Computer Modeling in Engineering & Sciences

Tech Science Press

https://doi.org/10.32604/cmes.2026.084489

ARTICLE

Strength Prediction of Ultra-High Performance Concrete (UHPC) Based on BOHB-XGBOOST Algorithm

Ling Wang[1,2], **Mohammad Faizuddin Md Noor**[2,*] **and Yanan Zhang**[3,*]

[1]School of Artificial Intelligence and Electronic Information, Nantong Vocational University, Nantong, China
[2]Malaysian Institute of Information Technology, Universiti Kuala Lumpur, Kuala Lumpur, Malaysia
[3]College of Water Conservancy and Hydropower Engineering, Hohai University, NanJing, China

*Corresponding Authors: Mohammad Faizuddin Md Noor. Email: mfaizuddin@unikl.edu.my;
Yanan Zhang. Email: zynl0615@163.com

Received: 23 April 2026; Accepted: 27 July 2026; Published: 28 August 2026

ABSTRACT: Ultra-high-performance concrete (UHPC) relies on multivariable mix design and curing regimes, which makes empirical estimation of compressive strength increasingly unreliable when material systems vary. In this context, unlike previous studies that only applied standard eXtreme Gradient Boosting (XGBoost), this study introduces an advanced hybrid optimization strategy, Bayesian Optimization and Hyperband (BOHB), which combines the sample efficiency of Bayesian optimization with the resource allocation mechanism of Hyperband, and incorporates SHapley Additive exPlanations (SHAP) for influencing factor analysis, thereby proposing a BOHB-XGBoost framework integrated with SHAP analysis. The proposed model demonstrates excellent predictive accuracy and stability, achieving a coefficient of determination R^2 of 0.967 and a Mean Absolute Percentage Error (MAPE) of 4.86% on the test set. Comparative experimental results indicate that this model shows improved and more stable performance than mainstream machine learning methods, including standard XGBOOST, Random Forest, Gradient Boosting Decision Trees, Support Vector Regression and a Multilayer Perceptron across multiple evaluation metrics. This fully demonstrates the significant advantages of the BOHB optimization framework in processing structured tabular data. By balancing exploration and exploitation during the hyperparameter search process, this strategy effectively overcomes the limitations of traditional manual tuning and significantly enhances the model's generalization capability. Furthermore, the SHAP-based interpretability analysis identifies the key factors influencing UHPC strength, such as the water-to-cement ratio and steel fiber content. These findings are highly consistent with the experimental results conducted in this study, thereby validating the model's physical plausibility and engineering applicability. This research not only provides an efficient and reliable method for predicting UHPC strength but also offers a transparent and interpretable scientific basis for the rational optimization of mix proportion design.

KEYWORDS: Ultra-high performance concrete (UHPC); BOHB-XGBOOST; compressive strength; predictive modeling; machine learning; SHAP; mix design

1 Introduction

Ultra-high-performance concrete (UHPC) has been developed to achieve high mechanical resistance and improved durability, and its application has expanded across structural and infrastructure projects [1–3]. However, the material cost remains non-negligible, and the adjustment of mixture constituents and curing regimes continues to be a practical constraint in engineering deployment [4,5]. Among the properties used to assess mixture design, compressive strength is commonly treated as a primary indicator for performance evaluation and process control. Understanding how compositional variables and preparation procedures govern strength development, and establishing quantitative prediction tools, is therefore of interest for mixture optimization.

Recent studies have incorporated magnetite [6], recycled constituents [7], and other replacement strategies to balance performance demands and economic considerations. These investigations indicate that strength is shaped by both macroscopic factors such as silica-based additions and fiber reinforcement [6–8] and microstructural evolution arising from particle packing, interfacial transition zones, and hydration reactions. Variables including fiber fraction, aspect ratio, and incorporation method, as well as the dosage and reactivity of silica-based powders, have been shown to modify strength responses in non-linear fashions [8–10]. Microstructural observations, enabled by Low-Field Nuclear Magnetic Resonance (LF-NMR), Backscattered Electron (BSE), X-Ray Diffraction (XRD), Scanning Electron Microscopy (SEM) and related techniques, have clarified how fibers and supplementary cementitious materials affect pore connectivity, crack-bridging, and packing density. The interaction between mixture composition and microstructure introduces multi-scale coupling, which complicates attempts to associate single variables with macroscopic strength outcomes [11,12]. Empirical strength equations remain in use because of their interpretability, and high coefficients of determination are often attainable when the formulation domain is narrow (e.g., fixed curing or porosity ranges). Nevertheless, such relations rarely extend across different material systems, and their reliance on a limited set of descriptors restricts applicability when mixtures incorporate multiple fine-scale adjustments. Therefore, to extrapolate across a broader range of UHPC formulations, modeling approaches that combine several variables and heterogeneous datasets are needed. In this context, data-driven regression including ensemble machine learning and neural networks has already been used to predict mechanical performance in cement-based materials [13–15]. More recently, optimized interpretable deep learning models have demonstrated their effectiveness for compressive strength prediction of geopolymer and rubberized concrete [16–18], highlighting the growing role of hybrid optimization and attention mechanisms. Complementing these purely data-driven efforts, Discrete Element Method-driven investigation and Automated Machine Learning-enhanced prediction for cemented composites [19,20] incorporate physics-informed features, further enriching the modeling toolkit. These methods, however, generally operate on structured compositional tables rather than image-based microstructural information, and classical algorithms have difficulty integrating high-dimensional image data without additional feature extraction. Tree-based learning algorithms such as eXtreme Gradient Boosting (XGBOOST) offer an advantage when only tabulated mixture descriptors are available: they process heterogeneous variables, include regularization mechanisms that limit overfitting, and remain effective with modest sample sizes. For UHPC research, experimental datasets typically comprise only a few hundred mixture samples. These properties are particularly relevant. In addition, XGBOOST has been adopted in other engineering domains, where structured regression problems dominate.

Unlike previous studies that merely applied standard XGBoost, the study innovatively adopts the Bayesian Optimization and Hyperband (BOHB) strategy for hyperparameter tuning and employs SHAP analysis for result interpretation. This strategy integrates the sample efficiency of Bayesian optimization with the resource allocation mechanism of Hyperband and incorporates SHAP for analyzing influencing factors, thereby proposing an integrated BOHB-XGBOOST framework incorporating SHAP analysis. The findings lay the groundwork for the future integration of macro- and micro-scale input data and provide a basis for applying machine learning to mixture design and optimization under limited-sample conditions.

2 Data Sources and Development of the Prediction Model

To support the development of an XGBOOST regression framework for UHPC compressive strength, this section outlines the data sources, variable preprocessing procedures, and configuration steps used in model construction and evaluation. The aim is to clarify how the training and testing datasets were assembled and how key hyperparameters and implementation choices were determined to ensure reproducible performance.

2.1 Data Sources and Processing

2.1.1 Data Sources

To examine whether the XGBoost regression model can reproduce the strength variation across different ultra-high performance concrete (UHPC) mixture designs, this study compiled 838 mixture records from published experiments in Table 1 [21–63]. These reports provide mixtures with varying binder compositions, silica-based additives, fiber dosages, and curing durations. Although the resulting dataset does not follow a single design protocol, its heterogeneity provides a basis for testing the model's performance under a range of mixture configurations. Based on the initially compiled 838 records, a rigorous data screening procedure was implemented: a record was excluded if it was missing any of the nine core input variables, if the reported compressive strength fell outside the physically reasonable range of 20–250 MPa, or if there was a clear inconsistency between the mixture proportions and the reported strength value. After this screening, 780 valid mixture records were retained for model development and evaluation. It should be noted that the compiled data originate from studies with varying experimental protocols, such as different curing temperatures, loading rates, and specimen preparation methods. Such heterogeneity may introduce hidden biases, and the model's generalization across unseen protocols should be further validated.

Table 1: Data distribution of compressive strength for UHPC.

Data Source	Numbers of Data Sets	Data Source	Numbers of Data Sets	Data Source	Numbers of Data Sets
Tian et al. [21]	4	Wang et al. [36]	2	Fan a et al. [51]	9
Liu et al. [22]	9	Ocelić et al. [37]	24	Lin et al. [52]	8
Zhao et al. [23]	6	Zhang et al. [38]	8	Wang et al. [53]	10
Wang et al. [24]	12	Hiew et al. [39]	12	Zhang et al. [54]	3
Xue et al. [25]	9	Khaksefidi et al. [40]	6	Wu et al. [55]	12
Tang et al. [26]	4	Wang et al. [41]	6	Qian et al. [56]	15
Yu et al. [27]	48	Jiao et al. [42]	30	Zhang et al. [57]	15
Hassan and Wille [28]	4	Wang et al. [43]	12	Sadrmomtazi et al. [58]	42
Wang et al. [29]	3	Ouyang et al. [44]	10	Zhang et al. [59]	12
Reddy and Ramadoss [30]	18	Arora et al. [45]	12	Yang et al. [60]	48
Bae and Pyo [31]	15	Alkaysi and El-Tawil [46]	26	Xuan et al. [61]	16
Yang et al. [32]	18	Sun et al. [47]	18	Guo et al. [62]	12
Fan et al. [33]	36	Abuodeh et al. [48]	110	Luo and Matsumoto [63]	26
Ragalwar et al. [34]	60	Smarzewski [49]	10		
Liu et al. [35]	52	Li et al. [50]	26		

2.1.2 Data Processing

Because the reviewed UHPC studies employ different binders, sands, silica-based additives and admixtures, treating each raw material as an explicit input would confine the model to a particular material system and reduce its transferability. To avoid this, the present work represents mixture composition using dimensionless ratios that follow standard mix-design practice. This approach allows mixtures based on different commercial sources to be compared on a common basis.

Cement, water, silica fume, silica sand and superplasticizer were therefore expressed relative to cement mass and used as the core compositional descriptors. Fiber reinforcement, which contributes to strength through crack bridging and load transfer, was included through the volume fractions of steel fibres and other fibres such as carbon. Although geometry and aspect ratio also influence fibre efficiency, systematic reporting is limited in the available literature, which prevents morphology from being treated as a standalone predictor. Curing age, given its direct influence on hydration and strength development, was retained as an independent variable. Specimen geometry (cube vs. cylinder) was also encoded, since published strength values depend on the testing form. These predictors were standardised prior to training to minimise numerical scale effects.

Variables such as fiber aspect ratio and aggregate gradation are known to influence strength but were reported inconsistently across the source studies which are available in less than 20% of records. Their exclusion is a limitation of the current model, and future datasets should include them for improved prediction.

In total, nine predictors were used: the five cement-normalised compositional ratios, the two fibre volume fractions, curing age and specimen shape. Together, these form the nine-dimensional input space for model training and evaluation. It is worth emphasizing that these input variables exhibit significant differences in dimensional scales. Directly applying them to model training could lead to degraded predictive performance. Therefore, all continuous input variables were standardized. To prevent data leakage between training and evaluation phases, the standardization parameters (μ and σ) were computed exclusively from the training set in each cross-validation fold or train-test split. These same parameters were then applied to standardize the corresponding validation or test samples. For a given sample, the standardization was carried out using the following formula:

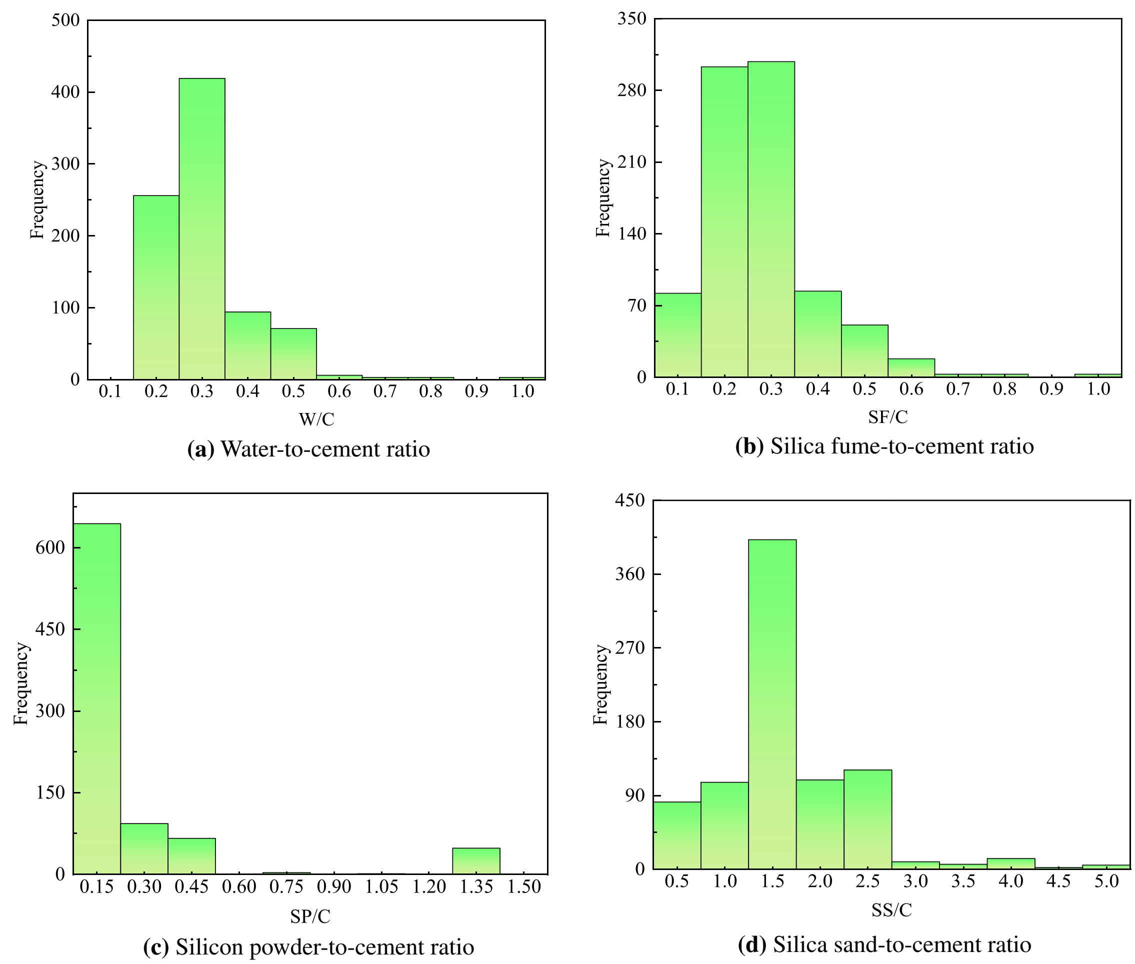
$$Z_i = \frac{x_i - \mu}{\sigma} \tag{1}$$

In the standardization step, Z_i denotes the transformed value, x_i is the original entry, and μ and σ are the mean and standard deviation estimated from the training data only. Categorical variables, specifically specimen shape (0 = cube, 1 = cylinder), were not standardized but directly fed into the model as integer indicators.

To illustrate the procedure, one record in the dataset contained the following inputs: curing age (CA), carbon-fiber volume fraction (CFV), specimen geometry, the water-to-cement ratio (W/C), silica sand-to-cement ratio (SS/C), silicon powder-to-cement ratio (SP/C), silica fume-to-cement ratio (SF/C), the superplasticizer-to-cement ratio (S/C) and steel fiber volume fraction (SFV). The raw values and the corresponding means and standard deviations used for scaling are reported in Table 2 and Fig. 1, and substitution into Eq. (1) yields the standardized set employed for model training.

Table 2: Example of data standardization processing.

Features	Maximum	Minimum	Mean	Standard Deviation
W/C	1	0.1092	0.2655	0.1021
SF/C	1	0	0.2360	0.1439
SP/C	1.25	0	0.1362	0.1621
SS/C	4.95	0	1.3956	0.8573
S/C	0.2816	0.0009	0.04	0.0282
CA (d)	730	1	24.39	31.85
SFV (%)	4	0	0.6353	1.0163
CFV (%)	3	0	0.0891	0.3842
Shape	1	0	0.1064	0.3084



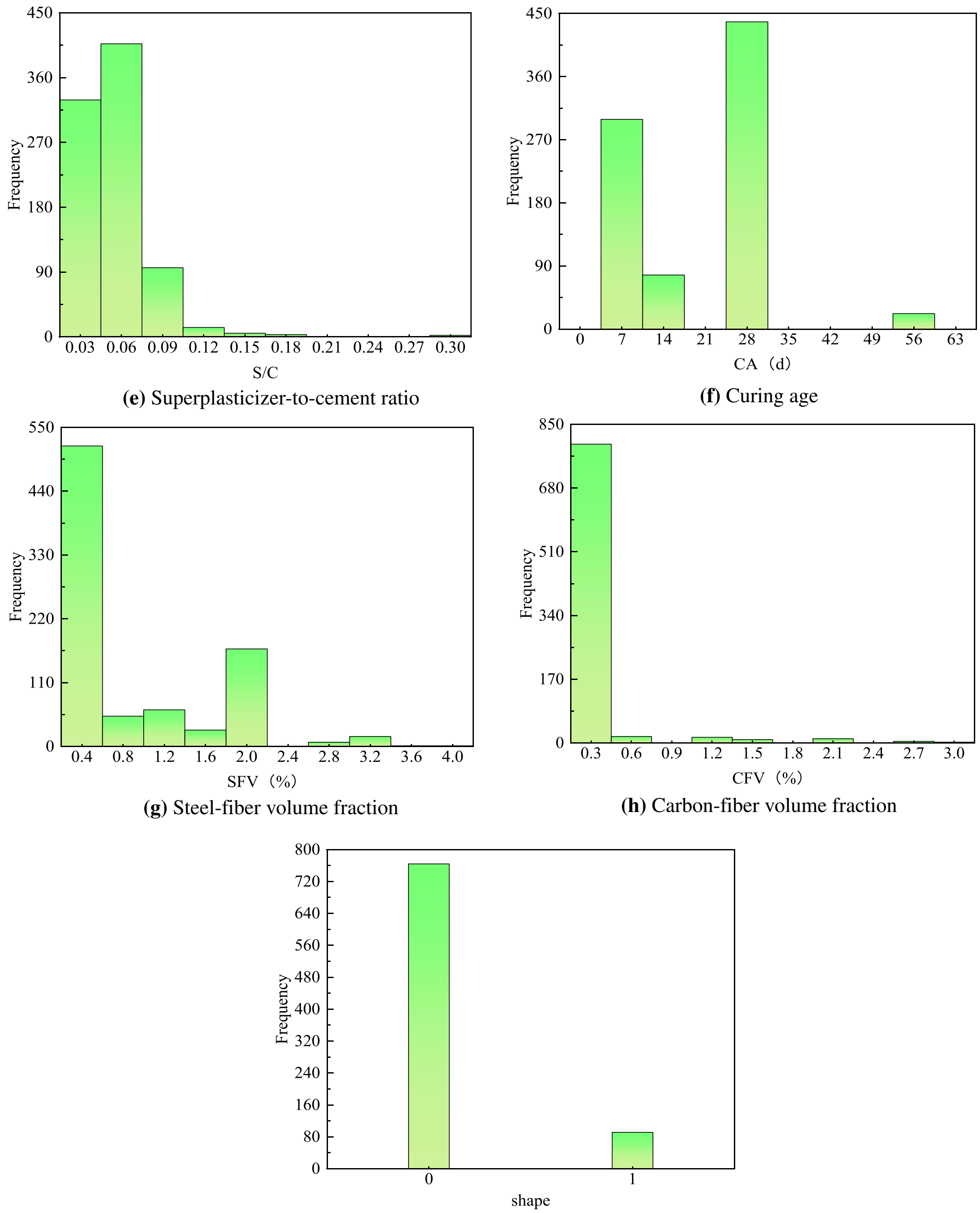
(a) Water-to-cement ratio

(b) Silica fume-to-cement ratio

(c) Silicon powder-to-cement ratio

(d) Silica sand-to-cement ratio

Figure 1: (Continued)



(e) Superplasticizer-to-cement ratio

(f) Curing age

(g) Steel-fiber volume fraction

(h) Carbon-fiber volume fraction

(i) Specimen shape (0: cube, 1: cylinder)

Figure 1: Distribution of input variables.

The pairwise correlation structure among the nine input variables is illustrated in Fig. 2, which presents the Pearson correlation coefficient matrix in the form of a heatmap. The analysis reveals that most mixture proportion variables exhibit only weak to moderate intercorrelations, with the absolute coefficient values predominantly remaining below 0.5.

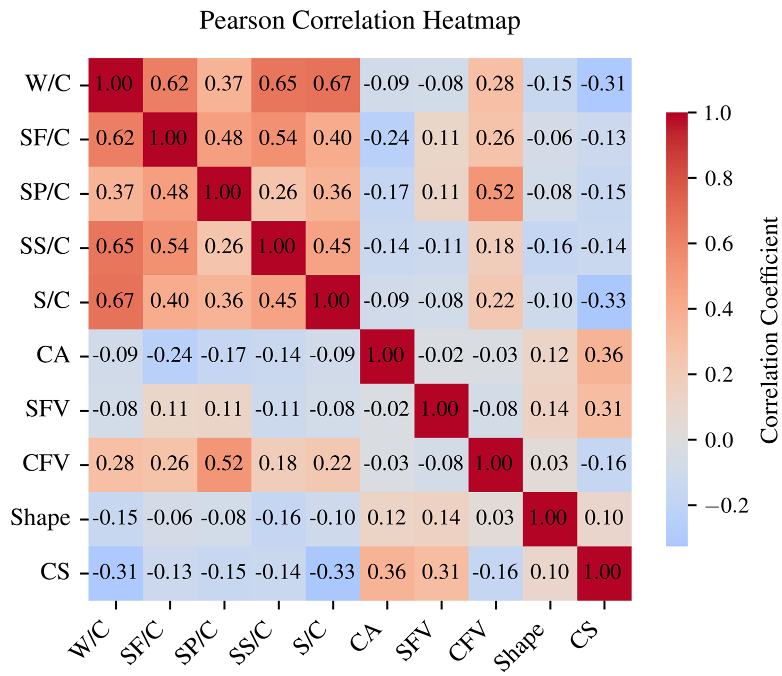


Figure 2: Correlation coefficient.

In Fig. 2, the absolute values of the correlation coefficients between most variable pairs are below 0.5, indicating that only weak to moderate linear correlations exist among the UHPC mix proportion parameters. Several variable pairs exhibit moderate correlations approaching or exceeding 0.4. For example, the correlation coefficient between SF/C and SP/C is approximately 0.42, and that between S/C and SP/C is approximately 0.44. These reflect the synergistic adjustment trends among different components in UHPC mix design. On the other hand, CFV shows correlation coefficients close to zero with almost all other variables, indicating that it has nearly no linear association with the main mix proportion parameters. Furthermore, the Pearson correlation coefficient only measures linear dependence; for factors such as steel fiber content, which may exhibit nonlinear or threshold effects with compressive strength, this coefficient cannot fully capture the strength of their association.

2.2 Development of the Prediction Model

2.2.1 Improvement of the XGBoost Model

XGBOOST iteratively trains multiple weak learners, where each iteration focuses on learning and optimizing the residual of the previous model [1]. By constructing new decision trees through gradient descent, the algorithm progressively improves predictive accuracy [48]. The prediction of XGBOOST is implemented via an objective function, as shown in Eq. (2).

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

In the formula, $\sum_{i=1}^{n} L(y_i, \hat{y}_i)$ is the loss function, such as mean squared error (MSE) or LogLoss. $\sum_{k=1}^{k} \Omega(f_k)$ is the regularization term, commonly including L1 and L2 regularization, which controls the model's complexity.

XGBoost's hyperparameters, such as tree depth, number of trees, and learning rate, directly affect the model's fitting and generalization ability. Improper parameter settings can lead to overfitting or underfitting. Manual tuning or grid search is inefficient, especially when the parameter space is large. Bayesian optimization guides the search by constructing and iteratively updating a probabilistic surrogate model of the objective function, enabling efficient identification of the global optimal solution within a limited number of evaluations [21]. By balancing exploration and exploitation, it significantly reduces computational costs and rapidly identifies near-optimal hyperparameter configurations. Tree-structured Parzen Estimator (TPE) or Tree-structured Parzen Estimator is an intuitive implementation of Bayesian optimization. Following the standard TPE formulation, the threshold is set as the γ-th quantile (γ is 0.15) of the observed loss values. This divides the historical evaluations into two groups: configurations with loss below the threshold are considered 'good', and those above the threshold are 'bad'. The algorithm then constructs kernel density estimates g(x) for the 'bad' group and l(x) for the 'good' group, and maximizes the ratio l(x)/g(x) to propose promising new hyperparameter configurations.

However, Bayesian optimization typically evaluates configurations sequentially, and the computational overhead can remain substantial when each evaluation is expensive. Hyperband uses successive halving to dynamically allocate resources and early-stop poor trials, balancing trial count and budget for faster hyperparameter search, but its random sampling cannot exploit prior knowledge.

To combine the advantages of both methods, researchers have proposed the BOHB method, a hybrid of Bayesian optimization and Hyperband. BOHB represents a significant methodological innovation, combining the strengths of two powerful optimization paradigms. It uses Bayesian optimization to build a probabilistic model of the objective function (model performance given hyperparameters), guiding the search toward promising regions of the hyperparameter space. Simultaneously, it leverages the Hyperband algorithm (a multi-armed bandit strategy) to dynamically allocate resources (e.g., boosting rounds) to promising configurations and aggressively terminate underperforming trials. This synergistic combination offers several key advantages: (1) By concentrating computational resources on promising trials, it substantially reduces the computational cost of hyperparameter search. (2) It intelligently explores the search space, utilizing prior knowledge to find near-optimal configurations faster than random or grid search. (3) The combined method is less prone to getting stuck in local optima compared to pure Bayesian optimization and is more sample-efficient than pure Hyperband. The overall workflow of the proposed BOHB-XGBOOST framework is schematically shown in Fig. 3.

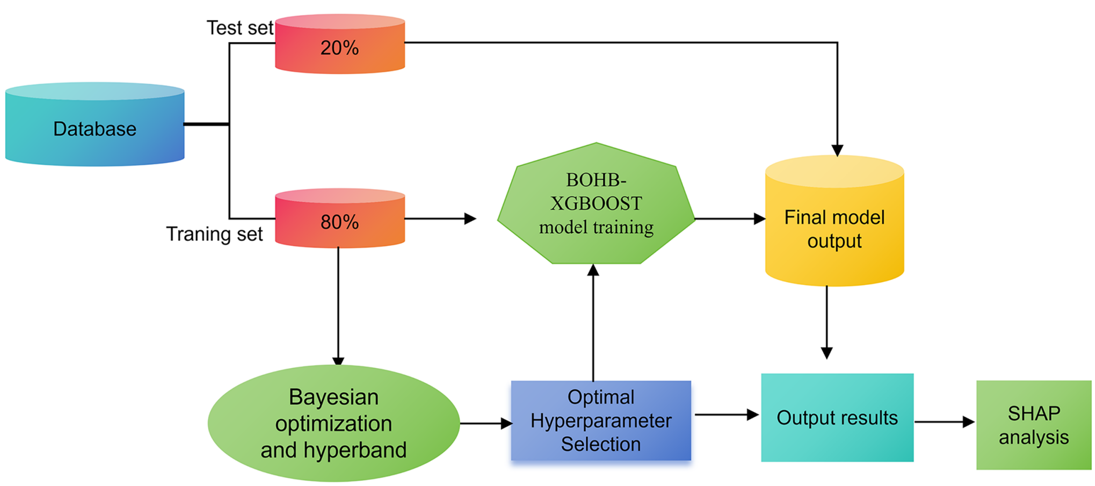


Figure 3: Algorithm flowchart of the proposed BOHB-XGBoost model.

2.2.2 Model Structure

The BOHB-XGBOOST model is implemented as an additive ensemble of decision trees. The input vector contains the selected mixture descriptors such as binder ratios, water content, granular constituents, and admixture dosage, and each tree partitions the feature space to fit a portion of the response. In the boosting sequence, later trees are trained on the residual errors of earlier ones, allowing the ensemble to iteratively correct its predictions. The final output is obtained by summing the contributions of all trees, yielding the estimated compressive strength for a given UHPC mixture.

2.2.3 Model Configuration Details

In addition to input variables, a series of hyperparameters must be configured when constructing the BOHB-XGBOOST model. The most critical parameters include the loss function, learning rate, and maximum number of training iterations, all of which directly affect the training process and final model performance.

The loss function measures the discrepancy between predicted and true compressive strength values of UHPC. The mean squared error (MSE) was employed as the loss function due to its differentiability and fast convergence properties, which are advantageous for gradient-based boosting., defined as:

$$L^N = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{3}$$

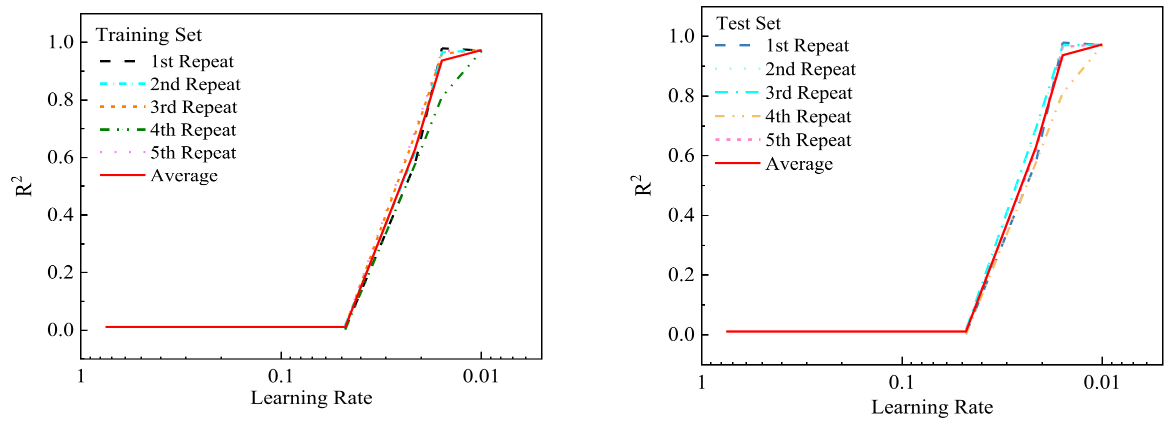
where N denotes the total number of samples, y_i is the true value of the i-th sample, and $\hat{y}_i$ is the predicted value of the i-th sample.

After establishing mean squared error (MSE) as the optimization objective, the BOHB algorithm was used to automatically tune the core hyperparameters of the XGBoost model, including the learning rate and the number of training iterations. The key BOHB parameters were configured as follows. The resource budgets, specifically a minimum resource of 50, a maximum resource of 500, and an eta of 3, were selected according to the default heuristics of the BOHB algorithm, thereby ensuring that at least 50 boosting rounds are allocated to each configuration while allowing sufficient budget for promising ones. A bandwidth factor of 3 is the BOHB default and was kept unchanged. The algorithm was run for 20 BOHB iterations, with 10 initial random exploration configurations, a maximum of 50 configurations per round, an early stopping patience of 50 rounds, and a minimum bandwidth of 0.001. These settings collectively govern resource allocation, the balance between exploration and exploitation, and the early stopping mechanism, thereby ensuring efficient convergence and stability of the BOHB-XGBOOST model.

The complete hyperparameter search space was defined as follows: learning rate [0.001, 0.1], max depth [3, 10], subsample ratio [0.5, 1.0], colsample_bytree [0.5, 1.0], reg_alpha [1e^{-5}, 10], and reg_lambda [1e^{-5}, 10]. After BOHB optimisation, the final selected hyperparameters were learning rate = 0.01, max depth = 6, subsample ratio = 0.8, colsample_bytree = 0.8, reg_alpha = 0.1, and reg_lambda = 1.0. A fixed random seed of 42 was used for all experiments to ensure reproducibility.

To prevent data leakage, we performed an 80/20 train-test split using stratified sampling by literature source, ensuring all mixtures from a given reference remained entirely in one set. A 5-fold cross-validation in Fig. 4 was applied to the full dataset of 780 records; within each fold, standardization parameters were computed solely from the training portion and applied to the validation portion. We chose 5-fold for its optimal trade-off between computational efficiency and estimation stability—offering lower variance than 3-fold while being less intensive than 10-fold. Based on learning-rate experiments, the final XGBoost parameters were set to a learning rate of 0.01, a subsample ratio of 0.8, and 500 iterations, consistent with [48].

The 5-fold cross-validation gave an average performance of R^2 (0.964 ± 0.008) and RMSE (6.7 ± 0.4 MPa). Subsequently, the BOHB-XGBoost model on the independent test set showed only a small difference from the validation results, which is within an acceptable tolerance, thereby confirming the model's stability.



(a) training set **(b)** Test set

Figure 4: Influence of learning rate on BOHB-XGBOOST model fit.

2.2.4 Model Evaluation Metrics

To comprehensively evaluate the predictive performance of the BOHB-XGBOOST model, the coefficient of determination (R^2) was adopted as the core metric for assessing overall prediction accuracy. The value of R^2 is calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^{M}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{M}(y_i - \overline{y}_i)^2} \tag{4}$$

where $\hat{y}_i$ represents the predicted compressive strength of the i-th sample, y_i is the corresponding measured compressive strength, and $\overline{y}_i$ denotes the mean of the measured compressive strengths across all samples.

In addition to reporting R^2, three error-based indicators were used to assess the dispersion of the predictions: the root mean squared error (RMSE), the mean absolute percentage error (MAPE), and the mean absolute error (MAE). These quantities were computed using the standard definitions shown below:

$$\mathrm{MAPE} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}\left|\frac{\hat{\mathrm{y}}_{\mathrm{i}} - \mathrm{y}_{\mathrm{i}}}{\mathrm{y}_{\mathrm{i}}}\right| \times 100\% \tag{5}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{6}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{7}$$

3 Results and Discussion

3.1 The Predictive Performance of the BOHB-XGBOOST Model

The same train-test split was used across repeated runs. As illustrated in Fig. 4, reducing the learning rate to 0.01 produced stable behaviour, with five independent trials yielding goodness-of-fit values above 0.95 for both subsets. Representative performance statistics from one run are summarised in Table 3.

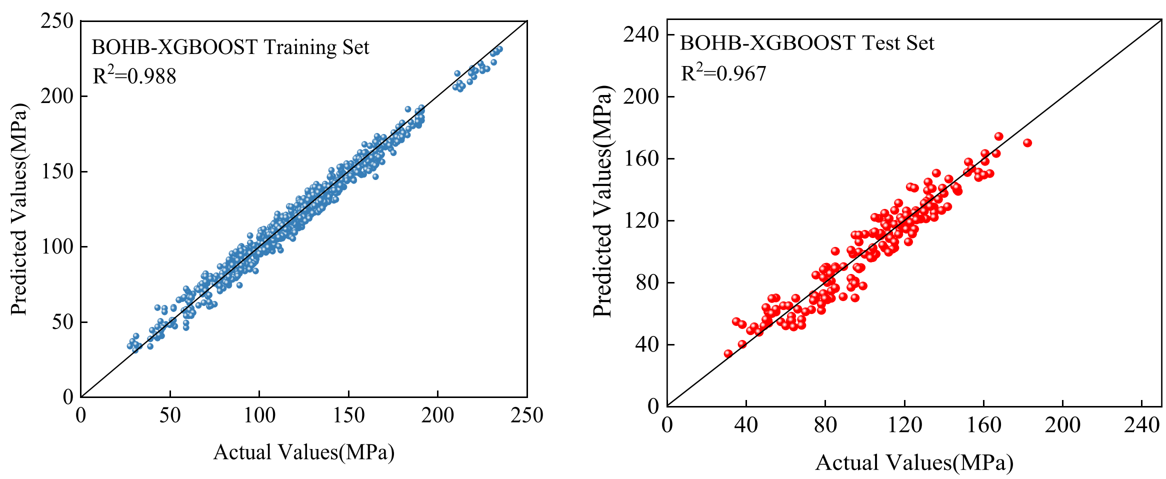
Table 3: Performance metrics of the BOHB-XGBOOST model on the training and testing datasets.

Dataset	Performance Metrics			
	R^2	RMSE (MPa)	MAPE (%)	MAE (MPa)
Training set	0.988	4.71	2.24	2.46
Testing set	0.967	6.5	4.86	4.8

Table 3 indicates that the BOHB-XGBOOST model reproduced most of the observed variance with limited scatter. On the training subset, the model achieved an R^2 of 0.988 and a mean absolute percentage error (MAPE) of 2.24%. When evaluated on the withheld test set, the corresponding values were an R^2 of 0.967 and a MAPE of 4.86%. This indicates that the increase in error from the training set to the test set is modest, and that the model retains good predictive capability outside the calibration set.

Although the training R^2 is close to unity, the evidence does not suggest severe overfitting: the testing results remain consistent, and the increase in percentage error between the two partitions is within an expected range for mixture datasets of this size. Regularization built into the XGBOOST objective (via L_1 and L_2 penalties), together with the small learning rate and capped iteration count, constrain tree growth and act as informal early-stopping criteria. Standardisation of input variables also limits numerical imbalance and reduces noise sensitivity. Nevertheless, the analysis would benefit from future validation using repeated or stratified k-fold procedures, particularly if a larger database becomes available.

Fig. 5 compares predicted and measured strengths. The majority of points lie near the 1:1 line for both the training and testing data, reinforcing that the model captures the main trends governing compressive strength within the sampled mixtures.



(a) Training set **(b)** Test set

Figure 5: Correlation between predicted and measured compressive strength for the training and testing datasets.

3.2 Comparative Analysis with Other Models

To place the XGBOOST results in context, four additional regression approaches were examined: Random Forest (RF), Gradient Boosting Decision Trees (GBDT), Support Vector Regression (SVR) and a multilayer perceptron (MLP). Each method was trained on the same 780 mixtures and evaluated using an identical 80/20 split to avoid differences arising from data partitioning. All model training and hyperparameter optimization were conducted on a workstation equipped with an Intel Core i7 12700K CPU (12 cores, 20 threads), 32 GB of DDR4 RAM, and an NVMe SSD, running Ubuntu 22.04 as the operating system. The models were implemented using Python 3.10 along with the XGBoost 1.7, scikit-learn 1.2, and BOHB libraries, utilizing all physical cores during parallel execution.

Table 4 shows that the BOHB-XGBOOST produced the strongest overall fit. Relative to RF, the improvement in R^2 was about 0.9% and the RMSE was lower by roughly 8%. When compared with GBDT, the gains are consistent with the incorporation of second-order gradients and explicit regularisation, which help control tree growth and reduce bias. Existing empirical formulas for UHPC [62] are only valid within the range of a single factor and specific processing conditions, and their generalization ability drops sharply when multiple factors are coupled. In contrast, the BOHB-XGBOOST method proposed in this paper can better handle multi-variable interactions and offers stronger prediction robustness.

Table 4: Comparative predictive performance of different machine learning algorithms.

Algorithm	R^2	RMSE (MPa)	MAPE (%)	MAE (MPa)
BOHB-XGBOOST	0.967	6.5	4.86	4.8
XGBOOST	0.961	7.6	5.14	5.1
Random Forest	0.952	8.3	5.80	5.7
GBDT	0.956	7.9	5.60	5.4
SVR	0.938	9.8	7.50	7.3
MLP	0.942	9.3	7.10	7.0

The tree-based models such as XGBOOST, RF and GBDT performed better than the MLP and the SVR. Given the tabular nature and modest size (780 samples) of the dataset, tree-based ensemble methods are generally more effective than neural networks or kernel-based methods, as they are less prone to overfitting and can naturally handle heterogeneous features and non-linear interactions without extensive preprocessing.

All baseline models (RF, GBDT, SVR, MLP) were also optimized using the BOHB algorithm with the same resource budget and search space tailored to each model's hyperparameters. The results reported in Table 4 therefore reflect each model's best performance under comparable optimization effort.

Fig. 6 compares predicted and measured strengths for the five models on the test subset. The outputs from the tree-based methods cluster closely around the 1:1 line, whereas SVR and the MLP show noticeably wider scatter, especially at higher strength values. This behaviour suggests that ensemble trees capture the nonlinear response of the mixture data more effectively than the neural-network and kernel approaches used here.

Fig. 7 summarises the error statistics and highlights the differences among the tested approaches. XGBOOST produced the RMSE (7.6 MPa) and the MAE and MAPE (5.1 MPa and 5.14%), indicating fewer large deviations than the other models. Random Forest trained more quickly (8.5 s) but at a modest cost in accuracy, whereas GBDT approached the performance of XGBOOST but required additional computation. SVR and the MLP showed both higher error levels and longer runtimes, suggesting that they are less suited to the type of tabulated mixture data considered here. Overall, each method yielded usable predictions, but XGBOOST offered the most balanced outcome combining fit quality, numerical stability and reasonable computational effort.

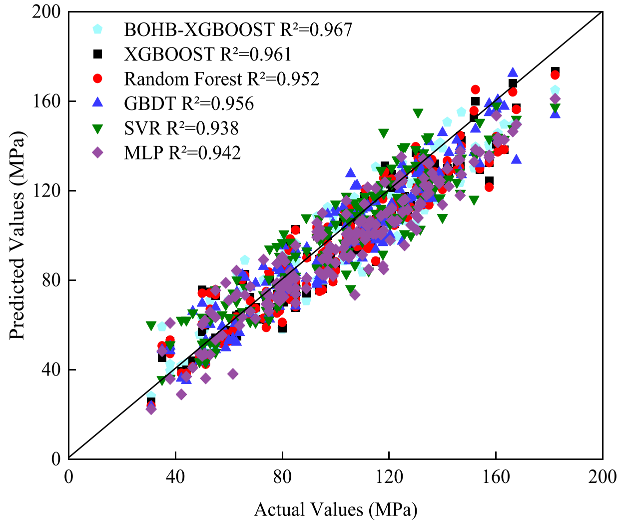
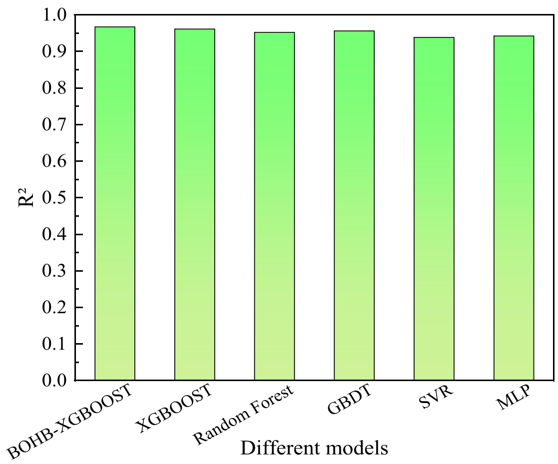
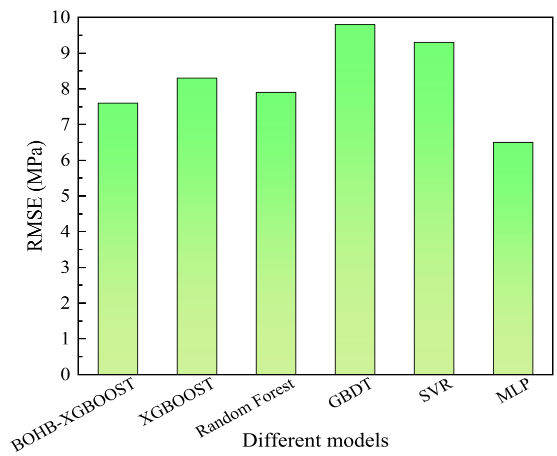


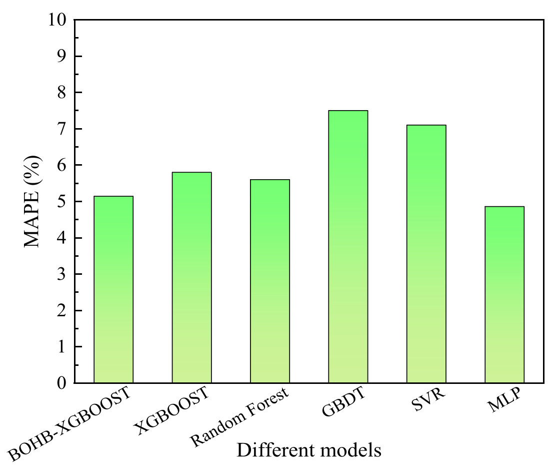
Figure 6: Comparison between predicted and actual values by different methods in the testing set.



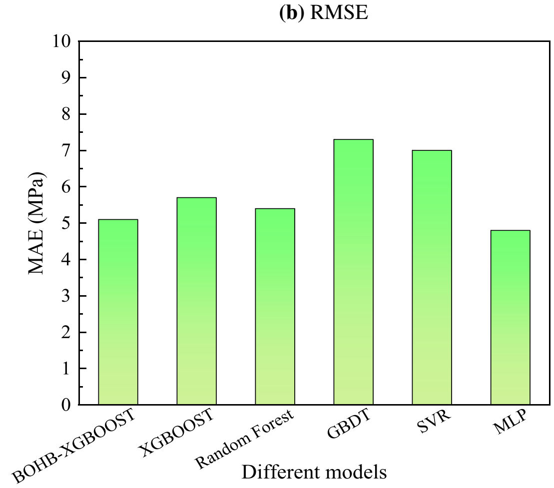
(a) R2



(b) RMSE

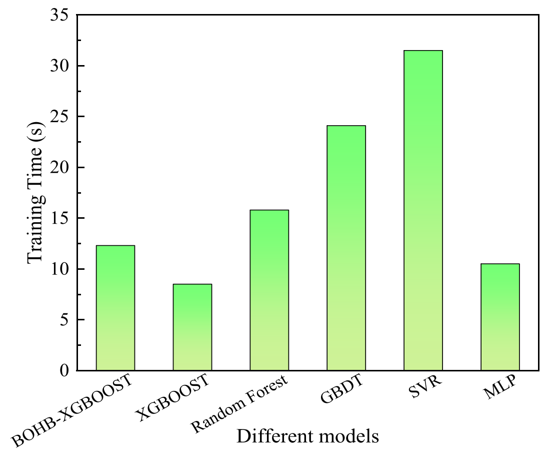


(c) MAPE



(d) MAE

Figure 7: (Continued)



(e) TIME

Figure 7: Comparative performance of XGBOOST and other methods.

It is worth noting that the XGBOOST model optimized with BOHB (R^2 is 0.967) was significantly superior to the standard XGBOOST model without complex optimization (R^2 is 0.961). The 0.6% improvement in R^2 and the reduction in RMSE from 7.6 to 6.5 MPa directly quantify the practical value brought by the BOHB optimization strategy, demonstrating its effectiveness in automatically identifying better hyperparameter configurations and thereby unlocking the model's maximum potential.

These results indicate that boosted tree ensembles provide a practical regression tool for UHPC mixture datasets and can support strength estimation in design-oriented applications.

3.3 Interpretability Analysis Based on SHAP

To examine how individual predictors contributed to the XGBOOST output, Shapley Additive Explanations (SHAP) were used to compute feature attributions for each mixture. The SHAP routines were applied to the trained model, and the resulting ranking of variable influence is shown in Fig. 8.

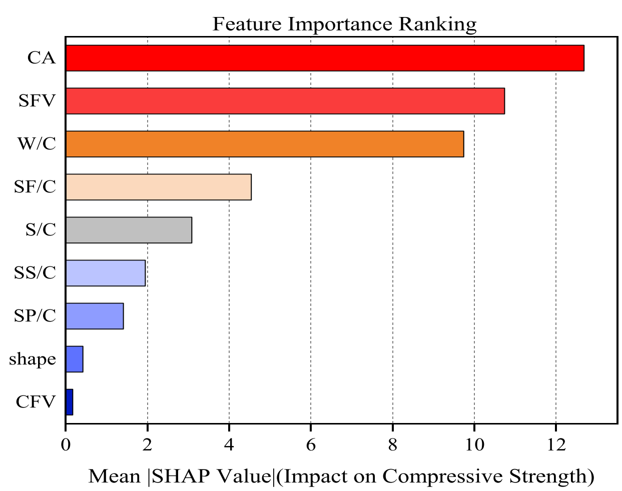


Figure 8: Ranking of input feature contributions in the model.

As shown in Fig. 8, the input variables are ranked by their mean absolute SHAP values. Five features exhibit SHAP values exceeding 3, indicating that they collectively dominate the model's predictions. Specifically, curing age, steel fiber volume fraction, water-to-cement ratio, silica fume-to-cement ratio, and superplasticizer-to-cement ratio are the primary drivers of compressive strength. In contrast, the carbon fiber volume fraction has a SHAP value of only 0.17, which is almost negligible, suggesting that its limited occurrence in the dataset restricts its statistical influence. Given its sparse non-zero occurrences in the dataset, future studies may consider removing this feature to simplify the model. The remaining variables such as silica sand-to-cement ratio, silica powder-to-cement ratio, and specimen shape contribute modestly but still meaningfully to the output, justifying their retention in the current feature set.

To examine the distribution and direction of these effects, SHAP summary diagrams were generated, allowing the influence of individual predictors on compressive strength to be visualized as shown in Fig. 9.

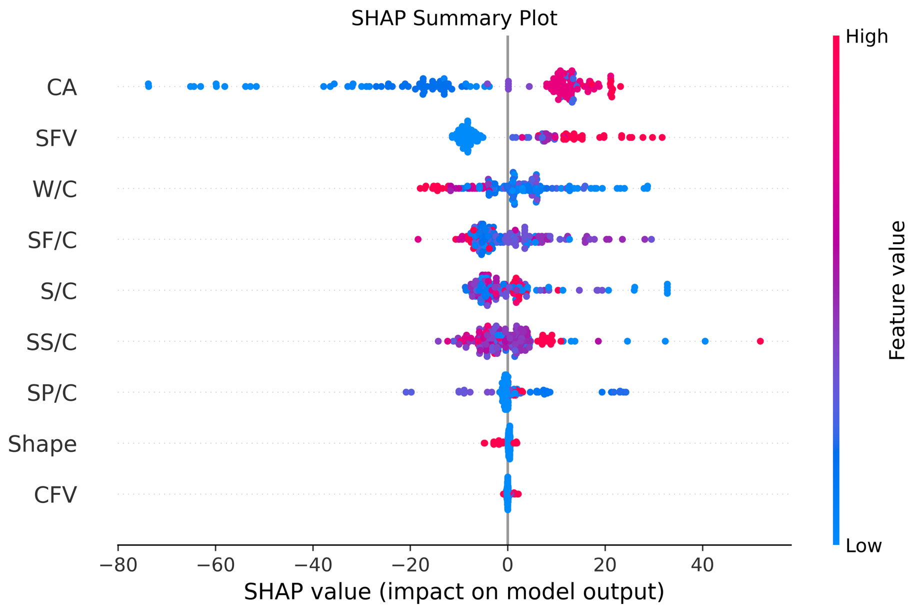


Figure 9: SHAP summary plot.

In the SHAP diagrams, warmer colours represent larger feature values and cooler colours represent smaller ones. Points with high feature values plotted on the positive side of the SHAP axis indicate that increasing that variable tends to raise the predicted strength, while the opposite pattern suggests a negative contribution.

From the SHAP summary plot in Fig. 9, the direction of influence of each feature can be observed. Curing age, steel fiber volume fraction, silica fume-to-cement ratio, superplasticizer-to-cement ratio (at moderate dosages), silica sand-to-cement ratio, and silica powder-to-cement ratio show positive influences: their high-value points are mainly distributed on the positive side of the SHAP axis, indicating that increasing these variables generally raises the predicted strength. This finding is consistent with the strengthening mechanisms of hydration maturity, fiber bridging, pozzolanic activity, and particle packing density on UHPC strength. In contrast, the water-to-cement ratio exhibits a negative influence such as higher values lead to lower predicted strength and specimen shape also contributes negatively, with cylindrical specimens showing lower strength than cubes, which agrees with conventional testing experience. The nonlinear SHAP pattern of the superplasticizer-to-cement ratio reflects a trade-off: low dosages lead to poor particle dispersion and reduced strength, while excessive dosages may cause segregation or air entrainment, both detrimental to mechanical performance. Overall, eight of the nine input variables exert a non-negligible influence on the model predictions, and the directions of influence are highly consistent with the known material behavior of UHPC. The SHAP analysis not only validates the reliability of the model but also provides a transparent and quantifiable measure of the contribution of each mix proportion parameter, transforming the "black-box" model into a practical decision-making tool for mix design.

To illustrate the model's interpretability at the individual sample level, a representative case from the test set is examined. The waterfall plot in Fig. 10 shows how each feature contributes to the deviation of the predicted strength from the baseline value. For this sample, the negative contributions from curing age, steel fiber volume fraction, and the silica powder-to-cement ratio collectively lower the prediction, consistent with the material's expected behavior. Notably, the model exhibits higher sensitivity to the silica powder-to-cement ratio than to curing age, indicating that even small deviations in this ratio from the optimal range can have a pronounced negative effect on strength.

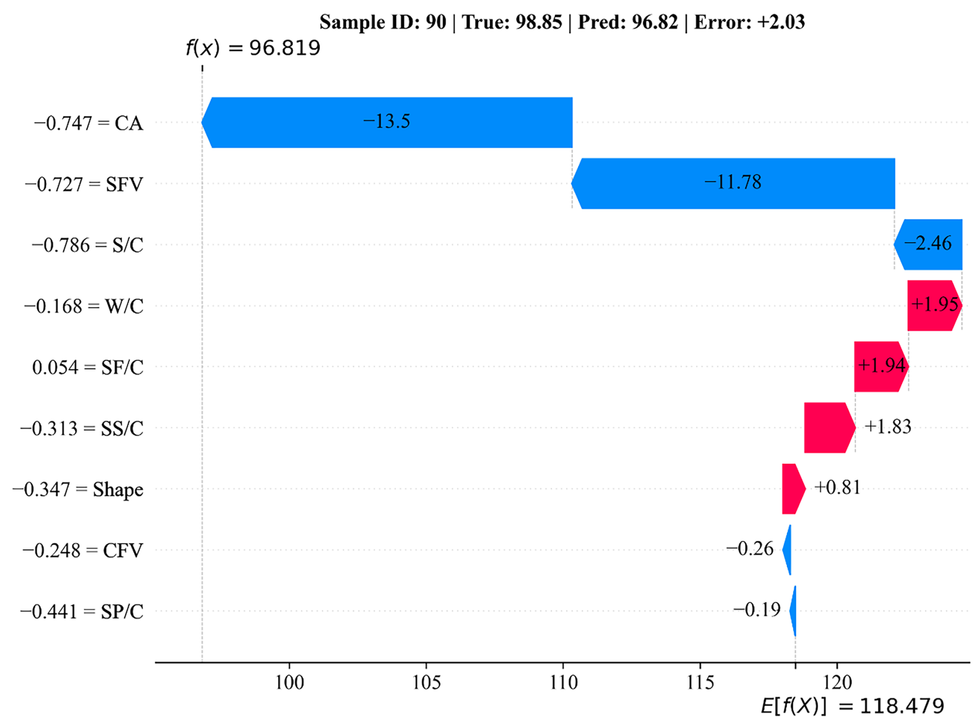


Figure 10: SHAP waterfall plot for a single sample.

3.4 Stability of the BOHB-XGBOOST Model

Direct comparison with other published models is still difficult because of differences in data provenance, curing practices and performance criteria. Future efforts could benefit from shared benchmark datasets and common validation procedures so that predictive approaches for UHPC strength can be compared more meaningfully across studies.

To further validate the generalization ability of the model under real experimental conditions, this study independently conducted a series of UHPC mix proportion experiments as shown in Table 5, obtaining 20 new sets of compressive strength data. These data were not used in model training and were solely employed for the final validation of predictive performance.

Table 5: Mixing ratio and experimental grouping (kg/m^3).

Group	Cement	Silica Sand	Silica Fume	Silica Powder	Steel Fibre	Water Reducer	Water	Curing Age (d)	Compressive Strength (MPa)
1	788.52	867.4	197.1	236.6	157	19.7	201	3	95.2
2	788.52	867.4	197.1	236.6	157	19.7	201	7	112.3
3	788.52	867.4	197.1	236.6	157	19.7	201	14	126.5
4	788.52	867.4	197.1	236.6	157	19.7	201	28	140.7
5	788.52	867.4	197.1	236.6	157	19.7	181	28	154.3
6	788.52	867.4	197.1	236.6	157	19.7	212	28	126.7
7	788.52	867.4	197.1	236.6	142	19.7	201	28	133.9
8	788.52	867.4	197.1	236.6	173	19.7	201	28	146.3
9	788.52	867.4	165.6	236.6	157	19.7	201	28	135.8
10	788.52	867.4	181.4	236.6	157	19.7	201	28	139.2
11	788.52	867.4	212.9	236.6	157	19.7	201	28	144.7
12	788.52	867.4	228.7	236.6	157	19.7	201	28	145.6
13	788.52	867.4	197.1	205.0	157	19.7	201	28	137.4
14	788.52	867.4	197.1	220.8	157	19.7	201	28	139.3
15	788.52	867.4	197.1	252.3	157	19.7	201	28	142.8
16	788.52	867.4	197.1	268.1	157	19.7	201	28	141.5
17	788.52	709.7	197.1	236.6	157	19.7	201	28	135.7
18	788.52	788.5	197.1	236.6	157	19.7	201	28	138.6
19	788.52	906.2	197.1	236.6	157	19.7	201	28	141.6
20	788.52	946.2	197.1	236.6	157	19.7	201	28	139.3

The UHPC mix design adopted in this study was developed based on the absolute volume method. The composition and proportions of raw materials are represented by the benchmark group as shown in Table 5 as follows: the binder system consists of 788.52 kg/m^3 of P·II 52.5 ordinary Portland cement, 197.1 kg/m^3 of industrial by-product silica fume, and 236.6 kg/m^3 of silica powder. High-purity silica sand with a particle size gradation of 0.1–1.2 mm was used as fine aggregate, incorporated at 867.4 kg/m^3 to optimize the particle packing density. To ensure adequate workability under a low water-to-binder ratio, 19.7 kg/m^3 of polycarboxylate-based high-performance water reducer was added. The reinforcement phase consisted of straight steel fibers with a length of 13 mm, aimed at enhancing crack resistance, ductility, and post-cracking load-bearing capacity. Potable tap water was used for mixing. The steel fiber dosage and water content were 157 and 201 kg/m^3, respectively. Test raw materials are shown in Fig. 11.

First, cement, silica fume, and silica sand were dry-mixed until uniform. Subsequently, mixing water (including the water reducer) was added for wet mixing. Finally, steel fibers were incorporated and mixed until a homogeneous mixture was achieved. After vibration molding, the specimens were allowed to stand until initial setting. Subsequently, they were demolded and then followed by transfer to a steam curing chamber at 90°C for continued curing to accelerate the hydration process.

For the compressive strength test, standard 70.7 mm × 70.7 mm × 70.7 mm cubic specimens were employed. Upon reaching the specified curing age, loading tests were conducted using a computer-controlled pressure testing machine as shown in Fig. 12. The tests were rigorously performed in accordance with relevant standards and specifications, applying a constant loading rate until final specimen failure. The maximum failure load was recorded, and the compressive strength of the matrix was calculated. For each test group, at least three parallel specimens were tested, and the arithmetic mean of the results was taken as the representative strength value for that group.

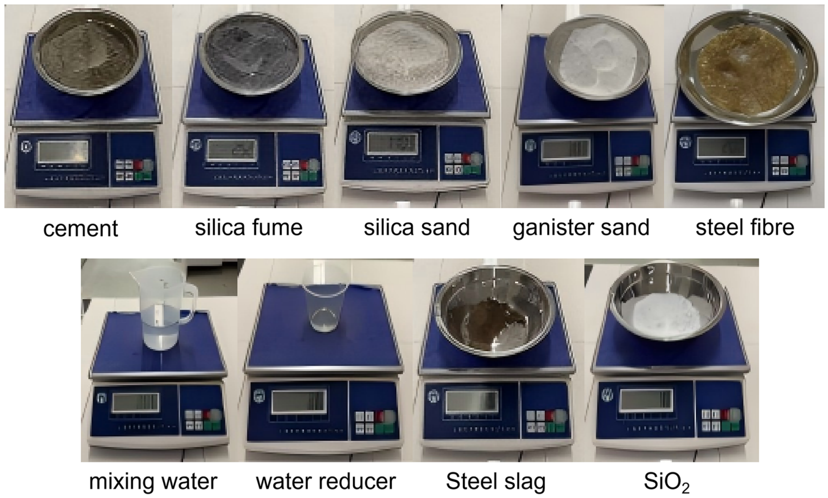


Figure 11: Test raw materials.

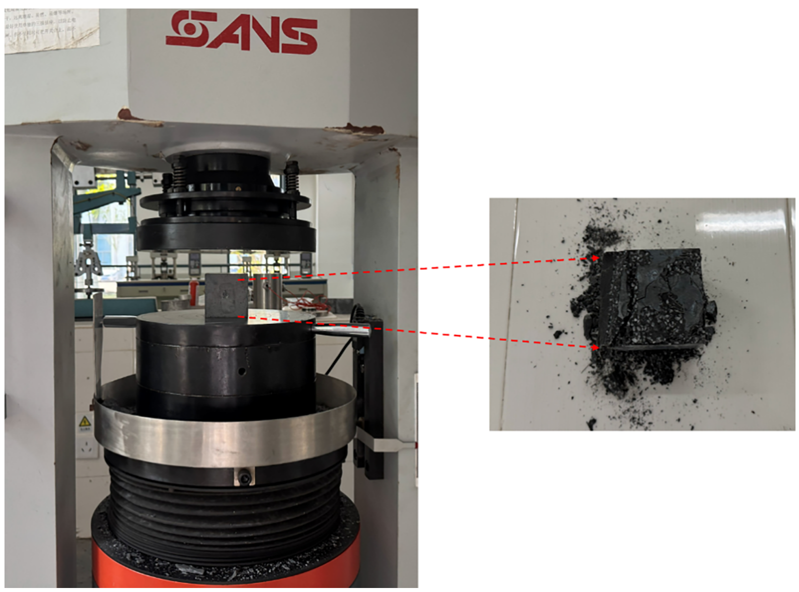


Figure 12: Diagram of cube compressive strength test setup.

Fig. 13 presents the comparison between the predicted and measured compressive strength values of the BOHB-XGBoost model on the testing set. The data points are tightly clustered around the 1:1 diagonal line, indicating a high degree of agreement between the model predictions and the true observations. First, as the curing age increases, the strength curves predicted by the model exhibit a typical logarithmic growth trend, with a rapid increase during the early stages and a gradual flattening in the later stages in Fig. 14. This accurately reflects the evolution of the cement hydration process and the pozzolanic reaction. Second, concerning the water-to-cement ratio, the model captures its negative correlation: as the water-to-binder ratio decreases, the compressive strength of UHPC increases substantially, following a generally monotonic trend. This is attributed to the enhanced particle packing density and reduced porosity under a low water-to-binder ratio. Finally, the model accurately characterizes the strengthening effect of steel fiber content: as the volume fraction of steel fibers increases, the predicted strength rises steadily. Although the gain slightly diminishes at high fiber dosages (reflecting the potential influence of fiber agglomeration or non-uniform distribution), the overall trend is highly consistent with the micromechanism of fiber bridging cracks and enhancing toughness.

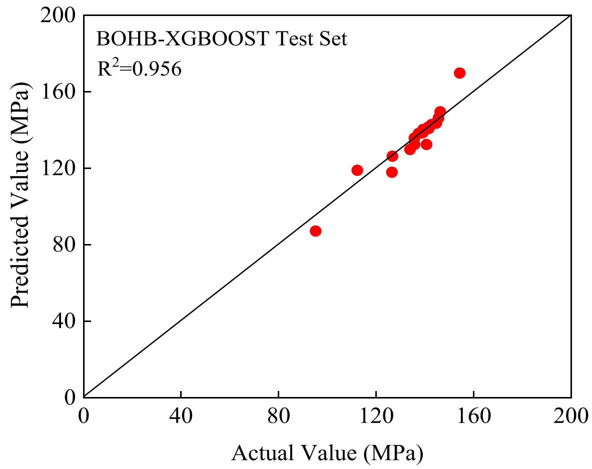


Figure 13: Comparison of actual and predicted values.

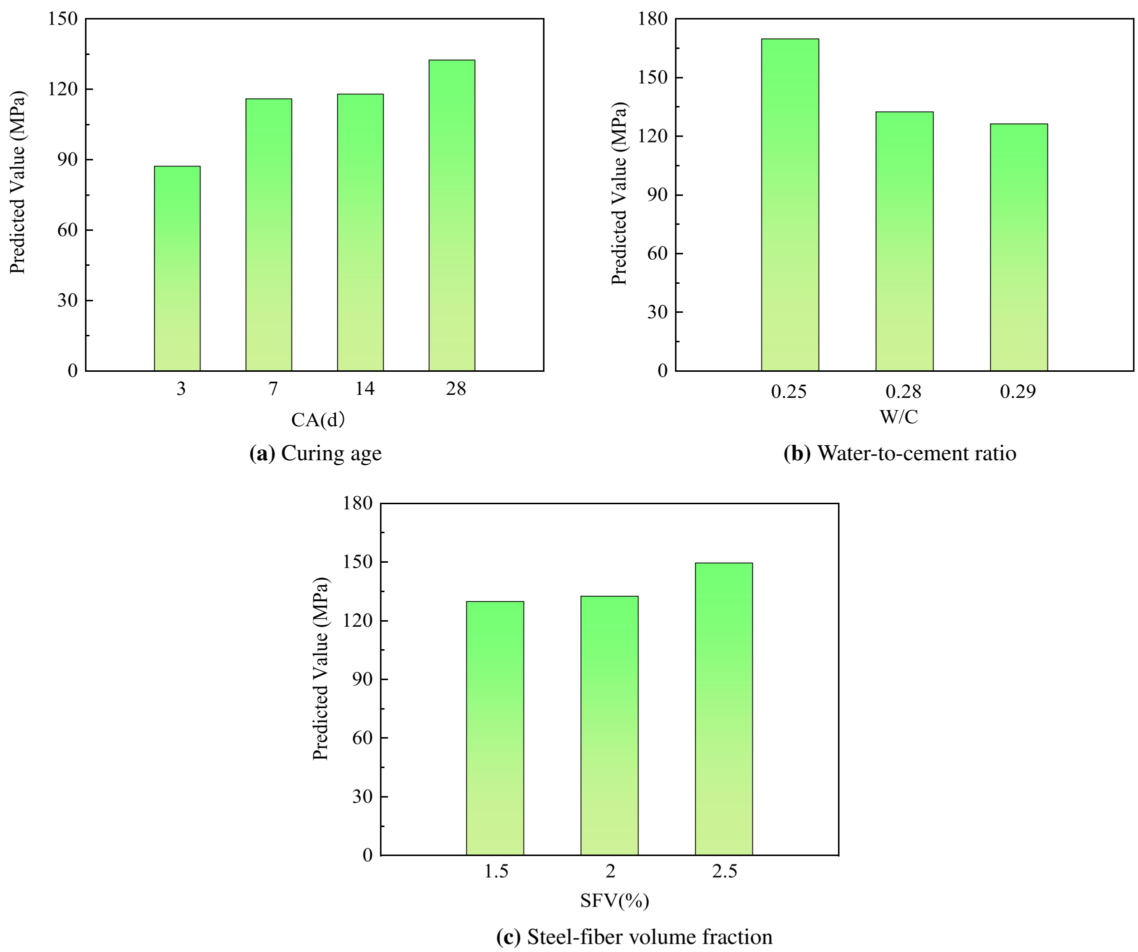


Figure 14: Sensitivity analyses.

To further examine the model's response to key variables, partial dependence-type trend analyses were conducted for curing age, water-to-binder ratio, and steel fiber volume fraction. The predictive response of the BOHB-XGBOOST model not only maintained good monotonicity and smoothness, but all predicted values also fell closely within the distribution envelope of the original training data. This result fully demonstrates that within the explored parameter space, the model predictions exhibit reasonable stability and physically consistent trends. However, it should be emphasized that this experimental validation was based on only 20 mixtures from a single design protocol, and the literature data used for training are heterogeneous. Therefore, the model's extrapolation capability to significantly different material systems or production conditions remains to be further validated.

4 Conclusion

This study developed a machine learning framework for predicting the compressive strength of ultra-high-performance concrete (UHPC) based on nine macroscopic mixture descriptors. By integrating the BOHB hybrid optimization algorithm with the XGBOOST regressor, the proposed approach addresses key challenges in hyperparameter tuning and model generalization for structured tabular data in civil engineering materials research. The main contributions and findings are summarized as follows:

This study introduces the BOHB algorithm, including a hybrid of Bayesian optimization and Hyperband for automatic hyperparameter search of the XGBOOST model. This methodological innovation efficiently balances exploration and exploitation within the high-dimensional parameter space, leveraging prior knowledge while dynamically allocating computational resources to promising configurations. The BOHB-XGBOOST framework overcomes the inefficiency of manual tuning and the limitations of grid search or random sampling, establishing a robust and automated pipeline for model development.

Trained and evaluated on a comprehensive dataset of 780 UHPC mixtures compiled from published literature, the BOHB-XGBOOST model achieved outstanding predictive performance, with a coefficient of determination (R^2) of 0.967 and a mean absolute percentage error (MAPE) of 4.86% on the testing set. Comparative analysis demonstrated that the performance of the proposed model was more improved and stable than standard machine learning algorithms, including Random Forest, GBDT, SVR, and MLP. Notably, the performance gain over standard XGBOOST (R^2 improvement from 0.961 to 0.967, RMSE reduction from 7.6 to 6.5 MPa) directly quantifies the added value of the BOHB optimization strategy, confirming its effectiveness in enhancing both accuracy and robustness for tabular mixture data.

SHAP analysis provided transparent and quantifiable insights into the contribution of each input feature. The analysis confirmed the dominant influence of curing age, steel fiber volume fraction aligning with established physical mechanisms of UHPC strength development. This interpretability not only validates the model's internal consistency but also offers quantitative guidance for mixture design optimization, transforming the predictive model into a practical decision-support tool.

In practical applications, the BOHB-XGBOOST model can quickly estimate strength, analyze mixture proportion sensitivity, and provide a preliminary assessment of alternative binder systems or fiber types. The experimental validation was based only on 20 mixtures from a single design protocol, and the literature data used are heterogeneous, which cannot represent all real-world production conditions. Furthermore, the current feature set lacks information on chemical composition, which can affect strength. Future work should focus on integrating microstructural descriptors via multi-modal learning; employing active learning to prioritize experimental design in regions of high model uncertainty; extending the framework to predict other properties such as tensile strength, durability indicators, and high-temperature resistance; while expanding the compositional variability of the dataset to enhance extrapolation capability.

Acknowledgement: Not applicable.

Funding Statement: This research was funded by the Natural Science Foundation of Jiangsu Province (Grant No. BK20240942), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. 24KJB130005), Nantong Science and Technology Project (MSZ2025084).

Author Contributions: Conceptualization, Ling Wang; methodology, Ling Wang; validation, Mohammad Faizuddin Md Noor; investigation, Ling Wang; resources, Ling Wang; writing—original draft preparation, Ling Wang; writing—review and editing, Mohammad Faizuddin Md Noor; visualization, Yanan Zhang; funding acquisition, Ling Wang. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The dataset compiled in this study is derived from the references [21–63] listed in the reference. The processed dataset, including mixture proportions and corresponding compressive strength values, is available from the corresponding authors upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Alabduljabbar H, Khan M, Awan HH, Eldin SM, Alyousef R, Mohamed AM. Predicting ultra-high-performance concrete compressive strength using gene expression programming method. Case Stud Constr Mater. 2023;18(7):e02074. doi:10.1016/j.cscm.2023.e02074.
2. James IE, Okafor FO, Mama BO, Ezihe JC. Design, modelling and optimisation of ultra high-performance fibre reinforced concrete incorporating waste materials. Discov Civ Eng. 2024;1(1):96. doi:10.1007/s44290-024-00100-z.
3. Wang Z, Wu J, Su L, Gao Z, Yin C, Ye Z. Optimization of ultra-high performance concrete based on response surface methodology and NSGA-II. Materials. 2024;17(19):4885. doi:10.3390/ma17194885.
4. Wang L, Liu L, Dai D, Liu B, Cheng Z. Predictive modeling of UHPC compressive strength: integration of support vector regression and arithmetic optimization algorithm. Appl Sci. 2024;14(17):8083. doi:10.3390/app14178083.
5. Zhu F, Wu X, Zhou M, Sabri MMS, Huang J. Intelligent design of building materials: development of an AI-based method for cement-slag concrete design. Materials. 2022;15(11):3833. doi:10.3390/ma15113833.
6. Xie C, Wang Y, Guan J, Zhang M, Li L, Guo F. From microstructure to macro-performance: the role of magnetite in the mechanical enhancement and conductive boosting of UHPC. Constr Build Mater. 2026;522(11):146136. doi:10.1016/j.conbuildmat.2026.146136.
7. Jiang H, Luo L, Hou Y, Yang Y. Mechanical and microstructural performance of UHPC with recycled aggregates modified by basalt fiber and nanoalumina at high temperatures. Materials. 2025;18(5):1072. doi:10.3390/ma18051072.
8. Li S, Chen Y, Brouwers HJH, Yu Q. Effects of amorphous silica on mechanical contribution of coarse aggregates in UHPC: from micromechanics to mesoscale fracture and macroscopic strength. Cem Concr Compos. 2023;142:105225. doi:10.1016/j.cemconcomp.2023.105225.
9. Choi D, Hong K, Ochirbud M, Meiramov D, Sukontaskuul P. Mechanical properties of ultra-high performance concrete (UHPC) and ultra-high performance fiber-reinforced concrete (UHPFRC) with recycled sand. Int J Concr Struct Mater. 2023;17(1):67. doi:10.1186/s40069-023-00631-2.
10. Huang Y, Kong D, Li Y, Zhou S, Shu J, Wu B. Effect of different shapes of steel fibers and palygorskite-nanofibers on performance of ultra-high-performance concrete. Sci Rep. 2024;14(1):8224. doi:10.1038/s41598-024-59020-8.
11. Hussain A, Yu T, Zou F. Durability analysis of nanocarbon black-based self-strain sensing ultra-high-performance concrete (UHPC): effects of moisture content and sea-salt solution exposure. Constr Build Mater. 2025;473(6):141031. doi:10.1016/j.conbuildmat.2025.141031.
12. Hussain A, Xiang Y, Yu T, Zou F. Nanocarbon black-based ultra-high-performance concrete (UHPC) with self-strain sensing capability. Constr Build Mater. 2022;359(3):129496. doi:10.1016/j.conbuildmat.2022.129496.

13. Wan X, Abumalek A. New insights into compressive strength of cement-based materials. Appl Mater Today. 2026;48(2):103083. doi:10.1016/j.apmt.2025.103083.
14. Çiftçioğlu AÖ, Unverdi M, Mardani N, Mardani A. A novel ensemble regressor and causal analysis for predicting the compressive strength of steel fiber-reinforced concrete. Ain Shams Eng J. 2026;17(7):104196. doi:10.1016/j.asej.2026.104196.
15. Elalaoui O, Abdennaji TS, Alanazi AO, Fatnassi A, Tipu RK, Pandya KS. A systematic and critical review of ensemble, hybrid, and stacked learning models for predicting the compressive strength of cement-based mortars. Arch Comput Meth Eng. 2026;6(3):1–46. doi:10.1007/s11831-026-10662-z.
16. Khaoula E, Amine B, Mostafa B, Deifalla A, El-Said A, Salama M, et al. Machine learning-based prediction of torsional behavior for ultra-high-performance concrete beams with variable cross-sectional shapes. Case Stud Constr Mater. 2025;22(6):e04136. doi:10.1016/j.cscm.2024.e04136.
17. Yu Y, Jayathilakage R, Liu Y, Hajimohammadi A. Intelligent compressive strength prediction of sustainable rubberised concrete using an optimised interpretable deep CNN-LSTM model with attention mechanism. Appl Soft Comput. 2025;185:113993. doi:10.1016/j.asoc.2025.113993.
18. Yu Y, Al-Damad IMA, Foster S, Nezhad AA, Hajimohammadi A. Compressive strength prediction of fly ash/slag-based geopolymer concrete using EBA-optimised chemistry-informed interpretable deep learning model. Dev Built Environ. 2025;23(1):100736. doi:10.1016/j.dibe.2025.100736.
19. Shafaie V, Movahedi Rad M. Multi-objective genetic algorithm calibration of colored self-compacting concrete using DEM: an integrated parallel approach. Sci Rep. 2024;14(1):4126. doi:10.1038/s41598-024-54715-4.
20. Shafaie V, Movahedi Rad M. Dem-driven investigation and AutoML-enhanced prediction of macroscopic behavior in cementitious composites with variable frictional parameters. Mater Des. 2025;254:114069. doi:10.1016/j.matdes.2025.114069.
21. Tian X, Fang Z, Zhou T, Xiang Y. Behavior and constitutive model of ultra-high-performance concrete under monotonic and cyclic tensile loading. Constr Build Mater. 2023;389(7):131634. doi:10.1016/j.conbuildmat.2023.131634.
22. Liu K, Song R, Li J, Guo T, Li X, Yang J, et al. Effect of steel fiber type and content on the dynamic tensile properties of ultra-high performance cementitious composites (UHPCC). Constr Build Mater. 2022;342(2):127908. doi:10.1016/j.conbuildmat.2022.127908.
23. Zhao S, Liu R, Liu J, Yang L. Comparative study on the effect of steel and plastic synthetic fibers on the dynamic compression properties and microstructure of ultra-high-performance concrete (UHPC). Compos Struct. 2023;324(1):117570. doi:10.1016/j.compstruct.2023.117570.
24. Wang JJ, Zhang SS, Nie XF, Yu T. Compressive behavior of FRP-confined ultra-high performance concrete (UHPC) and ultra-high performance fiber reinforced concrete (UHPFRC). Compos Struct. 2023;312:116879. doi:10.1016/j.compstruct.2023.116879.
25. Xue J, Aloisio A, Contento A, Yang Z, Hu N, Briseghella B. A piece-wise linear model for modelling the tension-stiffening effect in steel-reinforced UHPC composites. Compos Struct. 2024;345(5):118373. doi:10.1016/j.compstruct.2024.118373.
26. Tang Z, Fang Z, Peng F, Luo H, Liu X, Liu Q. Behavior of ultra-high-performance concrete under combined tensile and bending action. J Build Eng. 2024;89(1):109388. doi:10.1016/j.jobe.2024.109388.
27. Yu R, Liu K, Yin T, Tang L, Ding M, Shui Z. Comparative study on the effect of steel and polyoxymethylene fibers on the characteristics of ultra-high performance concrete (UHPC). Cem Concr Compos. 2022;127(6):104418. doi:10.1016/j.cemconcomp.2022.104418.
28. Hassan M, Wille K. Direct tensile behavior of steel fiber reinforced ultra-high performance concrete at high strain rates using modified split Hopkinson tension bar. Compos Part B Eng. 2022;246(1):110259. doi:10.1016/j.compositesb.2022.110259.
29. Wang Z, Yan J, Lin Y, Fan F. Influence of shear connectors on the ultimate capacity of steel-UHPC-steel slabs subjected to concentrated loads. Eng Struct. 2021;231(3):111763. doi:10.1016/j.engstruct.2020.111763.
30. Reddy GGK, Ramadoss P. Influence of alccofine incorporation on the mechanical behavior of ultra-high performance concrete (UHPC). Mater Today Proc. 2020;33(2):789–97. doi:10.1016/j.matpr.2020.06.180.

31. Bae Y, Pyo S. Effect of steel fiber content on structural and electrical properties of ultra high performance concrete (UHPC) sleepers. Eng Struct. 2020;222(3):111131. doi:10.1016/j.engstruct.2020.111131.
32. 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(5):118157. doi:10.1016/j.jclepro.2019.118157.
33. Fan D, Yu R, Shui Z, Wu C, Wang J, Su Q. A novel approach for developing a green ultra-high performance concrete (UHPC) with advanced particles packing meso-structure. Constr Build Mater. 2020;265(10):120339. doi:10.1016/j.conbuildmat.2020.120339.
34. Ragalwar K, Heard WF, Williams BA, Ranade R. Significance of the particle size distribution modulus for strain-hardening-ultra-high performance concrete (SH-UHPC) matrix design. Constr Build Mater. 2020;234(4):117423. doi:10.1016/j.conbuildmat.2019.117423.
35. Liu K, Yu R, Shui Z, Li X, Guo C, Yu B, et al. Optimization of autogenous shrinkage and microstructure for ultra-high performance concrete (UHPC) based on appropriate application of porous pumice. Constr Build Mater. 2019;214(7):369–81. doi:10.1016/j.conbuildmat.2019.04.089.
36. Wang JJ, Nie XF, Yang L, Li WG, Zhang SS. Compressive behavior of FRP-UHPC-steel double-skin tubular columns. Eng Struct. 2025;325:119451. doi:10.1016/j.engstruct.2024.119451.
37. Ocelić A, Baričević A, Smrkić MF. Synergistic integration of waste fibres and supplementary cementitious materials to enhance sustainability of ultra-high-performance concrete (UHPC). Case Stud Constr Mater. 2024;20(1):e02772. doi:10.1016/j.cscm.2023.e02772.
38. Zhang SS, Wang JJ, Lin G, Yu T, Fernando D. Stress-strain models for ultra-high performance concrete (UHPC) and ultra-high performance fiber-reinforced concrete (UHPFRC) under triaxial compression. Constr Build Mater. 2023;370(9):130658. doi:10.1016/j.conbuildmat.2023.130658.
39. Hiew SY, Bin Teoh K, Raman SN, Hung CC, Chaen YX, Kong D, et al. A unified tensile constitutive model for mono/hybrid fibre-reinforced ultra-high-performance concrete (UHPC). Cem Concr Compos. 2024;150:105553. doi:10.1016/j.cemconcomp.2024.105553.
40. Khaksefidi S, Ghalehnovi M, de Brito J. Bond behaviour of high-strength steel rebars in normal (NSC) and ultra-high performance concrete (UHPC). J Build Eng. 2021;33(1):101592. doi:10.1016/j.jobe.2020.101592.
41. Wang S, Xu L, Yin C, Cui K, Chi Y. Constitutive behavior of ultra-high-performance steel fiber reinforced concrete under monotonic and cyclic tension. J Build Eng. 2023;68(6):105991. doi:10.1016/j.jobe.2023.105991.
42. Jiao Y, Zhang Y, Guo M, Zhang L, Ning H, Liu S. Mechanical and fracture properties of ultra-high performance concrete (UHPC) containing waste glass sand as partial replacement material. J Clean Prod. 2020;277(5):123501. doi:10.1016/j.jclepro.2020.123501.
43. Wang X, Yu R, Song Q, Shui Z, Liu Z, Wu S, et al. Optimized design of ultra-high performance concrete (UHPC) with a high wet packing density. Cem Concr Res. 2019;126(6):105921. doi:10.1016/j.cemconres.2019.105921.
44. Ouyang X, Shi C, Wu Z, Li K, Shan B, Shi J. Experimental investigation and prediction of elastic modulus of ultra-high performance concrete (UHPC) based on its composition. Cem Concr Res. 2020;138(2):106241. doi:10.1016/j.cemconres.2020.106241.
45. Arora A, Almujaddidi A, Kianmofrad F, Mobasher B, Neithalath N. Material design of economical ultra-high performance concrete (UHPC) and evaluation of their properties. Cem Concr Compos. 2019;104:103346. doi:10.1016/j.cemconcomp.2019.103346.
46. Alkaysi M, El-Tawil S. Factors affecting bond development between ultra high performance concrete (UHPC) and steel bar reinforcement. Constr Build Mater. 2017;144(5):412–22. doi:10.1016/j.conbuildmat.2017.03.091.
47. Sun Y, Yu R, Shui Z, Wang X, Qian D, Rao B, et al. Understanding the porous aggregates carrier effect on reducing autogenous shrinkage of ultra-high performance concrete (UHPC) based on response surface method. Constr Build Mater. 2019;222(7):130–41. doi:10.1016/j.conbuildmat.2019.06.151.
48. 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(2):106552. doi:10.1016/j.asoc.2020.106552.
49. Smarzewski P. Mechanical properties and durability of ultra-high performance concrete containing steel fibers. Compos Struct. 2025;371(11):119471. doi:10.1016/j.compstruct.2025.119471.

50. Li W, Li W, Lu Y, Hu B, Zhou Y, Wu H, et al. Axial compressive behavior and failure mechanism of CFRP partially confined ultra-high performance concrete (UHPC). Constr Build Mater. 2024;426:136104. doi:10.1016/j.conbuildmat.2024.136104.
51. Fan D, Lu JX, Liu K, Ban J, Yu R, Poon CS. Multi-scale design of ultra-high performance concrete (UHPC) composites with centroplasm theory. Compos Part B Eng. 2024;281(5):111562. doi:10.1016/j.compositesb.2024.111562.
52. Lin JX, Luo RH, Su JY, Guo YC, Chen WS. Coarse synthetic fibers (PP and POM) as a replacement to steel fibers in UHPC: tensile behavior, environmental and economic assessment. Constr Build Mater. 2024;412(3):134654. doi:10.1016/j.conbuildmat.2023.134654.
53. Wang S, Xu L, Chi Y, Cui K, Yin C, Li B. Cyclic tensile behavior of ultra-high performance concrete with hybrid steel-polypropylene fiber: experimental study and analytical model. Compos Struct. 2023;321(12):117255. doi:10.1016/j.compstruct.2023.117255.
54. Zhang Y, Zhu Y, Qiu J, Hou C, Huang J. Impact of reinforcing ratio and fiber volume on flexural hardening behavior of steel reinforced UHPC beams. Eng Struct. 2023;285(2):116067. doi:10.1016/j.engstruct.2023.116067.
55. 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(3):148–57. doi:10.1016/j.cemconcomp.2017.02.010.
56. Qian D, Yu R, Shui Z, Sun Y, Jiang C, Zhou F, et al. A novel development of green ultra-high performance concrete (UHPC) based on appropriate application of recycled cementitious material. J Clean Prod. 2020;261(3):121231. doi:10.1016/j.jclepro.2020.121231.
57. Zhang H, Ji T, Zeng X, Yang Z, Lin X, Liang Y. Mechanical behavior of ultra-high performance concrete (UHPC) using recycled fine aggregate cured under different conditions and the mechanism based on integrated microstructural parameters. Constr Build Mater. 2018;192(2):489–507. doi:10.1016/j.conbuildmat.2018.10.117.
58. Sadrmomtazi A, Tajasosi S, Tahmouresi B. Effect of materials proportion on rheology and mechanical strength and microstructure of ultra-high performance concrete (UHPC). Constr Build Mater. 2018;187(5):1103–12. doi:10.1016/j.conbuildmat.2018.08.070.
59. Zhang WF, Gu XW, Qiu JP, Liu JP, Zhao YQ, Li XH, et al. Effects of iron ore tailings on the compressive strength and permeability of ultra-high performance concrete. Constr Build Mater. 2020;260(4):119917. doi:10.1016/j.conbuildmat.2020.119917.
60. Yang R, Yu R, Shui Z, Guo C, Wu S, Gao X, et al. The physical and chemical impact of manufactured sand as a partial replacement material in ultra-high performance concrete (UHPC). Cem Concr Compos. 2019;99(1):203–13. doi:10.1016/j.cemconcomp.2019.03.020.
61. Xuan MY, Wang YS, Lin R, Kwon SJ, Wang XY. Strength enhancement–carbon emission reduction–autogenous shrinkage mitigation of cement-slag-limestone ultra-high-performance concrete by using natural wollastonite. J Build Eng. 2025;115(7):114598. doi:10.1016/j.jobe.2025.114598.
62. Guo LK, Chen LB, Chen ZW, Wang JY. Enhancing UHPC tensile performance using polystyrene beads: significant improvements and mechanisms. Materials. 2024;17(11):2479. doi:10.3390/ma17112479.
63. Luo X, Matsumoto T. Tensile strength estimation of UHPFRC based on predicted cracking location using deep learning. Materials. 2025;18(10):2237. doi:10.3390/ma18102237.