Original/Research Article

OPEN ACCESS

Advances in Engineering and Intelligence Systems

e-ISSN:2821-0263
Volume 05, Issue 02, 2026

Journal Web Page: https://aeis.bilijipub.com

Statistical and AI/ML Evaluation of Coarse Aggregate Size Effects on Compressive Strength of High-Performance Concrete

Warqaa Muhammad Bahaaddin [1,*], Serwan Kamal Jalal [2], Naema Sadeeq Saleh [3], Lawend Kamal Askar [3], Sevar Dilkhaz Salahaddin [2], Ahmed Mohammed Ahmed [2], Salim T. Yousif [4]

[1] *Department of Building Construction Technology, Dohuk Polytechnic University, KRG, Iraq*
[2] *Department of Civil and Environmental Engineering, University of Zakho, KRG, Iraq*
[3] *Technical College of Engineering, Duhok Polytechnic University, KRG, Iraq*
[4] *Department of Civil Engineering, Nawroz University, KRG, Iraq*

Highlights

- Aggregate size effects on the strength of high-performance concrete are evaluated.
- Robust ANOVA confirms significant main effects of mix, curing age, and size.
- A supplementary case-level dataset improves reproducibility and AI scope.
- Gradient Boosting predicts strength with an RMSE of 4.92 MPa and an R^2 of 0.885.
- Maximum strength of 83.48 MPa is achieved with 12.5 mm aggregate at 28 days.

Article Info

Received: 28 January 2026
Received in revised: 23 April 2026
Accepted: 09 May 2026
Available online: 30 May 2026

Keywords

High-Performance Concrete
Compressive Strength
Aggregate Size
Machine Learning
ANOVA
Predictive Modeling

Abstract

This study evaluates how coarse aggregate size and combined admixture regimes influence compressive strength in high-performance concrete. A total of 120 cube-level observations were assembled from four mixes, five aggregate sizes of 38, 25, 19, 12.5, and 9.5 mm, and two curing ages, 7 and 28 days, with three specimens per design cell. The mixes followed a 1:1.16:2.02 proportion, cement: sand: aggregate and a nominal water-to-cement ratio of 0.32, while Mix2-Mix4 incorporated 5/0.5, 7/0.7, and 10/0.9 silica-fume/superplasticizer regimes, respectively. Classical and HC3-robust ANOVA confirmed statistically significant effects of mix regime, curing age, and aggregate size on compressive strength, while residual diagnostics indicated acceptable normality but some heteroscedasticity, motivating the robust inference check. To address reproducibility and AI scope, the revised study also includes a case-level supplementary dataset and an AI/ML prediction module. Under GroupKFold validation by design cell, Gradient Boosting achieved the strongest predictive performance, MAE 3.46 MPa, RMSE 4.92 MPa, R2 0.885. Variable Importance Permutation indicated the important parameters as silica fume ratio, curing period, superplasticizer ratio, and aggregate size. The highest measured strength was 83.48 MPa at 28 days for the 12.5 mm aggregate with the 10% silica fume and 0.9% superplasticizer regime. Because silica fume and superplasticizer were co-varied in the experimental design, Mix2-Mix4 are interpreted as combined admixture regimes rather than isolated single-factor effects.

1. Introduction

1.1. Basic Review

High-Performance Concrete (HPC) has become a central material in contemporary structural engineering because current design priorities extend beyond compressive strength alone to include durability, permeability control, workability retention, and long-term service performance under demanding environmental and loading conditions. Recent review evidence has shown that modern HPC mixture design is therefore governed by a coupled optimization problem rather than by a single target property [1]. Alongside this broader materials perspective, recent review work on soft-computing applications has shown that predictive modeling is now increasingly expected to complement experimental characterization in HPC and Ultra High-Performance Concrete (UHPC) research. Explainable Machine Learning (ML) workflows, transparent

* Corresponding Author: Warqaa Muhammad Bahaaddin
Email: Warqaa.bahaadin@gmail.com

https://doi.org/10.22034/aeis.2026.572527.1472

validation, and variable-importance analysis are being used not only for prediction accuracy but also for design-oriented interpretation [2]. Recent comprehensive reviews of Ultra High Performance (UHP) and high-performance cementitious systems have further emphasized that particle packing, matrix densification, and aggregate configuration remain decisive design variables even when binder technology is advanced. In dense matrices, the presence, absence, size, and volume fraction of coarse aggregate affect homogeneity, microcracking tendency, and the balance between strength gain and paste demand [3]. At the microstructural level, recent review evidence on Interfacial Transition Zone (ITZ) has clarified that aggregate-related behavior cannot be explained from bulk proportions alone. ITZ thickness, porosity, continuity, and bond quality govern how loads are transferred through the matrix, meaning that aggregate-size effects must be interpreted together with interface development and microstructural refinement [4]. Recent work on cement-free and low-cement high-performance systems has also shown that performance optimization is increasingly assessed in relation to sustainability and matrix quality at the same time. This illustrates that mixture properties like aggregate size, binder modification, and workability control ought to be considered collectively as opposed to being treated separately during designing [5]. Experimental results on aggregate volume fraction and aggregate size in UHPC have shown that any change in aggregate packing will affect both the fresh and hardened concrete properties significantly. Even in very dense matrices, aggregate-size-related variation remains important because it affects internal structure, particle interaction, and stress-transfer mechanisms [6]. The role of supplementary cementitious materials has also been revisited in recent review studies. These works show that materials such as metakaolin and silica-based modifiers influence performance not only through chemical reaction but also through filler effects, packing refinement, and the modification of pore and interface structures in dense cementitious systems [7]. Recent reviews of lightweight and specialized UHPC systems further indicate that even when mixture objectives change, the governing design questions remain closely linked to particle arrangement, interface quality, and the trade-off between matrix density and practical workability. This supports the continued relevance of aggregate-size investigation in advanced concrete systems [8]. Recent review evidence on based on ML prediction of concrete mechanical performance has also highlighted the growing importance of reproducible workflows, explicit validation strategies, and interpretable model behavior. In this context, case-level data quality and transparent predictor-response relationships are now considered essential parts of credible AI-assisted materials research [9]. Recent studies on sustainable HPC have likewise emphasized that aggregate characteristics, binder modification, and environmental design goals should be studied together rather than separately. This strengthens the need for integrated experimental and data-driven frameworks in which practical mixture variables are linked directly to measured structural performance indicators [10].

1.2. Literature Review

Recent studies have shown that the influence of coarse aggregate size in HPC should not be interpreted only as a geometric variable. Wang et al. [11] reported that aggregate size affects ITZ thickness, nano-scale mechanical properties, and concrete strength, especially when coupled with water-to-cement ratio and interfacial development. This indicates that aggregate-size effects should be understood from a microstructural perspective rather than from size alone. A second body of literature has focused on matrix refinement through silica fume and related supplementary materials. Ji et al. [12] showed that both the physical filler effect and the chemical reactivity of silica fume contribute to UHPC matrix refinement. Lekhya and Kumar [13] similarly reported that combining ultrafine fly ash with silica fume can improve HPC performance through enhanced packing and densification. Amin et al. [14] further demonstrated that silica-fume-based modification has a strong influence on the mechanical, durability, and microstructural behavior of UHP mixtures. Wang et al. [15] confirmed that silica fume can improve ITZ even when unconventional aggregates are used, while Smarzewski [16] showed that condensed silica fume significantly improves both the mechanical and microstructural properties of HPC. In addition, Šoukal et al. [17] reported that both silica-fume content and superplasticizer type affect slump flow, density, strength, and microstructure in dense cementitious systems. Taken together, these studies show that admixture-driven matrix refinement has a direct influence on how aggregate size contributes to compressive strength.

Another important line of work has addressed the role of coarse aggregate in highly refined cementitious systems. Zhong et al. [18] found that coarse aggregate in UHPC affects workability, compressive strength, and flexural behavior in a coupled manner. Shen et al. [19] showed that coarse aggregate characteristics and fiber dispersion influence flexural performance beyond simple size effects. Tajasosi et al. [20] also highlighted the importance of balancing workability, packing, and strength when natural coarse aggregates are used in UHPC. These findings support the view that nominal maximum aggregate size remains an active and practically relevant design variable in advanced concrete systems. A more recent research direction concerns AI and based on ML prediction of HPC compressive strength. Zhang et al. [21] proposed an interpretable ML framework for HPC strength prediction and showed that prediction accuracy can be combined with variable-level interpretation. Muhammad et al. [22] compared ensemble models and neural networks for HPC compressive-strength prediction and reported reliable predictive performance under explicit validation settings. Tipu et al. [23] developed a hybrid stacked ML model for HPC strength prediction, and later improved predictive performance for both workability and compressive strength through an extended dataset and improved model configuration [24]. Al Yamani et al. [25] emphasized that variable sensitivity itself can provide useful engineering insight, while Islam et al. [26] showed that deep learning methods can provide practical nonlinear prediction capability. Lee et al. [27] demonstrated that super-learner algorithms improve predictive stability, and Kang et al. [28] confirmed that regression and ensemble methods can provide useful strength estimates from mixture variables. Kashem et al. [29] further showed that interpretable approaches such as SHapley Additive exPlanations (SHAP) and partial-dependence analysis can reveal meaningful nonlinear predictor effects. Wang et al. [30] likewise confirmed that recent ensemble and hybrid deep learning models can improve HPC strength prediction while maintaining practical design value.

1.3. Research Gap

Although the recent literature shows strong progress in three major directions, namely aggregate-size and ITZ mechanics, admixture-driven matrix refinement, and AI-based strength prediction, these directions are still commonly treated in isolation. Materials-focused studies often stop at experimental or inferential interpretation of aggregate size and admixture effects, whereas AI-oriented studies frequently rely on large compiled datasets that are detached from a single controlled laboratory program. As a result, there remains a lack of studies that combine controlled specimen-level experimentation, aggregate-size variation, combined silica-fume and superplasticizer regimes, robust statistical inference, and interpretable AI-based prediction within one unified and reproducible HPC framework. The specific gap addressed in the present study is therefore the absence of a specimen-level framework that links coarse aggregate size, combined silica-fume and superplasticizer regimes, and curing age to compressive strength through both robust statistical analysis and interpretable ML prediction. By integrating these elements within a single controlled dataset, the present study aims to provide a more connected understanding of how experimental factors, inferential evidence, and predictive modeling can jointly support HPC mixture evaluation.

2. Materials

2.1. Concrete

The Portland cement was used as the main binder material. The hydraulic binder undergoes hydration and gains strength even when mixed with water. The Portland cement is the adhesive that binds the aggregates and any other additional materials together to form a solid mass. The material mixes with water to produce a dense material that makes the concrete gain strength and durability, as per the requirements set out in ASTM C150 [31].

Concrete aggregates refer to the coarse materials such as sand, gravel or crushed stone that make up concrete. These aggregates are graded according to their sizes after sieving; those that pass through 4.75 mm are referred to as fine aggregate, and those retained on the sieve are called coarse aggregates. Proper grading and particle size distribution is important because it ensures packing efficiency in order to allow the smaller particles to fill in the spaces created by the larger aggregates in order to minimize the amount of cement paste required, which leads to the production of stronger and more durable concrete. Aggregate constitutes approximately 75% or more of concrete and therefore determines many of its properties. Aggregates that make high-quality concrete should be strong and clean, and free of impurities such as silts, oils and other organic materials since these affect the hydration of the cement and reduce the bond strength between the aggregates and cement paste. Maximum coarse aggregate size also dictates the maximum achievable strength; concretes with compressive strengths of up to 70 MPa are made using aggregates ranging from 20 mm to 28 mm, but those with strengths of approximately 100 MPa are produced using aggregates sized 10 mm to 20 mm. Very strong concretes can be made using aggregates ranging from 10 mm to 14 mm. For proper concrete production, clean water was used in this experiment. The amount of water used has a large effect on concrete strength and durability [32]. The lower the amount of water, the denser and stronger the concrete. Excess water results in segregation, bleeding and shrinkage and reduces concrete quality. The raw materials used for making concrete are shown in Fig. 1 below.

The physical characteristics of the materials used in the above experiment have been determined through the use of standardized methods in the laboratory. These values have been presented in Table 1 below. From the values shown, the cement sample is 21.6% (cm^2/g) in fineness and 3.20 in specific gravity, which show that it is good quality and suitable for concrete preparation. The sand (fine aggregate) is 2.76 in fineness modulus, 2.60 in specific gravity, 3.9% in moisture content, and 2.04% in water absorption, making it good quality and suitable for concrete formation. The gravel (coarse aggregate) showed 7.20 in fineness modulus, 2.67 in specific gravity, 1.0% in moisture content, and 0.5% in water absorption.

2.2. Admixtures Used in HPC

In this study, micro-silica (silica fume) and a melamine-based superplasticizer are used for enhancing the properties of high-performance concrete. Silica fume, illustrated in Fig. 2 (a), is a very fine, amorphous, glass-like powder made up of spherical particles having an average size of about 0.15 microns, which is roughly two orders of magnitude less than the size of cement particles. Specific Gravity of Silica Fume: The specific gravity of silica fume lies between 2.2 to 2.3, while the density of silica fume is 130 kg/m3 to 600 kg/m3. Its density depends on the compaction and varies between 130 and 600 kilograms per cubic meter. The other substance used for improving concrete properties is sulfonated melamine formaldehyde (SMF) superplasticizer, illustrated in Fig. 2(b).

Fig. 1. Raw Materials Used for Concrete Mix.

Fig. 2. (a) Silica Fume (b) Superplasticizer used in the study.

Table 1. Physical Properties of Materials Used in the Study.

Materials	Tests	Value
Cement	F.M	21.6%
	Specific gravity	3.2
Sand	F.M	2.76
	Specific gravity	2.6
	Moisture content	3.9%
	Absorption	2.04 %
Gravel	F.M	7.2
	Specific gravity	2.67
	Moisture content	1%
	Absorption	0.5%

2.3. Mix Design

The concrete mixtures used in this study were designed to produce HPC under a fixed matrix composition so that the effects of aggregate size, curing age, and admixture regime could be examined within one controlled experimental framework. The water to cement ratio was 0.32 while the proportion of cement, fine aggregates, and coarse aggregates on a mass basis was 1:1.16:2.02. The selected w/c ratio should be interpreted as a practical low-w/c design value for the present HPC program rather than as a universally optimal value. It was retained to promote a dense high-strength matrix while preserving workable casting conditions, which is consistent with conventional HPC proportioning practice. In the present laboratory program, this choice was supported by practical workability screening, and the control mixture exhibited slump values in the range of 28-44 mm across the tested aggregate sizes, indicating that the selected w/c level remained within a castable range for the intended specimens. The adopted base proportion was retained as a fixed reference mixture so that changes in compressive strength could be associated more clearly with aggregate size, curing age, and the combined silica-fume/superplasticizer regimes rather than with multiple simultaneous shifts in matrix composition. Accordingly, the admixture-bearing mixtures were organized as progressive combined regimes rather than as a full optimization study. Silica fume content in Mix2 was 5%, whereas its superplasticizer content was 0.5%. Silica fume content in Mix3 was 7%, while its superplasticizer content was 0.7%. Silica fume content in Mix4 was 10%, while its superplasticizer content was 0.9%. These regimes were selected to represent practically increasing levels of matrix modification and water-reduction support, thereby allowing the strength response to be examined under low, moderate, and higher combined admixture intensities within the same experimental program. The nominal maximum aggregate sizes of 38, 25, 19, 12.5, and 9.5 mm were selected to span a practically meaningful range from relatively coarse to relatively fine aggregate conditions in HPC. This range was used to capture how changes in particle size may alter packing density, paste demand, interfacial development, and load-transfer efficiency within a dense cementitious matrix. The inclusion of both larger sizes such as 38 and 25 mm and smaller sizes such as 12.5 and 9.5 mm therefore allowed the study to compare mixtures that differ not only in grading condition but also in their likely balance between internal void structure and paste-aggregate bonding quality. For each mix, 30 cubes were cast, corresponding to six cubes per aggregate size. The six cubes were evenly divided between 7-day and 28-day testing, yielding three specimens per age within each design cell and a total of 120 specimen-level compressive-strength observations.

These case-level observations are supplied as supplementary CSV/XLSX material for reproducibility.

Table 2. Mix Composition Variables and Cube Testing Details.

Mix ID	S.F (%)	S.P (%)	Aggregate Size (mm)	Cubes per Size	Total Cubes
Mix 1	0	0	38, 25, 19, 12.5, 9.5	6	30
Mix 2	5	0.5	38, 25, 19, 12.5, 9.5	6	30
Mix 3	7	0.7	38, 25, 19, 12.5, 9.5	6	30
Mix 4	10	0.9	38, 25, 19, 12.5, 9.5	6	30

3. Method

3.1. Experimental Test

The concrete specimens were cast as cubes with dimensions of 150 × 150 × 150 mm, with 30 cubes prepared for each mix. After casting, the specimens were left to set and then cured under standard conditions to ensure proper hydration. Compressive strength was tested at curing ages of 7 and 28 days. All the specimens during the curing period are shown in Fig. 3.

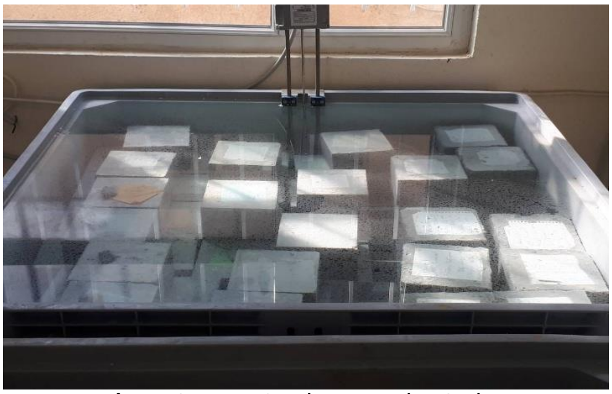


Fig. 3. Concrete Specimens During Curing.

The compressive strength test for concrete samples was done using the compression testing machine as shown in Fig. 4. Cube samples of dimensions 150 x 150 x 150 mm were tested at curing age of 7 days and 28 days. Before performing the test, the faces of the samples were cleaned and made even. The sample was placed in the center of the load frame and loaded at a constant rate until failure. The peak value of the load attained at failure was taken into account and the corresponding value of compressive strength was obtained by dividing the load value by the cross-sectional area of the sample.

Fig. 4. Compressive Strength Testing Machine.

The compressive strength of each specimen was computed according to Eq. (1), consistent with standard compressive-strength testing practice for hardened concrete [33].

$$f_c = \frac{P}{A} \qquad (1)$$

Where f_c is the compressive strength, P is the maximum load at failure, and A is the loaded cross-sectional area of the specimen. In the present study, cube specimens of size 150 × 150 × 150 mm were tested, and the measured strength values reported in the Results section were obtained from this relationship.

3.2. Statistical and AI/ML Analysis

The compressive-strength dataset was analyzed at the specimen level rather than only through average values. Each record contained Mix_{ID}, aggregate size, silica fume percentage, superplasticizer percentage, curing age, and measured compressive strength. The revised statistical workflow combined descriptive summaries, classical factorial ANOVA, HC3-robust inference, residual diagnostics, and an AI/ML prediction stage. This structure was adopted to strengthen reproducibility, provide a transparent validation path, and address the intelligent-systems scope requested by the editor. The ANOVA model was fitted using Mix_{ID}, curing age, and aggregate size, together with their two-way interactions. Mix_{ID} was used intentionally because silica fume and superplasticizer were changed together across Mix2-Mix4; accordingly, the ANOVA terms for Mix_{ID} should be interpreted as combined admixture-regime effects rather than isolated main effects of each admixture. Residual normality was checked using Shapiro-Wilk and Jarque-Bera tests, whereas variance behavior was examined using Levene's test and the Breusch-Pagan test. Because mild heteroscedasticity was detected, HC3-robust inference was additionally reported alongside the classical ANOVA results.

3.3. Case-Level Dataset and AI/ML Prediction Workflow

A specimen-level prediction system was implemented to determine the mapping function between the study variables and compressive strength, thus serving as a complement to the experiment and inference parts. The predictor variables include $AggregateSize_{mm}$, $SilicaFume_{pct}$, $Superplasticizer_{pct}$, and Age_{days}, with $CompressiveStrength_{MPa}$ being the target variable. The current study improved upon its predecessor through providing the complete case-level data file in CSV/XLSX format as additional files. Each row represents a test cube and includes Mix_{ID}, $AggregateSize_{mm}$, $SilicaFume_{pct}$, $Superplasticizer\ pct$, Age_{days}, and $CompressiveStrength_{MPa}$ as variables. Three regressors were compared, namely Ridge Regression, Random Forest, and Gradient Boosting. For the first method of validation, GroupKFold by design cell was selected, where the design cell is specified by a unique combination of Mix_{ID}, $AggregateSize_{mm}$, and Age_{days}. In this way, there will be no replicate specimen cases of the same experimental conditions in either train or test folds. In addition, another RepeatedKFold analysis was performed to verify the robustness of model performance. The evaluation metrics include MAE, RMSE, and R². To improve interpretability, permutation importance and partial-dependence analysis were applied to the best-performing model. The regression metrics used in this study are given in Eqs. (2)-(4) [21]:

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

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \quad (4)$$

Where y_i is the observed compressive strength of the i-th specimen; $\hat{y}_i$ is its prediction by the model; $\bar{y}$ is the mean observed compressive strength; n is the total number of evaluated specimens. A lower value of MAE and $RMSE$ indicates a smaller prediction error, whereas R^2 approaching 1 means a higher explanatory power/prediction consistency of the model. In addition, the permutation importance and partial dependence techniques were utilized for interpretation of input variables' effects on the predictions made by the best-performing model.

4. Results and Discussion

4.1. Experimental Compressive-strength Response

The revised case-level dataset contained 120 specimen-level observations arranged in a balanced experimental structure of 40 design cells, with three replicate cubes for each combination of mix regime, aggregate size, and curing age. Across all mixes, compressive strength increased from 7 to 28 days, confirming the expected contribution of continued hydration to strength development. The control mixture without admixtures provided the baseline mechanical response of the studied system. At 7 days, the control strength ranged from 32.95 MPa to 37.16 MPa, while the 28-day values ranged from 40.74 MPa to 44.28 MPa, depending on aggregate size. Among the control conditions, the 19 mm aggregate produced the highest average strength at both ages, indicating that under the unmodified matrix condition, this size provided the most favorable balance between internal packing and load transfer. The addition of the combined silica-fume and superplasticizer regimes substantially increased compressive strength relative to the control condition. For Mix2, which contained 5% silica fume and 0.5% superplasticizer, the 28-day mean strength increased to 61.70-73.96 MPa. For Mix3, containing 7% silica fume and 0.7% superplasticizer, the corresponding 28-day range increased further to 64.07-79.43 MPa. The highest overall response was obtained for Mix4, containing 10% silica fume and 0.9% superplasticizer, where the 28-day mean strength ranged from 68.52 MPa to 83.48 MPa. These results confirm that progressive matrix refinement through the combined admixture regimes produced clear strength gains at both early and later curing ages.

The aggregate-size effect also changed with admixture level. In the control mix, 19 mm gave the highest mean strength, whereas in the admixture-bearing mixes the best performance shifted toward the finer sizes. In Mix2, the highest 28-day mean strength was obtained at 9.5 mm, reaching 73.96 MPa. In Mix3 and Mix4, the optimum shifted to 12.5 mm, where the 28-day mean strengths reached 79.43 MPa and 83.48 MPa, respectively. By contrast, the 38 mm aggregate consistently showed the lowest or among the lowest strengths in the modified mixtures, especially at 28 days. This pattern indicates that as the binder system becomes denser and more refined, relatively smaller nominal maximum aggregate sizes become more favorable for compressive-strength development. Compared with the control condition at the same aggregate size and curing age, the strength gains were substantial. For example, at 28 days, Mix4 exceeded the control by about 68.2% at 38 mm, 83.4% at 25 mm, 76.9% at 19 mm, 99.2% at 12.5 mm, and 98.0% at 9.5 mm. The largest benefits therefore occurred in the finer aggregate systems, especially at 12.5 mm and 9.5 mm. From a materials perspective, these trends are consistent with denser particle packing, lower internal void content, and improved paste-aggregate load transfer in the finer aggregate mixtures. In high-performance concrete, where the matrix is already refined by silica fume and aided by superplasticizer, smaller aggregates can help produce a more homogeneous internal structure and a more favorable ITZ. However, aggregate size alone should not be interpreted as the only governing variable, since local surface texture, aggregate shape, and unmeasured interface characteristics may also contribute to the observed differences.

Table 3 summarizes the mean compressive strength values of all four mix regimes at both curing ages for each nominal maximum aggregate size. The table shows that the control mixture remained within a relatively narrow strength range, whereas the admixture-bearing mixes produced much larger gains, especially at 28 days. It also shows clearly that the aggregate-size effect depended on the mix regime, with 19 mm performing best in the control condition and 12.5 mm or 9.5 mm becoming more favorable in the modified mixtures.

Table 3. Mean compressive strength of the four mix regimes at 7 and 28 days for each nominal maximum aggregate size.

Aggregate size (mm)	Mix1 7d	Mix1 28d	Mix2 7d	Mix2 28d	Mix3 7d	Mix3 28d	Mix4 7d	Mix4 28d
38	32.953	40.737	53.118	61.698	50.273	64.071	54.612	68.524
25	33.694	41.345	57.060	62.545	59.323	67.194	64.832	75.811
19	37.160	44.278	58.769	64.071	59.923	71.261	65.124	78.348
12.5	33.441	41.907	58.383	69.834	66.709	79.426	68.536	83.482
9.5	33.023	41.046	51.215	73.959	58.895	70.198	65.880	81.277

As seen in Table 3, the transition from the control condition to Mix2, Mix3, and Mix4 was associated with a systematic increase in compressive strength across almost all aggregate sizes and both curing ages. The strongest responses were consistently observed in the finer-aggregate systems once silica fume and superplasticizer were introduced, confirming that the influence of aggregate size cannot be interpreted independently of matrix refinement.

Fig. 5 shows the 28-day mean compressive strength comparison across the four mix regimes as a function of aggregate size. The figure makes the upward shift in strength with increasing admixture intensity immediately visible and also highlights the change in the optimal aggregate size. While the control mixture reaches its highest value at 19 mm, the admixture-bearing mixes show superior performance at 12.5 mm and 9.5 mm, with Mix4 producing the highest overall response. This graphical comparison supports the interpretation that matrix refinement amplifies the benefit of relatively smaller nominal maximum aggregate sizes.

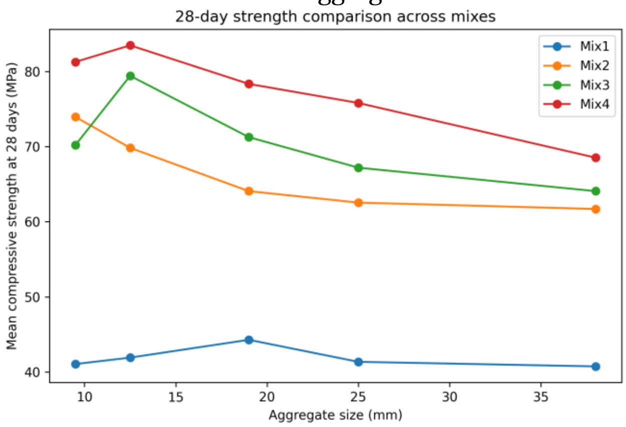


Fig. 5. Comparison of 28-day mean compressive strength versus aggregate size for the four mix regimes.

4.2. Statistical Assessment of Factor Effects

To verify that the observed strength differences were not attributable to random variation alone, the case-level dataset was analyzed using factorial ANOVA. Residual-diagnostic checks showed that the normality assumption was well satisfied. The Shapiro-Wilk test returned a p-value of 0.9781, and the Jarque-Bera test returned a p-value of 0.9510, indicating no evidence of problematic non-normality in the model residuals. At the same time, the Breusch-Pagan test indicated heteroscedasticity at the model level with a p-value equal to 0.0057, and Levene's test by Mix_{ID} also showed a significant result with a p-value equal to 0.0264. In contrast, Levene's tests by curing age and aggregate size were not significant. Taken together, these results suggest acceptable residual normality together with mild heteroscedasticity, which justifies reporting HC3-robust ANOVA in addition to the classical formulation. Under the HC3-robust ANOVA, all three main factors remained statistically significant. Mix regime was the dominant source of explained variation, with F1 equal to 1124.61, and the p-value is less than 0.001. Curing age also had a very strong effect, with F1 equal to 261.62, and the p-value is less than 0.001. Aggregate size retained a statistically significant but comparatively smaller effect, with F1 equal to 11.54 and the p-value is less than 0.001. These findings confirm that the compressive-strength response was governed primarily by the combined admixture regime and curing age, while aggregate size remained an important secondary design variable.

The interaction terms were also informative. The interaction between Mix_{ID} and Age_{days} was significant, with the p-value equal to 0.0047, indicating that the magnitude of strength development between 7 and 28 days depended partly on the admixture regime. The interaction between Mix_{ID} and $AggregateSize_{mm}$ was likewise significant, with the p-value less than 0.001, showing that the most favorable aggregate size was not identical across all mix regimes. In addition, the Age_{days} by $AggregateSize_{mm}$ interaction was significant at p-value equal to 0.0339, suggesting that the aggregate-size effect does not remain perfectly constant over curing time. These interaction results are consistent with the experimental trends, where the control mixture favored 19 mm, whereas the admixture-bearing mixes favored 12.5 mm and 9.5 mm. From a practical standpoint, the statistical results reinforce the experimental interpretation rather than replacing it. The significance of aggregate size should not be read only as a numerical p-value result; rather, it reflects the fact that aggregate size changes the internal mechanical organization of the mixture through packing density, paste requirement, and the quality of the paste–aggregate interface. Likewise, the very large effect of the mix regime is consistent with the strong matrix refinement introduced by the combined silica-fume and superplasticizer dosages.

The residual-diagnostic and variance-behavior checks used to assess the suitability of the ANOVA framework are summarized in Table 4. These tests were included to verify whether the main assumptions of the inferential analysis were sufficiently satisfied before interpreting factor significance.

Table 4. Residual-diagnostic and variance-behavior checks used to assess the suitability of the ANOVA framework.

Test	Statistic	p-value
Shapiro-Wilk residual normality	0.9958	0.9781
Jarque-Bera residual normality	0.1004	0.9510
Breusch-Pagan heteroscedasticity	49.1754	0.0057
Levene by Mix_{ID}	3.1883	0.0264
Levene by Age_{days}	1.2963	0.2572
Levene by $AggregateSize_{mm}$	0.8093	0.5217

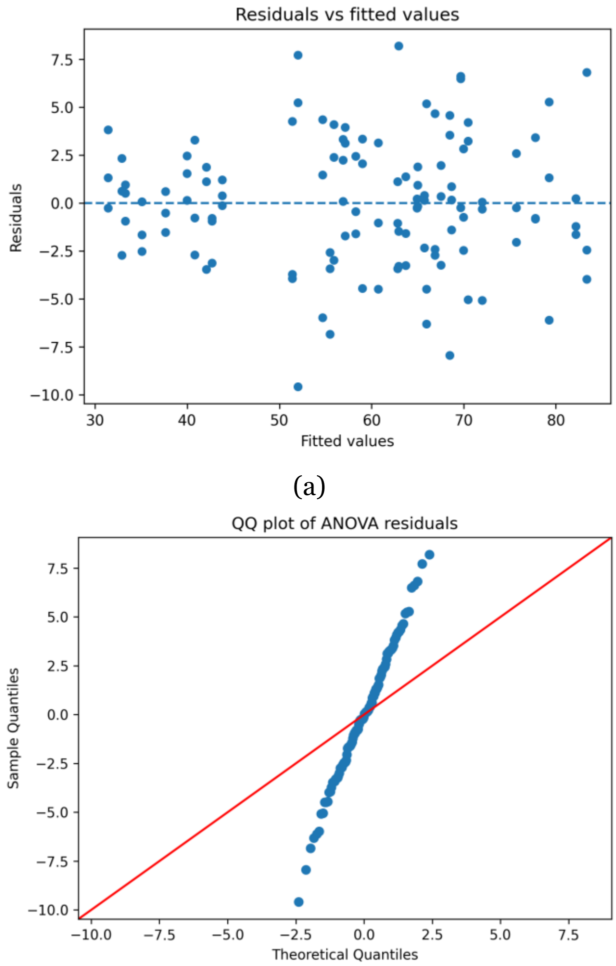
Table 5. Residual-diagnostic and variance-behavior checks used to assess the suitability of the ANOVA framework.

Term	df	F
Mix_{ID}	3	1124.61
Age_{days}	1	261.62
$AggregateSize_{mm}$	4	11.54
$Mix_{ID} \times Age_{days}$	3	4.62
$Mix_{ID} \times AggregateSize_{mm}$	12	3.54
Term	df	F

As reported in Table 4, both the Shapiro-Wilk and Jarque-Bera tests support acceptable residual normality, whereas the Breusch-Pagan result and the Levene result by Mix_{ID} indicate some heteroscedasticity. This combination of findings supports the use of HC3-robust inference as a precaution against variance non-uniformity. Table 5 reports the HC3-robust ANOVA results for the main factors and their two-way interactions. This table provides the formal statistical evidence supporting the experimental trends described in Section 4.1.

Table 5 confirms that mix regime, curing age, and aggregate size all exert statistically significant effects on compressive strength. It also shows that the interaction structure is not negligible, particularly for $Mix_{ID} \times Age_{days}$ and $Mix_{ID} \times AggregateSize_{mm}$, which means that the strength benefit of a given aggregate size depends partly on the admixture condition and curing stage.

Fig. 6 presents the residual diagnostics of the ANOVA model. In Fig. 6(a), the residuals-versus-fitted plot shows no severe systematic distortion, although some spread variation is visible, which is consistent with mild heteroscedasticity rather than severe model failure. In Fig. 6 (b), the Q-Q plot shows that the residuals follow the reference line closely, supporting the conclusion that the normality assumption is acceptable for practical inference. Together, these two panels justify the decision to retain the ANOVA framework while also reporting HC3-robust results.



(a)

(b)

Fig. 6. Residual diagnostics for the ANOVA framework: (a) residuals versus fitted values and (b) normal Q-Q plot of model residuals.

4.3. Data-driven predictive modeling and model interpretation

As an additional component to supplement the experimental and inferential analyses, a predictive modeling framework was developed using four variables, $AggregateSize_{mm}$, $SilicaFume_{pct}$, $Superplasticizer_{pct}$, and Age_{days} as inputs, with the response variable being $CompressiveStrength_{MPa}$. In the case of model validation, GroupKFold by design cell was used to guarantee that no replicate specimens from identical experimental conditions will ever be in the training set and in the testing set simultaneously. In terms of the prediction performance, Gradient Boosting model showed the best results. For out-of-fold predictions based on the main GroupKFold approach, it resulted in 3.46 MPa Mean Absolute Error (MAE), 4.92 MPa Root Mean Squared Error (RMSE), and 0.885 R^2. The averages across all folds indicated similarly good results: mean MAE was 3.455, mean RMSE was 4.821, and mean R^2 was 0.880. Random Forest took the second place, while Ridge regression had the worst performance of all considered models. Such results suggest that there is enough nonlinearity present in case-level data to warrant a useful data-driven estimation of the compressive strength of concrete. In addition, the relationship between observed and predicted compressive strength values was almost perfectly aligned with the one-to-one line, indicating that selected predictors were able to account for a considerable amount of variability. The model performance is remarkable considering the relatively small number of observations in the dataset and the designed experimental nature. Instead of replacing the experimental analysis, this additional AI-based module enables developing a reproducible estimation stage that can support screening studies or mixture-design workflows in the future. In order to further interpret the model results, permutation feature importance and partial dependence profiles were used. Permutation feature importance indicated that the largest impact had the presence of silica fume (i.e., higher silica-fume content increased the compressive strength of concrete), followed by curing age, superplasticizer dosage, and aggregate size. It is physically consistent since silica fume and curing time impact matrix densification and hydration reactions progress, while the superplasticizer makes mixing process easier. Although the aggregate size also had a significant impact on the compressive strength, the contribution of this predictor was not as large as those of admixture and age. At the same time, these importance rankings should be interpreted with appropriate caution. Because silica fume and superplasticizer were co-varied by design across Mix2-Mix4, their importance values are predictive rather than causal. The model can identify which variables are most useful for estimating strength in the present dataset, but it cannot isolate the fully independent physical effect of silica fume relative to superplasticizer without additional single-admixture experimental arms. The comparative predictive performance of the evaluated regression models under the primary GroupKFold validation scheme is summarized in Table 6. This comparison was included to determine whether the case-level dataset contains enough structure to support accurate data-driven estimation of compressive strength.

As shown in Table 6, Gradient Boosting produced the best overall performance, with the lowest mean MAE and RMSE and the highest mean R^2 among the evaluated models. Random Forest ranked second, whereas Ridge provided the weakest predictive accuracy. These results indicate that nonlinear ensemble learners are better suited than the linear baseline for capturing the specimen-level structure of the present dataset.

Table 6. Predictive performance of the evaluated models under GroupKFold validation.

Model	Mean MAE (MPa)	Mean RMSE (MPa)	Mean R^2
GradientBoosting	3.455	4.821	0.880
RandomForest	3.920	5.177	0.859
Ridge	4.129	5.249	0.865

Fig. 7 shows the observed-versus-predicted compressive-strength relationship for the best-performing model under GroupKFold validation. The points cluster reasonably closely around the one-to-one line, indicating that the model reproduces the measured strength pattern with good overall agreement. The absence of severe systematic deviation also supports the suitability of the selected input variables for predictive modeling.

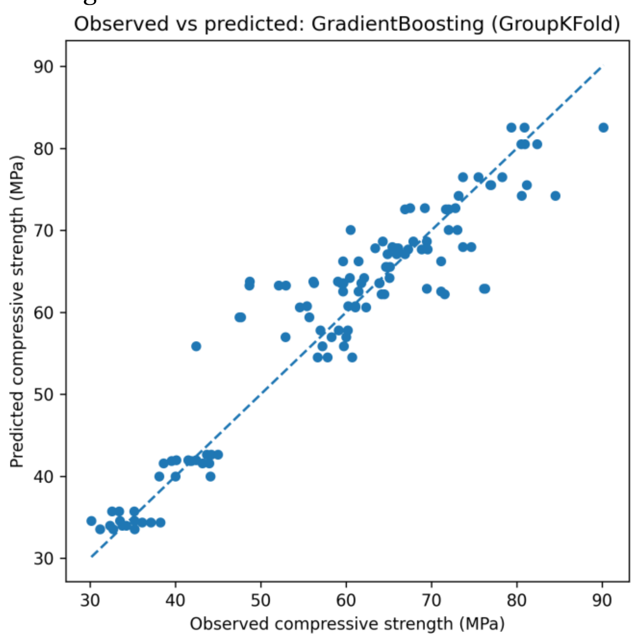


Fig. 7. Observed versus predicted compressive strength for Gradient Boosting under GroupKFold validation.

Fig. 8 presents the permutation-importance ranking of the predictor variables for the best-performing model. The figure shows that silica-fume content contributed the strongest predictive signal, followed by curing age, superplasticizer dosage, and aggregate size. This ranking is consistent with the experimental trends, although the importance values should be interpreted as predictive rather than causal because silica fume and superplasticizer were co-varied by design.

Fig. 9 displays the partial dependence plots corresponding to the four predictor variables. These plots show the trend that the output follows in relation to each input variable, namely, aggregate size, silica-fume percentage, superplasticizer dosage, and curing time.

They are consistent and realistic, which further validates the integrity of the AI/ML approach and adds credibility to the results produced by the model.

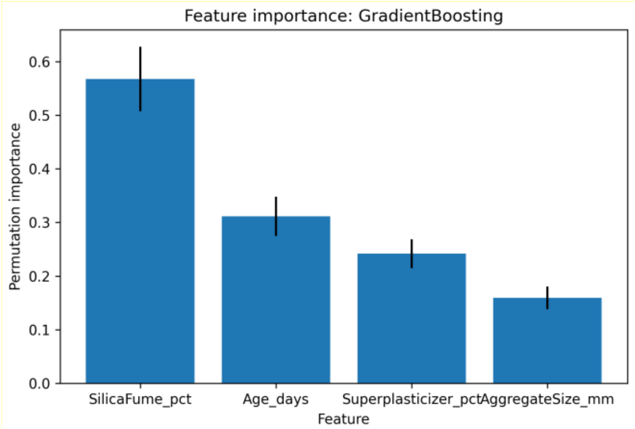


Fig. 8. Permutation importance ranking of the input variables for the best-performing predictive model.

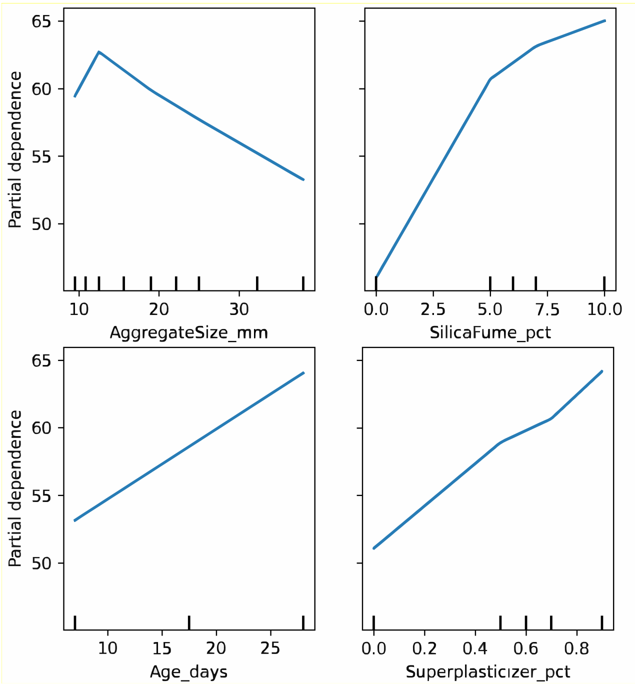


Fig. 9. Partial-dependence plots of the main predictors for compressive strength.

4.4. Integrated Discussion and Practical Interpretation

When the experimental, statistical, and predictive results are interpreted together, a consistent picture emerges. Under the control condition, the 19 mm aggregate provided the strongest compressive response, suggesting that in the absence of matrix-modifying admixtures, this size offered the best compromise between packing quality and paste demand. Once the matrix was refined through silica fume and superplasticizer, however, the optimum shifted toward 12.5 mm and 9.5 mm. This indicates that the preferred aggregate size in HPC is not fixed in isolation, but depends on the quality and refinement level of the surrounding cementitious matrix. The results, therefore, support a practical design implication: when the objective is to maximize compressive strength within a dense admixture-bearing HPC system, 12.5 mm and 9.5 mm nominal maximum aggregate sizes are promising candidates within the tested range. However, this implication should remain bounded by the scope of the present experimental program. The study was limited to compressive strength at 7 and 28 days, and it did not include silica-fume-only mixtures, superplasticizer-only mixtures, direct ITZ measurements, durability indicators, or long-term environmental exposure testing. Accordingly, the present findings should be interpreted as strong comparative evidence within a controlled laboratory framework rather than as a complete formulation rule for all HPC applications.

5. Conclusion

This study investigated the effects of aggregate size, curing age, and combined silica-fume/superplasticizer regimes on the compressive strength of high-performance concrete and extended the analysis through a specimen-level AI/ML framework. Experimentally, the highest strength was 83.48 MPa at 28 days for the 12.5 mm aggregate under the 10% silica fume and 0.9% superplasticizer regime, while the admixture-bearing mixes consistently outperformed the control mixture. Statistically, both classical and HC3-robust ANOVA confirmed that mix regime, curing age, and aggregate size all had significant effects on compressive strength. From the data-driven perspective, the AI/ML module provided a reproducible predictive layer, with Gradient Boosting yielding the strongest performance under GroupKFold validation. Permutation importance further showed that silica fume content, curing age, superplasticizer dosage, and aggregate size were the main predictive variables within the present dataset. From a practical standpoint, the results suggest that 12.5 mm and 9.5 mm nominal maximum aggregate sizes are promising choices when high compressive strength is prioritized in admixture-bearing HPC within the tested design envelope. However, Mix2-Mix4 should be interpreted as combined admixture regimes rather than isolated single-factor effects, and the present findings should not be extended to long-term durability or later-age performance without additional experiments. Future work should therefore include single-admixture mixtures, later-age testing, and durability-oriented measurements to establish a more complete performance framework.

Conflicts of Interest

The authors declare no conflict of interest.

Funding

No funding was received for this study.

Acknowledgments

Not applicable.

Availability of Data and Materials

The datasets used during the current study are publicly available.

Using Artificial Intelligence Chatbots

None.

References

[1] Chahar AS, Pal P. A review on various aspects of high performance concrete. Innovative Infrastructure Solutions 2023 8:6 2023;8:175-. doi:10.1007/s41062-023-01144-3.

[2] Kumar R, Rai B, Samui P. Prediction of mechanical properties of high-performance concrete and ultrahigh-performance concrete using soft computing techniques: A critical review. Structural Concrete 2025;26:1309–37. doi:10.1002/suco.202400188.

[3] El-Abbasy AA. Production, behaviour and mechanical properties of ultra-high-performance fiber concrete – A comprehensive review. Case Studies in Construction Materials 2022;17:e01637. doi:10.1016/j.cscm.2022.e01637.

[4] Chen Q, Zhang J, Wang Z, Zhao T, Wang Z. A review of the interfacial transition zones in concrete: Identification, physical characteristics, and mechanical properties. Eng Fract Mech 2024;300:109979. doi:10.1016/j.engfracmech.2024.109979.

[5] Mostafaei H, Bahmani H, Mostofinejad D, Wu C. A novel development of HPC without cement: Mechanical properties and sustainability evaluation. Journal of Building Engineering 2023;76:107262. doi:10.1016/j.jobe.2023.107262.

[6] Matiushin E, Sizyakov I, Shvetsova V, Soloviev V. The Properties and Behavior of Ultra-High-Performance Concrete: The Effects of Aggregate Volume Content and Particle Size. Buildings 2024, Vol 14, Page 2891 2024;14:2891. doi:10.3390/buildings14092891.

[7] Mohamed AM, Tayeh BA. Metakaolin in ultra-high-performance concrete: A critical review of its effectiveness as a green and sustainable admixture. Case Studies in Construction Materials 2024;21:e03967. doi:10.1016/j.cscm.2024.e03967.

[8] Ramezani A, Shafei B. Advances in lightweight ultra-high performance concrete: Mixture requirements, strength characteristics, and durability properties. Journal of Building Engineering 2025;112:113681. doi:10.1016/j.jobe.2025.113681.

[9] Xu Y, Zhuge W, Wang J, Yu X, Wu K. Review of the Mechanical Performance Prediction of Concrete Based on Artificial Neural Networks. Structural Durability & Health Monitoring (SDHM) 2025;19:1507. doi:10.32604/sdhm.2025.069021.

[10] Chahar AS, Pal P. Effects of sustainable materials on high-performance concrete. Discover Applied Sciences 2025 7:10 2025;7:1110-. doi:10.1007/s42452-025-07535-8.

[11] Wang Y, Zhang W, Wang J, Huang R, Lou G, Luo S. Effects of coarse aggregate size on thickness and micro-properties of ITZ and the mechanical properties of concrete. Cem Concr Compos 2024;154:105777. doi:10.1016/j.cemconcomp.2024.105777.

[12] Ji X, Han F, Pan T, Zhao W, Sha J, Liu J. Decoupling the physical and chemical effects of silica fume in ultra-high performance concrete (UHPC). Constr Build Mater 2024;444:137851. doi:10.1016/j.conbuildmat.2024.137851.

[13] Lekhya A, Kumar NS. A study on the effective utilization of ultrafine fly ash and silica fume content in high-performance concrete through an experimental approach. Heliyon 2024;10. doi:10.1016/j.heliyon.2024.e39678.

[14] Amin M, Zeyad AM, Tayeh BA, Saad Agwa I. Effect of ferrosilicon and silica fume on mechanical, durability, and microstructure characteristics of ultra high-performance concrete. Constr Build Mater 2022;320:126233. doi:10.1016/j.conbuildmat.2021.126233.

[15] Wang X, Li X, Zhong Y, Li H, Wang J. Properties and Microstructure of an Interfacial Transition Zone Enhanced by Silica Fume in Concrete Prepared with Coal Gangue as an Aggregate. ACS Omega 2023;9:1870–80. doi:10.1021/acsomega.3c08560.

[16] Smarzewski P. Mechanical and Microstructural Studies of High Performance Concrete with Condensed Silica Fume. Applied Sciences 2023, Vol 13, Page 2510 2023;13:2510. doi:10.3390/app13042510.

[17] Šoukal F, Bocian L, Novotný R, Dlabajová L, Šuleková N, Hajzler J, Koutný O, Drdlová M. The Effects of Silica Fume and Superplasticizer Type on the Properties and Microstructure of Reactive Powder Concrete. Materials 2023, Vol 16, Page 6670 2023;16:6670. doi:10.3390/ma16206670.

[18] Zhong R, Pan M, Wu H, Cheng Z, Liu J, Wang J, Yao Y, Ma H. Effect of coarse aggregate on the stability and mechanical performance of ultra-high performance concrete (UHPC). Compos B Eng 2025;297:112210. doi:10.1016/j.compositesb.2025.112210.

[19] Shen C, Wang H, Tang D, Lyu Z, Zhang H, Yu L, Yang C. Investigating the influence of coarse aggregate characteristics and fiber dispersion on the flexural performance of ultra-high performance concrete (UHPC). Journal of

Building Engineering 2024;97:110738. doi:10.1016/j.jobe.2024.110738.
[20] Tajasosi S, Barandoust J, Saradar A, Mohtasham Moein M, Rigby SE, Karakouzian M. Investigation of Mechanical and Fresh Properties of Ultra-High-Performance Concrete Incorporating Second-Generation Superplasticizers. Applied Sciences 2025, Vol 15, Page 5133 2025;15:5133. doi:10.3390/app15095133.
[21] Zhang Y, Ren W, Chen Y, Mi Y, Lei J, Sun L. Predicting the compressive strength of high-performance concrete using an interpretable machine learning model. Scientific Reports 2024 14:1 2024;14:28346-. doi:10.1038/s41598-024-79502-z.
[22] Muhammad UJ, Aminu II, Mahmoud IA, Aliyu UU, Usman AG, Jibril MM, Malami SI, Abba SI. An improved prediction of high-performance concrete compressive strength using ensemble models and neural networks. AI in Civil Engineering 2024 3:1 2024;3:21-. doi:10.1007/s43503-024-00040-8.
[23] Tipu RK, Suman, Batra V. Development of a hybrid stacked machine learning model for predicting compressive strength of high-performance concrete. Asian Journal of Civil Engineering 2023 24:8 2023;24:2985–3000. doi:10.1007/s42107-023-00689-z.
[24] Tipu RK, Suman, Batra V. Enhancing prediction accuracy of workability and compressive strength of high-performance concrete through extended dataset and improved machine learning models. Asian Journal of Civil Engineering 2023 25:1 2023;25:197–218. doi:10.1007/s42107-023-00768-1.
[25] Al Yamani WH, Ghunimat DM, Bisharah MM. Modeling and predicting the sensitivity of high-performance concrete compressive strength using machine learning methods. Asian Journal of Civil Engineering 2023 24:7 2023;24:1943–55. doi:10.1007/s42107-023-00614-4.
[26] Islam N, Kashem A, Das P, Ali MN, Paul S. Prediction of high-performance concrete compressive strength using deep learning techniques. Asian Journal of Civil Engineering 2023 25:1 2023;25:327–41. doi:10.1007/s42107-023-00778-z.
[27] Lee S, Nguyen NH, Karamanli A, Lee J, Vo TP. Super learner machine-learning algorithms for compressive strength prediction of high performance concrete. Structural Concrete 2023;24:2208–28. doi:10.1002/suco.202200424.
[28] Kang J, Gao J, Liu X. Machine learning for predicting the compressive strength of high-performance concrete materials. Multiscale and Multidisciplinary Modeling, Experiments and Design 2025 8:9 2025;8:403-. doi:10.1007/s41939-025-00981-2.
[29] Kashem A, Karim R, Malo SC, Das P, Datta SD, Alharthai M. Hybrid data-driven approaches to predicting the compressive strength of ultra-high-performance concrete using SHAP and PDP analyses. Case Studies in Construction Materials 2024;20:e02991. doi:10.1016/j.cscm.2024.e02991.
[30] Wang Z, Sun T, Sun Y, Liu N. Evaluating the strength properties of high-performance concrete in the form of ensemble and hybrid models using deep learning techniques. Scientific Reports 2025 15:1 2025;15:25453-. doi:10.1038/s41598-025-10860-y.
[31] ASTM C. 150 Standard Specification for Portland Cement, ASTM International. West Conshohocken, Pennsylvania, United States 2007.
[32] ASTM C. 1602/C 1602M-18 Standard Specification for mixing Water Used in the Production of hydraulic Cement Concrete. ASTM International, USA 2018.
[33] Standard B. Testing hardened concrete. Compressive Strength of Test Specimens, BS EN 2009:12390–3.