

UNIFYING CONVENTIONAL AND LIGHTWEIGHT CONCRETE: AN EXPLAINABLE AI WEB APPLICATION FOR COMPRESSIVE STRENGTH PREDICTION

Ahsan Ali Buriro¹, Ubaidullah Khan^{*2}, Irfanullah Khan³, Raheel Noor Chandio⁴,
Muhammad Aqib⁵, Shahbaz Shaikh⁶

^{1, *2,5,6}Department of Civil Engineering, Quaid-e-Awam University of Engineering, Science and Technology, Nawabshah, Sindh, Pakistan

³Higher Education Commission of Pakistan, Islamabad, Pakistan

⁴People's University of Medical & Health Sciences for Women, Nawabshah, Pakistan

^{*2}ubaidullahkhan@quest.edu.pk

DOI: <https://doi.org/10.5281/zenodo.22933138>

Keywords

concrete; foamed concrete; compressive strength; machine learning; explainable artificial intelligence

Article History

Received: 01 June 2026

Accepted: 23 August 2026

Published: 24 September 2026

Copyright @Author

Corresponding Author: *

Ubaidullah Khan

Abstract

Accurate prediction of concrete compressive strength is important for mixture development, quality control, and efficient selection of construction materials. Conventional strength determination requires specimen preparation, curing, and destructive testing, which can make extensive mixture evaluation time-consuming. This study develops an explainable machine-learning framework for predicting the compressive strength of plain and foamed concrete using Random Forest and Extreme Gradient Boosting (XGBoost). A database containing 1,373 observations was compiled, comprising 1,030 plain-concrete and 343 foamed-concrete mixtures. Separate models were developed for the two material classes using an 80:20 training-to-testing division, while five-fold cross-validation was used during model development and tuning. Model performance was evaluated using the coefficient of determination, mean absolute error, and root mean square error. For plain concrete, XGBoost achieved a testing coefficient of determination of 0.937, compared with 0.883 for Random Forest. For foamed concrete, the corresponding testing values were 0.563 and 0.561, respectively. Shapley Additive Explanations (SHAP) were subsequently used to interpret the contribution of mixture variables to model predictions. The results show that prediction of plain-concrete strength is substantially more reliable within the present database, whereas prediction of foamed-concrete strength remains constrained by material heterogeneity, limited sample representation, and the absence of direct pore-structure descriptors among the principal model inputs. Because plain and foamed concrete are frequently used together within the same modern construction project or structural element, the underlying motivation for the study is a single, hybrid web-based application that predicts compressive strength for both material types from one interface, rather than through separate, disconnected tools.

1. INTRODUCTION

Concrete remains one of the most widely used construction materials, and compressive strength is a fundamental parameter for assessing its mechanical performance and suitability for structural and non-structural applications. Foamed concrete is a lightweight cementitious material produced by incorporating preformed foam into a cementitious paste or mortar. The resulting cellular structure reduces density and can provide useful thermal and acoustic properties, making the material attractive for lightweight blocks, void filling, insulation, backfilling, and related applications (Amran et al., 2015; Ramamurthy et al., 2009; Raj et al., 2019). Its performance, however, is strongly affected by density, pore structure, binder composition, foam characteristics, and curing conditions, so changes in mixture proportions can produce substantial changes in strength.

The distinction between plain and foamed concrete is particularly important for predictive modelling. Plain concrete generally derives its strength from the combined effects of binder hydration, water content, aggregate skeleton, admixtures, and curing. Foamed concrete additionally contains a deliberately introduced air-void system, and the size, distribution, stability, and volume of those voids can alter density and mechanical response. Experimental work on foamed-concrete blocks has demonstrated the practical potential of lightweight cellular mixtures, including the use of supplementary cementitious materials, while review studies have highlighted the importance of constituent selection, production method, and air-void structure (Jose et al., 2020; Jones & McCarthy, 2005; Ramamurthy et al., 2009).

The use of supplementary cementitious materials, particularly fly ash, has also received considerable attention. Kearsley and Wainwright (2001) investigated the effect of high fly-ash contents on the compressive strength of foamed concrete, while Kearsley (1999) examined high-volume ungraded fly-ash systems over a range of casting densities. Related work has shown that fly ash and

silica fume can modify hardened properties and strength development (Gökçe et al., 2019). Earlier work by Jones et al. (2003) also emphasised the potential for increased utilisation of fly ash in concrete. These studies demonstrate that mixture composition and curing history interact in ways that are not always adequately represented by simple empirical relationships.

Traditional empirical and response-surface approaches have been used to relate foamed-concrete composition to density and strength. Nambiar and Ramamurthy (2006), for example, developed models connecting mixture composition with density and strength. Such relationships are useful within the ranges represented by the underlying experiments, but broader databases can contain nonlinear interactions among many mixture variables. This has encouraged the use of machine-learning approaches capable of approximating complex relationships without imposing a predetermined functional form.

The application of machine learning to concrete-strength prediction has developed from early artificial-neural-network studies to ensemble tree methods and explainable approaches. Yeh (1998) demonstrated that nonlinear data-driven models can capture relationships among concrete constituents and compressive strength. More recent studies have compared several tree-based and boosting algorithms and have reported strong predictive performance for high-strength concrete (Shaaban et al., 2025; Elhishi et al., 2023; Abioye et al., 2025). In addition to predictive accuracy, interpretability is important in engineering because a numerical prediction alone does not explain why a model produces a particular result. Explainable methods based on Shapley values can quantify the contribution of individual input variables and provide a bridge between machine-learning predictions and engineering interpretation (Elhishi et al., 2023; Abioye et al., 2025).

Foamed-concrete prediction has received increasing attention. Ullah et al. (2022) evaluated individual and ensemble machine-learning

approaches for sustainable foamed concrete and found that ensemble learning can represent nonlinear relationships among cement, sand, water-to-cement ratio, and foam volume. Salami et al. (2022) similarly compared neural, genetic, and gradient-boosting approaches using a database of 232 foamed-concrete results and reported strong validation performance for gradient boosting. These studies show that data-driven models can reduce the need for repeated trial batches, but their reliability remains dependent on database size, variable coverage, and the range of material systems represented.

More recent research has strengthened the role of explainable machine learning for foamed concrete. Zhang et al. (2024) evaluated eight machine-learning approaches using a database of 320 fly-ash composite foamed-concrete samples and reported that support vector regression produced the strongest test performance, while Shapley-based interpretation highlighted density, cement content, sand content, foam content, and water-to-binder ratio as influential variables. Yang et al. (2025) compared five machine-learning models for foamed concrete and reported strong performance from Bayesian-optimised Gaussian Process Regression, with density, water-to-cement ratio, age, fine-aggregate-to-binder ratio, superplasticizer, and supplementary-cementitious-material replacement among the assessed factors. Ullah et al. (2026) further applied explainable models to fly-ash foam composite concrete and combined model evaluation with Shapley analysis, individual conditional expectation, and partial-dependence plots. Together, these studies indicate that both prediction and interpretation are important when machine learning is applied to cellular cementitious materials.

Despite these developments, direct comparison of plain and foamed concrete within a unified data-driven framework remains useful because the two material classes have markedly different constituent ranges and physical mechanisms. A model trained across both classes without accounting for material type may learn cross-class differences rather than the within-class

relationships required for practical mixture assessment. The present study therefore develops separate Random Forest and XGBoost models for plain and foamed concrete using a combined database of 1,373 observations. The study evaluates predictive performance through an 80:20 train-test division and five-fold cross-validation, and it applies Shapley Additive Explanations to identify influential variables. The study also extends the modelling workflow into a web-based prediction application.

Beyond the comparative modelling exercise itself, the broader motivation for this work is practical. In modern construction, plain and foamed concrete are frequently used together within the same project, or even within the same structural element – for example, foamed concrete cast as an insulating or void-filling layer over or around conventional structural members, in composite sandwich panels, or as lightweight backfill placed around plain-concrete foundations and retaining structures (Amran et al., 2015; Jones & McCarthy, 2005; Ramamurthy et al., 2009). An engineer assessing such a composite system therefore needs to predict the compressive strength of both material types together, at the same time and from the same workflow, rather than consulting two separate, disconnected calculators or spreadsheets. This is the reasoning behind moving the underlying predictive models out of the research environment and into a single, hybrid web-based application that a designer, contractor, or researcher can use directly: rather than building two disconnected tools, the intention is one unified interface capable of predicting the compressive strength of either plain or foamed concrete, selected by the user at the point of use, so that both materials – as they actually occur together on site – can be screened side by side within one digital workflow. The comparative machine-learning analysis reported in Sections 3 and 4 is therefore presented not only as a benchmarking exercise between two material classes and two algorithms, but as the evidential basis for that hybrid application: it identifies which algorithm and which input variables should

drive each material-specific model within the deployed tool, and it makes transparent – through the SHAP-based interpretation in Section 4.5 – where the resulting predictions can and cannot be trusted.

2. Literature Review

2.1 Behaviour and composition of foamed concrete

Foamed concrete differs from conventional concrete principally because a controlled cellular system is intentionally introduced into the cementitious matrix. Reviews by Amran et al. (2015), Ramamurthy et al. (2009), and Raj et al. (2019) show that density, pore structure, constituent composition, production procedure, and curing conditions strongly influence its mechanical and functional properties. The strength-density relationship is especially important because increasing the volume of entrained air generally reduces the solid load-bearing fraction of the material. The same reduction in density that makes foamed concrete attractive for lightweight construction can therefore impose a penalty on compressive strength.

The production of stable foamed concrete also involves interactions that are difficult to describe using a single mixture ratio. Foam stability, mixing procedure, binder composition, fine-aggregate content, water demand, and curing conditions can influence the resulting air-void system. Ramamurthy et al. (2009) highlighted the importance of foaming-agent selection, foam preparation, mixing, and air-void distribution. Jones and McCarthy (2005) similarly discussed the behaviour and assessment of foamed concrete for construction applications. Experimental work on building blocks has demonstrated that foam volume and binder-related variables can be manipulated to produce materials suitable for different applications (Jose et al., 2020).

2.2 Fly ash and supplementary cementitious materials

The use of fly ash in foamed concrete is relevant to both material performance and sustainability. Kearsley and Wainwright (2001) showed that high fly-ash contents can be incorporated into foamed concrete while altering strength development. Kearsley (1999) investigated classified and unclassified fly ash at substantial replacement levels and documented relationships among density, porosity, and strength. Gökçe et al. (2019) further demonstrated that supplementary cementitious materials can modify hardened properties. These findings imply that a predictive model for foamed concrete should account for interactions among binder composition, water, foam, and curing age rather than relying on a single explanatory variable.

2.3 Machine learning for concrete strength prediction

Yeh (1998) provided an early demonstration that artificial neural networks could represent nonlinear relationships in concrete strength data. The subsequent literature has increasingly used ensemble methods because tree-based models can capture nonlinear effects and interactions without requiring extensive assumptions about the response surface. Elhishi et al. (2023) evaluated multiple machine-learning algorithms for concrete strength and combined model prediction with explainable analysis. Shaaban et al. (2025) and Abioye et al. (2025) similarly reported strong performance from tuned ensemble and boosting methods and emphasised the importance of validation and model interpretation.

2.4 Machine learning for foamed concrete

Foamed-concrete studies have increasingly adopted machine learning because laboratory determination of compressive strength is relatively slow and mixture development may require repeated trial batches. Ullah et al. (2022) used a database of 191 lightweight foamed-concrete observations and found strong performance from ensemble learning. Salami et al. (2022) used 232

experimental results and compared artificial neural networks, gene-expression programming, and gradient boosting. Zhang et al. (2024) expanded the comparison to eight algorithms using 320 fly-ash composite foamed-concrete observations and incorporated Shapley-based interpretation. Yang et al. (2025) used six principal mixture variables plus age and compared five models with hyperparameter optimisation. These studies collectively show that predictive performance varies with the database and input variables, and that reported performance should therefore be interpreted in relation to the data distribution and validation strategy rather than treated as an intrinsic property of an algorithm.

2.5 Research gap and contribution

The literature demonstrates strong progress in predicting compressive strength, but several issues remain relevant. First, plain and foamed concrete are often investigated independently, making direct comparison of model behaviour difficult.

Second, reported foamed-concrete datasets are generally smaller than conventional-concrete databases. Third, some studies use density as a direct input, whereas other datasets rely primarily on mixture quantities and foam content. Finally, high predictive accuracy does not by itself establish engineering interpretability. The present study addresses these issues by separating the plain and foamed subsets, using the same modelling framework for both, evaluating both hold-out and cross-validation performance, and interpreting the fitted XGBoost models using Shapley Additive Explanations.

3. Materials and Methods

The research methodology consisted of sequential stages involving data acquisition, preprocessing, correlation analysis, model development, statistical performance evaluation, sensitivity analysis, and deployment into a web-based estimator. An overview of the complete workflow is presented in Figure 1.

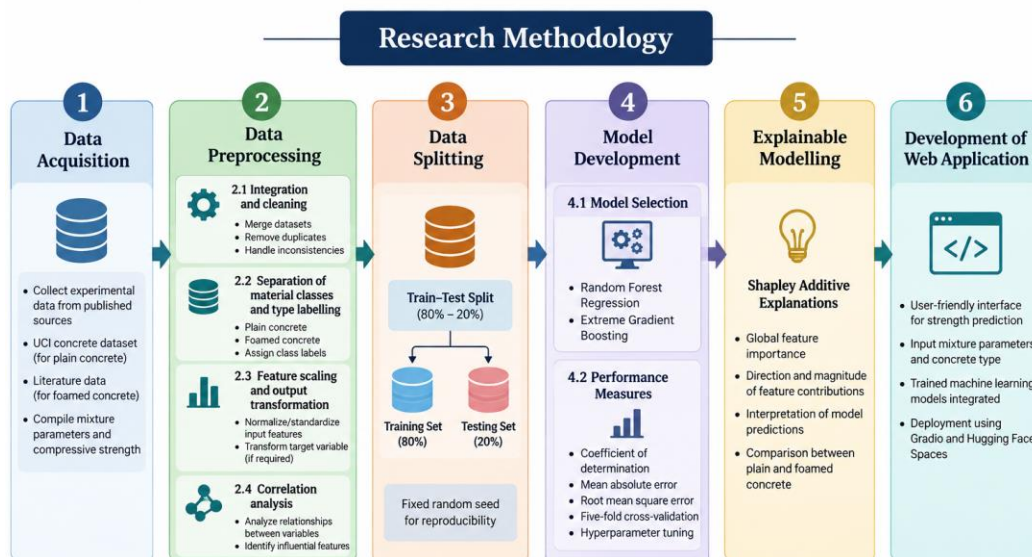


Figure 1. Overview of the Research Methodology.

3.1 Data acquisition

The combined database contains 1,373 observations. The plain-concrete subset contains 1,030 observations derived from the UCI concrete-strength database associated with Yeh

(1998). The foamed-concrete subset contains 343 observations, comprising 283 observations associated with the ACI 523.3R-14 cellular-concrete guide (American Concrete Institute, 2014) and 60 observations from Kearsley's (1999)

doctoral research. Each record contains mixture quantities and a measured 28-day compressive strength. The principal variables available in the combined database are cement content, blast-furnace slag content, fly-ash content, water

content, superplasticizer content, coarse-aggregate content, fine-aggregate content, foam volume, curing age, and compressive strength. Table 1 summarises the three data sources.

Table 1. Summary of data acquisition sources

Data source	Description	Number of samples
UCI concrete dataset (compiled by Yeh, 1998)	Plain (conventional) concrete mixtures	1,030
ACI 523.3R-14 cellular-concrete guide (American Concrete Institute, 2014)	Foamed concrete (with sand) mixes	283
Kearsley (1999) doctoral thesis	Foamed concrete (high-volume fly ash)	60

3.2 Data preprocessing

3.2.1 Integration and cleaning

Data collected from the three sources were integrated into a single table. Units were standardised (kilograms per cubic metre for masses, megapascals for strength) and column

names were harmonised. Minor inconsistencies were resolved and duplicate records were removed. No significant missing values were present, so imputation was not required. Table 2 summarises the statistical distribution of the combined dataset.

Table 2. Statistical description of the combined dataset

Variable	Mean	Std. dev.	Minimum	Maximum	Unit
Cement	304.92	102.51	102.00	595.00	kg/m ³
Blast-furnace slag	55.48	81.30	0.00	359.40	kg/m ³
Fly ash	67.89	158.39	0.00	1,037.00	kg/m ³
Water	182.47	37.66	52.38	399.00	kg/m ³
Superplasticizer	4.66	5.83	0.00	32.20	kg/m ³
Coarse aggregate	729.87	426.68	0.00	1,145.00	kg/m ³
Fine aggregate	742.99	222.55	0.00	1,389.47	kg/m ³
Foam volume	5.64	11.44	0.00	40.50	m ³
Curing age	47.02	69.67	1.00	365.00	days
Compressive strength	29.12	18.84	1.70	82.60	MPa

The dataset was assembled from heterogeneous sources, so units and column names were harmonised before modelling. The combined

dataset spans a broad strength range, from 1.70 to 82.60 MPa. The broad ranges of mixture constituents provide useful diversity for model

training but also increase the need for validation that is sensitive to data partitioning.

3.2.2 Separation of material classes and type labelling

Each record was labelled according to material class, but the material-class label was not used as a direct predictor in the final separate models. Instead, the database was divided into plain-concrete and foamed-concrete subsets. This approach was adopted because the two material systems differ substantially in constituent ranges, density, pore structure, and strength-development mechanisms. The plain-concrete model used cement content, blast-furnace-slag content, fly-ash content, water content, superplasticizer content, fine-aggregate content, coarse-aggregate content, and curing age. The foamed-concrete model used cement content, fly-ash content, water content, superplasticizer content, fine-aggregate content, foam volume, and curing age.

3.2.3 Feature scaling and output transformation

Tree-based models do not require conventional feature scaling because they operate through threshold-based partitions rather than distance metrics. Nevertheless, foam-volume values were normalised to a 0–1 range in the reported workflow, and the distributions of the input variables were checked for anomalous observations. Curing age was retained in days. The target variable, compressive strength, was retained in megapascals without logarithmic transformation, because the strength values were not severely skewed and interpretability is greater when predictions are reported in the original engineering units.

3.2.4 Correlation analysis

Correlation analysis was performed separately for the plain-concrete and foamed-concrete subsets to avoid masking within-class relationships through cross-class variability, since the two material systems differ significantly in their constituent proportions, density, pore structure, and

hydration patterns. For plain concrete, all conventionally used mix parameters were retained, namely cement, blast-furnace slag, fly ash, water, superplasticizer, fine aggregate, coarse aggregate, and curing age; foam volume was excluded because normal mixes do not contain foam. The resulting correlation heatmap (Figure 2) shows that cement content and curing age exhibit the strongest positive relationship with compressive strength, consistent with established hydration and strength-gain mechanisms, while water content shows a moderate negative correlation, reflecting its dilution effect on paste density when binder content is fixed. Fly ash and fine aggregate show weaker associations, indicating a secondary influence relative to binder mass and curing duration.

For foamed concrete, the analysis was restricted to the variables actually present in such mixes; coarse aggregate and blast-furnace slag were excluded because they do not appear in the foamed-concrete compositions in the combined dataset. The retained variables were cement, fly ash, water, superplasticizer, fine aggregate, foam volume, curing age, and compressive strength. The corresponding heatmap (Figure 3) reveals a strong negative correlation between foam volume and compressive strength, reaffirming the dominant role of entrained foam in reducing density and mechanical capacity. Cement content and curing age maintain positive correlations with strength, but with smaller magnitudes than in plain concrete, indicating that porosity introduced by foam tends to outweigh the influence of binder constituents. Water content again shows an inverse relation, consistent with its role in modifying pore-water distribution and mixture density.

Correlation was used as an exploratory tool rather than as a substitute for the nonlinear ensemble models: it captures only marginal, pairwise, linear association, whereas the fitted models and the subsequent Shapley analysis (Section 4.5) can represent nonlinear responses and interaction effects among several variables simultaneously.

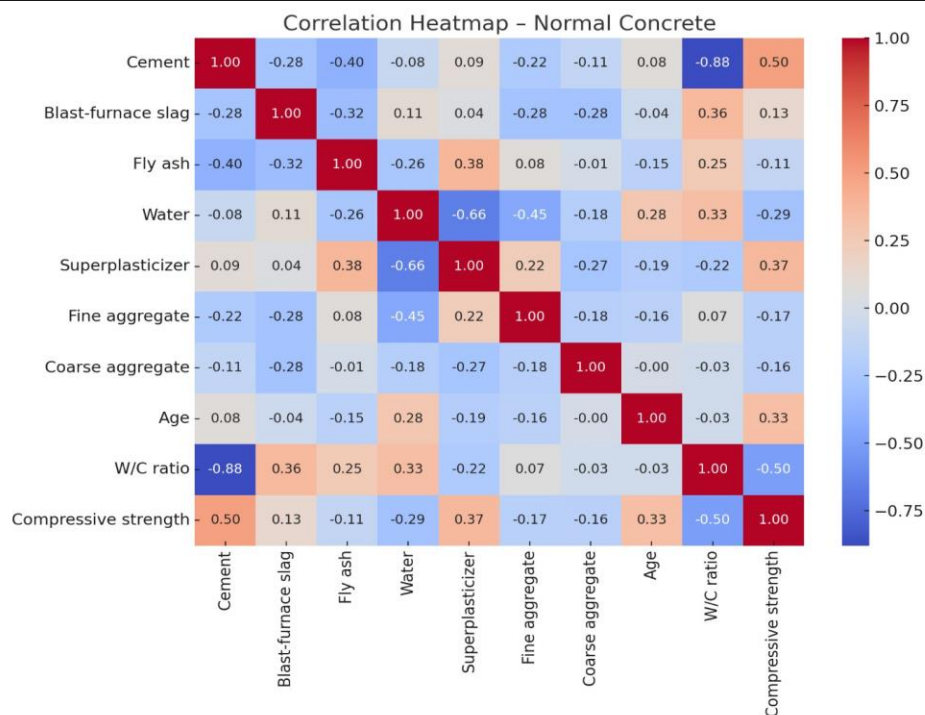


Figure 2. Correlation heatmap of mix-design parameters and compressive strength for plain concrete.

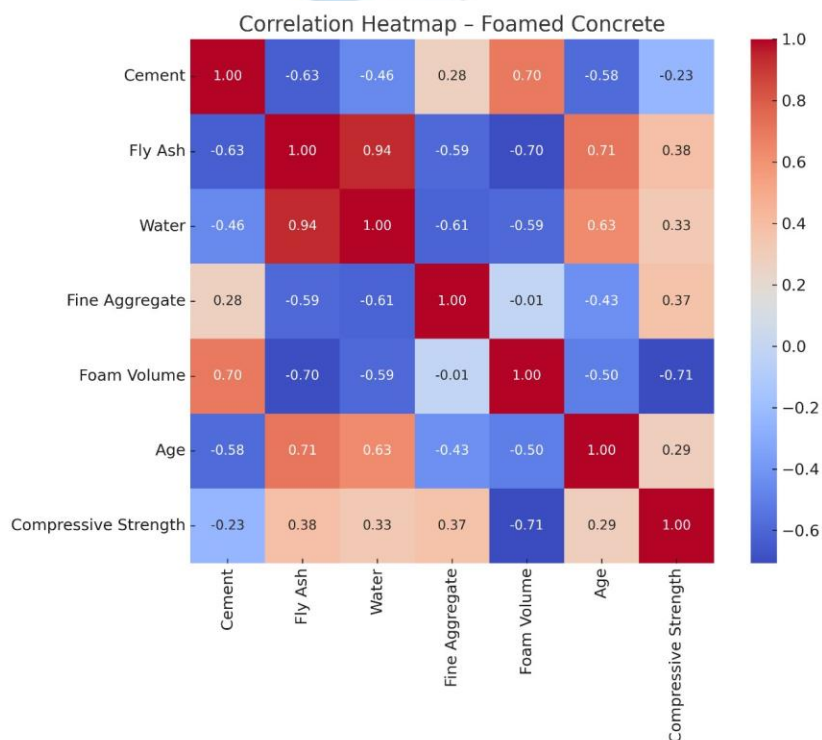


Figure 3. Correlation heatmap of mix-design parameters and compressive strength for foamed concrete.

3.3 Data splitting

Each material subset was randomly divided into 80% training data and 20% testing data using a fixed random seed to ensure reproducibility. The plain-concrete subset therefore contained 824 training observations and 206 testing observations. The foamed-concrete subset contained 274 training observations and 69 testing observations. The split preserved the distributional range of both the input variables and the corresponding compressive-strength values in each set, which is particularly important for the smaller foamed-concrete dataset. Five-fold cross-validation was applied to the training data during hyperparameter tuning: in each fold, four portions were used for model fitting and the remaining portion was used for validation, with the process repeated so that each portion served once as validation data.

3.4 Model development

3.4.1 Model selection

Random Forest Regression (bagging) and XGBoost Regression (boosting) were chosen because prior studies show that these ensemble algorithms offer high accuracy and robustness for concrete-strength prediction (Ullah et al., 2022; Salami et al., 2022; Zhang et al., 2024). Random Forests reduce variance through bootstrap sampling and averaging, while XGBoost reduces bias through sequential tree construction and regularisation.

Random Forest Regression constructs an ensemble of decision trees using bootstrap samples and random subsets of input variables. For regression, the individual tree predictions are averaged, which reduces the variance associated with a single decision tree and allows nonlinear interactions to be represented. Random Forest models were developed separately for the plain and foamed subsets. The reported baseline configuration used 100 trees, while grid search considered the number of trees, maximum tree depth, and minimum samples per leaf. Out-of-bag error and cross-validation results were used to monitor model stability.

XGBoost (Extreme Gradient Boosting) constructs decision trees sequentially, with each new tree designed to reduce the errors of the preceding trees. Regularisation, parallelised tree construction, and subsampling reduce excessive model complexity and help prevent overfitting. Separate XGBoost models were trained for plain and foamed concrete. The reported tuning process considered learning rate, tree depth, and row and column subsampling using five-fold cross-validation, with the objective of obtaining a model that maintained low validation error without relying solely on training performance.

3.4.2 Performance measures

Model performance was assessed using the coefficient of determination, mean absolute error, and root mean square error. The coefficient of determination was calculated as

$$R^2 = 1 - [\sum (y_i - \hat{y}_i)^2] / [\sum (y_i - \bar{y})^2]$$

where y_i is the observed compressive strength, $\hat{y}_i$ is the predicted compressive strength, and $\bar{y}$ is the mean observed compressive strength. Mean absolute error was calculated as

$$MAE = (1/n) \sum |y_i - \hat{y}_i|$$

Root mean square error was calculated as

$$RMSE = \sqrt{(1/n) \sum (y_i - \hat{y}_i)^2}$$

The coefficient of determination describes the proportion of variance in compressive strength represented by the model, while the two error measures describe the magnitude of prediction error in the original strength units (MPa). Lower error values and higher coefficient-of-determination values indicate closer agreement between predicted and observed compressive strength.

3.5 Explainable modelling using Shapley Additive Explanations

Shapley Additive Explanations (SHAP) were used to interpret the fitted XGBoost models, which achieved the strongest hold-out performance for both material classes (Section 4). For each observation, the method assigns a contribution value to each input variable based on cooperative game theory, describing its contribution to the

model output relative to a reference prediction. Global feature importance was examined using the mean absolute SHAP value across the dataset, while the sign and spread of the individual SHAP values were used to establish whether high or low values of a variable tend to increase or decrease the predicted compressive strength, and how strongly a variable interacts with others. This analysis was conducted separately for the plain-concrete and foamed-concrete XGBoost models, using a global importance bar chart and a corresponding beeswarm (summary) plot for each material class.

3.6 Web application

As set out in Section 1, the practical objective motivating this study is a single, hybrid web-based application that predicts compressive strength for both plain and foamed concrete within one interface, rather than two separate tools for the two material classes. The trained models were accordingly integrated into a web-based compressive-strength estimator, implemented in Python with Gradio serving as the front-end interaction framework. The interface accepts mixture parameters, material type (plain or foamed), curing age, and model selection (Random Forest or XGBoost), and returns a predicted compressive strength together with the corresponding training and testing performance indicators for the selected model. Selecting the material type at the point of use is what makes the tool hybrid in this sense: the same deployed application routes the request to the material-specific Random Forest or XGBoost model reported in Section 4, without requiring the user to consult separate software for plain and for foamed mixtures. The trained preprocessing and model objects are loaded without retraining, allowing rapid prediction. The application was deployed on Hugging Face Spaces, which removes the need for local installation of Python packages, machine-learning libraries, or development tools and automatically handles dependency

management, version control, and hosting. This extends the modelling workflow described above into an accessible decision-support tool for students, researchers, and practising engineers; it is intended to support preliminary mixture screening and is not a substitute for code-required laboratory testing or engineering judgement.

4. Results and Discussion

This section presents the comprehensive performance assessment of the machine-learning models developed to estimate the compressive strength of plain and foamed concrete. The evaluation draws on results from both hold-out testing (80:20 split) and five-fold cross-validation for Random Forest and XGBoost. Since the mechanical behaviour and density characteristics of the two concrete types differ significantly, separate models were trained for each material class to avoid cross-class bias and to improve predictive stability. Two performance dimensions were analysed for each model: (i) statistical evaluation using mean absolute error, root mean square error, and the coefficient of determination, and (ii) interpretability evaluation using SHAP analysis to examine feature contributions.

4.1 Overall model performance

Table 3 through Table 6 summarise the training, testing, and cross-validated performance of the Random Forest and XGBoost models for the plain-concrete and foamed-concrete subsets. The results show a marked difference between the prediction of plain and foamed concrete. The plain-concrete models achieved high testing performance, whereas the foamed-concrete models showed substantially lower testing coefficients of determination. This difference is important because it demonstrates that algorithm selection alone cannot overcome limitations imposed by database size, material heterogeneity, and missing descriptors of the internal pore structure.

4.2 Compressive strength prediction of plain concrete

4.2.1 Random Forest model

Table 3. Performance of the Random Forest model for plain concrete

Evaluation type	MAE (MPa)	RMSE (MPa)	R ²	No. of samples
Training	1.276	1.948	0.987	824
Testing	3.764	5.497	0.883	206
5-fold cross-validation	-	-	0.900 ± 0.016	824 (across folds)

The Random Forest model achieved strong predictive accuracy during training, with a mean absolute error of 1.276 MPa and a root mean square error of 1.948 MPa. The coefficient of determination was exceptionally high ($R^2 = 0.987$), indicating that the model successfully captured the nonlinear effects of cement content, water quantity, aggregates, and admixtures on the resulting compressive strength. Performance on the independent testing dataset also remained strong: the testing mean absolute error increased to 3.764 MPa and the root mean square error to

5.497 MPa, as expected when applying the model to unseen data, while R^2 stayed comparatively high at 0.883, demonstrating that the model generalised well.

Five-fold cross-validation yielded a cross-validated R^2 of 0.900 ± 0.016 , confirming the stability of the model; the small standard deviation across folds indicates that predictions were not strongly affected by how the dataset was partitioned. Figure 4 illustrates the corresponding training and testing metrics.

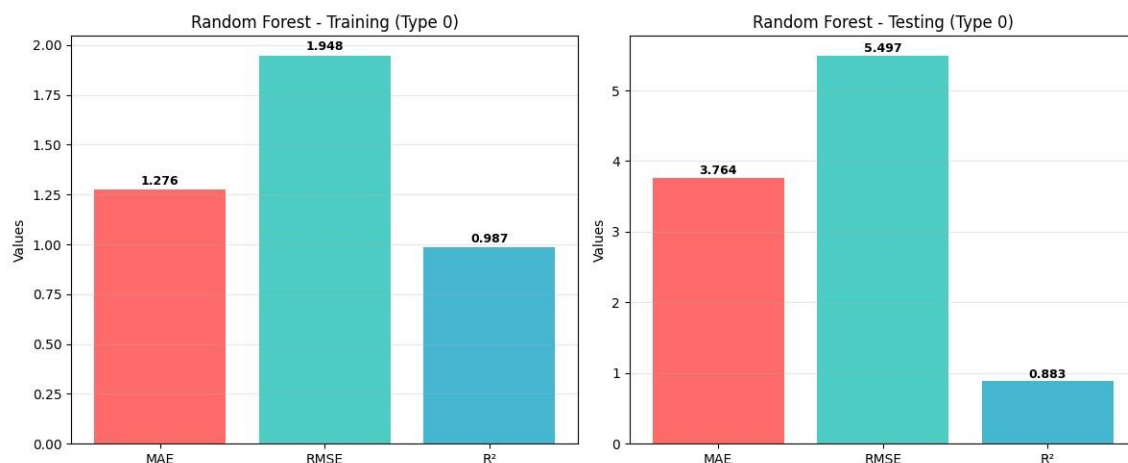


Figure 4. Training and testing performance metrics of the Random Forest model for plain concrete.

4.2.2 XGBoost model

Table 4. Performance of the XGBoost model for plain concrete

Evaluation type	MAE (MPa)	RMSE (MPa)	R ²	No. of samples
Training	0.444	1.203	0.995	824
Testing	2.596	3.840	0.937	206

Evaluation type	MAE (MPa)	RMSE (MPa)	R ²	No. of samples
5-fold cross-validation	–	–	0.916 ± 0.015	824 (across folds)

XGBoost produced still higher plain-concrete training performance, with a coefficient of determination of 0.995, mean absolute error of 0.444 MPa, and root mean square error of 1.203 MPa, reflecting the ability of the boosted-tree structure to capture nonlinear interactions among cement, water, aggregates, and admixtures. On the independent testing data, the coefficient of determination was 0.937, mean absolute error was 2.596 MPa, and root mean square error was 3.840 MPa, indicating that the fitted boosting model generalised effectively within the range of the database. Five-fold cross-validation produced a coefficient of determination of 0.916 ± 0.015 .

The difference between training and testing performance nevertheless indicates that the models fit the training data more closely than unseen observations, which is expected in supervised learning, particularly for flexible ensemble models. The cross-validation results provide additional evidence that performance was not produced by a single favourable train-test partition; however, because the database contains many closely related conventional-concrete mixtures, the reported performance should not be interpreted as universal accuracy for all concrete mixture designs. Figure 5 illustrates the corresponding training and testing metrics.

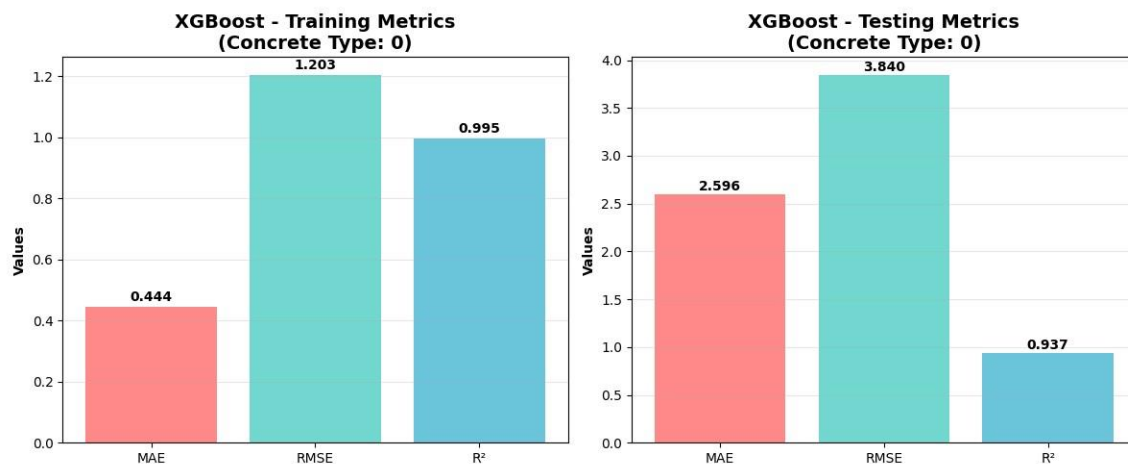


Figure 5. Training and testing performance metrics of the XGBoost model for plain concrete.

4.3 Compressive strength prediction of foamed concrete

4.3.1 Random Forest model

Table 5. Performance of the Random Forest model for foamed concrete

Evaluation type	MAE (MPa)	RMSE (MPa)	R ²	No. of samples
Training	0.593	2.213	0.898	274
Testing	1.991	4.732	0.561	69
5-fold cross-validation	–	–	0.776 ± 0.060	274 (across folds)

For foamed concrete, Random Forest achieved a training mean absolute error of 0.593 MPa, root mean square error of 2.213 MPa, and R^2 of 0.898 — indicating that the model learned general trends in the data, but with notably higher prediction error than for plain concrete. The lower R^2 reflects the inherent complexity and variability of foamed-concrete mixes, in which changes in foam volume, density, and pore structure introduce substantial nonlinearity and noise.

Testing performance decreased further, to a mean absolute error of 1.991 MPa, root mean square error of 4.732 MPa, and R^2 of 0.561, highlighting

the difficulty of modelling compressive strength in foamed concrete, where strength is highly sensitive to foam stability, the uniformity of air voids, and microstructural inconsistencies. Five-fold cross-validation produced a coefficient of determination of 0.776 ± 0.060 ; the larger dispersion relative to the plain-concrete model indicates sensitivity to the composition of the training folds, which is expected given the smaller sample size and greater variability of the foamed-concrete subset. Figure 6 illustrates the corresponding training and testing metrics.

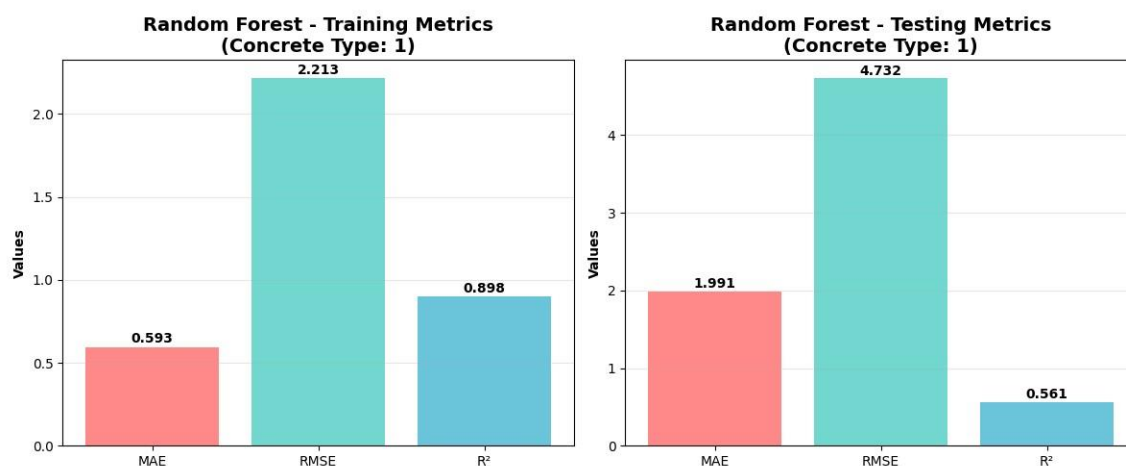


Figure 6. Training and testing performance metrics of the Random Forest model for foamed concrete.

4.3.2 XGBoost model

Table 6. Performance of the XGBoost model for foamed concrete

Evaluation type	MAE (MPa)	RMSE (MPa)	R^2	No. of samples
Training	0.599	2.215	0.898	274
Testing	1.959	4.721	0.563	69
5-fold cross-validation	–	–	0.796 ± 0.064	274 (across folds)

XGBoost produced a training mean absolute error of 0.599 MPa, root mean square error of 2.215 MPa, and R^2 of 0.898, again reflecting the difficulty of modelling a material in which foam volume, air-void stability, and density changes strongly influence compressive behaviour. Testing performance was a mean absolute error of 1.959

MPa, root mean square error of 4.721 MPa, and R^2 of 0.563. Five-fold cross-validation produced a coefficient of determination of 0.796 ± 0.064 , again showing considerable dispersion across folds. The two algorithms therefore produced almost identical hold-out testing coefficients of determination for foamed concrete (0.561 for

Random Forest and 0.563 for XGBoost), despite the slightly lower error values reported for XGBoost. Figure 7 illustrates the corresponding training and testing metrics.

The relatively modest testing performance for foamed concrete should not be interpreted as evidence that tree-based machine learning is unsuitable for this material. Instead, it indicates that the present input representation and database do not fully capture the mechanisms controlling strength. Foamed-concrete strength is strongly influenced by the cellular structure, including void volume, distribution, stability, and density; the present foamed-concrete inputs contain foam volume but do not directly describe pore-size distribution or measured density as a principal predictor. This limitation is important when comparing the present results with studies that explicitly include density – for example, Zhang et al. (2024) used wet density as an input, while Yang

et al. (2025) included foamed-concrete density, water-to-cement ratio, and other ratio-based descriptors, so their reported performance should not be compared directly with the present model without considering the different input representations and databases.

The database size is also relevant. The foamed subset contains only 343 observations, of which 274 were used for training and 69 for independent testing – considerably smaller than the 1,030-observation plain-concrete subset. In addition, the foamed data combine observations from different sources and mixture systems; such heterogeneity is valuable for broad applicability but can also increase unexplained variance. Salami et al. (2022), Ullah et al. (2022), and Zhang et al. (2024) all demonstrate that reported predictive performance depends strongly on the selected variables, data range, and validation design, and the present results reinforce that point.

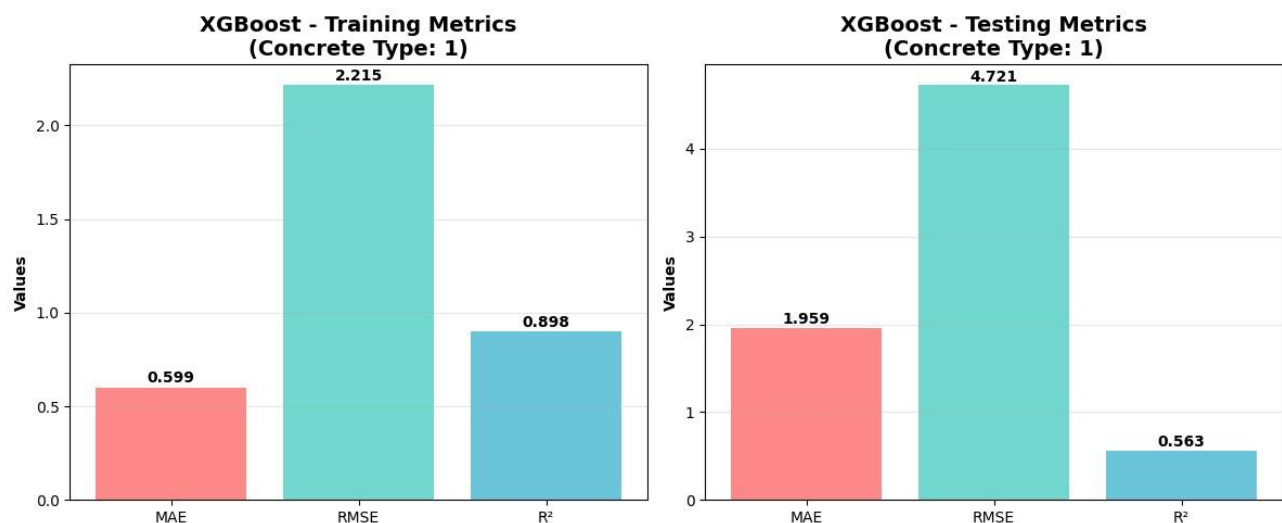


Figure 7. Training and testing performance metrics of the XGBoost model for foamed concrete.

4.4 Comparison with published foamed-concrete studies

The present results can be placed in the context of previous foamed-concrete machine-learning studies. Salami et al. (2022) reported strong validation performance for gradient boosting using density, water-to-cement ratio, and sand-to-cement ratio as principal inputs. Ullah et al.

(2022) also reported a strong coefficient of determination using ensemble learning and a database containing cement content, sand content, water-to-cement ratio, and foam volume. Zhang et al. (2024) used a larger 320-observation fly-ash composite foamed-concrete database and included wet density, cement, fly ash, sand, water-to-cement ratio, foam content, and curing age.

Yang et al. (2025) likewise used density and several ratio-based variables and reported high performance from Bayesian-optimised Gaussian Process Regression. These studies suggest that density and ratio-based variables may carry information that is not fully represented by raw constituent quantities alone.

A further relevant development is the use of explainable modelling. Ullah et al. (2026) demonstrated that explainable machine-learning models can combine high predictive performance with Shapley analysis and other interpretation tools for fly-ash foam composite concrete. The present study follows the same general principle but obtains substantially lower hold-out performance for foamed concrete. This difference should be treated as a result rather than concealed: it highlights the need to distinguish between a model's algorithmic capacity and the information content of its input database.

4.5 SHAP-based sensitivity analysis

To interpret the behaviour of the XGBoost models beyond aggregate error metrics, a sensitivity analysis was carried out using SHAP. For the plain-concrete model, the global importance ranking (Figure 8) shows that a small set of mix parameters dominates the prediction of compressive strength: coarse-aggregate content, cement content, water content, curing age, and, to a lesser extent, the quantities of supplementary cementitious materials (blast-furnace slag and fly ash) and superplasticizer exhibit the largest mean absolute SHAP values. This pattern confirms classical mix-design theory – higher binder content and longer curing generally increase strength, whereas excessive water reduces it. The beeswarm plot (Figure 9) further clarifies these trends: points corresponding to high cement content and longer curing days are associated with positive SHAP values, indicating an upward contribution to predicted strength, while higher water contents tend to yield negative SHAP values. The spread of the SHAP values also highlights the nonlinear response of strength at extreme combinations of water and cement, which is not visible from correlation analysis alone.

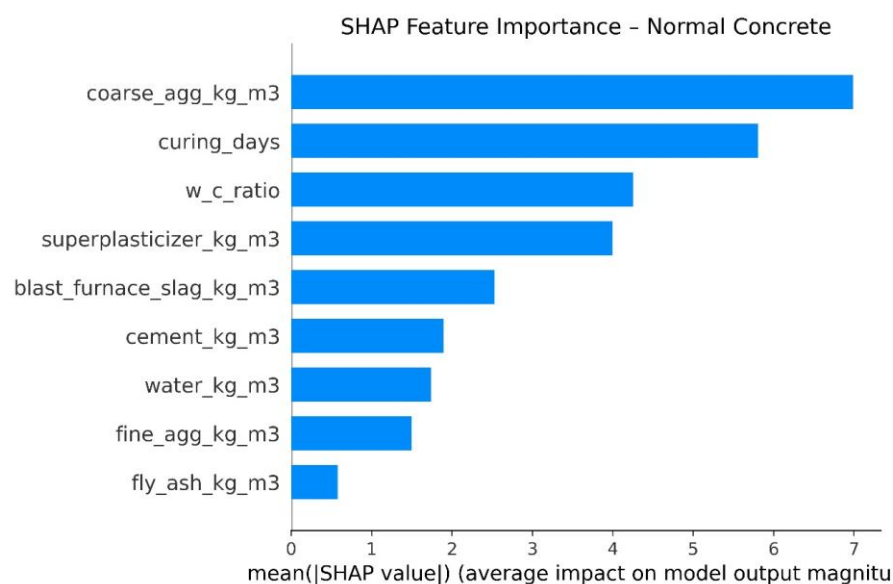


Figure 8. Mean absolute SHAP value (global feature importance) for compressive strength of plain concrete.

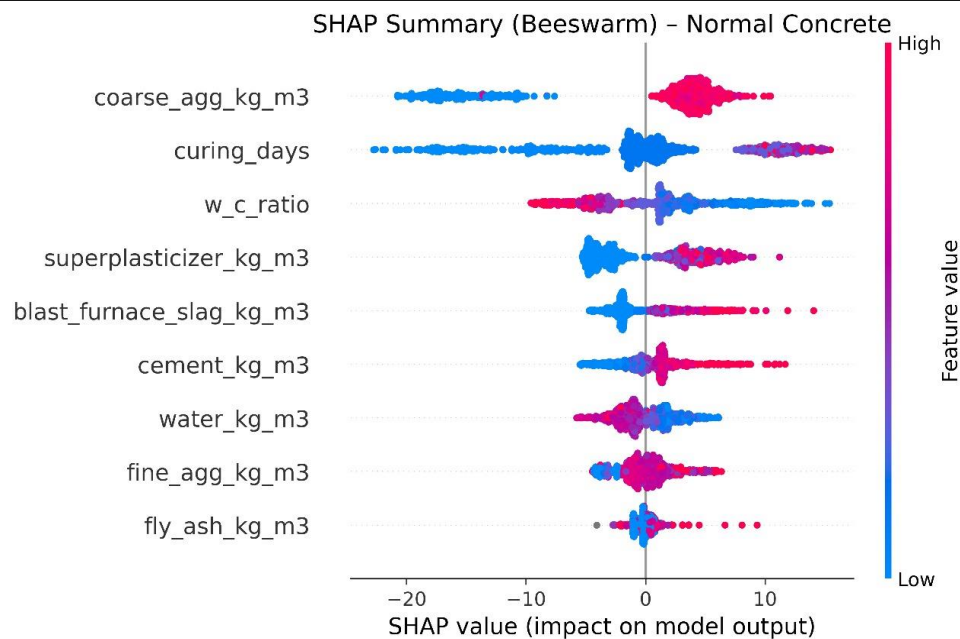


Figure 9. SHAP summary (beeswarm) plot for the XGBoost model of plain concrete.

The foamed-concrete model exhibits a different sensitivity pattern, reflecting the fundamentally different role of foam in governing mechanical behaviour. The SHAP summary (beeswarm) plot for foamed concrete (Figure 10) shows curing age, cement content, and water content among the most influential predictors, with foam volume contributing least despite its central role in the material. Higher cement contents and longer curing days are associated with predominantly

positive SHAP values, signifying a strengthening effect even in highly aerated matrices, while higher foam volumes are associated with negative SHAP values, consistent with the known relationship between air-void content, density, and mechanical capacity reported in the foamed-concrete literature (Amran et al., 2015; Ramamurthy et al., 2009; Raj et al., 2019).

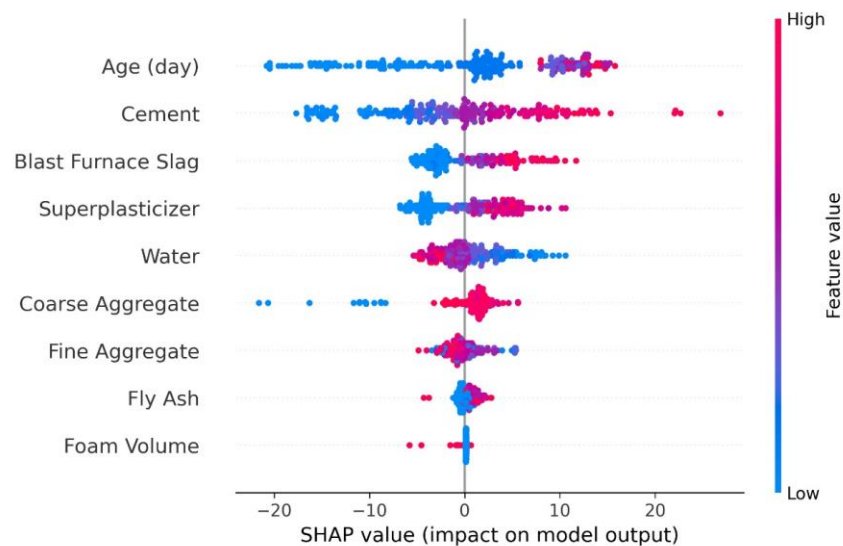


Figure 10. SHAP summary (beeswarm) plot for foamed concrete.

Overall, the SHAP-based sensitivity analysis demonstrates that the XGBoost models have learned physically meaningful relationships for both plain and foamed concrete. For plain concrete, strength is primarily governed by the classical cement–water–age relationship, modified by the presence of slag, fly ash, and admixtures. For foamed concrete, the dominant role shifts towards foam volume and density-related effects, with binder content and curing age still acting as key strengthening mechanisms. These insights

increase confidence in the predictive models and provide practical guidance for mix optimisation: adjusting the variables identified as most influential by SHAP offers an effective route for tailoring compressive strength in both material systems. For engineering interpretation, however, a SHAP contribution should not automatically be read as a causal influence – it describes how the trained predictive model uses the available variable within the observed database, not necessarily the underlying physical mechanism.

4.6 Practical application: a hybrid web-based estimator

The trained models were integrated into a fully functional software application (Section 3.6) designed to provide real-time prediction of compressive strength for both plain and foamed concrete mixes from a single, hybrid interface, translating the research outcomes into a practical decision-support tool for students, researchers, and practising engineers. This hybrid design – one application covering both material classes, rather than two separate tools – is the practical contribution that the comparative modelling in Sections 4.1 to 4.5 was carried out to support (Section 1). Figure 11 shows the graphical interface: users specify the material class, provide

the relevant mixture variables and curing age, select a trained model, and obtain a predicted compressive strength together with the corresponding training and testing performance indicators. The example shown corresponds to a foamed-concrete mixture with a cement content of 446 kg/m³, a water content of 160 kg/m³, a fine-aggregate content of 646 kg/m³, and a foam content of 28.21 kg/m³, for which the deployed XGBoost model predicted a compressive strength of 4.84 MPa.

This application is intended for preliminary assessment and research support rather than replacement of code-required laboratory testing or engineering judgement – a distinction that is particularly important for foamed concrete, given

that the present model's testing coefficient of determination is only about 0.56 and the database

does not directly encode several pore-structure descriptors.

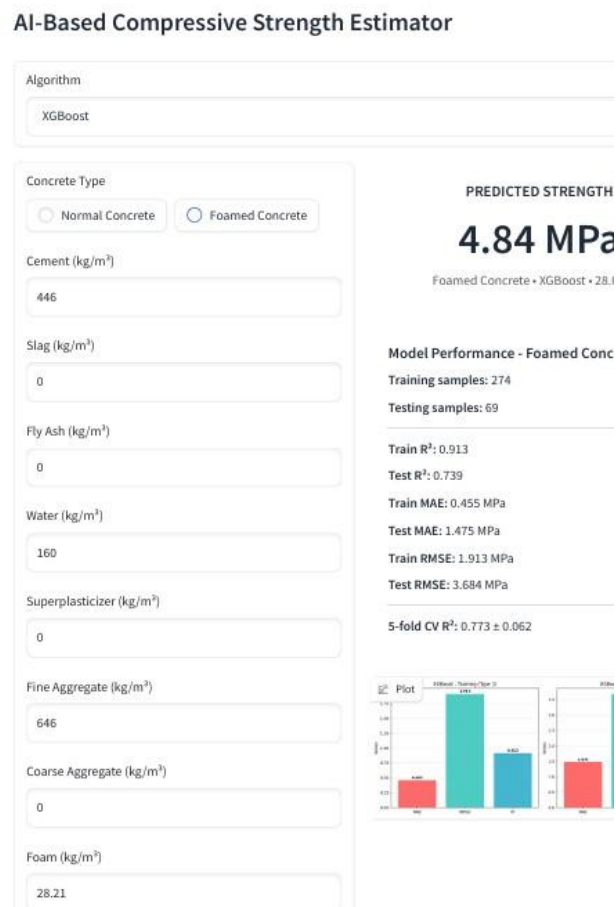


Figure 11. Graphical interface of the web-based compressive-strength estimator, showing a foamed-concrete prediction using the XGBoost model.

5. Conclusions

This study developed an explainable machine-learning framework for predicting the compressive strength of plain and foamed concrete. A database of 1,373 observations was separated into 1,030 plain-concrete and 343 foamed-concrete observations. Random Forest and XGBoost models were developed separately for the two material classes using an 80:20 train-test division and five-fold cross-validation.

For plain concrete, the models achieved strong independent testing performance: Random Forest produced a testing coefficient of determination of 0.883, while XGBoost produced 0.937, with lower

testing mean absolute error and root mean square error. The cross-validation values of 0.900 ± 0.016 for Random Forest and 0.916 ± 0.015 for XGBoost indicate reasonably stable model behaviour within the training database.

For foamed concrete, both algorithms produced substantially lower testing coefficients of determination – 0.561 for Random Forest and 0.563 for XGBoost. The five-fold cross-validation values were higher but showed greater dispersion, indicating sensitivity to data partitioning. These findings suggest that the principal limitation is not simply the choice between the two ensemble algorithms; instead, the smaller database,

heterogeneous source data, and limited representation of pore-structure-related characteristics constrain predictive generalisation. Shapley Additive Explanations provided engineering-oriented interpretation of the fitted XGBoost models. For plain concrete, coarse-aggregate content, cement content, water content, curing age, and superplasticizer content showed the largest global contributions. For foamed concrete, curing age, cement content, and water content were identified as the most influential variables, with foam volume contributing negatively as expected from established foamed-concrete behaviour.

The study demonstrates that machine learning can provide a useful rapid-assessment tool, but the results also show why explainability and validation are essential. The proposed web application can support preliminary mixture screening, education, and research; it should not be treated as a substitute for experimental verification, particularly for foamed-concrete mixtures outside the range represented in the training database.

Consistent with the motivating aim stated in Section 1, the resulting predictive models were unified within a single hybrid web-based application (Sections 3.6 and 4.6) that allows either plain or foamed concrete to be assessed from the same interface. Rather than requiring separate tools for conventional and lightweight cellular mixtures, the deployed application lets a user move between material classes within one workflow. This hybrid deployment – not only the comparative accuracy figures reported above – is intended as the primary practical contribution of the study, with the plain-versus-foamed modelling comparison serving as the evidential and interpretive basis for it.

6. Limitations and Future Research

Several limitations should be recognised. First, the foamed-concrete subset is comparatively small and combines data from different sources. Second, density and pore-structure descriptors are not represented in the same way as in several recent foamed-concrete studies. Third, the database is

assembled from published and code-based sources rather than generated under a single controlled experimental programme. Fourth, the present comparison is limited to Random Forest and XGBoost.

Future research should expand the foamed-concrete database with measured density, pore-size characteristics, foam-stability indicators, water-to-binder ratio, and other physically meaningful descriptors. Additional models, including Gaussian Process Regression, symbolic regression, and hybrid or physics-informed approaches, could then be assessed using identical validation protocols. Future experimental work should also include external validation using newly produced mixtures that are not present in the training database, which would provide a stronger test of engineering generalisation than random hold-out testing alone. If additional data become available, uncertainty quantification should be incorporated so that the application reports not only a point prediction but also an indication of prediction reliability – a development that would be particularly valuable for lightweight foamed concrete, where material variability can be large.

A larger, more diverse, and systematically documented foamed-concrete dataset – providing broader representation of density ranges, mixture proportions, curing conditions, foam characteristics, and material combinations – should be regarded as a fundamental prerequisite for obtaining more robust predictive performance and for establishing stronger academic and engineering confidence in machine-learning-based strength estimation. Increasing the number of independent observations would also reduce sensitivity to individual train-test partitions and improve the statistical reliability and generalisability of predictive models.

Acknowledgements

The authors acknowledge the assistance of colleagues and researchers who contributed to the development and review of this work, and gratefully acknowledge the underlying published

experimental and code-based sources from which the modelling database was compiled.

REFERENCES

- Abioye, S. O., Babatunde, Y. O., Abikoye, O. A., Shaibu, A. N., & Bankole, B. J. (2025). Optimized machine learning algorithms with SHAP analysis for predicting compressive strength in high-performance concrete. *AI in Civil Engineering*, 4, 16. <https://doi.org/10.1007/s43503-025-00061-x>
- American Concrete Institute. (2014). *ACI 523.3R-14: Guide for cellular concretes above 50 lb/ft³ (800 kg/m³)*. American Concrete Institute.
- Amran, Y. M., Farzadnia, N., & Ali, A. A. (2015). Properties and applications of foamed concrete: A review. *Construction and Building Materials*, 101, 990–1005. <https://doi.org/10.1016/j.conbuildmat.2015.10.112>
- Elhishi, S., Elashry, A. M., & El-Metwally, S. (2023). Unboxing machine learning models for concrete strength prediction using explainable artificial intelligence. *Scientific Reports*, 13, 19892.
- Gökçe, H. S., Hatungimana, D., & Ramyar, K. (2019). Effect of fly ash and silica fume on hardened properties of foam concrete. *Construction and Building Materials*, 194, 1–11.
- Jones, M., & McCarthy, A. (2005). Behaviour and assessment of foamed concrete for construction applications. In *Use of foamed concrete in construction: Proceedings of the international conference held at the University of Dundee, Scotland, UK on 5 July 2005* (pp. 61–88). Thomas Telford Publishing.
- Jones, M. R., McCarthy, M. J., & McCarthy, A. (2003). Moving fly ash utilisation in concrete forward: A UK perspective. *Proceedings of the 2003 International Ash Utilization Symposium*, 20–22.
- Jose, S. K., Soman, M., & Evangeline, S. Y. (2020). Experiments on foamed concrete for the development of building blocks. *International Journal of Recent Technology and Engineering*, 8(5), 2824–2829. <https://doi.org/10.35940/ijrte.C6250.018520>
- Kearsley, E. P. (1999). *The effect of high volumes of ungraded fly ash on the properties of foamed concrete* [Doctoral dissertation, University of Leeds]. White Rose eTheses.
- Kearsley, E., & Wainwright, P. (2001). The effect of high fly ash content on the compressive strength of foamed concrete. *Cement and Concrete Research*, 31(1), 105–112.
- Nambiar, E. K., & Ramamurthy, K. (2006). Models relating mixture composition to the density and strength of foam concrete using response surface methodology. *Cement and Concrete Composites*, 28(9), 752–760.
- Raj, A., Sathyan, D., & Mini, K. M. (2019). Physical and functional characteristics of foam concrete: A review. *Construction and Building Materials*, 221, 787–799. <https://doi.org/10.1016/j.conbuildmat.2019.06.052>
- Ramamurthy, K., Nambiar, E. K., & Ranjani, G. I. S. (2009). A classification of studies on properties of foam concrete. *Cement and Concrete Composites*, 31(6), 388–396. <https://doi.org/10.1016/j.cemconcomp.2009.04.006>
- Salami, B. A., Iqbal, M., Abdulraheem, A., Jalal, F. E., Alimi, W., Jamal, A., Tafsirojjaman, T., Liu, Y., & Bardhan, A. (2022). Estimating compressive strength of lightweight foamed concrete using neural, genetic and ensemble machine learning approach. *Cement and Concrete Composites*, 133, 104721. <https://doi.org/10.1016/j.cemconcomp.2022.104721>

- Shaaban, M., Amin, M., Selim, S., & Riad, I. M. (2025). Machine learning approaches for forecasting compressive strength of high-strength concrete. *Scientific Reports*, 15, 25567. <https://doi.org/10.1038/s41598-025-10342-1>
- Ullah, A., Yang, Y., Ullah, W., Ayub, B., Alzlfawi, A., & Iqbal, I. (2026). Toward transparent AI: Predicting strength of fly ash foam composite concrete using explainable machine learning models. *Structural Concrete*, 27, 595–624.
- Ullah, H. S., Khushnood, R. A., Farooq, F., Ahmad, J., Vatin, N. I., & Ewais, D. Y. Z. (2022). Prediction of compressive strength of sustainable foam concrete using individual and ensemble machine learning approaches. *Materials*, 15(9), 3166. <https://doi.org/10.3390/ma15093166> 3634.
- Yang, S., Zhong, J., Gan, B., Sun, Y., Bu, C., Zhang, M., Li, J., & Yu, Y. (2025). Prediction and sensitivity analysis of foam concrete compressive strength based on machine learning techniques with hyperparameter optimization. *Computer Modeling in Engineering & Sciences*, 144(3), 2943–2967. <https://doi.org/10.32604/cmescs.2025.067282>
- Yeh, I.-C. (1998). Modeling of strength of high-performance concrete using artificial neural networks. *Cement and Concrete Research*, 28(12), 1797–1808.
- Zhang, C., Zhu, Z., Shi, L., Kang, X., Wan, Y., Huo, W., & Yang, L. (2024). Compressive strength and sensitivity analysis of fly ash composite foam concrete: Efficient machine learning approach. *Advances in Engineering Software*, 192, 103634. <https://doi.org/10.1016/j.advengsoft.2024.103634>