Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Understanding the Governing Parameters of UHPC's Strength Using Explainable Machine Learning

Nafis Niaz Chowdhury
Ahsanullah University of Science and Technology
Syeda Tabassum
Khulna University of Engineering and Technology
Badhon Singha
badhon10776@gmail.com

Ahsanullah University of Science and Technology

Research Article

Keywords: UHPC, compressive strength, XGBoost, SHAP, explainable AI, machine learning

Posted Date: June 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9859596/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Understanding the Governing Parameters of UHPC's Strength Using Explainable Machine Learning

Nafis Niaz Chowdhury[1] orcid, Syeda Tabassum[2], and Badhon Singha[3] orcid

[1,3] Department of Civil Engineering, Ahsanullah University of Science and Technology, Dhaka, Bangladesh

[2] Department of Building Engineering and Construction Management, Khulna University of Engineering and Technology, Khulna, Bangladesh

Corresponding Author: Badhon Singha

Corresponding Author's Email: badhon10776@gmail.com

Abstract:

Ultra-High-Performance Concrete (UHPC) exhibits highly nonlinear compressive strength behavior governed by complex interactions among mixture constituents and curing conditions, which are difficult to capture using conventional empirical models. This study applies explainable machine learning techniques to predict and interpret UHPC compressive strength using a curated database of 810 experimental mixtures compiled from peer-reviewed literature. Four machine learning algorithms namely Support Vector Regression (SVR), Artificial Neural Network (ANN), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) were trained and evaluated using systematic hyperparameter optimization and five-fold cross-validation. Among the evaluated models, XGBoost achieved the highest predictive performance with a test R^2 of 0.9641, RMSE of 7.3237 MPa, and MAE of 5.0748 MPa, while cross-validation produced a mean R^2 of 0.970 ± 0.004, indicating strong generalization capability. Pearson correlation analysis showed that silica fume ($r = 0.521$), fiber content ($r = 0.440$), curing age ($r = 0.415$), and superplasticizer dosage ($r = 0.396$) had the strongest positive correlations with compressive strength. SHapley Additive Explanations (SHAP) were subsequently employed to interpret the governing parameters influencing model predictions. SHAP analysis identified fiber content and curing age as the most influential variables controlling UHPC compressive strength development. Local SHAP interpretation further demonstrated that fiber content and curing age contributed +14.82 MPa and +12.26 MPa, respectively, toward a representative predicted strength of 153.65 MPa. The explainability analysis revealed physically meaningful trends associated with hydration maturity, matrix densification, and fiber-reinforcement mechanisms, demonstrating that explainable machine learning can provide both high predictive reliability and engineering interpretability for UHPC strength assessment.

Keywords: UHPC, compressive strength, XGBoost, SHAP, explainable AI, machine learning.

1 Introduction

Ultra-high-performance concrete (UHPC) represents a major advancement in cementitious composites, characterized by an exceptionally dense microstructure and superior mechanical performance[1,2]. UHPC compressive strength typically exceeds 150 Mpa [3], with lab/specialized mixes (pressure-heat cured

reactive powder concrete) reaching up to 810 MPa [2]. UHPC's exceptional strength, durability, and energy absorption capacity makes it more appealing for demanding structural applications such as long-span bridges, high-rise structures, and protective infrastructure systems. Compressive strength is considered the most important mechanical property since it directly affects structural integrity, load-bearing capacity, safety, and long-term serviceability[4]. These exceptional properties arise from fundamental material science mechanisms, including microstructural refinement, reduced porosity, and enhanced homogeneity [4,5]. The scientific complexity of UHPC stems from its carefully engineered composition, which incorporates an ultra-low water-to-binder (w/b) ratio typically between 0.15 and 0.25 and a high content of cementitious materials combined with fine pozzolanic additives such as silica fume [6]. Strength development in UHPC is inherently time-dependent and governed by continuous hydration, as well as by the interaction between the dense cementitious matrix and micro-steel fibers that bridge microcracks and provide lateral confinement[7,8]. These coupled mechanisms generate a highly nonlinear and multidimensional performance space, in which the influence of individual parameters such as superplasticizer dosage is strongly modulated by factors including the specific surface area of fine powders and the w/b ratio. Consequently, UHPC exhibits interaction-driven behavior that fundamentally distinguishes it from conventional concrete, as variations in a single constituent can disproportionately affect overall performance [9,10].

Although UHPC exhibits advanced material properties, its compressive strength is still commonly assessed through destructive testing of standard specimens at fixed curing ages. This approach is time-consuming, material-intensive, and retrospective. To reduce experimental dependency, empirical and regression-based predictive models have been proposed as cost-effective alternatives. However, these formulations typically impose predefined functional structures, such as:

$$\hat{f}(\mathbf{x}) = \beta_0 + \sum_{i=1}^{n} \beta_i \, x_i \qquad 1$$

where β_i are regression coefficients estimated via ordinary least squares or maximum likelihood methods. Such linear or weakly nonlinear mappings are may not adequately capture of representing higher-order feature interactions, multiplicative coupling effects, and the threshold-sensitive nonlinearities intrinsic to UHPC hydration and microstructural systems. Polynomial extensions and interaction terms can partially mitigate these constraints, but they require manual feature engineering and remain vulnerable to multicollinearity among highly correlated mix design parameters such as the co-variance between cement content, silica fume dosage, and water-to-binder ratio[9,11–13] . Furthermore, widely adopted design standards, including ACI 318-19 and Eurocode 2 (EN 1992), lack explicit constitutive provisions for UHPC and rely on simplified stress–strain relationships calibrated for normal-strength concrete[9,14]. Existing code-based models cannot fully represent the complex failure behavior of UHPC under compression, leading to less accurate and less reliable predictions for different mix compositions.[15].

Machine learning (ML) offers a fundamentally different modeling paradigm that demonstrates improved interpretability by approximating nonlinear input-output mappings directly from experimental data, without imposing constitutive constraints:

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

where $x \in \mathbb{R}^d$ represents the d-dimensional feature vector comprising mixture parameters and curing conditions, $f(x)$ is a learned nonlinear function unconstrained by predefined analytical forms, and ε is the irreducible stochastic error term [16]. This formulation allows ML models to autonomously identify interaction effects, feature non-monotonicities, and latent structural patterns embedded within high-dimensional UHPC strength datasets. Early applications to UHPC compressive strength prediction primarily used standalone learners such as Artificial Neural Networks (ANNs), Support Vector Regression (SVR), and univariate Decision Trees. Although these models can approximate smooth nonlinear response surfaces, their predictive reliability is limited by sensitivity to hyperparameter settings, initialization in gradient-based optimization, and susceptibility to overfitting when trained on the sparse and imbalanced databases typical of UHPC research. The bias–variance trade-off inherent in single-estimator models is particularly problematic in low-data regimes, where high-variance learners tend to memorize noise rather than generalize underlying material behavior.

Ensemble learning methods improve prediction accuracy by combining the outputs of multiple base learners instead of relying on a single estimator:

$$\hat{y} = \sum_{k=1}^{K} w_k \; h_k(\mathbf{x}) \quad (3)$$

where $h_k(\mathbf{x})$ denotes the prediction of the k-th base estimator (typically a shallow decision tree), w_k is its associated weighting coefficient, and K is the total ensemble cardinality. Bagging-based methods such as Random Forest reduce prediction variance by constructing each h_k on bootstrap-resampled subsets of the training data combined with random feature subspace selection, thereby promoting decorrelation among base learners[17]. Boosting-based algorithms such as XGBoost and CatBoost, in contrast, operate through sequential residual correction, where each successive learner h_{k+1} is trained to approximate the negative gradient of a differentiable loss function with respect to the ensemble's current prediction effectively performing functional gradient descent in hypothesis space[18,19]. This hierarchical residual-fitting strategy enables improved variance reduction, enhanced generalization under distributional shift, and superior representation of nonlinear parameter interactions and locally varying feature importance, making ensemble methods particularly well-suited to the complex, interaction-driven compressive strength behavior of UHPC.

Although ensemble machine learning models provide highly accurate predictions, their complex structure and lack of simple mathematical equations make them difficult to interpret, verify physically, and widely

accept in structural practice. The opacity of high-capacity learners often described as the "black-box" problem renders it impossible to audit whether model predictions arise from physically grounded feature relationships or from spurious statistical correlations induced by data imbalances or distributional artifacts[20,21]. To address this critical limitation, Explainable Artificial Intelligence (XAI) methods were systematically integrated into the predictive framework as post-hoc interpretability tools. Among available XAI approaches, SHapley Additive exPlanations (SHAP) provide the most theoretically rigorous framework for local model explanation, grounded in axiomatic cooperative game theory. SHAP decomposes the prediction for any individual observation into additive feature contributions according to:

$$\hat{y} = \phi_0 + \sum_{i=1}^{n} \phi_i \qquad 4$$

where ϕ_0 denotes the baseline prediction corresponding to the mean model output over the training distribution, and ϕ_i represents the Shapley value for the i-th input feature defined formally as the weighted average of the marginal contribution of feature i across all possible feature coalitions $S \subseteq \mathcal{F} \setminus \{i\}$:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|\,S\,|\,!\;(|\,F\,| - |\,S\,| - 1)!}{|\,F\,|\,!} [\mathrm{f(S} \cup \mathrm{\{i\})} - \mathrm{f(S)}] \qquad 5$$

Where, S= subset of features F= full feature set $|\,S\,|$= cardinality. This formulation satisfies the classical Shapley axioms of efficiency, symmetry, dummy, and linearity, ensuring that attribution values are unique, consistent, and invariant to model architecture. The resulting additive decomposition enables both global feature importance ranking aggregated via mean absolute SHAP values across the dataset and local sample-level attribution, allowing examination of how individual mix design decisions influence a specific predicted strength outcome.

Previous SHAP-based UHPC studies have typically identified curing age and cement content as dominant they directly governing parameters because influence hydration kinetics and C–S–H formation, while water-to-binder ratio exerts a strong nonlinear negative effect on strength consistent with the Powers–Brownyard hydration theory [14]. However, the relative importance of individual parameters depends on mixture diversity, curing regime distribution, and dataset composition, which motivates the dataset-specific SHAP analysis performed in this work. In the present study, SHAP is applied to the best-performing XGBoost model to quantify the relative influence of mixture composition and curing conditions including fiber content, curing age, silica fume dosage, and superplasticizer content on UHPC compressive strength. This explainable ML framework provides engineering insight into the key governing parameters and supports performance-based UHPC mix design and optimization, directly aligning with the empirical findings reported in the subsequent sections.

2 Conceptual Framework

This study applies machine learning (ML) and explainable artificial intelligence (XAI) techniques to investigate the governing parameters influencing the compressive strength of Ultra-High-Performance Concrete (UHPC). Due to the highly nonlinear behavior of UHPC, conventional empirical models often fail to fully represent the combined effects of mixture composition and curing conditions. Machine learning models are used to capture complex relationships among input variables, while SHapley Additive Explanations (SHAP) are employed to interpret the relative influence of different parameters on compressive strength development. This approach improves model transparency and provides engineering insight into the material mechanisms governing UHPC strength behavior, including hydration, matrix densification, and fiber reinforcement effects.

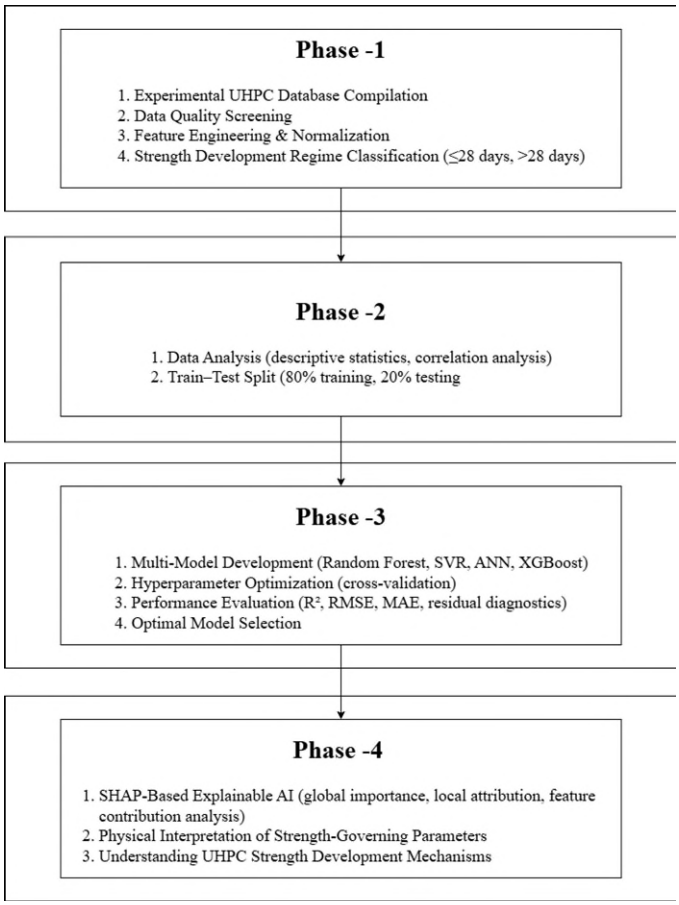


Figure 1: Data-Driven Predictive and Interpretability Framework for UHPC Compressive Strength

The end-to-end workflow follows a structured pipeline designed to ensure robustness and reliability. As illustrated in the flowchart Figure 1 the process begins with the compilation of a comprehensive experimental database derived from peer-reviewed literature, followed by rigorous data preprocessing to handle cleaning and normalization. After an exploratory correlation analysis, the data is divided into an 80/20 train-test split to facilitate multi-model learning using Random Forest, SVR, ANN, and XGBoost.

Each model is quantitatively evaluated through statistical metrics such as, RMSE, and MAE, alongside graphical diagnostics to assess stability and bias. Once the best-performing model (XGBoost) is identified, it is subjected to SHAP analysis to perform global and local feature attribution. This final step bridges the gap between raw data and engineering interpretation, providing a dependable tool for performance-based UHPC mixture design and optimization.

2.1 Experimental Data Acquisition and Curation

The foundation of this study is an experimentally derived database compiled from [22] Dataset. The dataset compiles experimental results from multiple peer-reviewed sources and represents a high degree of UHPC mix diversity, encompassing a broad range of binder systems, supplementary cementitious materials, fiber dosages, and various curing regimes. To ensure the credibility and reliability of the data-driven models, strict inclusion and exclusion criteria were implemented. The collected data were screened for completeness and physical consistency; records with missing key parameters, duplicate entries, or physically inconsistent values such as unrealistic material proportions or anomalous strength results were excluded to minimize noise and prevent information leakage.

Following systematic screening and quality control, the curated database comprised 810 unique UHPC mixtures, each associated with a measured compressive strength value. The dataset contains 13 physically measurable input variables and one output variable (compressive strength), ensuring statistical representativeness while preserving engineering interpretability.

2.2 Data Preprocessing and Strength-Development Regime Identification

The UHPC experimental database was carefully processed to create a reliable dataset for machine learning analysis. First, duplicate and incomplete records were removed to ensure that each mixture represented a unique experimental result. Mixtures with unrealistic material proportions or abnormal compressive strength values were also excluded to reduce noise and improve data quality. In addition, only directly measurable and physically meaningful parameters were selected as input features. This helped reduce redundancy, prevent overfitting, and ensure that the models learned from actual material behavior rather than misleading data patterns.

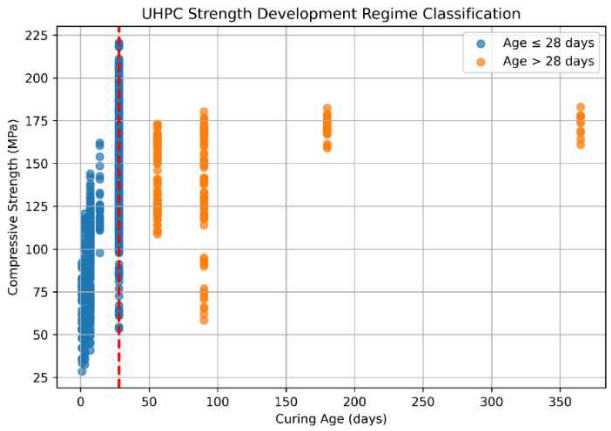


Figure 2: Classification of UHPC compressive strength development regimes based on curing age (from the algorithm)

An important part of the data processing was dividing the UHPC data into two curing-age groups because curing age strongly affects hydration and strength development. The dataset was classified into:

Early-age regime: Age ≤ 28 days

Later-age regime: Age > 28 days

Out of 810 samples, 628 samples (77.53%) belonged to the early-age group, while 182 samples (22.47%) were in the later-age group. The 28-day limit was used to separate the two stages of strength development, as shown in Figure 2. The early-age stage represents rapid strength gain due to active hydration and microstructural development, whereas the later-age stage reflects slower strength growth as the concrete approaches a stable structure. Including both stages helps the model learn short-term and long-term strength behavior, improving prediction reliability and reducing bias.

2.3 Statistical Description of Input Variables

The dataset consists of 13 input variables, including cement (C), sand (S), silica fume (SF), limestone powder (LP), quartz powder (QP), fine aggregate (FA), nano-silica (NS), aggregate(A), water (W), fiber content (Fi), curing temperature (T), and curing age (Age). The output variable is compressive strength (CS). These variables were selected to ensure mechanistic relevance and to minimize redundancy or information leakage. Table 1 represents the statistical summary of these input variables, including their minimum, maximum, mean, and standard deviation values. The statistics highlight the variability in UHPC mixture compositions and curing conditions across the compiled dataset.

Table 1: Statistical Summary of Input Variables (dataset:[22])

Variable	Mean	Std_Dev	Min	Max	Count
C	737.91	173.46	270.0	1251.2	810

S	25.19	74.37	0.0	375.0	810
SF	136.99	104.14	0.0	433.7	810
LP	41.93	133.13	0.0	1058.2	810
QP	33.27	79.67	0.0	397.0	810
FA	26.26	67.46	0.0	356.0	810
NS	3.64	7.78	0.0	47.5	810
A	1150.11	312.15	407.8	1992.0	810
W	179.89	25.57	90.0	272.6	810
Fi	56.04	75.23	0.0	234.0	810
SP	30.03	13.99	1.1	57.0	810
T	23.92	16.21	20.0	210.0	810
Age	37.09	53.12	1.0	365.0	810

The maximum curing temperature of 210 °C corresponds to accelerated heat-curing and autoclave regimes reported in selected UHPC studies. However, most samples in the compiled database were subjected to standard laboratory curing conditions (20-25 °C), which explains the relatively low mean curing temperature of 23.9 °C. The dataset therefore incorporates both ambient and elevated-temperature curing regimes to ensure broader generalization capability of the predictive model across diverse UHPC production practices.

2.4 Machine Learning Model Development and Evaluation Framework

To predict the compressive strength of Ultra-High-Performance Concrete (UHPC), four machine learning models Random Forest (RF), Support Vector Regression (SVR), Artificial Neural Network (ANN), and Extreme Gradient Boosting (XGBoost) were used. These models represent different machine learning approaches, including ensemble-based, kernel-based, and neural-network-based methods, allowing a balanced comparison of predictive performance. The prediction problem was formulated as a supervised regression model:

$$CS = f(X)$$

where $CS$is the compressive strength and Xrepresents the mixture composition and curing-related input variables.

Before training, all input variables were standardized to improve numerical stability and reduce scale-related bias. The dataset was divided into training and testing subsets using an 80/20 split, and the same data partition was applied to all models to ensure fair comparison.

To improve model reliability, hyperparameter optimization was performed using GridSearchCV and RandomizedSearchCV with five-fold cross-validation. The coefficient of determination (R^2) was used as the optimization metric. RandomizedSearchCV was applied to RF, ANN, and XGBoost, while GridSearchCV was used for SVR. In addition, random seeds were fixed at 42 to ensure reproducibility.

The optimized hyperparameter configurations obtained through systematic cross-validation-based tuning are summarized in Table 2. For ensemble-based models, the optimization process identified configurations that effectively balanced predictive accuracy and overfitting resistance. The Random Forest model achieved optimal performance using 500 estimators with unrestricted tree depth, while XGBoost performed best with 500 boosting estimators, a learning rate of 0.1, maximum tree depth of 6, and a subsampling ratio of 0.8. For SVR, the optimal radial basis function kernel parameters were identified as $C = 500$and $\gamma =$auto. The ANN model achieved the best performance with a hidden-layer architecture of (100, 50), a learning rate initialization of 0.01, and L2 regularization parameter $\alpha = 0.01$. These optimized configurations were subsequently used for final model training and evaluation.

Table 2: Hyperparameter Configuration of Machine Learning Models (from the algorithm)

Model	**Search Method**	**Search Space**	**Best Parameters**
Random Forest (RF)	RandomizedSearchCV	n_estimators: [100, 300, 500], max_depth: [10, 15, None], min_samples_split: [2, 5, 10]	n_estimators = 500, max_depth = None, min_samples_split = 2
SVR	GridSearchCV	C: [10, 100, 500], gamma: [0.01, 0.1, auto], kernel = rbf	C = 500, gamma = auto
ANN (MLPRegressor)	RandomizedSearchCV	hidden_layer_sizes: [(50,50), (100,50)], learning_rate_init: [0.001, 0.01], alpha: [0.0001, 0.01]	hidden_layer_sizes = (100,50), learning_rate_init = 0.01, alpha = 0.01
XGBoost	RandomizedSearchCV	n_estimators: [100, 300, 500], learning_rate: [0.01, 0.05, 0.1], max_depth: [4, 6, 8], subsample: [0.8, 1.0]	n_estimators = 500, learning_rate = 0.1, max_depth = 6, subsample = 0.8

Model performance was evaluated using a combination of quantitative statistical metrics and graphical diagnostics to ensure a rigorous and holistic assessment. The coefficient of determination (R^2), root mean

square error (RMSE), and mean absolute error (MAE) were employed to evaluate goodness-of-fit, prediction accuracy, and average error magnitude, respectively.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad 6$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad 7$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \quad 8$$

In addition to these numerical measures, actual-versus-predicted plots were used to assess the alignment of predictions with experimental values, residual scatter plots were examined to detect potential bias or heteroscedasticity, and error distribution histograms were analyzed to evaluate error concentration and tail behavior. This integrated evaluation strategy ensures that the selected optimal model demonstrates not only high predictive accuracy but also stability, robustness, and reliable generalization across the full range of UHPC strength development mechanism.

3 Statistical Characteristics and Correlation Analysis

The correlation heatmap (Figure 3) illustrates the statistical relationships between Ultra-High-Performance Concrete (UHPC) mixture parameters and compressive strength, providing an initial insight into the linear dependencies within the experimental database.

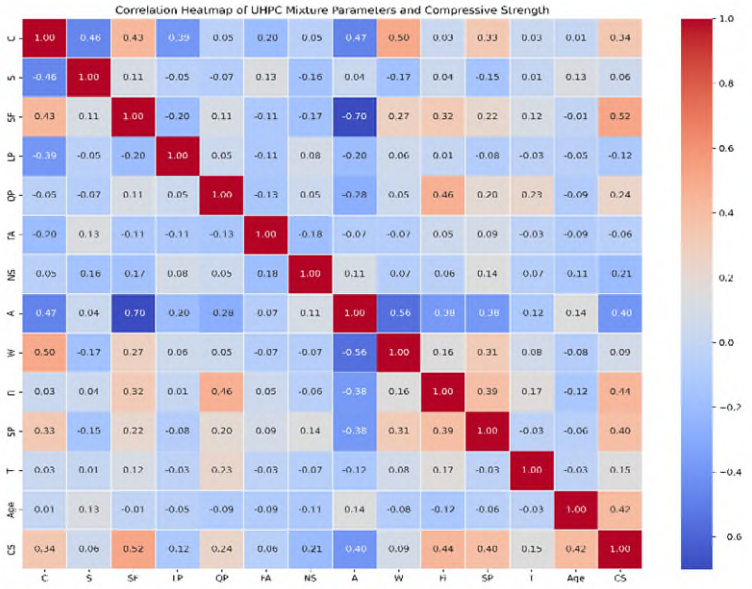


Figure 3: Correlation heatmap of UHPC mixture parameters and compressive strength

The correlation matrix shown in Figure 3 was computed from the curated UHPC experimental database to evaluate the linear relationships between mixture parameters and compressive strength (CS). The analysis provides preliminary insight into the governing variables influencing UHPC strength development and helps identify potential interactions among material constituents and curing conditions. Among all input variables, silica fume (SF) exhibited the strongest positive correlation with compressive strength ($r = 0.521$), indicating its significant role in enhancing matrix densification and pozzolanic reaction activity within UHPC systems. Fiber content (Fi) also showed a strong positive correlation ($r = 0.440$), reflecting the contribution of steel fibers to crack-bridging, confinement, and stress redistribution mechanisms. Curing age demonstrated a moderate positive correlation with compressive strength ($r = 0.415$), consistent with the continued hydration and microstructural refinement processes governing UHPC strength maturation.

In addition, superplasticizer dosage (SP) showed a positive association with compressive strength ($r = 0.396$), suggesting its influence on particle dispersion, workability improvement, and enhanced packing density. In contrast, the age-related parameter (A) exhibited a moderate negative correlation with compressive strength ($r = -0.399$), indicating an inverse relationship with strength development under the studied conditions.

Several other parameters, including cement content (C) and quartz powder (QP), demonstrated weaker positive correlations, whereas limestone powder (LP), nano-silica (NS), and fine aggregate (FA) exhibited relatively low or negative linear relationships with compressive strength. Overall, the correlation analysis indicates that UHPC compressive strength is governed by multiple interacting variables rather than a single dominant parameter, highlighting the necessity of nonlinear and explainable machine learning approaches for understanding complex UHPC strength behavior. It is important to note that the Pearson correlation and SHAP importance rankings differ because they measure fundamentally distinct quantities. Silica fume exhibited the highest Pearson correlation with compressive strength (r = 0.521), indicating a strong positive linear relationship when examined in isolation. However, within the multivariate XGBoost model, silica fume ranked third in global SHAP importance ($\bar{\varphi}$ = 6.53 MPa), because its contribution is strongly conditioned on cement content and water-to-binder ratio meaning its effect is partially absorbed by correlated features during ensemble learning. In contrast, curing age ranked first in SHAP importance ($\bar{\varphi}$ = 18.69 MPa) despite its moderate Pearson correlation (r = 0.415), reflecting its highly nonlinear, threshold-sensitive interaction with hydration kinetics and microstructural development an effect that bivariate linear correlation cannot detect. The Pearson analysis therefore serves as a useful screening tool for linear trends, while SHAP values provide the mechanistically superior representation of governing parameter influence in the nonlinear UHPC strength system.

3.1 Distribution Characteristics of Compressive Strength

To further examine dataset variability, the distribution of compressive strength (CS) values was analyzed. The histogram of CS (Figure 4) reveals a broad strength range spanning low to ultra-high-strength regimes,

confirming the dataset's suitability for training nonlinear predictive models. The distribution exhibits moderate right-skewness, reflecting a higher concentration of early-age strength measurements, consistent with the regime distribution analysis. The absence of extreme outliers indicates effective data curation and quality control.

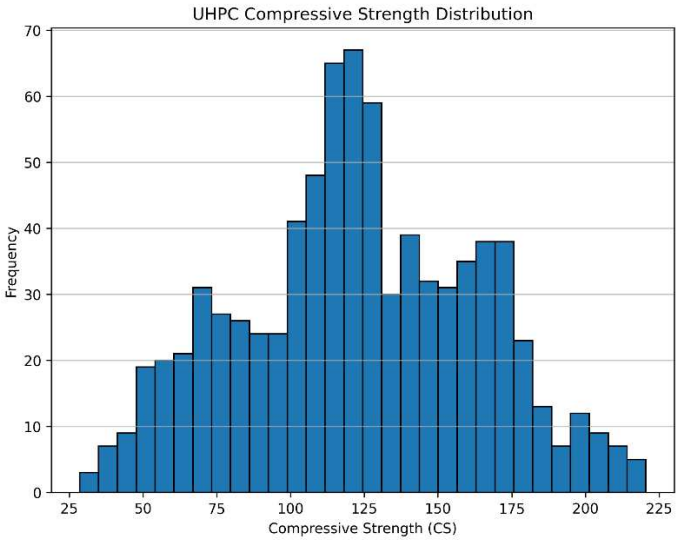


Figure 4: Histogram of UHPC compressive strength distribution

3.2 Comparative Performance of Machine Learning Models

The comparison between experimental and predicted Ultra-High-Performance Concrete (UHPC) compressive strength reveals clear differences in predictive performance among the four evaluated machine learning models. The Artificial Neural Network (ANN) model exhibited the weakest predictive capability, achieving the lowest test R^2value of 0.8678 along with the highest prediction errors (RMSE = 14.0472 MPa and MAE = 9.5168 MPa), as presented in Table 3. In the actual-versus-predicted plot, ANN predictions display considerable scatter around the 45° reference line, particularly at higher compressive strength levels, indicating limited generalization capability in capturing the highly nonlinear behavior of UHPC materials. This behavior is further reflected in the residual scatter plot (Figure 5), where ANN residuals show relatively wide dispersion with several deviations exceeding ±40 MPa. The corresponding residual distribution histogram also appears broader with noticeable tails, confirming reduced prediction stability across the full-strength range of the dataset.

In comparison, the Support Vector Regression (SVR) model demonstrated moderate predictive performance with a test R^2value of 0.9025, RMSE of 12.0635 MPa, and MAE of 8.0717 MPa. Although SVR provided improved predictions relative to ANN, the model still exhibited noticeable dispersion in both prediction and residual plots, suggesting limitations in representing the complex nonlinear interactions governing UHPC compressive strength development.

.

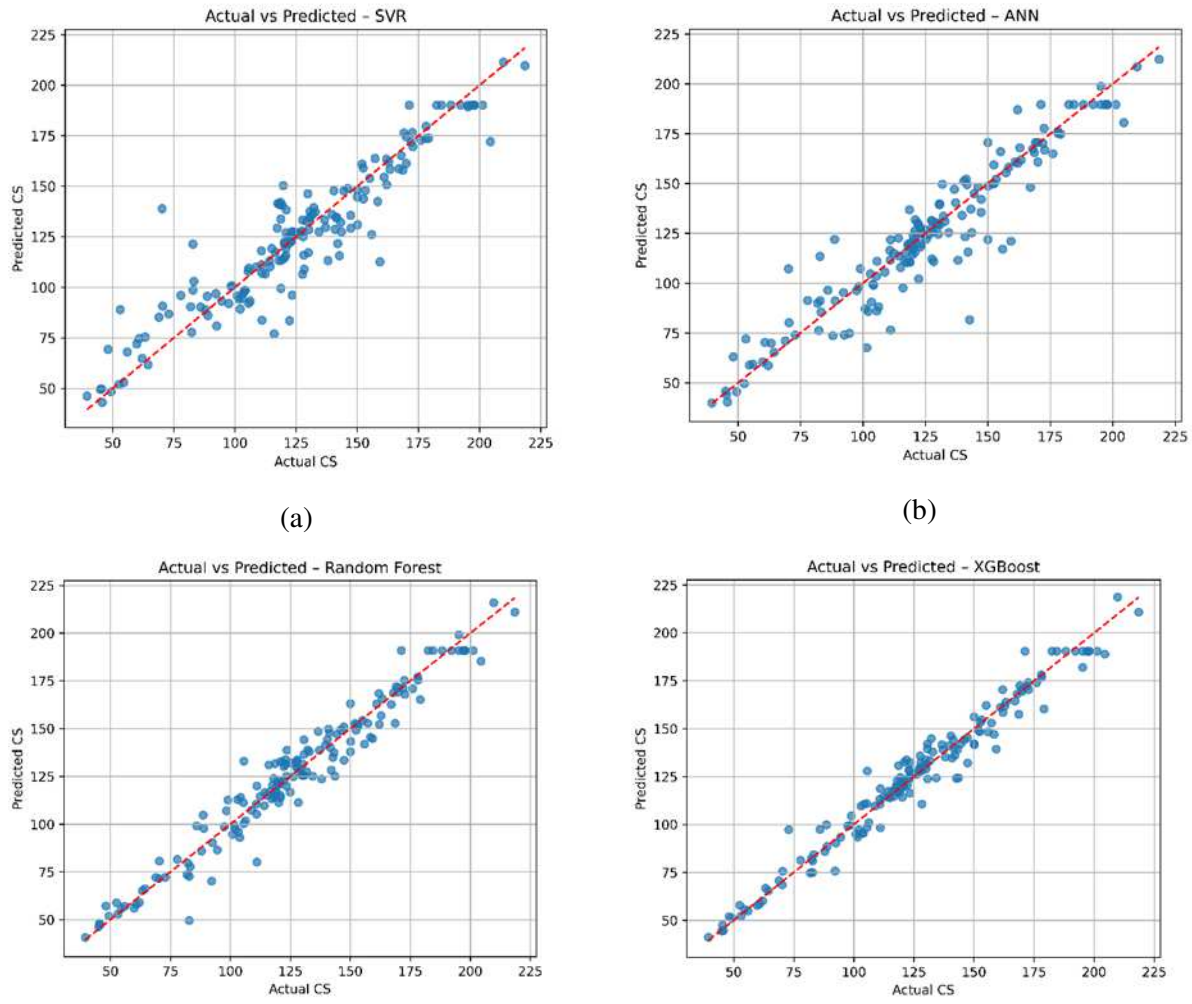


Figure 5: Actual versus predicted UHPC compressive strength for (a) SVR, (b) ANN, (c) Random Forest, and (d) XGBoost models.

The Artificial Neural Network (ANN) model demonstrated comparatively lower predictive performance than the ensemble-based approaches, achieving a test R^2 value of 0.8678, with an RMSE of 14.0472 MPa and an MAE of 9.5168 MPa. Although the ANN predictions generally followed the overall experimental strength trends, the actual-versus-predicted plot still exhibited noticeable dispersion at both lower and higher compressive strength ranges. Residual analysis further revealed relatively wide error dispersion across the full-strength range, with several residuals exceeding ±40 MPa.

The corresponding residual distribution histogram showed a broader spread and more pronounced tails compared with the ensemble-based models, indicating reduced prediction stability and intermittent large deviations. These results suggest that the ANN model had limited capability in consistently capturing the highly nonlinear and interaction-driven behavior governing UHPC compressive strength development.

Table 3: Comparison of Model Predictive Accuracy

Machine Learning Model	Train R^2	Test R^2	RMSE (MPa)	MAE (MPa)
Support Vector Regression (SVR)	0.9472	0.9025	12.0635	8.0717
Artificial Neural Network (ANN)	0.9410	0.8678	14.0472	9.5168
Random Forest (RF)	0.9891	0.9502	8.6173	6.3600
Extreme Gradient Boosting (XGBoost)	0.9929	0.9641	7.3237	5.0748

The Random Forest (RF) model showed excellent predictive performance, demonstrating the advantage of ensemble-based learning methods over standalone models. In the actual-versus-predicted plot, most RF predictions were closely distributed around the 45° reference line, indicating strong agreement between experimental and predicted values. The RF model achieved a high test R^2value of 0.9502 with relatively low prediction errors (RMSE = 8.6173 MPa and MAE = 6.3600 MPa). Residual analysis also showed improved stability, with most residuals distributed within ±15 MPa around the zero-error line. The residual histogram appeared relatively narrow and symmetric, indicating reduced prediction variance and improved robustness. Among all evaluated models, the Extreme Gradient Boosting (XGBoost) model achieved the best overall performance, with the highest test R^2value of 0.9641 and the lowest prediction errors (RMSE = 7.3237 MPa and MAE = 5.0748 MPa). The actual-versus-predicted plot showed very close agreement with the experimental results across the full-strength range. Residual analysis further confirmed the strong stability of XGBoost, with most residuals tightly concentrated around zero and generally remaining within ±10 MPa. In addition, the residual histogram for XGBoost was the most compact among all models, showing minimal skewness and very few extreme errors.

To further evaluate model reliability, five-fold cross-validation was performed for all machine learning models. The mean coefficient of determination (R^2) and corresponding standard deviation obtained from the cross-validation process are presented in Table 4.

Table 4: Five-Fold Cross-Validation Performance

Model	Mean R^2 (CV)	Std Dev
SVR	0.873	0.020
ANN	0.902	0.026
RF	0.959	0.007
XGBoost	0.970	0.004

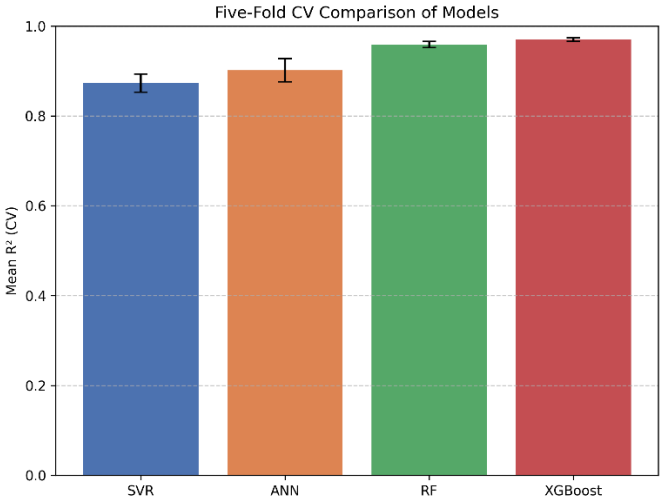


Figure 6: Five-fold cross-validation comparison of machine learning models (mean R² values)

Figure 6 presents the five-fold cross-validation results used to evaluate the generalization performance of the machine learning models. The results show that ensemble-based algorithms achieved better predictive performance than the other models. XGBoost obtained the highest mean R^2 value of 0.970 with the lowest standard deviation (±0.004), indicating strong prediction stability across different validation folds. Random Forest also demonstrated reliable performance with a mean R^2value of 0.959 ± 0.007.

In comparison, ANN and SVR produced lower mean R^2values along with relatively larger standard deviations, suggesting greater sensitivity to variations in the training and testing data partitions. Overall, the cross-validation results confirm that gradient boosting and ensemble learning methods are highly effective in capturing the complex nonlinear relationships governing UHPC compressive strength prediction.

3.3 Error Distribution Characteristics of the Developed Model

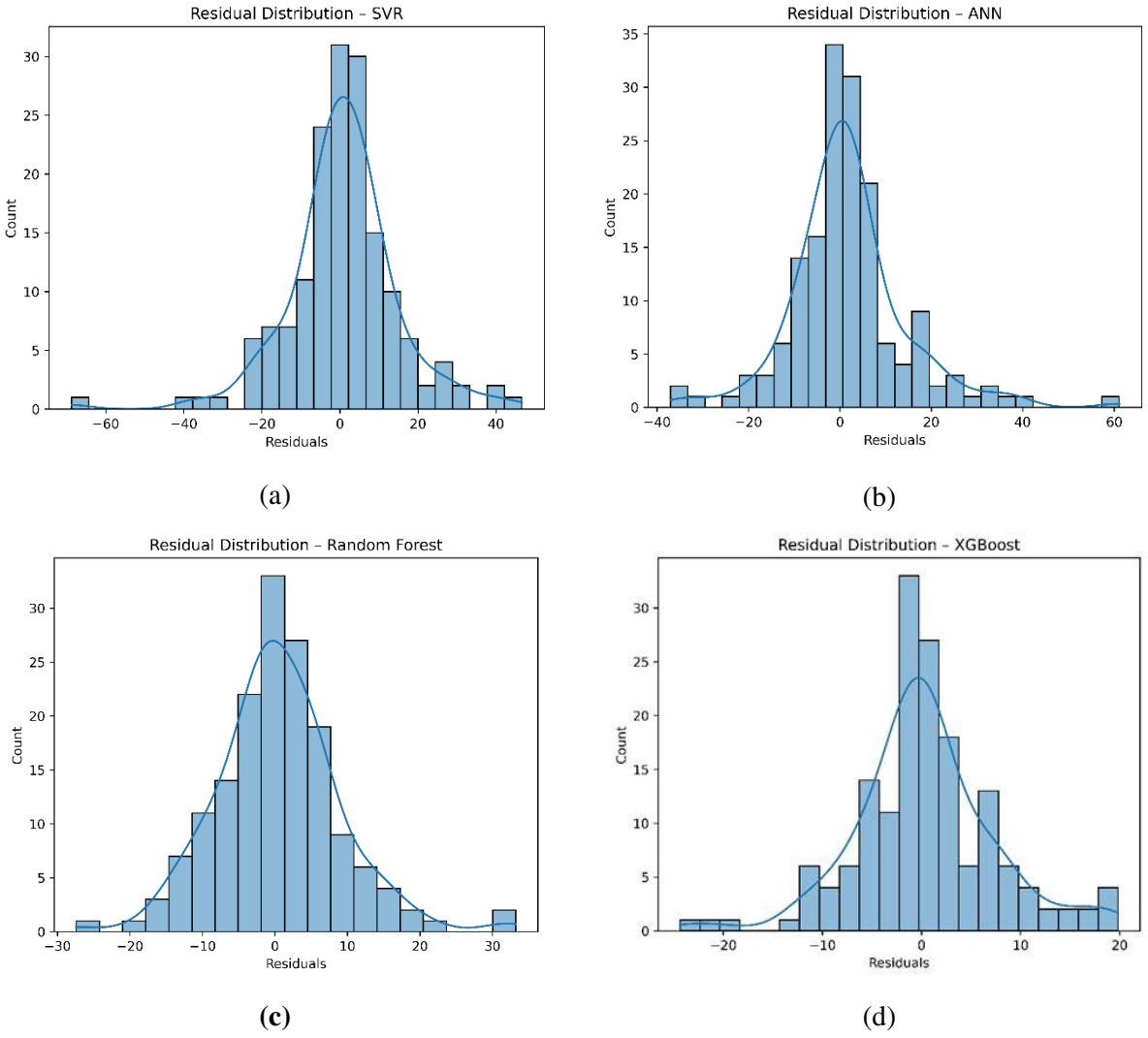


Figure 7: Residual Distribution Histograms of Machine Learning Models (a) SVR, (b) ANN, (c) Random Forest, and (d) XGBoost (from the algorithm)

Figure 7 illustrates the residual distribution histograms of the four machine learning models used for UHPC compressive strength prediction. The XGBoost model, identified as the optimal predictor, demonstrates strong predictive performance and error stability across the entire experimental range. Diagnostic analysis of the residuals reveals a high concentration around the zero-error axis, with the majority of prediction deviations tightly confined within a narrow interval of ±10 MPa.

The robustness of the proposed XGBoost model was further evaluated using five-fold cross-validation. The model achieved a mean R^2 of 0.970 with a standard deviation of 0.004, confirming stable predictive performance across different data partitions. Moreover, the residual distribution for XGBoost is the most compact and sharply peaked among all evaluated algorithms, characterized by minimal skewness and the

absence of extreme outliers. This superior error behavior underscores the model's ability to mitigate prediction uncertainty and provide consistent and reliable estimates, making it a robust tool for performance-based design and optimization of UHPC mixtures.

3.4 Explainability and Feature Influence Analysis



Figure 8: SHAP explanation of a single UHPC strength prediction, showing how feature contributions move the baseline to the final output.

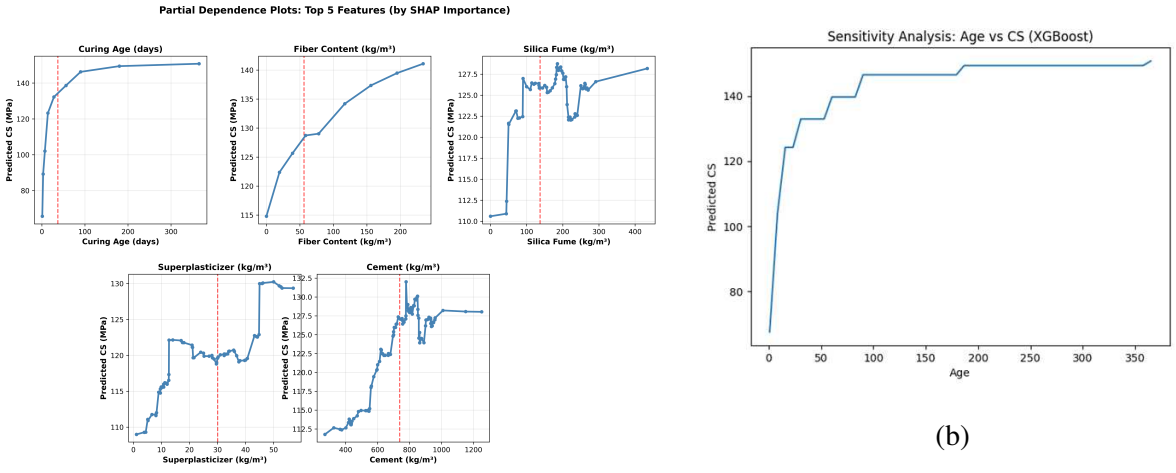
At the local level, the SHAP waterfall plot shown in Figure 8 explains how individual input variables contribute to a representative UHPC compressive strength prediction. Starting from a baseline expected value of approximately 126.7 MPa, the final predicted strength increases to 153.6 MPa through the combined effects of several influential features. Fiber content (Fi) provides the largest positive contribution (+14.82 MPa), followed by curing age (+12.26 MPa). Fine aggregate (FA) and cement content (C) also contribute positively with smaller effects of +1.93 MPa and +1.67 MPa, respectively.

In contrast, several parameters slightly reduce the predicted strength. Superplasticizer (SP), age-related parameter (A), silica fume (SF), nano-silica (NS), and water content (W) exhibit relatively small negative SHAP contributions ranging from approximately −0.34 MPa to −1.56 MPa. The remaining features collectively contribute only marginally to the final prediction.

The SHAP decomposition demonstrates that the XGBoost model predictions are governed primarily by curing age and fiber content, which is consistent with established UHPC hydration, matrix densification, and fiber-reinforcement mechanisms. The local interpretability provided by SHAP therefore improves transparency and increases confidence in the engineering reliability of the developed predictive framework.

3.5 Effect of Key Parameters and Sensitivity Analysis

To further examine the governing behavior captured by the predictive model, partial dependence plots (PDPs) and sensitivity analysis were used to evaluate the marginal effects of key parameters while averaging the influence of the remaining variables. The PDP for curing age shows a rapid increase in predicted compressive strength at early ages, followed by a gradual plateau beyond approximately 28 days.



(a) (b)

Figure 9: (a) Partial dependence plots showing the marginal effects of key variables on UHPC compressive strength. (b) Sensitivity analysis of curing age using the XGBoost model, illustrating the nonlinear strength development trend.

This trend is consistent with the transition from hydration-dominated strength development to slower long-term microstructural stabilization in UHPC. Fiber content exhibits a positive monotonic relationship with compressive strength, although the rate of increase becomes smaller at higher dosages, indicating diminishing marginal benefit at elevated fiber contents. Cement content also shows a moderate positive effect on predicted strength, suggesting its primary contribution is related to matrix densification and packing improvement rather than abrupt strength gain.

The sensitivity analysis for curing age further supports these observations by showing a smooth and monotonic increase in predicted strength with age. Sensitivity is pronounced up to approximately 28 days and then gradually levels off, indicating smaller strength increments at later ages. The absence of strong oscillations or unrealistic fluctuations suggests stable model behavior across the investigated curing range.

Overall, the effects identified through PDP and sensitivity analysis are consistent with established UHPC behavior, particularly hydration development, fiber bridging, and cementitious matrix densification. These results indicate that the model captures physically plausible nonlinear relationships among the governing parameters. Within the limits of the compiled dataset, the XGBoost-based model provides a transparent and reliable basis for interpreting UHPC strength behavior and supporting mix optimization and performance-based decision-making.

4 Conclusion

This study applied explainable machine learning techniques to investigate the governing parameters influencing the compressive strength of Ultra-High-Performance Concrete (UHPC) using a curated database of 810 experimental mixtures. Four machine learning models namely Support Vector Regression (SVR), Artificial Neural Network (ANN), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) were evaluated using systematic hyperparameter optimization and five-fold cross-validation.

Among all evaluated models, XGBoost achieved the highest predictive performance with a test coefficient of determination (R^2) of 0.9641, RMSE of 7.3237 MPa, and MAE of 5.0748 MPa, while cross-validation produced a mean R^2value of 0.970 ± 0.004, confirming strong prediction stability and generalization capability. In comparison, RF achieved a test $R^2$of 0.9502, whereas SVR and ANN produced lower predictive performance with test R^2values of 0.9025 and 0.8678, respectively.

Correlation analysis revealed that silica fume (SF) exhibited the strongest positive linear relationship with compressive strength ($r = 0.521$), followed by fiber content (Fi) ($r = 0.440$), curing age ($r = 0.415$), and superplasticizer dosage (SP) ($r = 0.396$). In contrast, the age-related parameter (A) showed a moderate negative correlation with compressive strength ($r = -0.399$). SHAP-based explainability analysis further demonstrated that fiber content and curing age were the most influential governing parameters affecting UHPC strength predictions. Local SHAP interpretation showed that fiber content increased the representative predicted strength by +14.82 MPa, while curing age contributed +12.26 MPa toward the final predicted compressive strength of 153.65 MPa.

Partial dependence and sensitivity analyses further confirmed physically realistic nonlinear strength-development trends associated with hydration maturity, matrix densification, and fiber-reinforcement mechanisms. Residual diagnostics showed that XGBoost produced the most stable prediction behavior, with most residuals concentrated within ±10 MPa and minimal systematic bias across the evaluated strength range.

The results demonstrate that explainable machine learning can effectively identify the governing parameters controlling UHPC compressive strength while maintaining high predictive reliability. The findings provide meaningful engineering insight into the relative influence of mixture constituents and curing conditions, supporting scientifically informed UHPC mix design and performance-based engineering applications.

List of Abbreviations

UHPC - Ultra-High-Performance Concrete

ML - Machine Learning

XAI - Explainable Artificial Intelligence

SHAP - SHapley Additive exPlanations

SVR - Support Vector Regression

ANN - Artificial Neural Network

MLP - Multilayer Perceptron

RF - Random Forest

XGBoost - Extreme Gradient Boosting

PDP - Partial Dependence Plot

CS - Compressive Strength

R^2 - Coefficient of Determination

RMSE - Root Mean Square Error

MAE - Mean Absolute Error

VIF - Variance Inflation Factor

w/b - Water-to-Binder Ratio

CFRP - Carbon Fiber Reinforced Polymer

DIC - Digital Image Correlation

C–S–H - Calcium Silicate Hydrate

NSGA-II - Non-dominated Sorting Genetic Algorithm II

ACI - American Concrete Institute

FHWA - Federal Highway Administration

EN - European Norm (Eurocode

Ethics approval and consent to participate

Not applicable.

Clinical trial number

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

The authors received no financial support for the research, authorship, or publication of this article.

Authors' contributions

Nafis Niaz Chowdhury and Badhon Singha were jointly responsible for conceptualization, methodology development, data curation, model implementation, statistical analysis, and manuscript preparation. Syeda Tabassum contributed to literature review, data organization and overall research supervision and coordination. She also provided valuable feedback during the interpretation and refinement of the study findings.

Acknowledgment

The authors express their gratitude to The Almighty and extend thanks to all those who provided support and contributed to the successful completion of this work. Their encouragement and assistance have been essential throughout the research process.

References

1. da Silva ML, Pepe M, Filho RDT, Martinelli E. The influence of materials on the mechanical properties of ultra-high-performance concrete (UHPC): A literature review. Materials. 2024;17(8):1801. doi:10.3390/ma17081801

2. Ullah R, Qiang Y, Ahmad J, Vatin NI, El-Shorbagy MA. Ultra-High-Performance Concrete (UHPC): A State-of-the-Art Review. Materials. 2022 Jun 10;15(12):4131. doi:10.3390/ma15124131

3. ACI Committee 239. 239R-18: Ultra-High Performance Concrete: An Emerging Technology Report [ACI Emerging Technology Report]. American Concrete Institute; 2018 Oct. p. 21. ACI Emerging Technology Report ACI 239R-18.

4. Hung CC, El-Tawil S, Chao SH. A Review of Developments and Challenges for UHPC in Structural Engineering: Behavior, Analysis, and Design. J Struct Eng. 2021 Sep;147(9):03121001. doi:10.1061/(ASCE)ST.1943-541X.0003073

5. Wang XQ, Chow CL, Lau D. Multiscale perspectives for advancing sustainability in fiber reinforced ultra-high performance concrete. npj Materials Sustainability. 2024 Jun 10;2(1):13. doi:10.1038/s44296-024-00021-z

6. Bahmani H, Mostofinejad D. Microstructure of ultra-high-performance concrete (UHPC) – A review study. Journal of Building Engineering. 2022 Jun;50:104118. doi:10.1016/j.jobe.2022.104118

7. Lim K, Kim K, Koh K, Ryu G. The Time-Dependent Effect in Ultra High-Performance Concrete According to the Curing Methods. Materials. 2022 Jul 21;15(14):5066. doi:10.3390/ma15145066

8. Mohebbi A, Graybeal B, Haber Z. Time-Dependent Properties of Ultrahigh-Performance Concrete: Compressive Creep and Shrinkage. J Mater Civ Eng. 2022 Jun;34(6):04022096. doi:10.1061/(ASCE)MT.1943-5533.0004219

9. Aylas-Paredes BK, Han T, Neithalath A, Huang J, Goel A, Kumar A, Neithalath N. Data driven design of ultra high performance concrete prospects and application. Sci Rep. 2025 Mar 18;15(1):9248. doi:10.1038/s41598-025-94484-2

10. El-Helou RG, Graybeal BA. Flexural Behavior and Design of Ultrahigh-Performance Concrete Beams. J Struct Eng. 2022 Apr;148(4):04022013. doi:10.1061/(ASCE)ST.1943-541X.0003246

11. Paruthi S, Verma R, Sharma N, Khan AH, Hasan MA. Ensemble machine learning models for predicting strength of concrete with foundry sand and coal bottom ash as fine aggregate replacements. Sci Rep. 2025 Nov 3;15(1):38331. doi:10.1038/s41598-025-22212-x

12. 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. Applied Sciences. 2024 Sep 9;14(17):8083. doi:10.3390/app14178083

13. Zhen L, Qu C, Tang ML, Yin J. High-Performance Concrete Strength Regression Based on Machine Learning with Feature Contribution Visualization. Mathematics. 2025 Dec 12;13(24):3965. doi:10.3390/math13243965

14. 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 Oct 4;17(19):4885. doi:10.3390/ma17194885

15. Xu L, Yu X, Zhu C, Wang L, Yang J. Prediction of Ultra-High-Performance Concrete (UHPC) Compressive Strength Based on Convolutional Neural Networks. Materials. 2025 Jun 17;18(12):2851. doi:10.3390/ma18122851

16. Zhu P, Cao W, Zhang L, Zhou Y, Wu Y, Ma ZJ. Interpretable Machine Learning Models for Prediction of UHPC Creep Behavior. Buildings. 2024 Jul 7;14(7):2080. doi:10.3390/buildings14072080

17. 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 [Internet]. San

Francisco California USA: ACM; 2016 [cited 2025 Dec 22]. p. 785–94. Available from: https://dl.acm.org/doi/10.1145/2939672.2939785 doi:10.1145/2939672.2939785

18. Afriadi I, Thongchom C, Kumar DR, Keawsawasvong S, Wipulanusat W. Ensemble boosting-based soft-computing models for predicting the bond strength between steel and CFRP plate. Sci Rep. 2025 Jul 1;15(1):20459. doi:10.1038/s41598-025-04866-9

19. Katlav M, Ergen F. Improved forecasting of the compressive strength of ultra-high-performance concrete (UHPC) via the CATBOOST model optimized with different algorithms. Structural Concrete. 2025 Feb;26(1):212–35. doi:10.1002/suco.202400163

20. Ekanayake IU, Meddage DPP, Rathnayake U. A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP). Case Studies in Construction Materials. 2022 Jun;16:e01059. doi:10.1016/j.cscm.2022.e01059

21. Le BA, Nguyen HQ, Tran BV, Vu TS, Bui TL. Symbolic regression model for predicting compression strength of steel fiber-reinforced concrete. Transportation Research Procedia. 2025;85:94–101. doi:10.1016/j.trpro.2025.03.138

22. Kashem A. Ultra-High-Performance Concrete (UHPC) [Internet]. Mendeley Data; 2023 [cited 2026 Mar 2]. Available from: https://data.mendeley.com/datasets/85r7bh4zsz/1 doi:10.17632/85R7BH4ZSZ.1