectively, while those for XGBoost were 6.671, 5.109, and 0.973, respectively. Therefore, Stacking-XGBoost and XGBoost proved to be accurate methods for forecasting the CS of UHPC.

Additionally, based on the SHAP analysis, the age of specimens was found to have the most significant influence on the CS of UHPC. The Stacking-XGBoost model's prediction accuracy is affected by the following input variables in decreasing order: fiber, silica fume, cement, sand, superplasticizer, water, relative humidity, and temperature. Among these factors, Age, Fiber, Silica fume, Cement, and Superplasticizer positively impact the CS of UHPC, with the latter decreasing as these inputs tend to increase. Conversely, relative humidity negatively affects the CS of UHPC. The remaining factors exhibit inconsistent effects on UHPC's CS, with both positive and negative impacts. Detailed SHAP dependence charts were provided for Age (X_{15}) and six other input parameters: fiber (X_{11}), silica fume (X_3), cement (X_1), sand (X_9), superplasticizer (X_{12}), and water (X_8). The analysis of these charts enables civil engineers to quantitatively estimate the CS of UHPC in practical applications.

Availability of data and material: Data will be made available on request.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

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

References

- [1] Abuodeh OR, Abdalla JA, Hawileh RA. Assessment of compressive strength of Ultra-high Performance Concrete using deep machine learning techniques. *Appl Soft Comput* 2020;95:106552.
- [2] Yu R, Spiesz P, Brouwers HJH. Mix design and properties assessment of ultra-high performance fibre reinforced concrete (UHPFRC). *Cem Concr Res* 2014;56:29–39.
- [3] Zhang J, Zhao Y, Li H. Experimental investigation and prediction of compressive strength of ultra-high performance concrete containing supplementary cementitious materials. *Adv Mater Sci Eng* 2017;2017:1–8.
- [4] Graybeal B. Design and construction of field-cast UHPC connections, United States. *Federal Highway Administration* 2014.
- [5] S. Eppers, C. Müller, Autogenous shrinkage strain of ultra-high-performance concrete (UHPC), in: Proceedings of the 2nd International Symposium on UHPC, Kassel, Germany, 2008: pp. 433–441.

- [6] Yunsheng Z, Wei S, Sifeng L, Chujie J, Jianzhong L. Preparation of C200 green reactive powder concrete and its static–dynamic behaviors. *Cem Concr Compos* 2008;30(9):831–8.
- [7] Zheng W, Luo B, Wang Y. Compressive and tensile properties of reactive powder concrete with steel fibres at elevated temperatures. *Constr Build Mater* 2013;41: 844–51.
- [8] Pishro AA, Feng X. Experimental Study on Bond Stress between Ultra High Performance Concrete and Steel Reinforcement, Civil. Engineering Journal 2018;3: 1235–46.
- [9] Zhang X, Liu Z, Wang F. Autogenous shrinkage behavior of ultra-high performance concrete. *Constr Build Mater* 2019;226:459–68.
- [10] Bache, H.H., Densified Cement/Ultra-Fine Particle-Based Materials. Presented at the 2nd International Conference on Superplasticizers in Concrete, Ottawa, 10-12 June, Published by Aalborg Cement, PO Box 165, DK-9100 Aalborg, Denmark, 12p. 18 (1981) 33–39.
- [11] Zhou M, Lu W, Song J, Lee GC. Application of ultra-high performance concrete in bridge engineering. *Constr Build Mater* 2018;186:1256–67.
- [12] Gu ChunPing, Ye G, Sun W. Ultrahigh performance concrete—properties, applications and perspectives. *Sci China Technol Sci* 2015;58(4):587–99.
- [13] Rndl N, Steiner T, Ofner S, Baumgartner E, Mészöly T. Development of UHPC mixtures from an ecological point of view. *Constr Build Mater* 2014;67:373–8.
- [14] Wille K, Naaman AE, Parra-Montesinos GJ. Ultra-High Performance Concrete with Compressive Strength Exceeding 150 MPa (22 ksi): A Simpler Way. *ACI Mater J* 2011;108.
- [15] Abbas S, Nehdi ML, Saleem MA. Ultra-high performance concrete: Mechanical performance, durability, sustainability and implementation challenges, *International Journal of Concrete. Struct Mater* 2016;10:271–95.
- [16] Soliman NA, Tagnit-Hamou A. Using glass sand as an alternative for quartz sand in UHPC. *Constr Build Mater* 2017;145:243–52.
- [17] Chen T, Gao X, Ren M. Effects of autoclave curing and fly ash on mechanical properties of ultra-high performance concrete. *Constr Build Mater* 2018;158: 864–72.
- [18] Yazici H, Yardimci MY, Yiğiter H, Aydn S, Türköl S. Mechanical properties of reactive powder concrete containing high volumes of ground granulated blast furnace slag. *Cem Concr Compos* 2010;32(8):639–48.
- [19] Wu Z, Shi C, He W. Comparative study on flexural properties of ultra-high performance concrete with supplementary cementitious materials under different curing regimes. *Constr Build Mater* 2017;136:307–13.
- [20] An LH. Behavior of Circular Steel Tube Confined UHPC and UHPFRC Columns Under Axial Compression. Kassel University Press; 2018.
- [21] Hannawi K, Bian H, Prince-Agbodjan W, Raghavan B. Effect of different types of fibers on the microstructure and the mechanical behavior of ultra-high performance fiber-reinforced concretes. *Compos B Eng* 2016;86:214–20.
- [22] Graybeal BA. Compressive behavior of ultra-high-performance fiber-reinforced concrete. *ACI Mater J* 2007;104:146.
- [23] Graybeal B. UHPC making strides. *Public Roads* 2009;72:17–21.
- [24] Schmidt M, Fehling E. Ultra-high-performance concrete: research, development and application in Europe. *ACI Special Publication* 2005;228:51–78.
- [25] Yoo D-Y, Banthia N. Mechanical properties of ultra-high-performance fiber-reinforced concrete: A review. *Cem Concr Compos* 2016;73:267–80.
- [26] Wang C, Yang C, Liu F, Wan C, Pu X. Preparation of ultra-high performance concrete with common technology and materials. *Cem Concr Compos* 2012;34(4): 538–44.
- [27] Yu R, Spiesz P, Brouwers HJH. Effect of nano-silica on the hydration and microstructure development of Ultra-High Performance Concrete (UHPC) with a low binder amount. *Constr Build Mater* 2014;65:140–50.
- [28] Wang D, Shi C, Wu Z, Xiao J, Huang Z, Fang Z. A review on ultra high performance concrete: Part II. Hydration, microstructure and properties, *Construction and Building Materials* 2015;96:368–77.
- [29] Alsalam A, Dang CN, Hale WM. Development of ultra-high performance concrete with locally available materials. *Constr Build Mater* 2017;133:135–45.
- [30] Li W, Huang Z, Cao F, Sun Z, Shah SP. Effects of nano-silica and nano-limestone on flowability and mechanical properties of ultra-high-performance concrete matrix. *Constr Build Mater* 2015;95:366–74.
- [31] Li W, Huang Z, Zu T, Shi C, Duan WH, Shah SP. Influence of nanolimestone on the hydration, mechanical strength, and autogenous shrinkage of ultrahigh-performance concrete. *J Mater Civ Eng* 2016;28:04015068.
- [32] Zhang X, Zhao S, Liu Z, Wang F. Utilization of steel slag in ultra-high performance concrete with enhanced eco-friendliness. *Constr Build Mater* 2019;214:28–36.
- [33] Yang R, Yu R, Shui Z, Gao X, Xiao X, Zhang X, et al. Low carbon design of an ultra-high performance concrete (UHPC) incorporating phosphorous slag. *J Clean Prod* 2019;240:118157.
- [34] Larsen IL, Thorstensen RT. The influence of steel fibres on compressive and tensile strength of ultra high performance concrete: A review. *Constr Build Mater* 2020; 256:119459.
- [35] Arora A, Yao Y, Mobasher B, Neithalath N. Fundamental insights into the compressive and flexural response of binder-and aggregate-optimized ultra-high performance concrete (UHPC). *Cem Concr Compos* 2019;98:1–13.
- [36] Marani A, Jamali A, Nehdi ML. Predicting Ultra-High-Performance Concrete Compressive Strength Using Tabular Generative Adversarial Networks. *Materials* 2020;13:4757.
- [37] Nguyen HQ, Ly H-B, Tran VQ, Nguyen T-A, Le T-T, Pham BT. Optimization of Artificial Intelligence System by Evolutionary Algorithm for Prediction of Axial Capacity of Rectangular Concrete Filled Steel Tubes under Compression. *Materials* 2020;13:1205.
- [38] Murad Y, Ashteyat A, Hunafat R. Predictive model to the bond strength of FRP-to-concrete under direct pullout using Gene expression programming. *J Civ Eng Manag* 2019;25:773–84.
- [39] Pham TA, Ly H-B, Tran VQ, Giap LV, Vu H-L-T, Duong H-A-T. Prediction of Pile Axial Bearing Capacity Using Artificial Neural Network and Random Forest. *Appl Sci* 2020;10:1871.
- [40] Ly H-B, Monteiro E, Le T-T, Le VM, Dal M, Régnier G, et al. Prediction and sensitivity analysis of bubble dissolution time in 3D selective laser sintering using ensemble decision trees. *Materials* 2019;12:1544.
- [41] Wilcox C, Woon WL, Aung Z. Applications of machine learning in environmental engineering, Technical. Report 2013.
- [42] Ly H-B, Nguyen MH, Pham BT. Metaheuristic optimization of Levenberg–Marquardt-based artificial neural network using particle swarm optimization for prediction of foamed concrete compressive strength. *Neural Comput & Applic* 2021;33(24):17331–51.
- [43] Le T-H, Nguyen H-L, Pham BT, Nguyen MH, Pham C-T, Nguyen N-L, et al. Artificial Intelligence-Based Model for the Prediction of Dynamic Modulus of Stone Mastic Asphalt. *Appl Sci* 2020;10:5242. <https://doi.org/10.3390/app10155242>.
- [44] Tran VQ, Mai H-V, To QT, Nguyen MH. Machine learning approach in investigating carbonation depth of concrete containing Fly ash. *Structural Concrete n/a* 2023;24 (2):2145–69.
- [45] Armaghani DJ, Asteris PG. A comparative study of ANN and ANFIS models for the prediction of cement-based mortar materials compressive strength. *Neural Comput & Applic* 2021;33(9):4501–32.
- [46] Asteris PG, Armaghani DJ, Hatzigeorgiou GD, Karayannis CG, Pilakoutas K. Predicting the shear strength of reinforced concrete beams using Artificial Neural Networks, *Computers and Concrete, An. Int J* 2019;24:469–88.
- [47] Apostolopoulou M, Asteris PG, Armaghani DJ, Douvika MG, Lourenço PB, Cavalieri L, et al. Mapping and holistic design of natural hydraulic lime mortars. *Cem Concr Res* 2020;136:106167.
- [48] Asteris PG, Lourenço PB, Roussis PC, Adami CE, Armaghani DJ, Cavalieri L, et al. Revealing the nature of metakaolin-based concrete materials using artificial intelligence techniques. *Constr Build Mater* 2022;322:126500.
- [49] Barkhordari MS, Armaghani DJ, Mohammed AS, Ulrich DV. Data-driven compressive strength prediction of fly ash concrete using ensemble learner algorithms. *Buildings* 2022;12:132.
- [50] Sadegh Barkhordari M, Jahed Armaghani D, Asteris P. Structural damage identification using ensemble deep convolutional neural network models. *Comput Model Eng Sci* 2023;134(2):835–55.
- [51] Cavalieri L, Barkhordari MS, Repapis CC, Armaghani DJ, Ulrich DV, Asteris PG. Convolution-based ensemble learning algorithms to estimate the bond strength of the corroded reinforced concrete. *Constr Build Mater* 2022;359:129504.
- [52] Ly H-B, Nguyen T-A, Tran VQ. Development of deep neural network model to predict the compressive strength of rubber concrete. *Constr Build Mater* 2021;301: 124081.
- [53] Khademi F, Jamal SM, Deshpande N, Londhe S. Predicting strength of recycled aggregate concrete using artificial neural network, adaptive neuro-fuzzy inference system and multiple linear regression, *International Journal of Sustainable. Built Environ* 2016;5(2):355–69.
- [54] Duan Z-H, Kou S-C, Poon CS. Prediction of compressive strength of recycled aggregate concrete using artificial neural networks. *Constr Build Mater* 2013;40: 1200–6.
- [55] Sahoo S, Mahapatra TR. ANN Modeling to study strength loss of Fly Ash Concrete against Long term Sulphate Attack. *Mater Today: Proc* 2018;5:24595–604.
- [56] Topcu İB, Sarıdemir M. Prediction of compressive strength of concrete containing fly ash using artificial neural networks and fuzzy logic. *Comput Mater Sci* 2008;41 (3):305–11.
- [57] Golafshani EM, Pazouki G. Predicting the compressive strength of self-compacting concrete containing fly ash using a hybrid artificial intelligence method. *Comput Concr* 2018;22:419–37.
- [58] O. Abuodeh, J.A. Abdalla, R.A. Hawileh, Prediction of compressive strength of ultra-high performance concrete using SFS and ANN, in: 2019 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO), IEEE, 2019, pp. 1–5.
- [59] Abellán García J, Fernández Gómez J, Torres Castellanos N, Torres Castellanos, Properties prediction of environmentally friendly ultra-high-performance concrete using artificial neural networks. *Eur J Environ Civ Eng* 2022;26(6):2319–43.
- [60] Yeh I-C. Modeling of strength of high-performance concrete using artificial neural networks. *Cem Concr Res* 1998;28(12):1797–808.
- [61] Özcan F, Atış CD, Karahan O, Uncuoğlu E, Tanyıldız H. Comparison of artificial neural network and fuzzy logic models for prediction of long-term compressive strength of silica fume concrete. *Adv Eng Softw* 2009;40(9):856–63.
- [62] Dao DV, Ly H-B, Trinh SH, Le T-T, Pham BT. Artificial intelligence approaches for prediction of compressive strength of geopolymers concrete. *Materials* 2019;12:983.
- [63] Ghafari E, Bandarabadi M, Costa H, Júlio E. Design of UHPC using artificial neural networks. *Woodhead Publishing* 2012;303:61–9.
- [64] Ahmed Ramadan Suleiman, Moncef L. Nehdi, Modeling Self-Healing of Concrete Using Hybrid Genetic Algorithm–Artificial Neural Network, Department of Civil and Environmental Engineering, Western University, London, ON N6A 5B9, Canada. 10 (2) (2017).
- [65] Fan D, Yu R, Fu S, Yue L, Wu C, Shui Z, et al. Precise design and characteristics prediction of Ultra-High Performance Concrete (UHPC) based on artificial intelligence techniques. *Cem Concr Compos* 2021;122:104171.
- [66] Chen T, Guestrin C, XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016. p. 187.

- [67] Wang C, Yang C, Liu F, Wan C, Pu X. Preparation of Ultra-High Performance Concrete with common technology and materials. *Cem Concr Compos* 2012;34: 538–44. <https://doi.org/10.1016/j.cemconcomp.2011.11.005>.
- [68] Alsalmi A, Dang CN, Micah Hale W. Development of ultra-high performance concrete with locally available materials. *Constr Build Mater* 2017;133:135–45. <https://doi.org/10.1016/j.conbuildmat.2016.12.040>.
- [69] Wu Z, Shi C, Khayat KH, Xie L. Effect of SCM and nano-particles on static and dynamic mechanical properties of UHPC. *Constr Build Mater* 2018;182:118–25. <https://doi.org/10.1016/j.conbuildmat.2018.06.126>.
- [70] Liang X, Wu C, Su Y, Chen Z, Li Z. Development of ultra-high performance concrete with high fire resistance. *Constr Build Mater* 2018;179:400–12. <https://doi.org/10.1016/j.conbuildmat.2018.05.241>.
- [71] Yoo D-Y, Shin H-O, Yang J-M, Yoon Y-S. Material and bond properties of ultra high performance fiber reinforced concrete with micro steel fibers. *Compos B Eng* 2014; 58:122–33. <https://doi.org/10.1016/j.compositesb.2013.10.081>.
- [72] Yu R, Spiesz P, Brouwers HJH. Development of Ultra-High Performance Fibre Reinforced Concrete (UHPFRC): Towards an efficient utilization of binders and fibres. *Constr Build Mater* 2015;79:273–82. <https://doi.org/10.1016/j.conbuildmat.2015.01.050>.
- [73] Wille K, Boisvert-Cotulio C. Material efficiency in the design of ultra-high performance concrete. *Constr Build Mater* 2015;86:33–43. <https://doi.org/10.1016/j.conbuildmat.2015.03.087>.
- [74] Wu Z, Shi C, He W, Wang D. Static and dynamic compressive properties of ultra-high performance concrete (UHPC) with hybrid steel fiber reinforcements. *Cem Concr Compos* 2017;79:148–57. <https://doi.org/10.1016/j.cemconcomp.2017.02.010>.
- [75] Song Q, Yu R, Shui Z, Wang X, Rao S, Lin Z. Optimization of fibre orientation and distribution for a sustainable Ultra-High Performance Fibre Reinforced Concrete (UHPFRC): Experiments and mechanism analysis. *Constr Build Mater* 2018;169: 8–19. <https://doi.org/10.1016/j.conbuildmat.2018.02.130>.
- [76] Kang S-H, Jeong Y, Tan KH, Moon J. The use of limestone to replace physical filler of quartz powder in UHPFRC. *Cem Concr Compos* 2018;94:238–47. <https://doi.org/10.1016/j.cemconcomp.2018.09.013>.
- [77] Rajasekar A, Arunachalam K, Kottaismayam M. Assessment of strength and durability characteristics of copper slag incorporated ultra high strength concrete. *J Clean Prod* 2019;208:402–14. <https://doi.org/10.1016/j.jclepro.2018.10.118>.
- [78] Yoo D-Y, Kim M-J. High energy absorbent ultra-high-performance concrete with hybrid steel and polyethylene fibers. *Constr Build Mater* 2019;209:354–63. <https://doi.org/10.1016/j.conbuildmat.2019.03.096>.
- [79] Li Y, Tan KH, Yang E-H. Synergistic effects of hybrid polypropylene and steel fibers on explosive spalling prevention of ultra-high performance concrete at elevated temperature. *Cem Concr Compos* 2019;96:174–81. <https://doi.org/10.1016/j.cemconcomp.2018.11.009>.
- [80] Kang S-H, Hong S-G, Moon J. The use of rice husk ash as reactive filler in ultra-high performance concrete. *Cem Concr Res* 2019;115:389–400. <https://doi.org/10.1016/j.cemconres.2018.09.004>.
- [81] Ghafari E, Costa H, Júlio E, Portugal A, Durães L. The effect of nanosilica addition on flowability, strength and transport properties of ultra high performance concrete. *Mater Des* 2014;59:1–9. <https://doi.org/10.1016/j.matedes.2014.02.051>.
- [82] Zhou H, Deng Z, Xia Y, Fu M. A new sampling method in particle filter based on Pearson correlation coefficient. *Neurocomputing* 2016;216:208–15. <https://doi.org/10.1016/j.neucom.2016.07.036>.
- [83] Feng D-C, Liu Z-T, Wang X-D, Chen Y, Chang J-Q, Wei D-F, et al. Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach. *Constr Build Mater* 2020;230:117000.
- [84] Opitz D, Maclin R. Popular ensemble methods: An empirical study. *J Artif Intell Res* 1999;11:169–98.
- [85] Li Q, Song Z. Prediction of compressive strength of rice husk ash concrete based on stacking ensemble learning model. *J Clean Prod* 2023;382:135279.
- [86] Cai R, Han T, Liao W, Huang J, Li D, Kumar A, et al. Prediction of surface chloride concentration of marine concrete using ensemble machine learning. *Cem Concr Res* 2020;136:106164.
- [87] Valkheria V, Gujar R. Prediction of compressive strength and portland cement composition using cross-validation and feature ranking techniques. *Constr Build Mater* 2019;225:292–301.
- [88] Alabdullah AA, Iqbal M, Zahid M, Khan K, Amin MN, Jalal FE. Prediction of rapid chloride penetration resistance of metakaolin based high strength concrete using light GBM and XGBoost models by incorporating SHAP analysis. *Constr Build Mater* 2022;345:128296.
- [89] Asadi Shamsabadi E, Roshan N, Hadigheh SA, Nehdi ML, Khodabakhshian A, Ghalehnoei M. Machine learning-based compressive strength modelling of concrete incorporating waste marble powder. *Constr Build Mater* 2022;324: 126592. <https://doi.org/10.1016/j.conbuildmat.2022.126592>.
- [90] R. K, A study of cross-validation and bootstrap for accuracy estimation and model selection, Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, 1995. 2 (1995) 1137–1143.
- [91] Xu Q-S, Liang Y-Z. Monte Carlo cross validation. *Chemom Intel Lab Syst* 2001;56: 1–11. [https://doi.org/10.1016/S0169-7439\(00\)00122-2](https://doi.org/10.1016/S0169-7439(00)00122-2).
- [92] Asteris PG, Lourenço PB, Hajihassani M, Adami C-E-N, Lemonis ME, Skentou AD, et al. Soft computing-based models for the prediction of masonry compressive strength. *Eng Struct* 2021;248:113276.
- [93] Asteris PG, Maraveas C, Chountalas AT, Sophianopoulos DS, Naveed Alam N. Fire resistance prediction of slim-floor asymmetric steel beams using single hidden layer ANN models that employ multiple activation functions. *Steel Compos Struct* 2022;44:755–74.
- [94] Asteris PG, Alexakis DE, Tsoukalas MZ, Gamvroula DE, Guney D. Machine learning approach for rapid estimation of five-day biochemical oxygen demand in wastewater. *Water* 2022;15:103.