

Quantifying the compressive strength of basalt fiber reinforced concrete using advanced hybrid machine learning models

Ibrahim Y. Hakeem^{1,2} , Abdullah Alzlfawi³ , Md. Habibur Rahman Sobuz^{4,6} , Ali Alhamami^{1,2}, Faisal Amsyar⁵, Md. Kawsarul Islam Kabbo⁴ , Abdulaziz Saleh Ahmed Alnasi^{1,2}, Md. Munir Hayet Khan⁶ 

¹Najran University, College of Engineering, Department of Civil Engineering. Najran, Saudi Arabia.

²Najran University, Science and Engineering Research Center. Najran, Saudi Arabia.

³Majmaah University, College of Engineering, Department of Civil and Environmental Engineering. 11952, Al Majmaah, Saudi Arabia.

⁴Khulna University of Engineering & Technology, Department of Building Engineering and Construction Management. 9203, Khulna, Bangladesh.

⁵Universiti Malaysia Sarawak, Faculty of Engineering, Department of Civil Engineering. 94300, Sarawak, Malaysia.

⁶INTI International University, Faculty of Engineering & Quantity Surveying. Negeri Sembilan, Malaysia.

e-mail: iyhakeem@nu.edu.sa, a.alzlfawi@mu.edu.sa, habib@becm.kuet.ac.bd, ahalhamami@nu.edu.sa, rfamsyar@unimas.my, kabbo@becm.kuet.ac.bd, Z2a2e23456@gmail.com, munir.hayetkhan@newinti.edu.my

ABSTRACT

Basalt Fiber Reinforced Concrete (BFRC) is being recognized as an eco-friendly advanced material with reduced environmental footprint, higher mechanical performance and long-term durability. However, its compressive strength prediction still appears to be a difficult problem due to the nonlinearity caused by the interaction of the mix components. This paper demonstrates a production-quality hybrid ML model to predict 28-day compressive strength of BFRC, providing an economical alternative to laborious and expensive laboratory-based testing. A well-defined database of 450 samples is generated involving the essential parameters, such as cement (440 kg/m³), SCMs (110 kg/m³), basalt fiber (0–4.25 kg/m³), fine aggregate (740 kg/m³), and coarse aggregate (975 kg/m³), water (121 kg/m³), superplasticizer (3.52 kg/m³), and curing period (3–365 days). The pre-processed and normalized data were partitioned into the training set (80%) and the test set (20%). Five ML models Gradient Boosting (GB), Cat Boost (CB), Light GBM (LGB), and their hybrid ensembles: GB+CB, GB+LGB were trained and compared using different metrics such as R², RMSE, MAE, MedAE, etc. Among all ML models, the GB+LGBM model showed the best performance with R² = 0.9445, RMSE = 5.99 MPa, MAE = 3.62 MPa, and MedAE = 2.26 MPa on the test set. SHAP analysis revealed that coarse aggregates (SHAP ≈ +8) and cement (SHAP ≈ +7) were the most influential factors, while the remaining water content and the excessive dosage of fiber were disadvantageous. Estimated compressive strength varied from 20 to 140 MPa. This study shows a novel approach by demonstrating the ability of ensemble ML models to capture complex concrete behavior, providing a data-driven approach for sustainable manufacturing of concrete. However, full reliance on the literature-based dataset still has significant limitations, which will increase noise, and these limitations can be further overcome by experimental validation in future studies.

Keywords: Basalt fiber; Compressive strength; Hybrid models; Concrete; Machine learning.

1. INTRODUCTION

Concrete is one of the most used materials in the global construction industry and the primary contributor of CO₂ from this particular field [1–5]. In addition, production of excessive cement and other construction materials causes significant consumption of natural resources, which is a notable environmental threat [6–10]. Nowadays, the use of basalt fiber in developing regions like Najran City for the production of concrete has increased and has developed a lot in the literature because of the positive effect on the mechanical and durability properties [11–15]. Abundant research reports suggest that including basalt fibers can notably increase the strength of concrete, especially when the dosages are not large or moderate [16, 17]. This reinforcement is mainly due to the ability of the fiber to bridge micro-cracks so that their size is restricted and, as a consequence, load carrying capacity is enhanced [18–21]. Level of increase in strength is also dictated by the volume and length of the fiber used. Generally, 0.1 to 0.5% (by volume) basalt-fiber inclusion induces advantageous properties; although,

higher volume fractions lead to fiber unserviceability due to clumping and can ruin the composite's strength and workability [15, 22–24]. Therefore, the fiber proportion is vital and it should be well optimized for better performance. In addition to the tensile strength, BF has also proved to improve the toughness and ductility of the concrete [25–34]. While basalt fibers are most effective in improving tensile and flexural properties, their implementation also results in a uniform stress distribution under compression, which contributes to delaying the structural failure [35–38].

Basalt rock, a natural material formed from solidified lava, melts at the temperature of 1500–1700°C [39–42]; to fabricate basalt fiber (BF), the raw material is melted in a furnace at about 1450–1500°C. Moreover, subsequently drawn through the nozzles made of platinum and rhodium by means of technology denoted as continuous spinning [43–45]. This fiber production process was developed for the first time in the Moscow Research Institute for Glass and Plastic in 1953–1954, and the first commercial furnace was put into operation in a Ukrainian glass fiber research lab in 1985 [46–49]. Thereafter, BF has become popular in the construction amendment use on account of the superior heat, corrosion, and chemical-resistivity as a reinforcing material for concrete structures [50–52]. In addition, the application of modified supplementary materials also helps to enhance the structural integrity of concrete construction [49, 53–56].

It has been widely reported that basalt fibers improve the mechanical properties of concrete. For instance, KHARUN and KOROTEEV [57] observed that the stability and durability of HSC with BF addition were enhanced. Their experiments demonstrated an increase in compressive and tensile strengths by 37% and more than 70%, respectively, by adding 1 vol% BF to CS. Similarly, YANG *et al.* [58] determined that the optimum content of BF in normal concrete contributes to a decrease in early age crack development and improvement of toughness properties. With the increase in fiber content, the larger surface cracks are prone to producing microcracks, and the structure is improved. Additionally, an additional study by BIRADAR *et al.* [59] found a 0.3% BF mix to be most effective, leading to improvements of 9.82% in compressive strength, 36.7% in tensile strength, and 18.83% in flexural strength. The flexural and tensile strength property was improved with the chopped BF by KATKHUDA and SHATARAT [60], whereas the BF did not achieve much gain in the compressive strengths. Basalt fiber reinforced concrete is commonly used in infrastructure based on its chemical resistance and durability [61, 62]. Nevertheless, brittleness and low tensile and flexural strength present problems. Incorporation of basalt fiber has shown its advantages. Increased compressive strength and ductility in BFRC has been shown by AYUB *et al.* [63], indicating that the replacement of 2% BF volume enhanced compressive strength. Figure Similarly, MOHAMMADI MOHAGHEGH *et al.* [64] discovered that BF at level of 1.33% based on BC maximum in compressive, tensile, and flexural strengths. A further investigation by KHARUN *et al.* [65] reported the positive effect of 1.2% BF on HPC properties, suggesting that an optimal BF content for HPC improvement generally exceeds 1%.

Outside of materials applications, there is growing uptake of machine learning (ML) across industries including manufacturing, medicine and civil engineering [66–70]. For materials science studies about construction, ML techniques were successfully used in the prediction of various properties of concretes, which can be seen as an economic substitution for some expensive laboratory experiments [71–75]. For example, SU *et al.* [76] employed several ML approaches such as linear regression, SVM, and ANN to predict mechanical characteristics of P-FRC, where SVM produced better results. Similarly, NGUYEN *et al.* [77] leveraged ML to screen 339 GPC mixes and select the most promising compressive strength predictions. ULLAH *et al.* [78] used LS-SVR and RF for the prediction of the strength of lightweight foam concrete with $R^2=0.95$ for RF, which demonstrates the good prediction ability. In his work, LIU [79] also applied SVR, RF, and XG-Boost to HPC prediction, and the results also provided evidence of the high MD accuracy by ML, including an XG-Boost R^2 of 0.99774. REZA KASHYZADEH *et al.* [80] employed a genetic algorithm (GA)-optimized BPNN to estimate the concrete strength through variables including the curing temperature and aggregate properties, and obtained good agreements with experiments.

ONYELOWE *et al.* [81] carried out a comprehensive investigation on the prediction of compressive strength of basalt fiber-reinforced polymer (BFRP) reinforced concrete utilizing a series of ML algorithms such as Random Forest, M5P, M5P-Stochastic, ANN, and Random Tree-based methods. The dataset includes different variables, such as cement, water, aggregate, fly ash superplasticizer, and BFRP content based on multiple research studies. Through the comparison of different models using different evaluation criteria, Random Forest showed the best predictive probability. PAL *et al.* [82] developed an ANN model to predict the strength of GFC and BFC at 28 days, taking into account parameters such as A/C, W/C, FA content, and fiber dimensions and properties, such as the density and the elastic modulus. NAJM *et al.* [83], in their mix proportion, about 80% of cement was replaced with SCMs. Of these components, GGBS and basalt fibers followed by cement, had the most influence on strength. ANN proved to work well in the prediction due to irregularities in the data. The model of the compressive strength of BFRC is predicted by six advanced ensemble and classification ML

algorithms. It also uses the sensitivity analysis of Hoffman and Gardner to determine the influence of different mix components. The aim of the work is to develop more sustainable and accurate predictive models to inform the optimized design and use of BFRC in the construction industry [84, 85].

Basalt fiber reinforced concrete (BFRC) has attracted much attention for its improvement in compressive strength, durability, and mechanical properties of concrete. Meanwhile, the available literature indicates that adding basalt fibers into the concrete mix design refines the internal structure of the concrete, limits the propagation of cracks, and improves stress distribution. Nonetheless, the performance of BFRC is affected by factors such as fiber dosage, length, and type. Performance is usually improved with a moderate level of fiber, but when these levels are too high, clumping can occur, and the ability to lay up is hindered. Optimum fiber contents have been studied and their effects on mechanical and durability performance have been reported by previous research, indicating the utility of basalt fibers in structural concrete applications.

Notwithstanding this knowledge, a comprehension of the BFRC's behavior remains difficult and mostly goes beyond that of standard experimental testing. Machine learning (ML) has been proposed as an attractive option to model these complexities, leading to the prediction of performance metrics [86–89]. Various machine learning (ML) algorithms, such as ANN, SVR, RF and PR, have been used for the prediction of compressive, flexural and tensile strengths. These strategies are effective in simulating the non-linearity of the behavior of concrete mixes. Hybrid and multiscale ML approaches have been employed in some cases to enhance predictive accuracy. Besides ML modeling, numerical simulation and other modeling approaches have also gained significant attention to predict and evaluate important properties of concrete and other materials [90–94]. Several studies reported advanced numerical analysis that can effectively predict material and structural behavior in the construction industry [95–99].

However, one important gap remains: studies often focus only on prediction without exploring the contribution of each mix component, using sensitivity analysis. Integrated studies evaluating the effect of each variable upon the BFRC performance through a comprehensive sensitivity analysis approach are scarce. Moreover, previous studies are concentrated on a small subset of ML models, and the comparisons among different conditions lead to the challenge of finding the most robust modelling for various datasets and environments. The work attempts to fill in these gaps using machine learning and advanced sensitivity analysis methods to analyze the level of influence of individual concrete components on the value of BFRC compressive strength. The study suggests a more comprehensive and sustainable forecasting framework by combining ensemble ML models with feature importance analysis. The novelty consists of combining predictive modelling with comprehensive input analysis as a tool in the optimization of the mix design of BFRC. This methodology will provide engineers with a superior tool for designing the environmental sustainability of Najran city, high-performance, and durable concrete materials.

1.1. Research objective

The objective of this study is to design and assess new artificial intelligence models for predicting the compressive strength of BFRC. More specifically, the study aims to identify a high-performance predictive model among various machine-learning models (i.e., Gradient Boosting, Cat-Boost; Light GBM; and a mix of these models) through comparison between them. Furthermore, the study aims to improve model performance with hyperparameter tuning by employing grid search CV. An important objective is to perform the sensitivity analysis of variations in mix components upon the compressive strength of BFRC. In addition, the proposed study intends to quantify a data-driven method for optimizing BFRC design and demonstrate the robustness and generality of models via uncertainty bounds. These aims are linked to the wider scope of materials science; they include: situating how machine learning can be leveraged for improvement in concrete mix design.

2. METHODOLOGY

2.1. Methods of research

AI models are trained on the available data and calibrated to and validated by laboratory test results. For the completion of complex engineering tasks, the ability of the AI process to identify trends reduces the complexity of real data into an understandable trend. To evaluate and compare the predictive accuracy of the network-based versus tree-based decision-making systems, several machine learning methods were employed in this work. Feature combination helps improve the prediction power of the models generated and ensembles the individual ML models.

2.2. Information gathering

We searched Google Scholar, Scopus, and Web of Science, as it is the most widely used and easiest to access academic literature. Google Scholar was the primary search engine used in searching. During the study, a

few important terms were encountered, such as sustainability, low carbon footprints, basalt fiber mechanical strength, and different types of SCM. Articles published in other languages were not considered in the search strategy. After that, we narrowed down our search and review to cover only basalt fiber reinforced concrete studies, which resulted in over 200 research papers. Moreover, the abstract has been chosen and reviewed in relevance with this study. A demonstration was heard on the impact of the various cement replacements (which are called supplementary cementitious materials, SCM) on social acceptance, the economy and the ecosystem without compromising compressive strength. Fifty articles were excluded from screening based on their content and annotation, only because they included several types of SCM.

The result factor was 28-d compressive strength, whereas the influencing factors were cement content, fine aggregates, SCM, density of concrete, and amount of water and foam; age and amount of superplasticizer; this study involved 450 different test points. All other compressive strength test findings in this study are related to cylinder specimens of 100 x 200 mm. The statistical description of the data is presented in Table 1.

2.3. Preprocessing data

Data pre-processing is an essential step in a series of raw data processing, which makes raw data available for the machine learning algorithm to predict accurately. The first part of this process is data cleaning, a portion of which involves missing data. One way of doing this is by removing the records that are not full, or by imputing values, like mean or median substitution. Outliers are detected and dismissed or modified as they may influence the model. There are two common methods for doing this.

Normalization (Z-score normalization):

It scales the data between the range [0, 1] using

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

This normalizes the smallest element to 0 and the largest one to 1, keeping the relative differences.

Min-Max normalization:

Here, μ and σ are the mean and standard deviation of the population, respectively. This idea is the light when the data follows a Gaussian distribution. Encoding methods (e.g., one-hot encoding) make qualitative variables quantitative by translating them to numbers. This category-level variable becomes available in the model by creating binary columns out of each category. This data split is finally partitioned into training and testing data, typically, 80:20, or 70:30. The samples in the training data are utilized to estimate the model's parameters, and the testing data is utilized to evaluate how well the model can generalize. By doing so, preprocessing makes the data in a good shape or easier for the model to learn, so that the data processing can also be completed faster, as well as more accurately finally, the model also gets better chances to generalize.

Table 1: The statistical description of the data.

FEATURE	UNIT	MEAN	STD	MIN	25%	75%	MAX
Cement	(kg/m ³)	440.0	0.0	440.0	440.0	440.0	440.0
SCMs (Silica fume, Fly ash)	(kg/m ³)	110.0	0.0	110.0	110.0	110.0	110.0
Basalt Fiber Quantity	(kg/m ³)	0.0	0.0	0.0	0.0	0.0	0.0
Fine Aggregate	(kg/m ³)	740.0	0.0	740.0	740.0	740.0	740.0
Coarse Aggregate	(kg/m ³)	975.0	0.0	975.0	975.0	975.0	975.0
Superplasticizer	(kg/m ³)	3.52	0.0	3.52	3.52	3.52	3.52
Water Quantity	(kg/m ³)	121.0	0.0	121.0	121.0	121.0	121.0
Curing Duration (Days)	(days)	3.0	0.0	3.0	3.0	3.0	3.0
Compressive Strength	(MPa)	73.185	0.0	73.185	73.185	73.185	73.185

Scales data to have a zero mean and a variance of one with the formula

$$x' = \frac{x - \mu}{\sigma} \quad (2)$$

This study used min-max normalization techniques as for this dataset the min-max normalization showed better performance than Z-score normalization.

The Pearson correlation coefficient (r) was also calculated to assess the degree of association between the variables. Using statistical software Minitab-16, a Pearson correlation (r) matrix of the variables was obtained. A correlation of 1 means that the input and output properties are strongly related, a correlation of 0 means they're not related at all. The single sign case corresponds to positive if and negative if. The general properties of the composite materials are shown in Table 2.

2.4. Machine learning modelling

The best fit machine learning (ML) model for predicting CS of basalt fiber-based concrete was derived based on a comparison of the model metrics, and a multivariate multi-class modelling approach was performed by applying seven different models [100]. Machine learning (ML) modelling uses algorithms to enable computers to discover patterns in data and make some kind of decision or prediction without explicitly programming them. The process usually consists of collecting the data, preprocessing (cleaning and transforming) the data, selecting a machine learning algorithm (e.g., neural networks, decision trees, linear regression), training the model on a dataset, evaluating performance with metric(s) (e.g., RMSE, accuracy), and finally recommending or deploying the model where it needs to be used. Machine learning (ML) modelling is commonly used for classification, regression, and clustering applications and in industries from healthcare to finance to marketing to manufacturing. The names of the models that we are using are Categorical boosting, Gradient boosting, Light GBM, GB+CB, GB+LGB. That is, 80% of the data is used for training, and the remaining 20% is used for testing. Figure 1 is the overall flowchart of the machine learning.

2.4.1. Cat boosting

Cat Boost was developed by Yandex and is supposed to be better at working with categorical data than other more traditional tree-based models such as Light GBM [101]. We can add in additional layers of refinement to improve the model's accuracy and generalization. The design of Cat Boost relies on the GBDT based approach. In essence, the algorithm refines predictions iteratively with its update rule, after accounting for the residual of past models.

$$F_t(x) = F_{t-1}(x) + \eta \cdot h_t(x) \quad (3)$$

here $h_t(x)$ is a new decision tree fitted to the gradients, $F_t(x)$ is the new prediction, and η is the learning rate.

$$\tilde{P}_{cat} = \frac{\sum_{i=1}^{k-1} y_i + \alpha \cdot P}{k - 1 + \alpha} \quad (4)$$

Table 2: The basic properties of the composite materials.

MATERIALS	PROPERTY
Ordinary Portland Cement (OPC)	EN197-1, Type 2, 42.5 OPC, ES 4756-1/2007.
Basalt fiber	Density: 2.65–2.80 g/cm ³ , Color: brown to dark brown, sometimes golden, Diameter of Filaments: Generally, 10–20 micro meters, Moisture Absorption: Low water absorption (~0.1–0.3%)
Fine aggregate	Maximum size 4.75 mm, FM value 2.75 to 3.09.
Superplasticizers	Poly-carboxylic ether, EN 934-2 (2009).
Water	Molecular Weight: 18.015 g/mol, Appearance: Colorless, odorless, tasteless liquid, State at Room Temperature: Liquid.

where P is a prior, α is a regularization term, k is the index of the current sample in random order, and y_i are the previous target values the model has seen. Moreover, Cat Boost grows symmetric trees, which also simplifies the model, accelerates the computation and reduces overfitting by splitting data on the same criterion at each level. A categorical boosting diagram is illustrated in Figure 1.

2.4.2. Gradients boosting

Consensus decision trees by default are weak learners in gradient boosting, an ensemble machine learning, which strengthens a model by repeatedly adding weak learners to it [102, 103]. To not lose too much of the training data, we begin by fitting the model with a constant. At each iteration, it calculates the pseudo-residuals by taking the negative of the gradients of the loss and uses them to determine the predictions for the current model. To predict this new set of residuals, we now train a new weak learner. We update the model by multiplying these predictions by the learning rate. Conclusion: The algorithm formulates the model in mathematical form.

$$F_m(x) = F_{m-1}(x) + v\gamma_m h_m(x) \quad (5)$$

Here $h_m(x)$ is the new learner, v is the learning rate, and the optimal step size for loss minimization is $h_m(x)$. Repeating this process until convergence or a fixed number of iterations leads to the final prediction function, which is a linear combination of all weak learners. Gradient boosting achieves model accuracy by optimizing to correct the mistakes of previous model. The principle of the gradient boosting model presented in Figure 2.

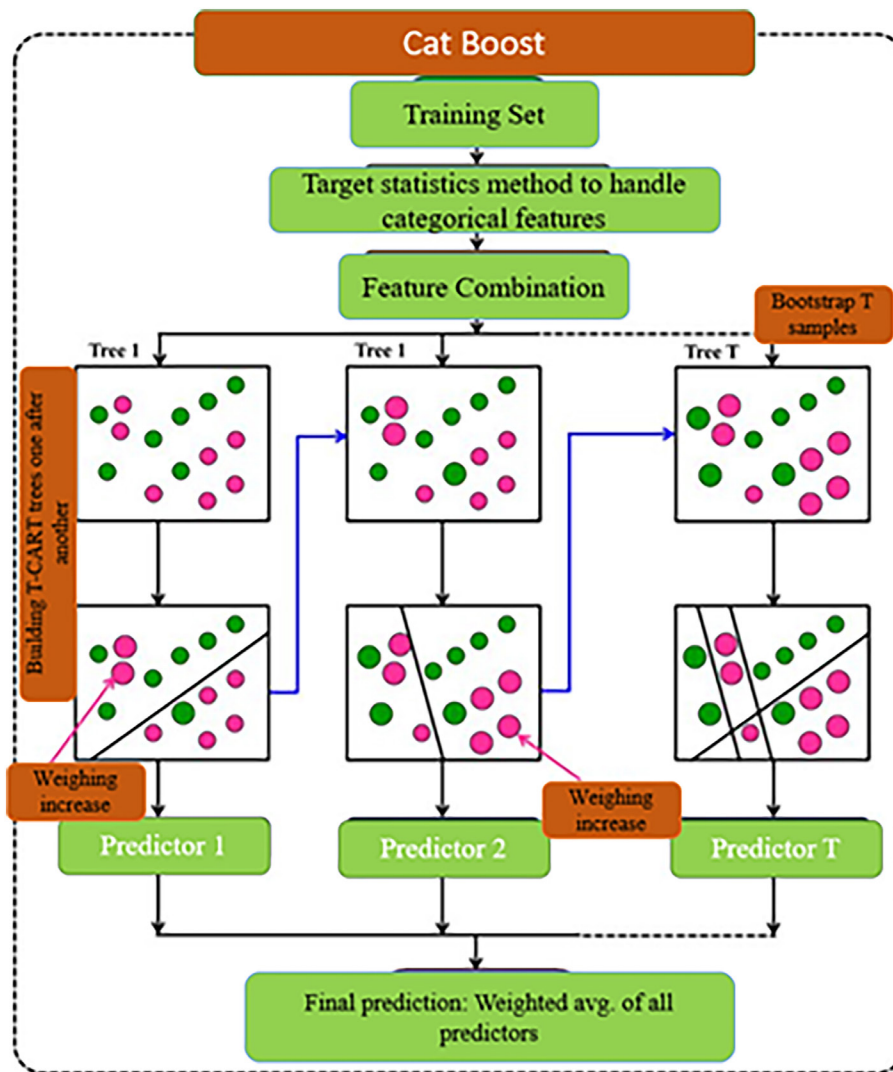


Figure 1: Categorical boosting diagram.

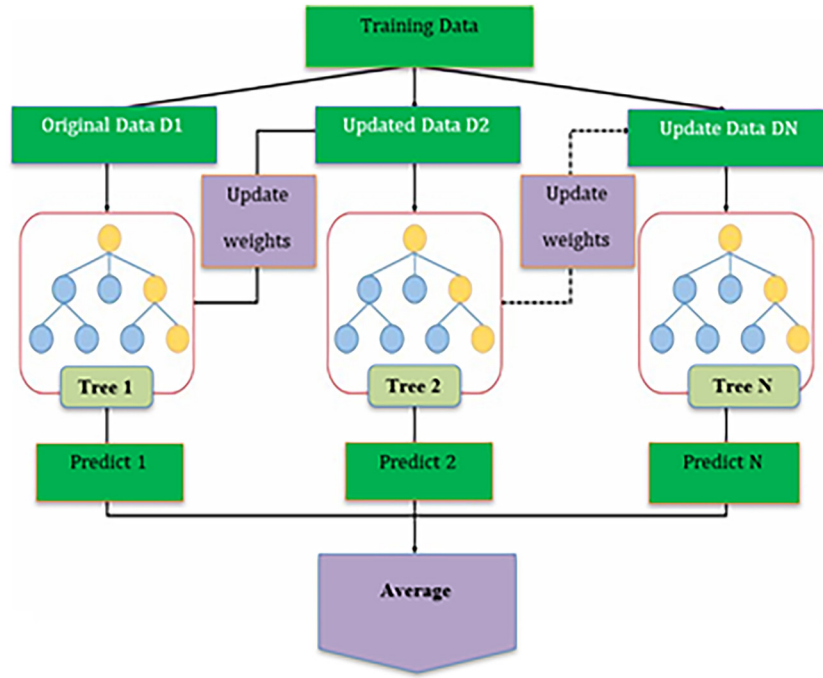


Figure 2: The principle of the gradient boosting model.

2.4.2.1. Mathematical formulation

Training data: $\{(x_i, y_i)\}_{i=1}^N$ and a differentiable loss function $L(y, F(x))$, where $F(x)$ is the current prediction function.

1. Initialization:

$$F_0(x) = \underset{\gamma}{\operatorname{arrgmin}} \sum_{i=1}^N L(y_i, \gamma) \quad (6)$$

For example, for squared error loss $L(y, F(x)) = (y - F(x))^2$, the initial prediction is the mean of y_i

$$F_0(x) = \frac{1}{N} \sum_{i=1}^N y_i \quad (7)$$

2. For each boosting iteration $m = 1, 2, \dots, M$:

a. Compute the pseudo-residuals:

$$r_{im} = - \left[\frac{\partial L(y_i, F(x))}{\partial F(x)} \right]_{F(x)=F_{m-1}(x)} \quad (8)$$

b. Fit a weak learner $h_m(x)$ to predict the residuals r_{im}

c. Compute the optimal step size γ_m

$$\gamma_m = \underset{\gamma}{\operatorname{arrgmin}} \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + \gamma h_m(x_i)) \quad (9)$$

d. Update the model

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \quad (10)$$

3. Final prediction after M iteration

$$\hat{y} = F_M(x) \quad (11)$$

2.4.3. Light GBM

Light GBM, meaning Light Gradient Boosting Machine, by Microsoft, is a gradient boosting framework for fast and high-performance efficiency and scalability [104]. Light GBM is based on decision tree algorithms in a supervised manner for classifications, regression, and ranking tasks. Light GBM is unique in that it uses leaf-wise (rather than level-wise) tree growth strategy, allowing it to reduce more loss and produce better accuracy. It supports parallel learning and GPU learning, making it possible to train fast on large scales [105, 106]. Light GBM builds trees by growing leaves vertically, as compared to horizontal construction by levels, and at each split, chooses the leaves that have the maximum gain. Light GBM applies two critical techniques: GOSS (Gradient-based One-Side Sampling) and EFB (Exclusive Feature Bundling). GOSS retains those data points which have a large gradient, while EFB reduces feature dimensions by bundling mutually exclusive features. GOSS and EFB improved training time and memory use, without sacrificing accuracy. The principle of the Light GBM model is illustrated in Figure 3.

2.4.4. GB+CB

Gradient boosting + Categorical boosting is a hybrid machine learning framework that combines the strength of gradient boosting with modern approaches to handling categorical data efficiently [107]. Gradient boosting is an ensemble approach that builds trees iteratively, which maximize prediction accuracy by addressing the model's prediction errors at each step. This generally has the downside of requiring preprocessing to transform categorical variables into numerical data. Categorical boosting will augment gradient boosting to use categorical features directly in an iterative framework through ordered boosting (to prevent overfitting) and target encoding (to give categorical variables a more meaningful numerical representation). The combination of these capabilities means that the model can effectively mitigate overfitting while reducing time to build more relevant predictive models. Hybrid approaches to machine learning, such as this one, are promising, particularly on datasets where the majority of the features are categorical, and may offer better performance faster on modelling real-world problems in areas that include finance, healthcare, or marketing [108, 109]. This encoder averages the target for the category taking into account only previous data to avoid overfitting (ordered boosting). Having a model that

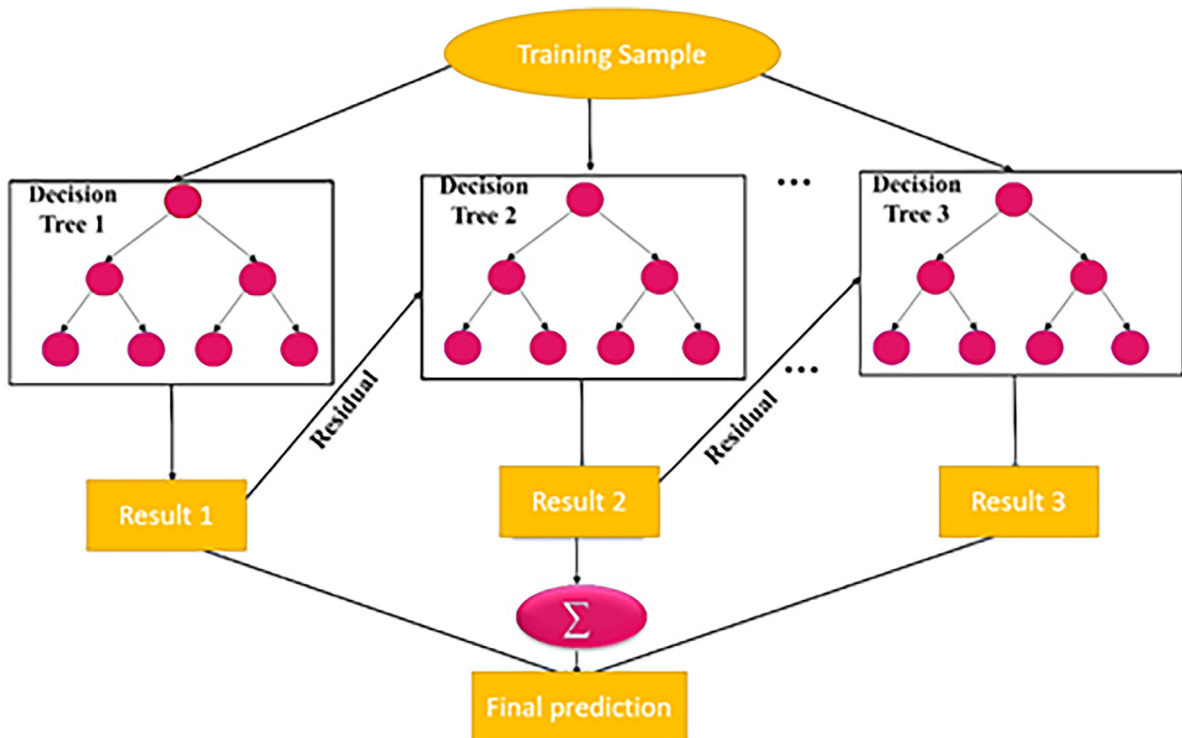


Figure 3: The principle of the light GBM model.

uses both methods, however, it can make reasonable decisions in both cases and hence can give us both numbers and relations at the same time.

$$\check{y}(x) = \sum_{m=1}^M \gamma_m h_m(x) \quad (12)$$

2.4.5. GB+LGB

Hybrid Learning GB+LB Hybrid learning based on gradient boosting + light gradient boosting which extends the traditional gradient boosting, and it is more efficient and scalable [110]. Gradient boosting proceeds by adding decision trees (which minimize the chosen loss function, using negative gradient descent) sequentially. Light GBM is a highly efficient implementation that further enhances the process using a leaf-wise tree-expanding rule instead of the level-wise approach [111, 112]. This method chooses the leaf that has the most information gain as the split point, enabling deeper and more concentrated trees and faster convergence. Figure 4 shows overall ML flowchart.

2.5. Parametric analysis

2.5.1. SHapley additive explanations (SHAP)

The SHAP framework offers a light to understand the behavior of a machine learning model by computing positive or negative values for each feature with regard to the prediction. It works such that SHAP values are computed to provide a unique, optimal interpretability of the outcome of a machine learning model that are grounded in coalition game theory [113]. The effect of a factor on the predictions are estimated by these numbers, taking into account not only the market share of all potential combinations, but also how these features interact [114]. SHAP, the most popular instrument in the field, is a highly robust software system that can explain the output of any machine learning model, deliver estimates on the relative importance of the minimal difference of each predictor in a prediction model, and assist in modeling feature prominence and importance [115]. The SHAP value of feature *iii* in a model prediction can be expressed mathematically as:

2.5.2. Partial dependency plot (PDP)

One of the tools that we use to understand the effect of a single property, or set of properties, on model prediction is the partial dependency plot (PDP) [116]. They can also be used to show the complexity or simplicity

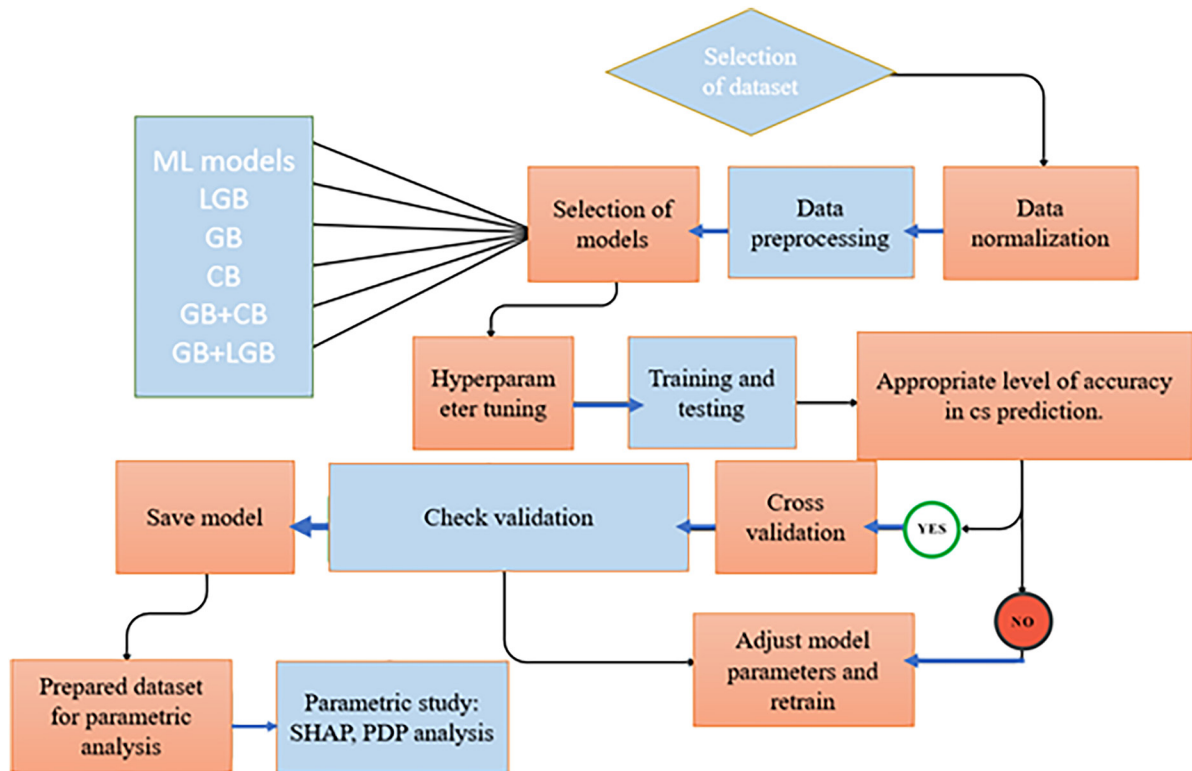


Figure 4: Overall ML flowchart.

of the interactions between an output parameter and the independent variables [117]. Since PDP is computed independently of the other factors, it gives a focused evaluation of the importance of a feature on predictions. By varying the value of one input the so-called one-dimensional PDP (or PDP) can be computed with fixed values of all the other inputs. Similarly, from two inputs, a 2D PDP can be computed [118]. PDP not the PD Plot The mean prediction of the model is calculated from all observations using the PDP technique with one predictor held at a fixed value. The predicted average response based on all data points is made with all of the other predictor variables being held fixed. a mean predicted value for each observation in the dataset, typically adjusted for each other predictor variable if any.

2.6. Evaluation of the performance of the ML model

The reliability of the prediction performance model was analyzed by four statistical methods over three subgroups (training set, testing set, and validation set). The statistical performance indicators that were used were root mean square error (RMSE), mean absolute error (MAE), MedAE and coefficient of determination (R^2). The average error of observed and predicted values was used to perform comparison using RMSE criterion and MAPE with percentage error.

$$R^2(y, \hat{y}) = 1 - \left(\sum_{i=1}^l (y_i - \hat{y}_i)^2 / \sum_{i=1}^l (y_i - \bar{y})^2 \right) \quad (13)$$

$$\text{MAE} = \frac{100}{n} \sum_{i=1}^n \frac{|a_i - \hat{a}_i|}{|\hat{a}_i|} \quad (14)$$

$$\text{RMSE} = (y, \hat{y}) = \sqrt{\frac{1}{l} \sum_{i=1}^l (y_i - \hat{y}_i)^2} \quad (15)$$

$$\text{MedAE} = \text{median}(|y_i - \hat{y}_i|) \quad (16)$$

3. RESULTS AND DISCUSSION

3.1. Heatmap visualization of compressive strength's relationship with various mix proportions, and curing conditions in basalt fiber reinforced concrete

The Figure 5 shows the influence of compressive strength (MPa) for various concrete mixtures and cure regimes. Figure 5(a) heatmap suggests that the quantity of cement is strongly positively correlated with compressive strength, especially if the amount of cement in the mix ranges from 400 to 600 kg/m³. Compressive strengths greater than 120 MPa are obtained, clearly indicating that cement is the main parameter contributing to high compression strengths. Figure 5(b) graph indicates that when the amount of SCMs is moderate, the compressive strength is the highest. Too much or too little SCMs lead to a decreased compressive strength, which indicates there might be an appropriate level of SCMs for strength development. Figure 5(c) demonstrates that compressive strength is less affected by basalt fiber quantity. There is moderate filling up to around 50–100 kg/m³ of basalt fiber, while the effect of basalt fiber is somewhat lower than cement or SCMs on strength. Figure 5(d) shows that fine aggregate has a positive correlation with the compressive strength. The maximum strength is achieved at ~700–800 kg/m³. But too much fine aggregate hurts compressive strength, and there exists an optimal content. Figure 5(e) shows that the correlation between these variables is positive. The maximum CS (close to 130 MPa) is achieved at the coarse aggregate content of approximately 900–1300 kg/m³. It indicates that a reasonable coarse aggregate content plays a key role in increasing compressive strength. Figure 5(f) shows that the reduction of compressive strength is influenced by the dosage of superplasticizer especially when its amount is within 2 and 6 kg/m³. At higher dosage (over 6 kg/m³), the strength reduction can be explained by the overdose or non-optimal mix conditions. Figure 5(g) illustrates that as the water-cement ratio increases (above 250 kg/m³), a considerable decrease in compressive strength is realized. The highest strengths are achieved with intermediate water contents (150–250 kg/m³), indicating that water must be carefully dosed. Figure 5(h) demonstrates a more complicated suggestion, where the greatest strength is had at 7–28 days of curing. The compressive strength build-up also gets reduced after certain curing times, and higher curing times do not

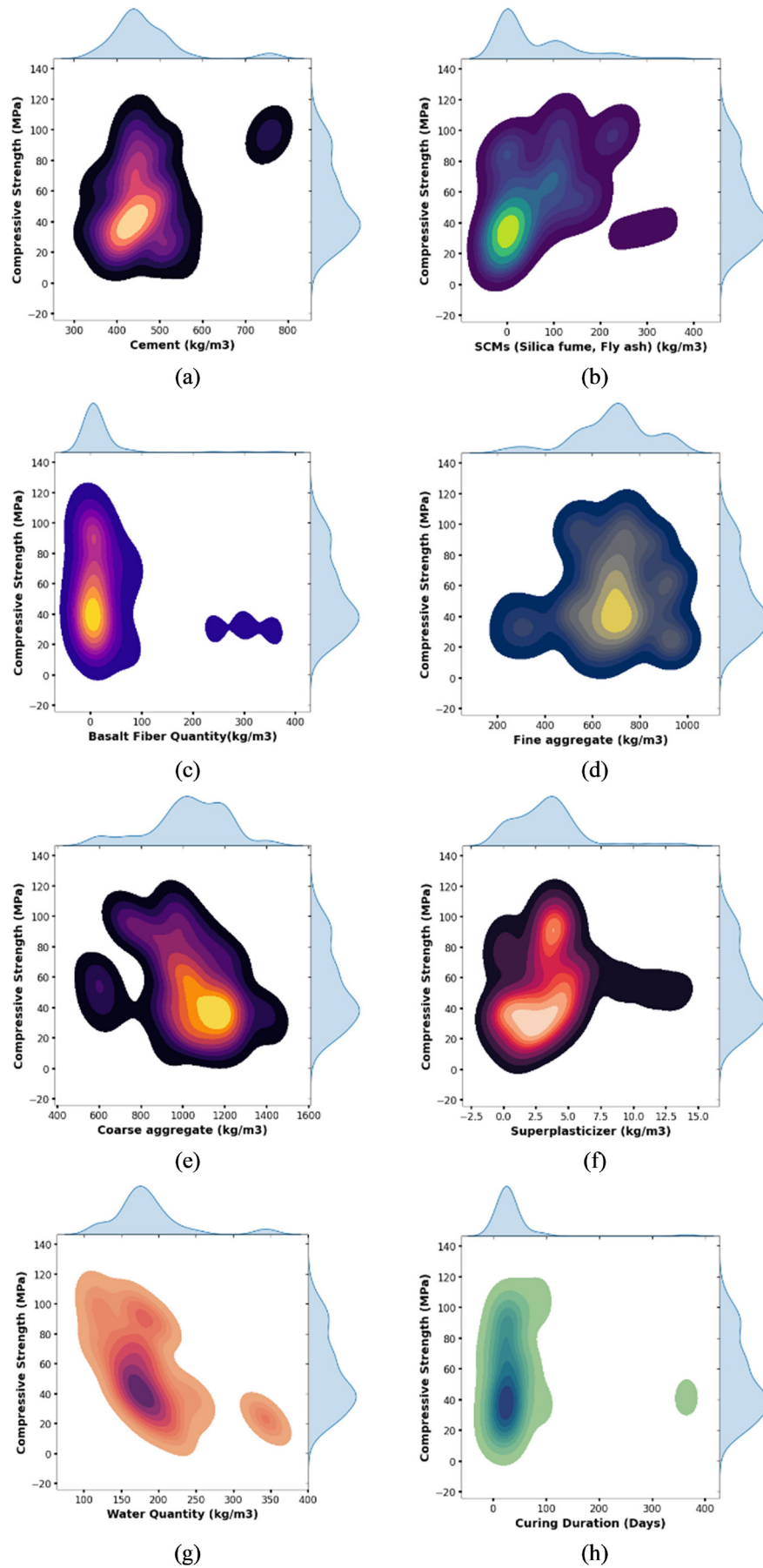


Figure 5: Heatmap analysis of compressive strength factors.

contribute to higher strength. According to the heatmaps, cement content (Figure 5(a)) and coarse aggregate (Figure 5(e)) are therefore considered to be the main properties influencing compressive strength. The content of cement is directly related to the strength, and the best range is 400–600 kg/m³. The influence of coarse aggregate amount is also positive on strength, when it is within the range of 900–1300 kg/m³, it gives the best results. The ideal model to obtain the highest compressive strength should prioritize a balance in these two aspects, namely, keeping the cement content in a proper range and the right proportions of coarse aggregates. While the superplasticizer and water content also have effects, these are relatively insignificant when compared to cement and coarse aggregate.

3.2. Correlation matrix and hierarchical clustering of concrete mix components and their impact on compressive strength in basalt fiber reinforced concrete

Figure 6 includes two main parts: one is a Correlation Matrix (Figure 6(a)) and a hierarchical clustering matrix (Figure 6(b)), all illustrating the links between different concrete mix proportions and compressive strength. This is the matrix of Pearson correlation coefficients between different concrete mix ingredients and the compressive strength (MPa). Cement has a moderate correlation with compressive strength as well (0.33), implying that the higher the cement content, the higher the compressive strength is, but not too high. SCMs have an R-value of 0.39 with compressive strength. Strength improves with the presence of SCMs, but to a weak extent. Basalt Fiber has a weak negative relationship of -0.19 , indicating that the higher the basalt fiber, the less the effect on compressive strength by Basalt fiber. Fine Aggregate has a negative correlation of -0.48 with compressive strength, indicating that when the ratio of fine aggregate increases, it may be possible to reduce compressive strength. For coarse aggregate, 0.49, which is a strong correlation towards compressive strength, suggests that a higher amount of coarse aggregate generally results in a higher compressive strength. Superplasticizer, it shows a positive correlation of 0.26 with compressive strength; however, the increase in compressive strength with an increase in superplasticizer is found to be mild. Water Content is negatively related to compressive strength, -0.56 , meaning that more water can reduce compressive strength. Curing Time has almost no correlation (0.02) in the dataset 6.2.

Fine aggregate, water quantity have a negative high correlation (-0.56 and -0.49 , respectively) with compressive strength and correspond to the next highest in cluster size. Cement and SCMs have a strong significant correlation. The relationship is close, which means the two closely relate to each other, while the sign of the relationship is positive, which means that they both have similar effects towards strength. From Figure 6(a), the main factors affecting the CS increase are coarse aggregate (0.49) and cement (0.33). Although the effect of cement is relatively modest, coarse aggregate seems to have the strongest positive relationship, which indicates that it can be the dominant factor contributing to the strength. The ideal mix for the best compressive strength with permeable pavers will, therefore, be the one in which the right quantity of the coarse aggregate (Figure 6(b)) where the same as for coarse aggregate clusters with strength improving factors, with focus allocated to

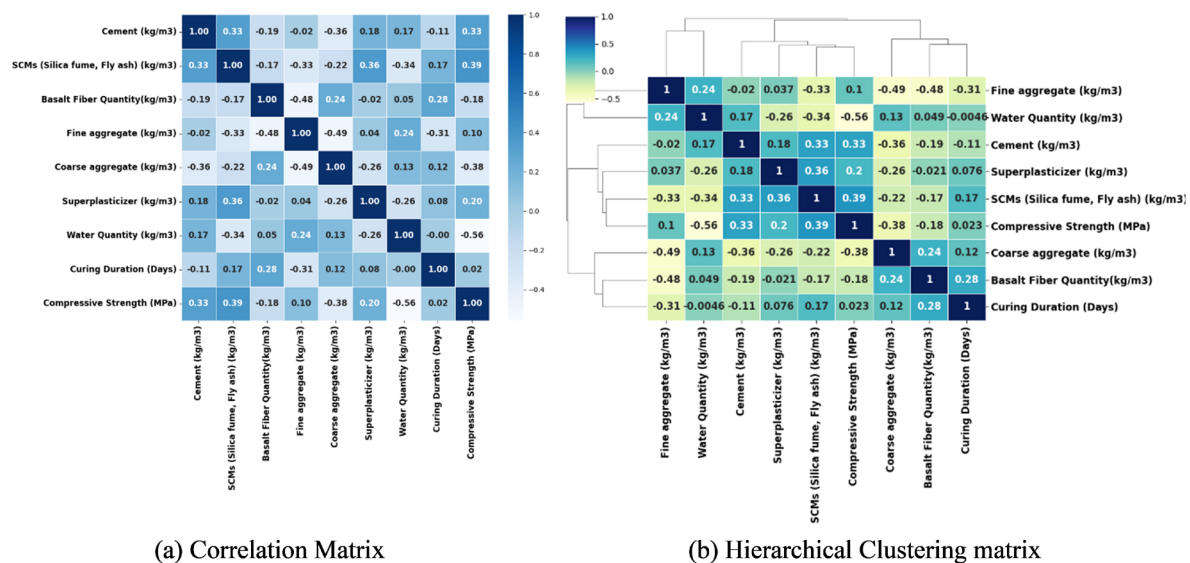


Figure 6: Correlation matrix of concrete mix proportions and compressive strength. (a) Correlation matrix; (b) Hierarchical clustering matrix.

cement content as well, even though it is not as powerful as coarse aggregate. Thus, the relative importance of compressive strength of coarse aggregate and cement is high in this dataset, whereas water amount and basalt fiber have lower than positive or negative effects. The clustering matrix is shown in Figure 6(b), which also integrates the relations and the clusters between these variables.

3.3. Correlation between compressive strength and concrete mix variables

The results presented examine the influence of the different materials and parameters on concrete strength. Each diagram demonstrates how varying a single factor be it a material component or step influences the concrete's strength.

In Figure 7(a), the effect of cement content upon compressive strength is positive, and an increase in the cement content within the range from 350 to 750 kg/m³ results in higher strength, falling in general from a value close to 40 MPa to greater than 100 MPa. This means that cement strongly contributes to increasing the strength of concrete and the more cement there is the less impact it will have salt. Similarly, in Figure 7(b), the compressive strength increased slightly with increasing SCMs, such as silica fume or fly ash although less obvious than that of cement. Conversely, Figure 7(c) Clear evidence of basalt fiber detrimental influence to concrete strength. The results showed that a decrease in compressive strength with increasing in quantity of basalt fiber is possibly attributed to break of structural cohesiveness or brittleness of the concrete. Figure 7(d) showing percentage of fine aggregate content. A previous positive trend is observed when the amount of fine aggregate is increased, that is the compressive strength and the fine aggregate increases, this could be due to the fact that fine aggregates contribute in shaping more denser matrix in the concrete. Figure 7(e) indicates the decrease in compressive strength with an increase of coarse aggregate content that could have been affected by the interference of big size aggregates on the bonding between aggregates and cement paste or increase the quantity of air voids. It is also consistent with the observations in several of the investigations mentioned above that coarse aggregates can adversely influence concrete strength by insufficient balance. On the other hand, Figure 7(f) shows that there is a good and significant relationship between the superplasticizer content and compressive strength, and the compressive strength increases from 20 MPa to approximately 85 MPa when the dosage of superplasticizer increases from 0 to 14 kg/m³. Superplasticizers contribute to water reduction together with workability retention, and they improve the interparticle bonding and enhance the strength of the concrete.

In Figure 7(g), the water content significantly negatively contributes and the increase decreases the compressive strength. This is to be expected as water in excess will dilute the cement paste and will facilitate the reduction of the overall strength of the concrete each time the original fewer are exceeded. Lastly, Figure 7(h), the graph of the curing time, we note that compressive strength does not change very much ($\approx 40\text{--}60$ MPa) hence indicating that over a certain curing time frame, further providing of the same would not derive a large percentage improvement in strength.

The most sensitive variable of strength improvement for the concrete is, in general, Superplasticizer (Figure 7f) and, as a consequence, an appreciable strength enhancement is achieved. Cement and fine aggregate, as well as water and basalt fiber, also share positive but weaker relationships. Curing time appears to have little-to-no effect on strength in this data set, possibly with the exception of 1 day.

3.4. Impact of material and mix variables on concrete compressive strength

In Figure 8(a), for Silica Fume (Fly Ash) and compressive strength (MPa) of concrete. There's quite a bit of spread in the plotted data points, but there seems to be a general trend where increased amounts of silica fume/fly ash seem to be associated with decreased compressive strength, somewhere between 20 and 100 MPa. The hex bins show that the data points are heavily focused around lower amounts of silica fume, suggesting that adding the material does not have much impact on the strength of the mix. In Figure 8(b), the cement content and its impact on compressive strength. The plot indicates a positive correlation in that the more cement is used, the higher the compressive strength. So, the individual points vary from 20MPa to 100MPa, and the hex bins show that most points kind of congregate around the 400–600 kg/m³ region of cement content, indicating that more cement in a mix leads to stronger concrete.

Figure 8(c) comparison between basalt fiber content and compressive strength. The data shown here varies significantly, but in general, basalt fibers hurt compressive strength, and more of the fibers have led to lower compressive strength, especially to values in the range of 20–50 MPa. The hex bins demonstrate a clustering of lower-valued points and low compressive strength that does not reach 80 MPa even at higher fiber contents. Figure 8(d) shows the influence amount of fine aggregate on strength. The relationship is ambiguous because in some data points the molecular weights with a low percentage of fine aggregate are high, whereas those with a high percentage of fine aggregate are low. The hex bins are distributed over a wide range of fine aggregate quantities, and a clear positive or negative trend is not observed.

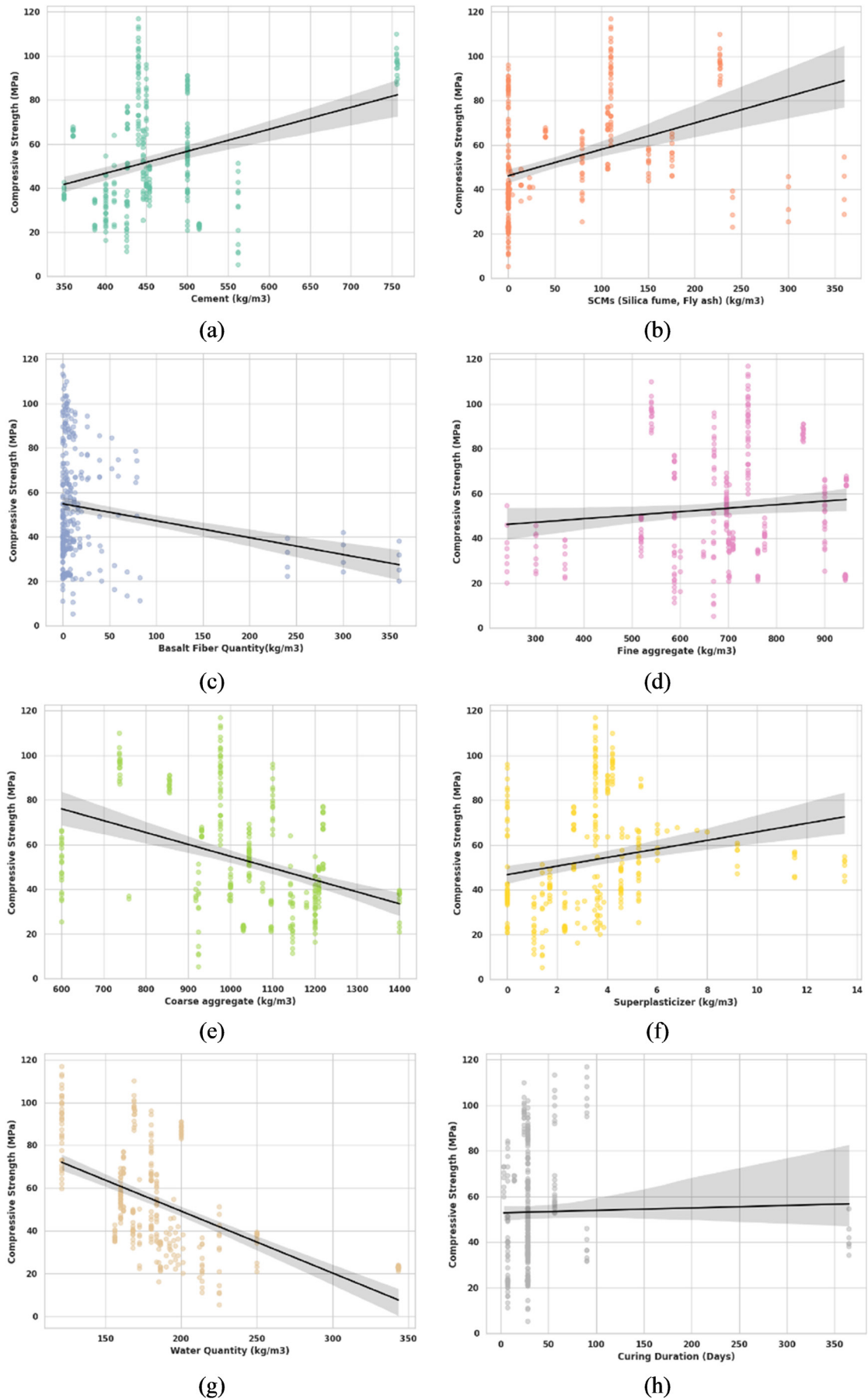


Figure 7: Effect of key factors on compressive strength of concrete.

In Figure 8(e), the coarse aggregate is analyzed, finding a slight negative association. For the strength properties, the co-combustible concrete with a high content of CA normally has the lower compressive strengths between 20 MPa and 100 MPa. The hex bins are hotspots in the low coarse aggregate range ($600\text{--}1100\text{ kg/m}^3$), indicating that an excessive coarse aggregate content is detrimental to the concrete composition. Figure 8(f) considers the effect of superplasticizer on the concrete mix. In this case, the positive correlation can be inferred, since both examined superplasticizers induce higher compressive strengths with a higher amount of superplasticizer. The data points scatter over 20 MPa to 100 MPa, while most hex bins mainly group even-distributed data points, suggesting that different dosages of superplasticizer have a positive effect on compressive strength. Figure 8(g) investigates the amount of water, and it is clear that there is a negative relationship between them on compressive strength. It can also be seen that with the increase of the water content under investigation, the number of data points in the hex bins for lower compressive strengths (lower than 60 MPa) is concentrated. This is logical since it is well known that too much water destroys the concrete.

Finally, Figure 8(h) examines cure time and shows a yielding positive relationship with compressive strength. Compressive strength slightly raises with curing age, yet data points indicate an extent of scatter. The hex bins reveal a high density within the 0–50 days domain and for most of the points, the compressive strength is between 30–80 MPa. There is a positive but declining trend with the increase of curing time, strengthening the compressive strength. According to these results, Figure 8(b), cement content seems to be the most significant factor for increasing compressive strength. The direct dependence of the cement consumption on strength shows that an increase in the amount of cement content leads to better strength in a more or less uniform manner, as other factors contribute to it. This makes the cement content a critical factor in the design of concrete mixtures for higher strength.

3.5. Influence of material constituents on the compressive strength of concrete

The box plots in Figure 9(a–h) show the effects of different mix design parameters on compressive strength (MPa) of concrete. Each subplot corresponds to a given input parameter, discretized in intervals with their respective compressive strength distributions. In Figure 9(a), there is a pattern of higher strength for higher cement content, where the $513.5\text{--}735.5\text{ kg/m}^3$ range has the highest median, above 90 MPa (a confirmation of the importance of cement for form development) in average compressive strength. Figure 9(b) indicates that ($103.5\text{--}130.0\text{ kg/m}^3$) of SCMs also provide high median compressive strength, at about 100 MPa, which indicates ideal SCM dosage increases strength. In Figure 9(c), the issues with the basalt fiber amount ($2.75\text{--}4.25\text{ kg/m}^3$) result in the best performance (which has a high median strength at about 90 MPa), indicating that the fibers improve the crack resistance and strength. Figure 9(d), the fine aggregate content, which is $790.0\text{--}870.0\text{ kg/m}^3$, provides the highest median compressive strength up to 100 MPa, and the higher and lower content cannot achieve such a performance. Figure 9(e) shows that ≥ 4 coarse aggregates exhibited a reduction of compressive strength, the case of the superlative performance being $699.99\text{--}737.26\text{ kg/m}^3$, with the median greater than 85 MPa. Figure 9(f), it is added that the workability and the compaction are significantly improved when superplasticizer dosage ($36.6\text{--}42.1\text{ kg/m}^3$) is increased in the ranges when the strength is over 100 MPa. In Figure 9(g), water of the lowest representative range ($120.99\text{--}160.0\text{ kg/m}^3$ (median)) yields the highest compressive strength (100 MPa), proving the principle of the water-cement ratio. Lastly, Figure 9(h), longer curing time greatly promotes the strength, for (96.0–365.0 days), the quantitative median strength was more than 90 MPa.

Overall, Figure 9(f) is the best-prediction performance which can be achieved in maximum and uniformity of compressive strength, followed by cement content (Figure 9a) and curing time (Figure 9h). This demonstrates that the correct use of superplasticizers, thus allowing best materials hydration and compaction with no surplus of water, is the single parameter that leads to the most significant increase in concrete strength.

3.6. Effect of material composition on compressive strength of concrete: cement, scms, basalt fiber, and fine aggregate

The violin plots in Figure 10(a–h) represent scatter diagrams of separating parameters with compressive strength value of concrete mixes. Starting with Figure 10(a), the strength of the CFRPs is directly proportional to the cement dosage, i.e., mixtures with $512.6\text{--}586.6\text{ kg/m}^3$ of cement afford the maximum strength values, which can even exceed 140 MPa, so cement is the dominant factor. In Figure 10(b), show similar high strength potential for SCMs of $172.8\text{--}246.8\text{ kg/m}^3$ (135 MPa), moderate strength for lower SCM contents. Figure 10(c) demonstrates that the basalt fiber amount of $1.1\text{--}2.7\text{ kg/m}^3$ corresponds to higher CS (120–130 MPa), although the scatter is more distributed as compared to cement. In Figure 10(d), maximum strengths ($\sim 130\text{ MPa}$) are obtained for fine aggregate contents of $1778.5\text{--}2030.0\text{ kg/m}^3$, while lower and mid-area contents result in dramatically decreasing strength.

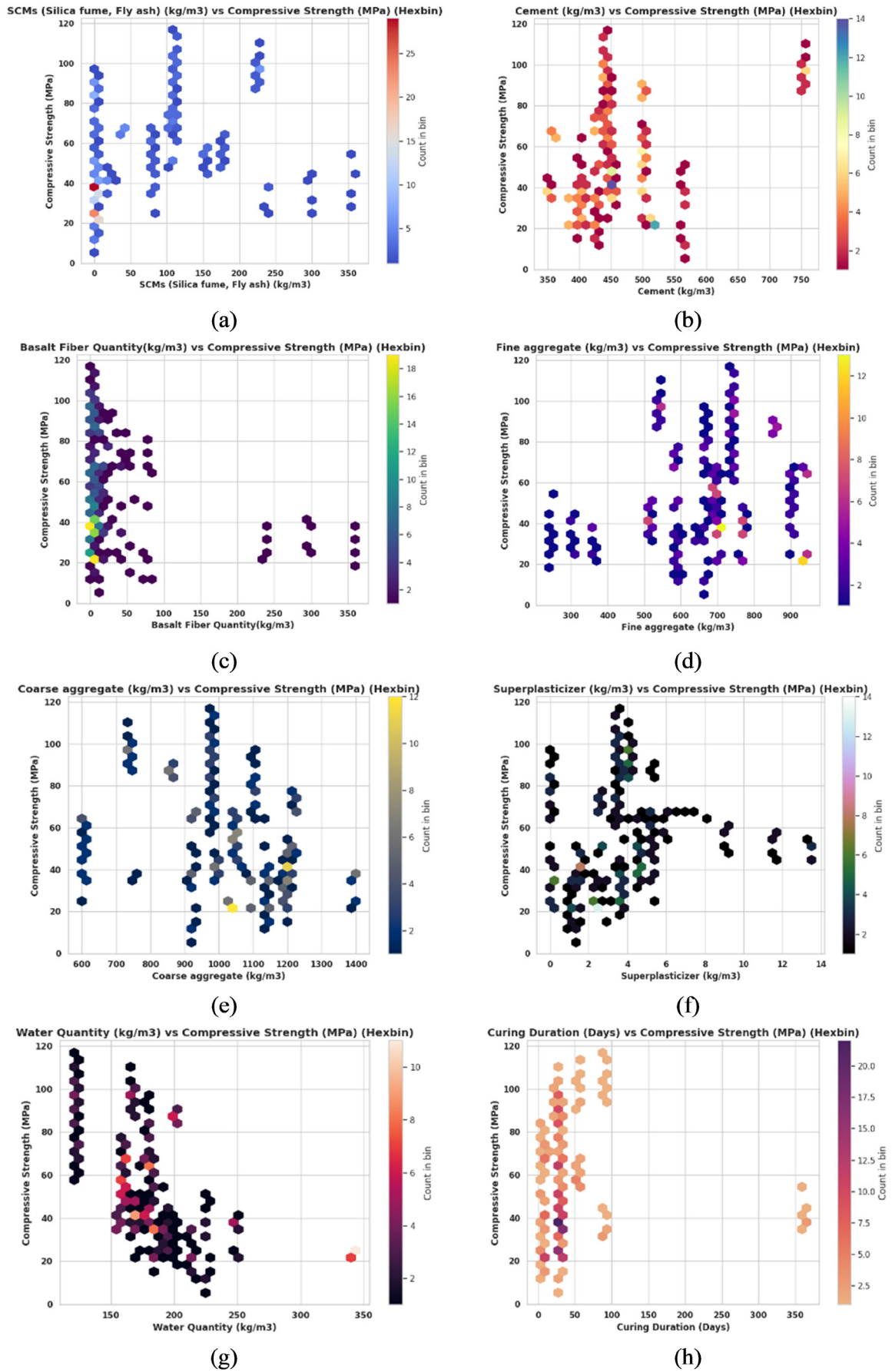


Figure 8: Hex bin analysis of concrete ingredients and their effect on strength.

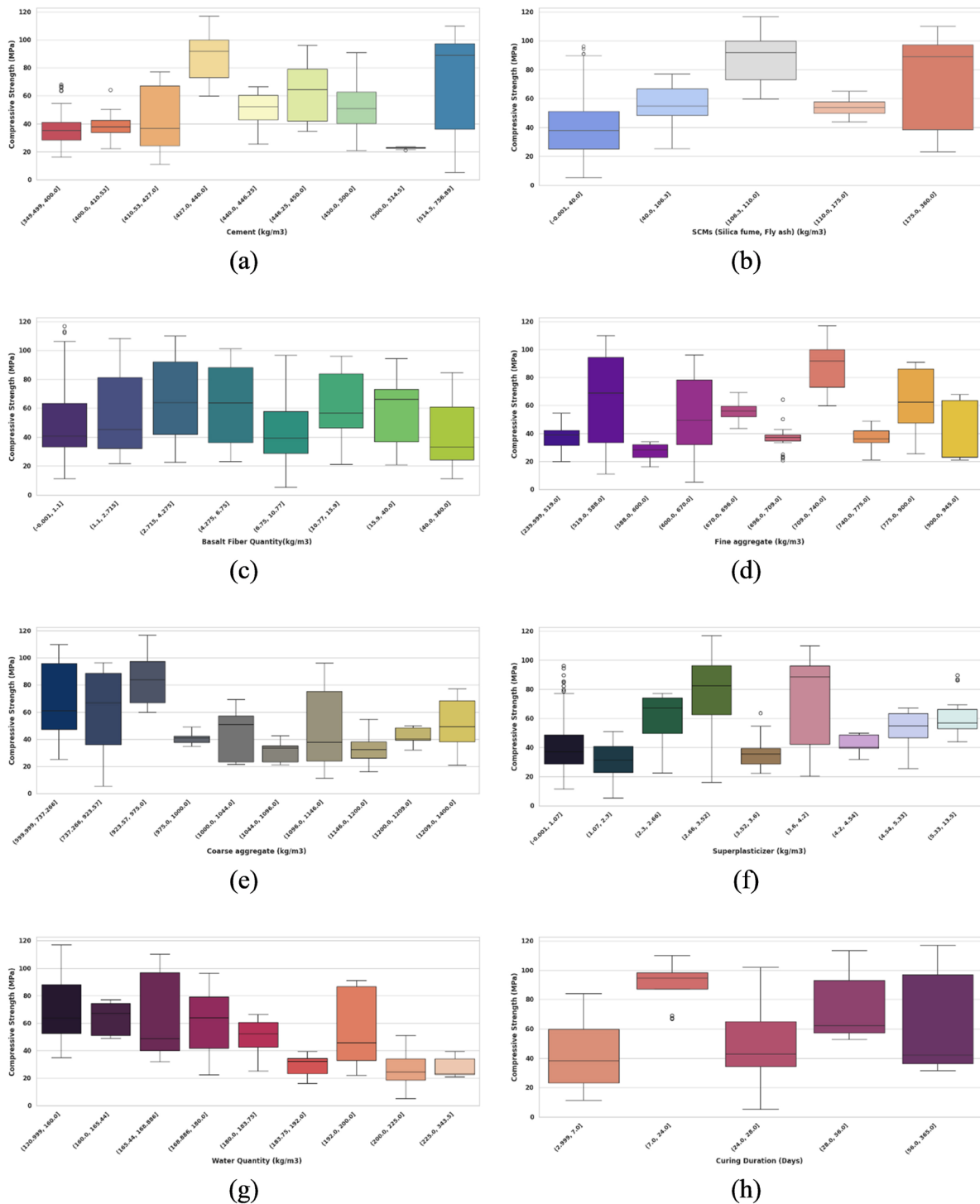


Figure 9: Box plot analysis of key mix design parameters affecting concrete strength.

Moving to Figure 10(e) also shows that with respect to water amount, the higher the water content, the lower the strength, when water content is within 139.0–160.0 kg/m³, the compressive strength of the mixture is higher than 120 MPa, but when the water content is 220.0–240.0 kg/m³, the strength of the mix is reduced to lower than 50 MPa. Figure 10(f) that curing duration has an important effect on strength: duration of 17.0–30.0 days and 46.0–90.0 days result in strength >130 MPa, which again highlights the crucial role of long-term curing. In Figure 10(g), the largest strength value (~130 MPa) is achieved with coarse aggregate contents of 973.6 to 1097.3 kg/m³, and the strength decreases with excessive aggregate content. Lastly, Figure 10(h) indicates that dosages of Superplasticizer ranging from 2.8–4.3 kg/m³ reach compressive strength values between 130–140 MPa, which are decreased for lower or for too high quantities. The cement content in Figure 10, where factor (a) is identified as the most important factor governing the behavior of compressive strength and the

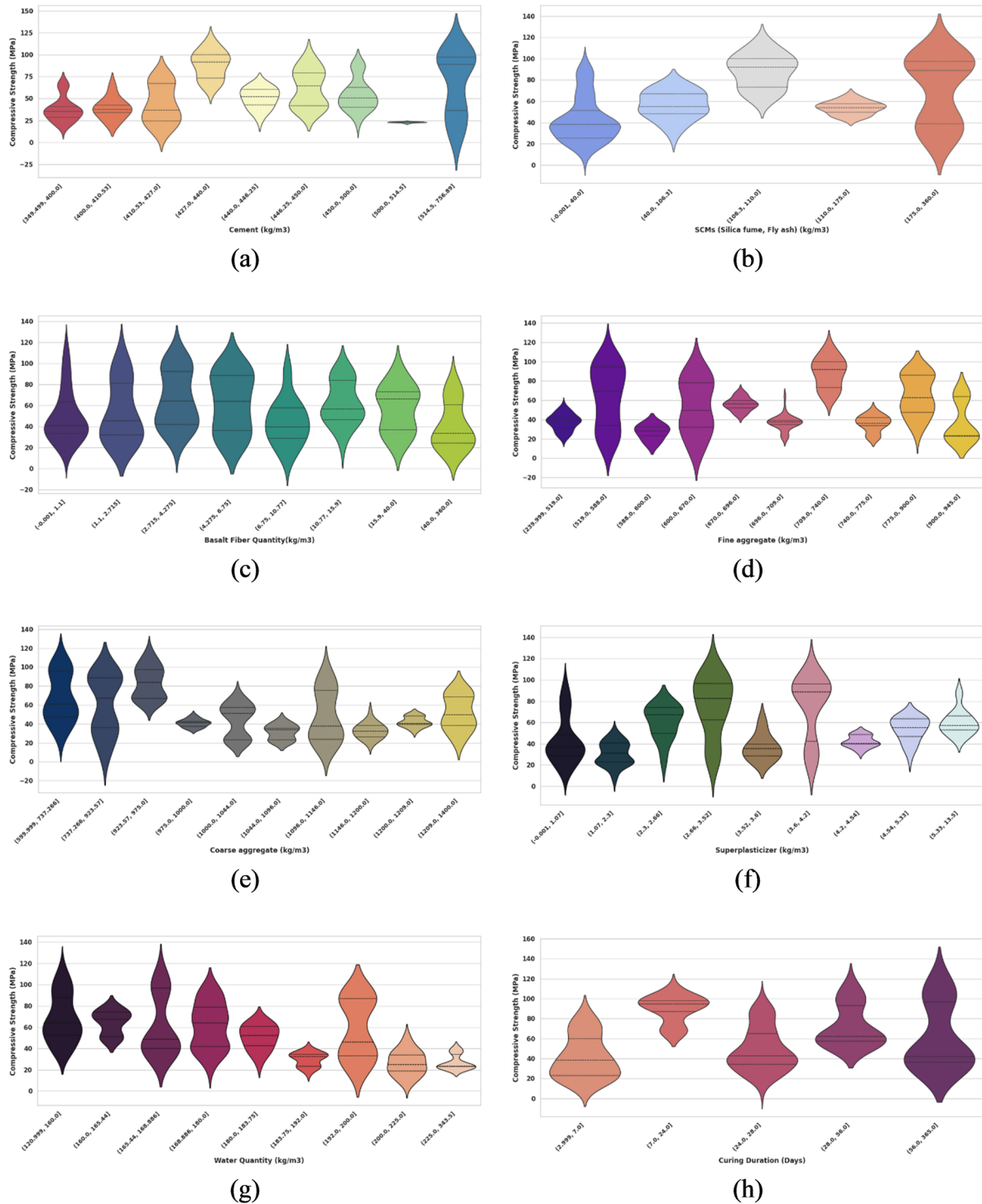


Figure 10: Impact of mix parameters on compressive strength: coarse aggregate, superplasticizer, water content, and curing duration.

maximum and most uniform strength values are obtained at high cement contents. This is expected as it is directly involved in the hydration of concrete and its matrix formation, which dictates the compressive performance of concrete.

3.7. Visualization of pairwise relationships between concrete mix features and compressive strength

The Figure 11 shows a pair plot which shows the association between different numerical variables with the Compressive Strength for the dataset. Each plot displays pairwise interactions between two features, as the diagonal histograms display each feature's distribution. The variables Cement and Coarse aggregate reveal a stronger linear association with Compressive Strength, which implies that they could be important predictors.

Other characters as Superplasticizer and Water, show more scattered data and less relation with compressive strength. Curing Period (Days) also appears to be correlated with compressive strength, but with a scattered behavior. This pair plot is useful for finding possible strong correlations and weak correlations between features and compressive strength, providing an interesting silhouette through which to proceed to further analysis and modelling.

3.8. Comparison of actual vs predicted concrete strength for different models

The Figure 12 shows Actual vs Predicted Compressive Strength for different models like Gradient Boosting, Cat Boost, Light GBM and ensemble models for predicting Concrete Compressive Strength. In Figure 12(a) (Gradient Boosting), in the training set R^2 is 0.9696, and in the test set it is 0.9437. The results demonstrate a very nice agreement between the true and predicted values, and most points closely follow the dashed line, which represents the ideal fitting. But there are some exceptions at the higher compressive strengths, especially in the test set, where the predicted values are slightly underestimated, and the errors reach up to 10 MPa.

Figure 12(b) (Cat Boost) follows the same pattern, the R^2 values for the training and test sets being, respectively, 0.9654 and 0.9367. Both for the training and the test set, the predicted values are in good agreement with the actual values, but the deviations are higher for the test set, which in large part displayed errors of up to around 10–15 MPa. Figure 12(c) (Light GBM) is R^2 is 0.9615 in training and 0.9194 in testing. The plot looks well-centered, but the residuals in the test set are a bit more spread, and we reach errors up and down to 15–20 MPa.

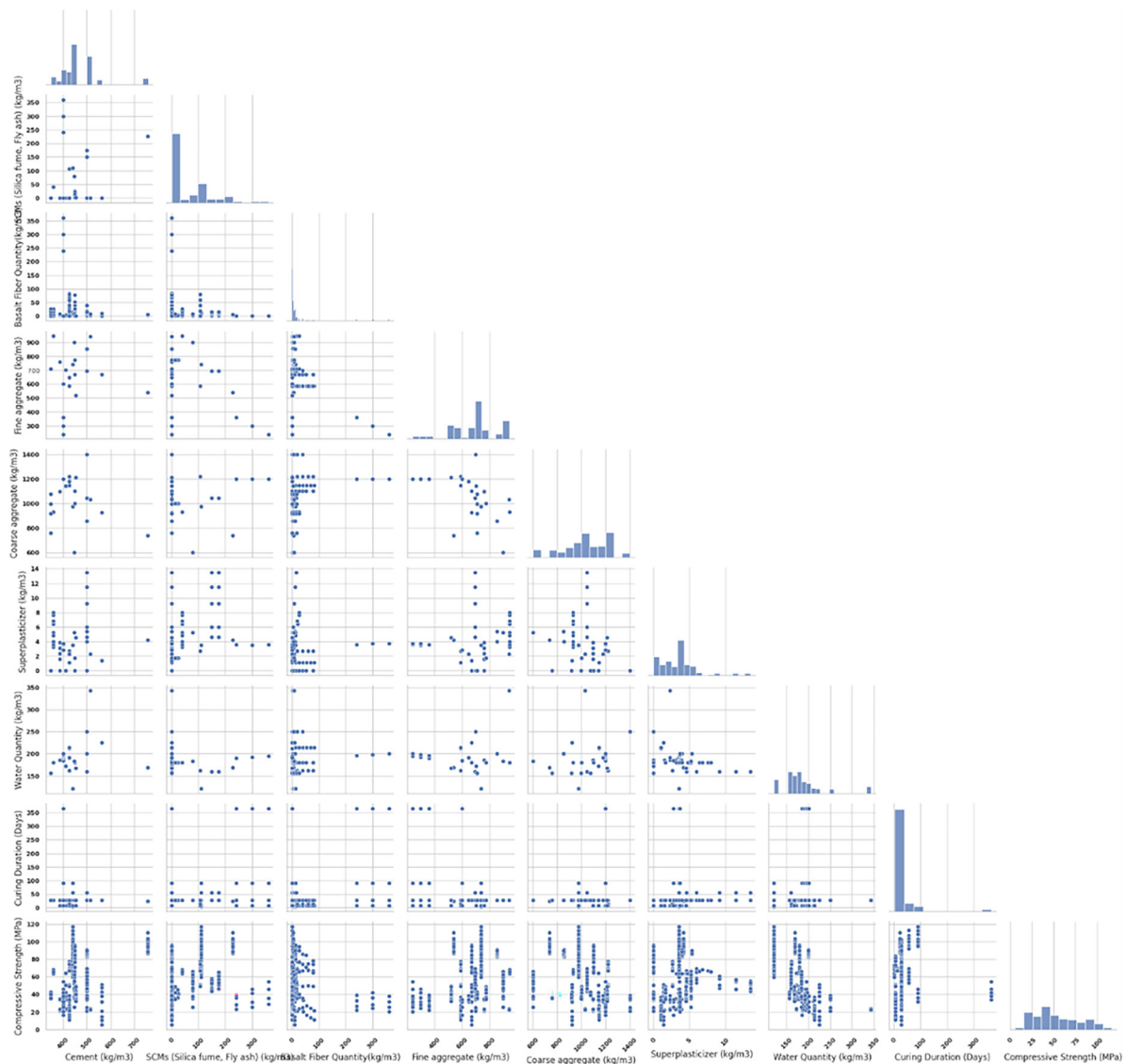


Figure 11: Pair plot of numeric features and their relationship with compressive strength.

In Figure 12(d) (GB+CB Ensemble), the R^2 for the training and test sets are 0.9671 and 0.9441, respectively. The actual and predicted values' alignment in the plot of XG Boost is better than Light GBM, with the test set having less strong readings, mostly between 10 MPa. Finally, Figure 12(e) (GB + LGBM Ensemble) achieved the best R^2 for the training (0.9672) and the test (0.9445). The actual response for both training and testing, and the predicted response, and we see how closely we get to the predictions we made. The majority of the data points are centered at the dashed ideal line and have an insignificant degree of ± 5 MPa deviation, confirming the excellent performance of the model. GB + LGBM has the best results among others due to higher R^2 and smaller deviations with respect to actual and predicted values. The accuracy of model predictions is excellent, with virtually no deviance in the test set. In contrast, Figure 12(c) (Light GBM) exhibits the worst performance, with the biggest residual spread, both successful and failing ones (especially on test).

3.9. Assessment of model residuals across predicted compressive strength values

Figure 13 illustrates the Residuals vs Predicted Compressive Strength plots for the test set in the four subplots (a) to (d), where the residual (actual–predicted strength in MPa) is plotted against the predicted Compressive strength.

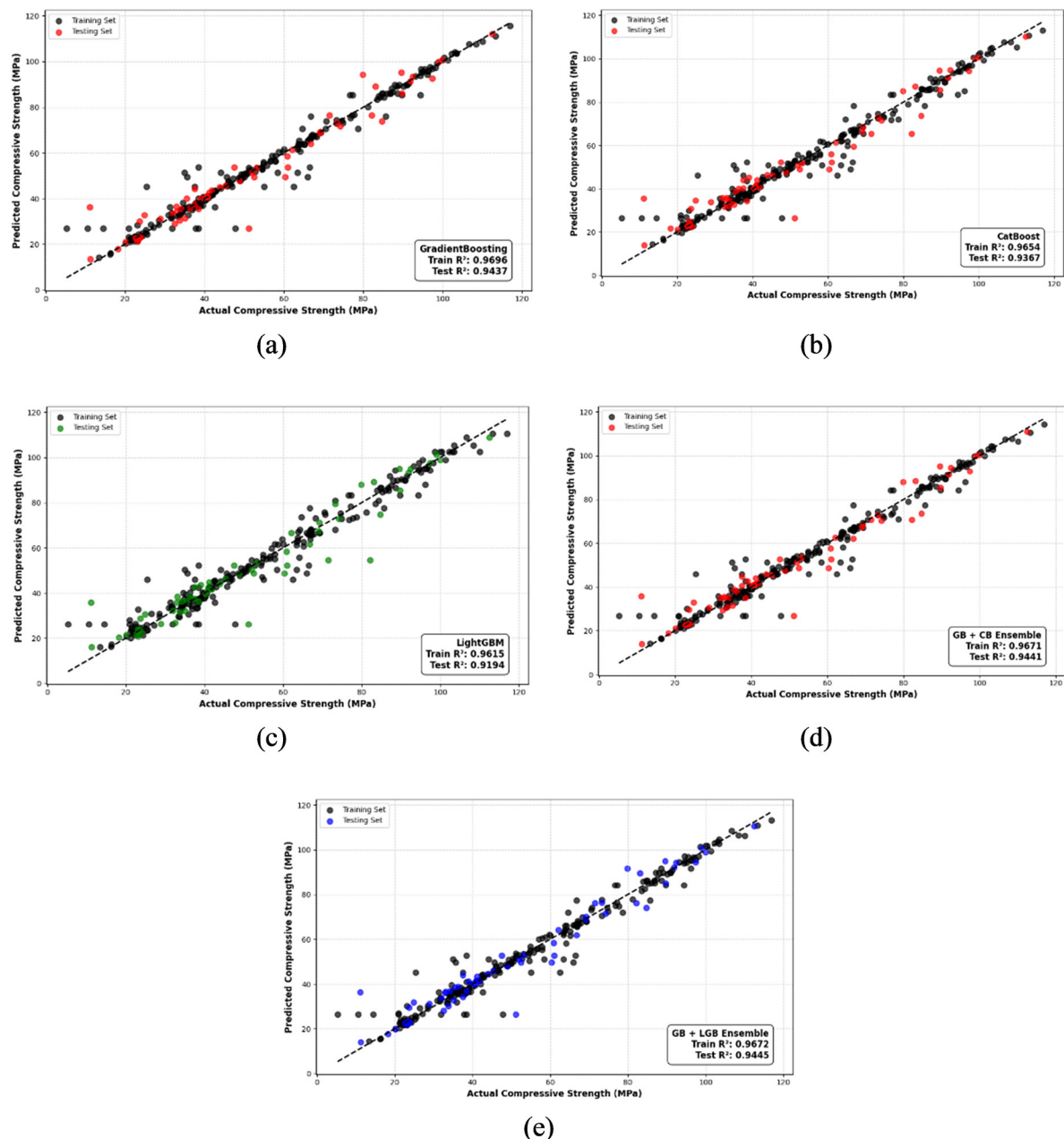


Figure 12: Performance of (a) GradientBoosting, (b) CatBoost, (c) LightGBM, (d) GB + CB ensemble and (e) GB + LGB ensemble models in predicting concrete compressive strength.

In Figure 13(a), the residuals are dispersed with few small trends and overall appear flat. A (dashed red) horizontal line (the '0 residual' line) is added to the plot, which indicates where all residuals should ideally terminate if the model is good. The residuals reach up to ± 20 MPa, meaning that for certain predictions the model brings large errors. The scatter of the residuals grows as a function of the predicted compressive strength, and there are no clear patterns that show that overfitting or underfitting could potentially have taken place. In Figure 13(b), it exhibits the same pattern is exhibited, but the residuals seem to be more evenly spread. The residuals thus still vary between -20 MPa and $+15$ MPa without showing any structure in the data. It appears that the model performs a little bit better than in Figure 13(a), but it deviates noticeably, especially for larger predicted strengths.

Figure 13(c) indicates more trend of homo scholasticate than those depicted in the previous figures. The majority of residuals lie in the interval of -10 MPa and $+10$ MPa, except for some outliers. This shows a better performance of the model, in the sense that errors are smaller and more evenly distributed over the various levels of the predicted compressive strength. There is no significant correlation in the residuals, indicating the model is well-calibrated. Finally, Figure 13(d) yields the best predictions as the residuals are concentrated around zero with a smaller spread than in the other figures. Most of the residuals are in the region of -5 MPa to $+5$ MPa, reflecting very high predictive accuracy. The non-existence of patterns and trends also verifies the good fit of the model. Figure 13(d) is the optimal model since it presents the smallest scatter of residuals, which are mainly concentrated around zero, except for very few large errors. In contrast, Figure 13(a) reveals the most dispersed residuals amounting to ± 20 MPa, indicating the least accuracy. Thus, Figure 13(d) shows the model which is the best at predicting since residuals are minimized over the entire range of predicted compressive strengths.

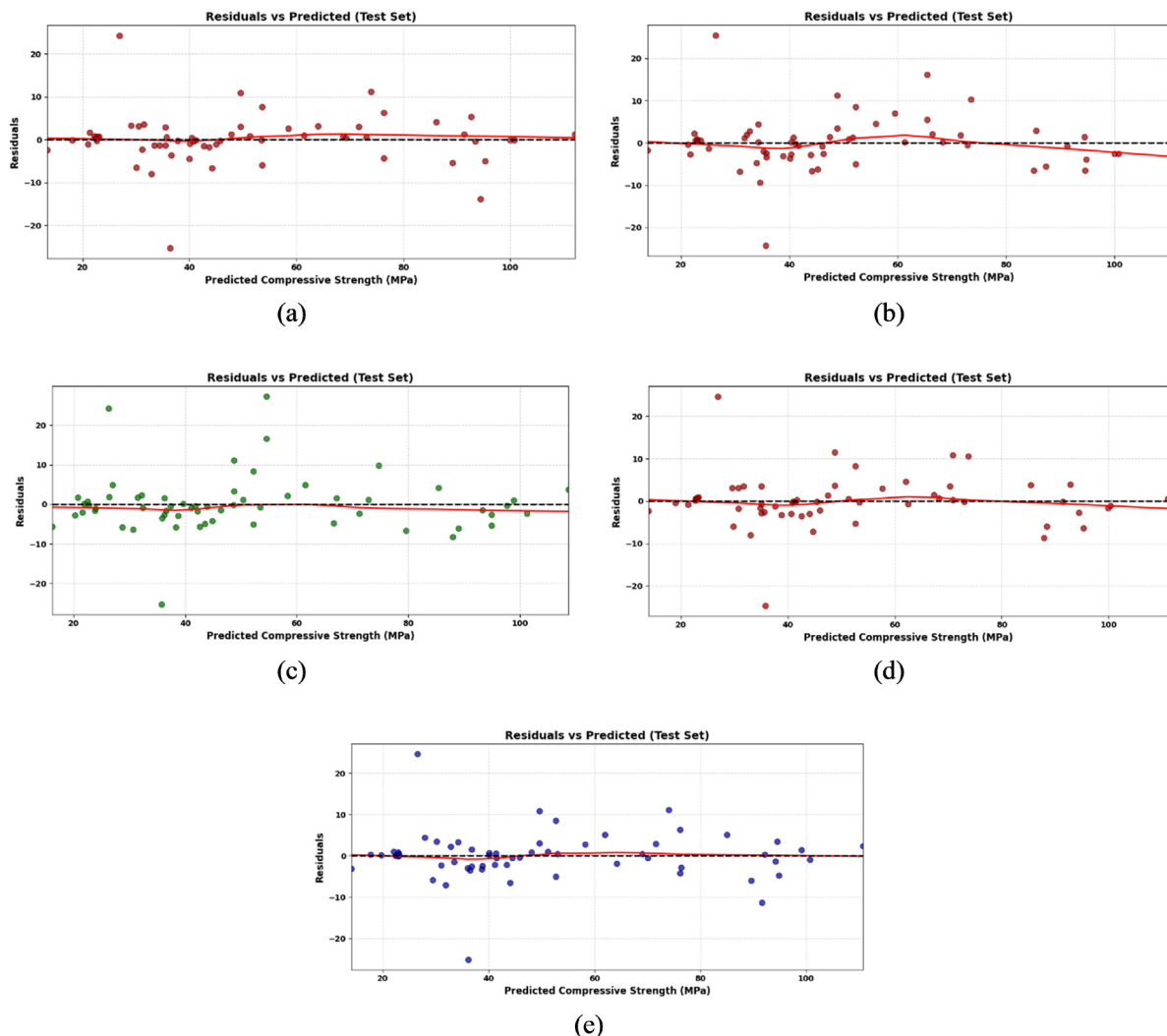


Figure 13: Residuals vs predicted compressive strength for test set using (a) GradientBoosting, (b) CatBoost, (c) LightGBM, (d) GB + CB ensemble and (e) GB + LGB ensemble models.

3.10. Comparison of actual and predicted concrete strength for train and test sets

In Figure 14(a), the true train values (red) and predicted train values (black) are in good agreement for the compressive strength 20–120 MPa. The approximately predicted test values (green) strongly differ from the actual test values (black), especially for the large data indices, with deviations up to 25 MP. This means the model is overfitting and is having trouble making good predictions on the test data. In Figure 14(b), the estimated test values (green) are more consistent with the test values (black), especially in the vicinity of 20–50 MPa. But the cases in which the difference between the two is still 10–15 MPa are not an exception. This can be considered as an improvement compared to Figure 14(a), however the model remains somewhat inadequate, especially in the higher compressive strength region. In Figure 14(c), it can be seen that the forecasted test values (green) are nearer to the test values themselves, with differences of 5–10 MPa in most cases. The range of predicted values for both training and test data is now more constrained, meaning the model is more generalizable than in the previous figures, but we are still seeing some larger errors occurring towards the end of the test data, going as high as 15 MPa in some locations. In Figure 14(d), there to the actual test data, and the precalculated values for the test set (green) are quite similar in that case, with a few larger differences on the order of 10–15 MPa. Predicted values are within a range of 30–100 MPa in the test set, and the model yields the compressive strength with reasonably good accuracy but fails on some (relatively higher) values. Finally, Figure 14(e) shows the optimum model performance observed. The predicted values of test were close to 5 MPa for most of the test points, with some evidence for large deviations at the extreme of the test data, Propanol values. The estimates vary from 20 MPa to 100 MPa, which corresponds to a nearly perfect match to the true values over the test data. Figure 14(e) is the best one with

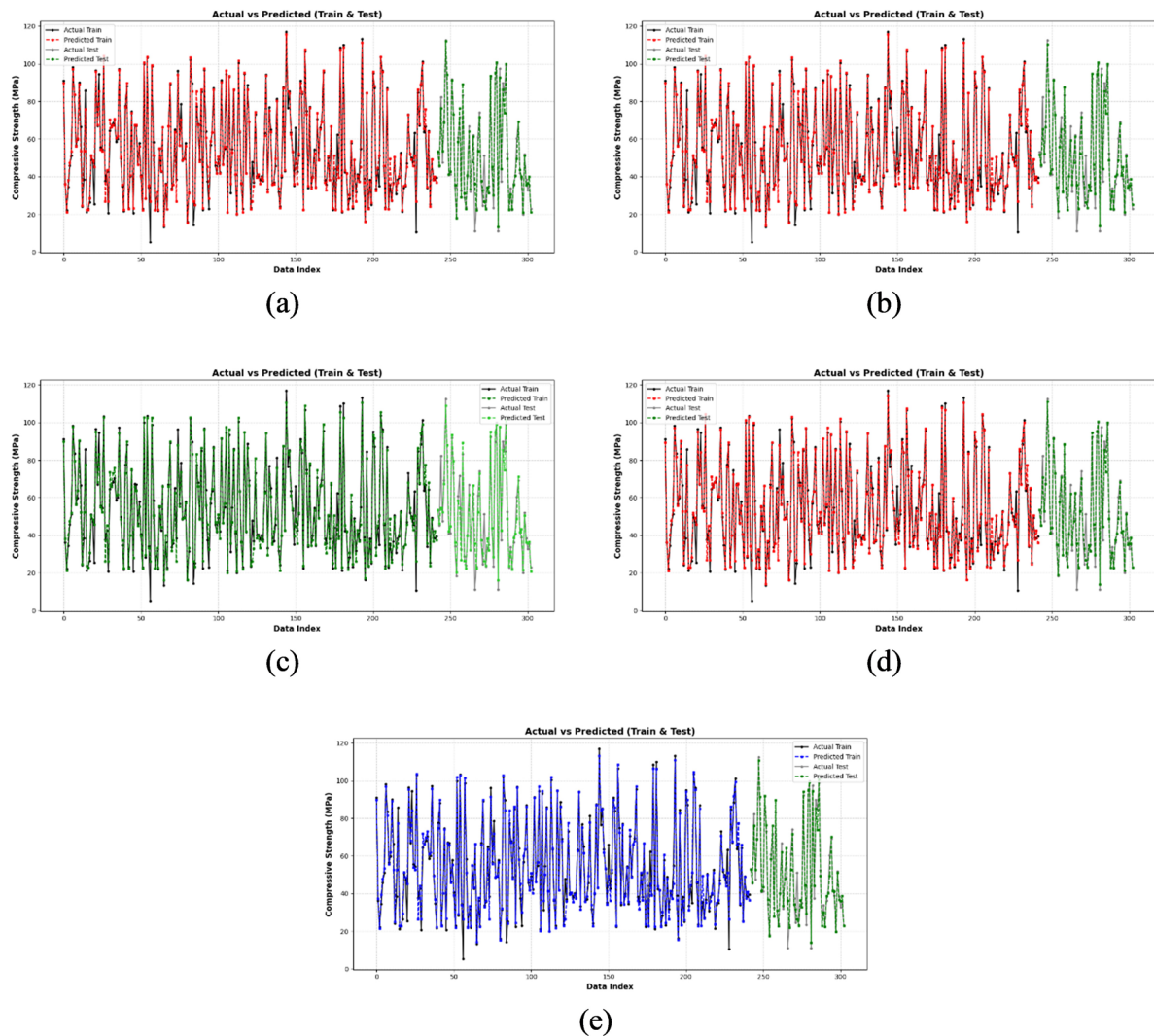


Figure 14: Visualization of actual vs predicted concrete strength across train and test data using (a) GradientBoosting, (b) CatBoost, (c) LightGBM, (d) GB + CB ensemble and (e) GB + LGB ensemble models.

the maximum errors (actual values minus the predicted ones) being not more than approximately 5 MPa. This shows that the predictions on the test set for the model are the most accurate among the models presented in the previous figures, which had errors up to 25 MPa.

3.11. Analysis of absolute prediction errors for different models on test set

The distribution of prediction errors on the test set for our various models (plotted in subplots (a)–(e)) is shown in the Figure 15. Furthermore, Table 3 highlighted the performance indices of all ML models. In Figure 15(a), the model is characterized by a high density of errors around zero, with the magnitude of the prediction error typically being less than 5 MPa. Some errors in the tens of minutes still exist, but this does not prevent them from conveying that some meaningful mispredictions were made. This distribution indicates a model that performs fairly well but sometimes makes large errors. Figure 15(b) presents a distribution of the error with relatively more balanced between low and high error, in which some samples have the absolute error around 5–10 MPa, although the majority still have lower errors. It indicates that the model does not perform too badly and has some varying levels of correctness. Figure 15(c) is also distributed similarly, with the majority of errors being very small and only a few larger errors, which demonstrates that it has good precision with sporadic large deviations.

Figure 15(d), which is a model with most of the prediction errors accumulated on the lower end, and indeed around zero, reflecting very good performance. There are still outliers with errors in the range 5–10 MPa yet, and these may correspond to specific cases or rather complex patters in the reference data not easily handled by the model. Finally, Figure 15(e) shows a histogram with most of the prediction errors within the range of 0–5 MPa and only very few errors beyond 10 MPa. It infer superiority in performance as the majority of the model's

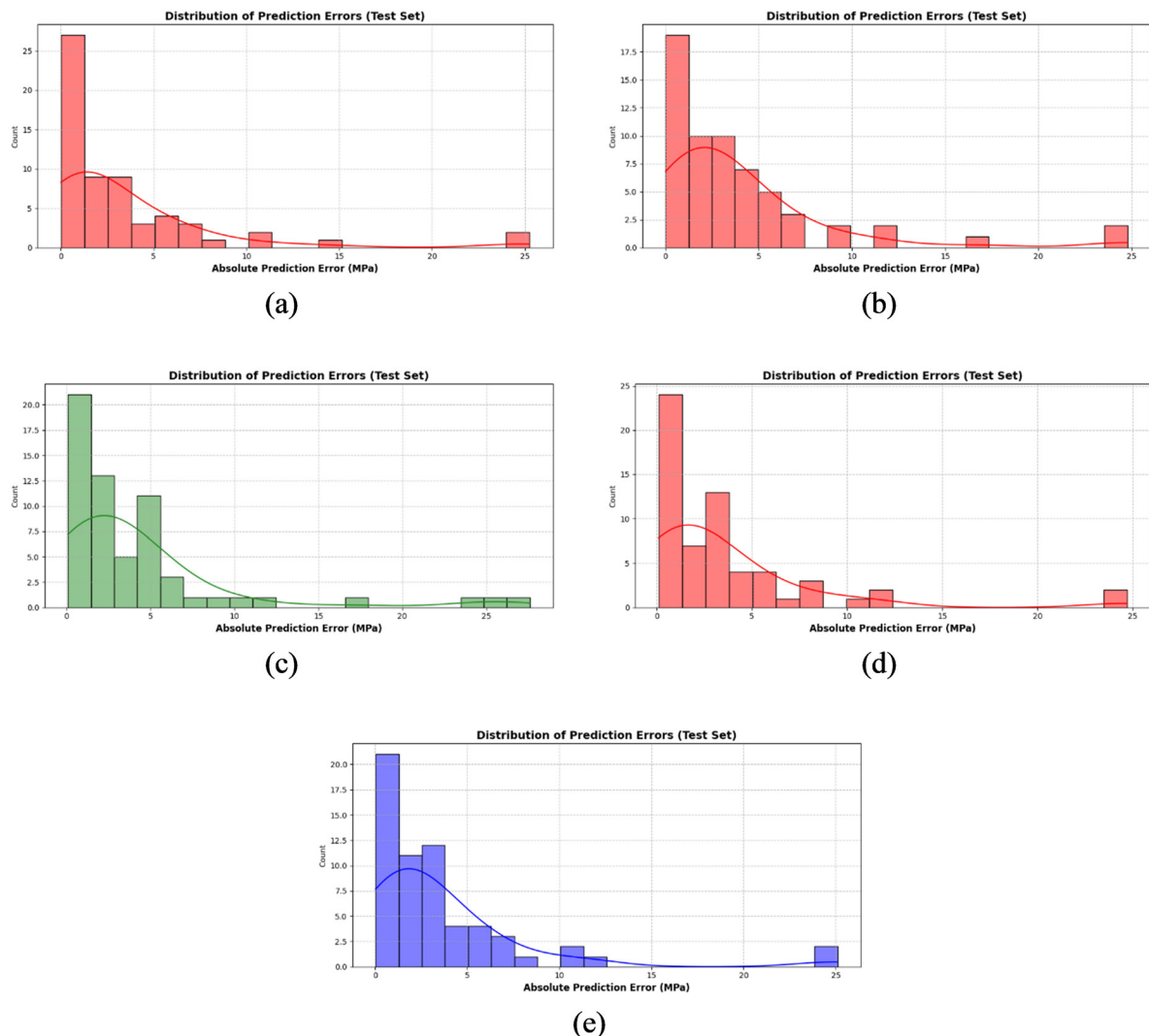


Figure 15: Distribution of prediction errors for test set across (a) GradientBoosting, (b) CatBoost, (c) LightGBM, (d) GB+CB ensemble and (e) GB + LGB ensemble models.

prediction errors are very small, and thus this is the only model displayed, which is the most accurate one. Figure 15(e) corresponds to the best model because it has the smallest prediction errors, where larger errors are, for the most part, limited to the 0–5 MPa range. This suggests that it also makes very accurate predictions compared to the rest of the models displayed.

3.12. Comparison of regression model metrics across train and test sets

The Figure16 shows the grouped bar plot for multiple models such as GB, CB, LGBM, GB + CB, and GB + LGBM, and several performance metrics (R^2 , RMSE, MAE, and MedAE) on training and testing data. At GB, the model presents an average performance a test R^2 of about 6 is even higher than the train value. But the RMSE for the test set is around 6, which is higher than the training performance from the model, which means its accuracy decreased for the test data. MAE and MedAE are relatively higher as well, but the train and test scores agree with each other. CB does a little bit worse, the R^2 of the test set is already higher, shooting around 5,

Table 3: Model performance indices.

MODEL	DATASET	R^2	RMSE	MAE	MEDAE
Gradient Boosting	Train	0.9696	4.4197	2.1629	0.8315
	Test	0.9437	6.0296	3.4779	1.5440
Cat Boost	Train	0.9654	4.7146	2.6973	1.2830
	Test	0.9367	6.3955	4.0353	2.6565
Light GBM	Train	0.9615	4.9736	3.1378	1.7556
	Test	0.9194	7.2177	4.3588	2.4124
GB + CB	Train	0.9671	4.5959	2.5069	1.2002
	Test	0.9441	6.0097	3.6091	2.0879
GB + LGB	Train	0.9672	4.5904	2.5772	1.3131
	Test	0.9445	5.9884	3.6235	2.261

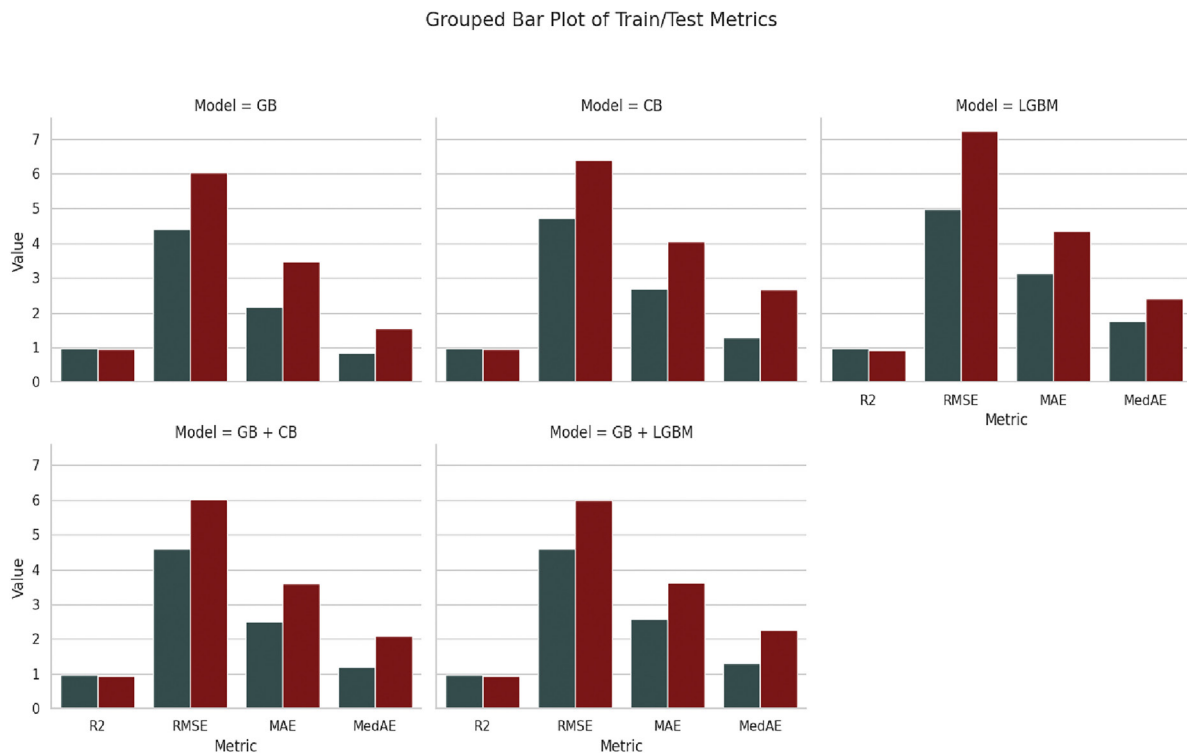


Figure 16: Grouped bar plot of model performance metrics (Train vs Test).

meaning it partly explains the test: moderate explanatory power. The RMSE of the test is ~ 5 (less than that of the train set), but it gives a greater error overall than GB. The values of both MAE and MedAE are also larger than GB, representing larger average and median error measures.

On LGBM, the model is underperforming across all metrics. The R^2 value is under 3, which means it does not explain a whole lot of variance, and the test set RMSE comes out pretty high at 6, which implies a poor predictive performance. Likewise, the MAE and MedAE are both high, meaning high prediction errors. The combined GB + CB outperforms the other two single models, indicating the ability of combining models for more explanatory power, with our R^2 on the test set raised to around 6. The RMSE is around 5, and the MAE and MedAE are also improved to 3. Finally, we obtain that the GB + LGBM mixed model always has superior performance than the others. It has the highest R^2 of 7, meaning that it explains the largest proportion of the variance in the data. It has the lowest average RMSE over the entire test set, which is around 4, indicating the best predictive power. They are also the lowest, of 2, and ~ 1 , respectively, hence indicating that they commit not many errors, on average, and not many errors at the median. GB + LGB is the best model, including all measures, achieving optimal prediction and error rates. It has constant best performance over other models and also combinations like GB + CB.

3.13. Evaluation of regression models: R^2 , RMSE, MAE, and MedAE across train and test sets

The Figure17 depicts different model performance measures for different ML models, evaluating them on training and testing data with different accuracy scores. The reported metrics are R^2 , RMSE, MAE, and MedAE for different models, as GB (Gradient Boosting), LGBM (Light GBM), and CB (Cat Boost), and linear combinations of these models (GB + CB, GB + LGBM).

GB + LGBM model is the best performing, with the highest R^2 for training as well as testing data which means this model has the capability of explaining more variation between test and train data. The LGBM model is the second best, while GB and CB have lower R^2 scores. GB + LGBM is the best according to RMSE, exhibiting the smallest RMSE values and indicating better prediction accuracy on the training and testing data, together. Despite that, the GB+LGBM model is better and has no residuals in terms of MAE, since for this model, the MAE values are lower, meaning smaller average absolute errors between true and predicted values. Next, LGBM and GB, while CB once more presents the largest MAE values, meaning greater errors on average.

GB + LGBM head with the smallest MedAEs, which is a measure of predicting the median of the error distribution. Also, here, CB trails again with elevated MedAE values, indicating that its model performance is more sensitive to the outliers compared to the others. The GB + LGBM model performs better than any other model in all the metrics, having the best results in R^2 , RMSE, MAE and MedAE. The performance of the CB model is typically the weakest, but LGBM and GB models are comparable, but weaker than the ensemble ones. Consequently, GB + LGBM is the most appropriate model for our purpose, due to the fact that it produces the lowest error and the highest predictive accuracy. CNN and DNN, as well as Random Forest (RF) and XG Boost were adopted, up to the CNN showed the best performance in the test phase (NSE = 0.95 for TDS, NSE = 0.96 for PS, NSE = 0.67 for SAR, and NSE = 0.93 for CI) [119]. We experimented with several models such as Linear Regression, K-nearest Neighbor, Support Vector Regression, Decision Tree, Random Forest, Extra Tree, and the Gradient boosting. He observed that among other techniques used for the prediction models, random forest gave the best prediction task with an R^2 of 0.926 and mean absolute error of 1.608 for the determination of the compressive strength [120].

3.14. Feature impact distribution in concrete mix predictions using SHAP

The SHAP Summary Plot for Concrete Mix Components in Figure18 graphically demonstrates the effect of individual features on the model response. The feature CA is the one that seems more influential, with a great level of SHAP values, with a standard deviation of -10 to 10 , which means it is considerably affecting the model predictions. This feature has more SHAP values for a greater quantity of coarse aggregate, i.e., a greater amount of coarse aggregate leads to a greater positive effect on the model. Water Quantity, when close to the range, the distribution of SHAP values is also similar, reaching our maximum at close to $+5$, and indicating that larger Water Quantities are contributing positive output predictions. Both release agents and SCMs (Silica fume, Fly ash) have a relatively weak effect, with SHAP values between 5 and -5 ; that is, SCMs are either not included in the model or their presence increases the model's output. Curing Duration (Days) has a weaker impact, with mostly small SHAP values in the range $[-2, 3]$. Cement, Superplasticizer (kg/m^3), and Basalt Fiber Quantity present higher concentration of SHAP values with a moderate influence on concrete mix properties. Finally, Fine aggregate (kg/m^3) is the least influential and so has SHAP value is negligible, showing it plays the minimum role in the model. Finally, Coarse aggregate and Water Quantity are the most and Fine aggregate is the least influential.

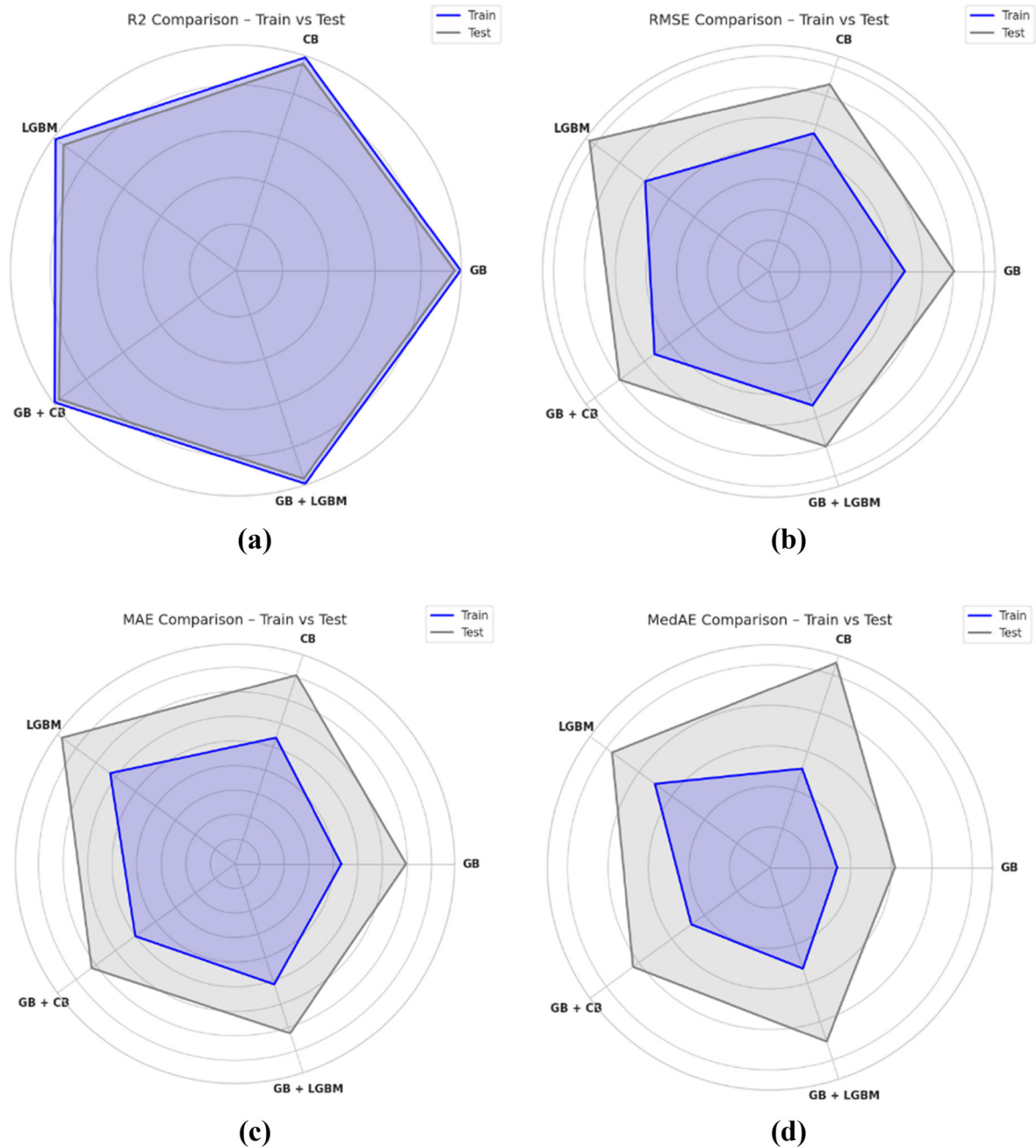


Figure 17: Comparison of model performance metrics (Train vs Test).

3.15. Feature importance of concrete mix variables based on SHAP values: impact on compressive strength prediction

Figure 19 depicts the SHAP feature importance across different components of a concrete mix. It's a point of explaining how much each feature is affecting the model's result. The dominant feature is Coarse aggregate, which has a SHAP value of nearly 8, such that it is most influential for the prediction of the model. Next is Water Quantity around 7, and on it, SCM (Silica fume, Fly ash) around 6. Other parameters such as Curing Days, Cement, and Superplasticizer demonstrate mild influences with 3 and 5 of them. The factors of Basalt Fiber and FA have a less obvious effect than that of FA. As a whole, Coarse aggregate is the most effective in controlling concrete mix characteristics, and Fine aggregate is the least effective.

3.16. SHAP dependence plots showing the impact of concrete mix variables (cement, SCMs, aggregates, water, and curing) on compressive strength

The Fig SHAP dependence plots are shown in Figs. (20) depict the connections between concrete ingredients (such as cement, SCMs, basalt fiber, water, aggregates, SP, and curing period) and their influences on compressive

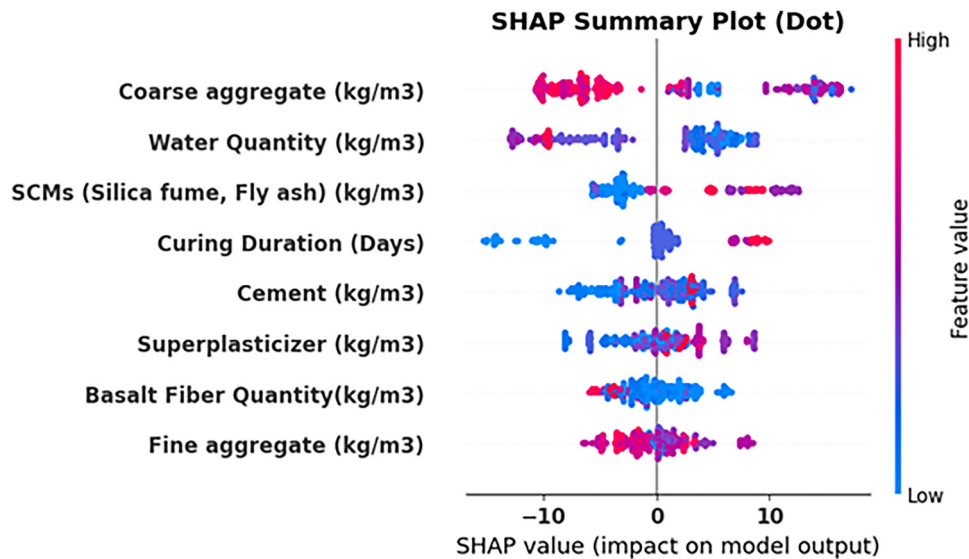


Figure 18: SHAP summary plot for concrete mix components.

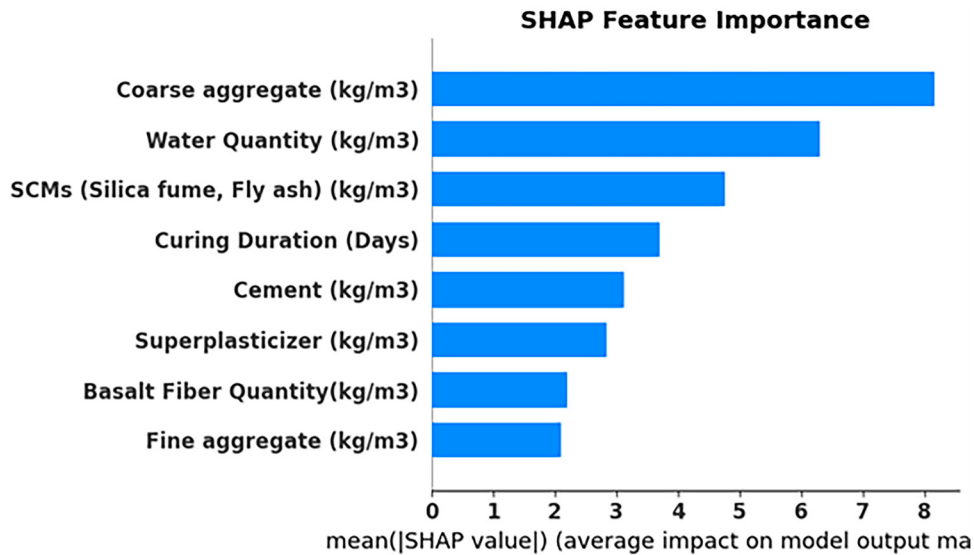


Figure 19: SHAP feature importance for concrete mix components.

strength. SHAP values represent the contribution of each feature to the predicted target (i.e., compressive strength), with positive values referring to an increased strength and negative values referring to a decreased strength.

Figure 20(a) shows that the SHAP values are considerably higher with higher cement content. The figure shows that the influence of cement on compressive strength is very strong and positive, with the SHAP values between -8 and $+8$. Increasing the content of coarse aggregates enhances the role of cement, which is indicated by the blue-to-red color in the graph. A linear relationship is shown, which represents the consistent effect of cement on the increase in compressive strength. Figure 20(b) shows that the SHAP values are between -5 and $+12.5$. The results in the plot reveal that SCMs have a dual impact, such that a greater amount of SCM addition increases compressive strength (SHAP values shift to the positive region). Values decrease sharply for lower contents of SCMs, implying that their favorable effect is only appreciable compared to their application in an amounts.

Figure 20(c) shows the SHAP value plot for basalt fiber revealing an insignificant difference; most of the data points are close to 0, indicating that the basalt fiber has little impact on the compressive strength. The SHAP values are negative or close to 0, which indicates that the amount of basalt fiber does not contribute to the compressive strength in this specific set of observations. Figure 20(d) reveals that SHAP values lie between

−6 and +6. The relationship becomes weakly adverse at the highest fine aggregate contents, i.e. with fine aggregate contents towards the higher end of the range, would be expected to reduce strength. Without much reverse, so it means that the level of proportion is very or moderate to low reverse and it is a moderate reverse between water absorption and compressive strength.

Figure 20(e) depicts coarse aggregate from −10 to +15 of SHAP values. From the plot, it can be seen that increasing the percentage of coarse aggregate results in a corresponding increase in compressive strength, and the correlation is uniformly upward in proportion to the percentage of coarse aggregate. Negative values obtained for lower percentages of coarse aggregate indicate a decrease in strength. Figure 20(f) demonstrates that the SHAP values change from negative correlation to positive correlation when the superplasticizer dosage reduces, suggesting that the increase in superplasticizer content it is sometimes beneficial on the compressive strength. The SHAP values are between −7 and +7, indicating an overall positive influence, but the influence is relatively weaker than that of cement and coarse aggregate. Figure 20(g) shows that the water quantity plot had a negative relation with compressive strength, reported SHAP from −10 to +5. Compressive strength reduces with an increase in water ratio in a manner consistent with the negative correlation established in previous assessments. The increased water content, that way, reduced the strength of the concrete. Figure 20(h) shows that the SHAP value increases significantly with curing period. The SHAP values for curing time are between −15 and +10, in an upward trend. This indicates that the curing period has had a strong positive influence on compressive strength, except in the very early days.

According to the SHAP dependence plots, cement (Figure 20(a)) and CA (Figure 20(e)), the most influential parameters affecting compressive strength seem to be Cs and G. Both features have a stable and strong positive effect on compressive strength, with cement exhibiting the biggest range of SHAP values reflecting its importance. Coarse aggregate comes second to it in effect, and has a clear and orderly positive connection with compressive strength. On the contrary, basalt fiber included (Figure 20(c)) and volume of water (Figure 20(g)) cause affecting on the compressive strength but less important on the compressive strength in comparison with the other mentioned parameters. Thus, the cement and the coarse aggregate are the major aggregates responsible for the high compressive strength of the concrete mix. The outcomes of the feature optimization approach are very significant for developed regions like Najran City to reduce the dependency on material-based trial and error process and directly predict the CS of concrete.

MILLER *et al.* [121] note that the SHAP value for the age feature is negative for the first 70 observations and very positive for the second half (71–140). For the coarse aggregate characteristic, the SHAP value is low in the cases of 20–115 and relatively higher in others. According to the SHAP studies of KASHEM *et al.* [122] the curing time has the most prominent effect on the mean cylinder compressive strength of UHPC, followed by fiber, cement, silica fume, and superplasticizer. Meanwhile, the effect of fly ash parameter is the smallest for the compressive strength of UHPC.

3.17. Partial dependence of concrete compressive strength on mix components and curing duration

Figure 21 shows Partial Dependence Plots based on the feature space affected by the prediction of compressive strength of concrete. Of all features, curing time presents the most dominant and constant positive effect on compressive strength. With the growth of curing time (from −0.6 to 0.9 standard units), the predicted value of the strength increases dramatically, and the maximum level is near the value of 60 MPa, which means the longer the curing time, the higher the strength. SCMs also exhibit a rising trend, with the strength predicted to increase from approximately 47 to over 57 MPa as their amount is increased, suggesting that they have an important contribution to strength gain.

The amount of water has the largest negative impact on the other side. With increasing water content (between 1 and 3 above and below standardized units), the predicted strength falls drastically from 58 to around 41 MPa, illustrating that, at excess water, the concrete strength tends to weaken. Cement and superplasticizer have mild beneficial effects, improving the strength up to around 54–55 MPa with the optimal dosages. Basalt fiber has a detrimental or almost neutral effect, with 45 MPa remaining the maximum compressive strength when used at any dosage. The curing period is the most important positive factor. Depending on the water/cement ratio, water content can be the most detrimental if it is not controlled. Thus, it is important to find a balance here to create as much impervious and strong concrete as can be achieved.

As far as PDP analysis is concerned, which is used in determining the optimal levels of the input factors in a mixture design, the PDP analysis can also be considered to improve the strength of UHPC by KASHEM *et al.* [122]. SHAFIGHFARD *et al.* [123] stated that the three most significant predictors in the PDP analysis are silica fume (SF), slag, and curing temperature. Moreover, in their PDP report, SOBUZ *et al.* [124] also disclosed that raising the cement content from 450 kg/m³ up to 475 kg/m³ can achieve the improvement of the compressive strength from 46 MPa up to 52.2 MPa, with approximately 17.8% increment.

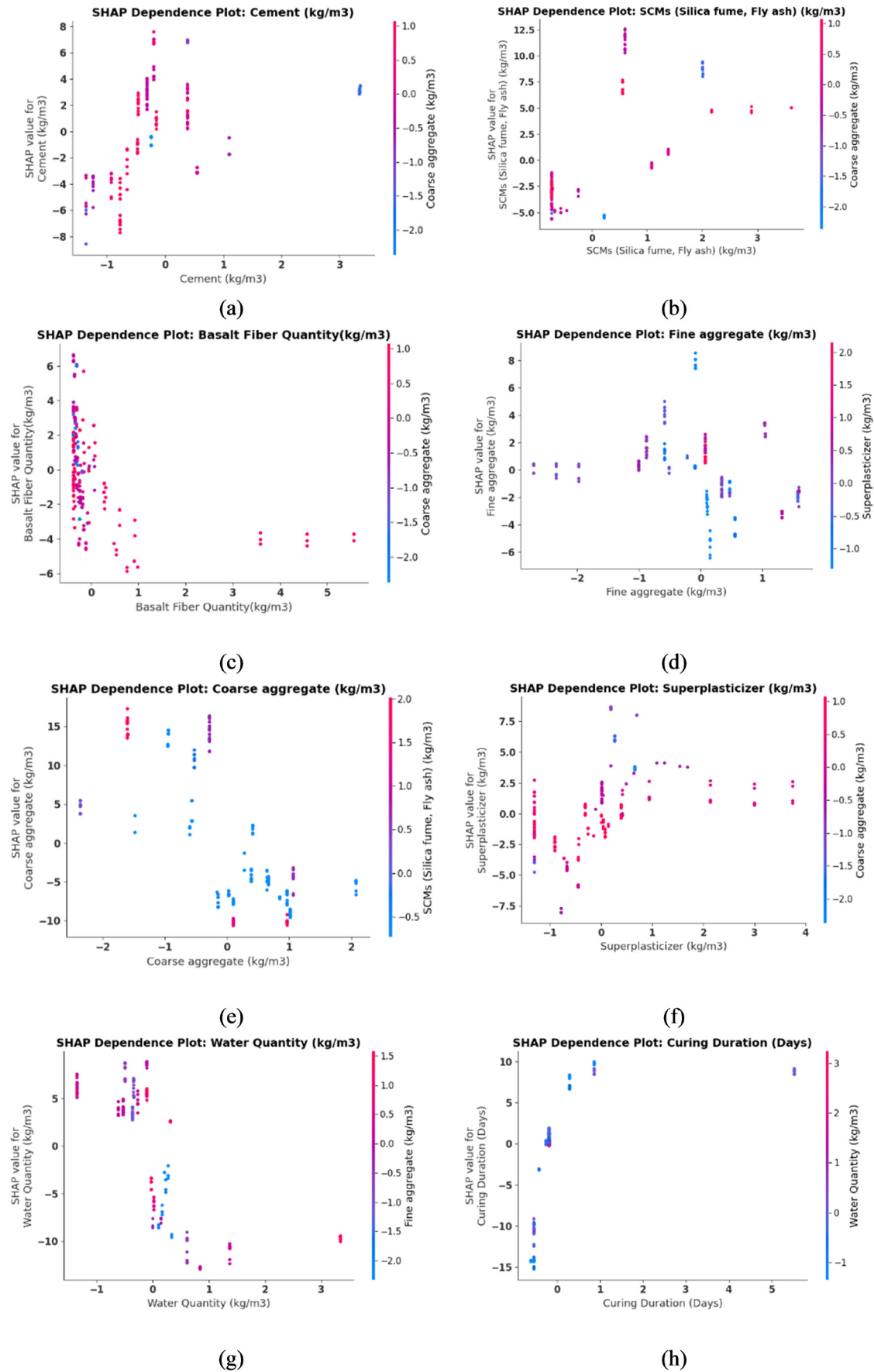


Figure 20: SHAP dependence plots for concrete mix components and compressive strength.

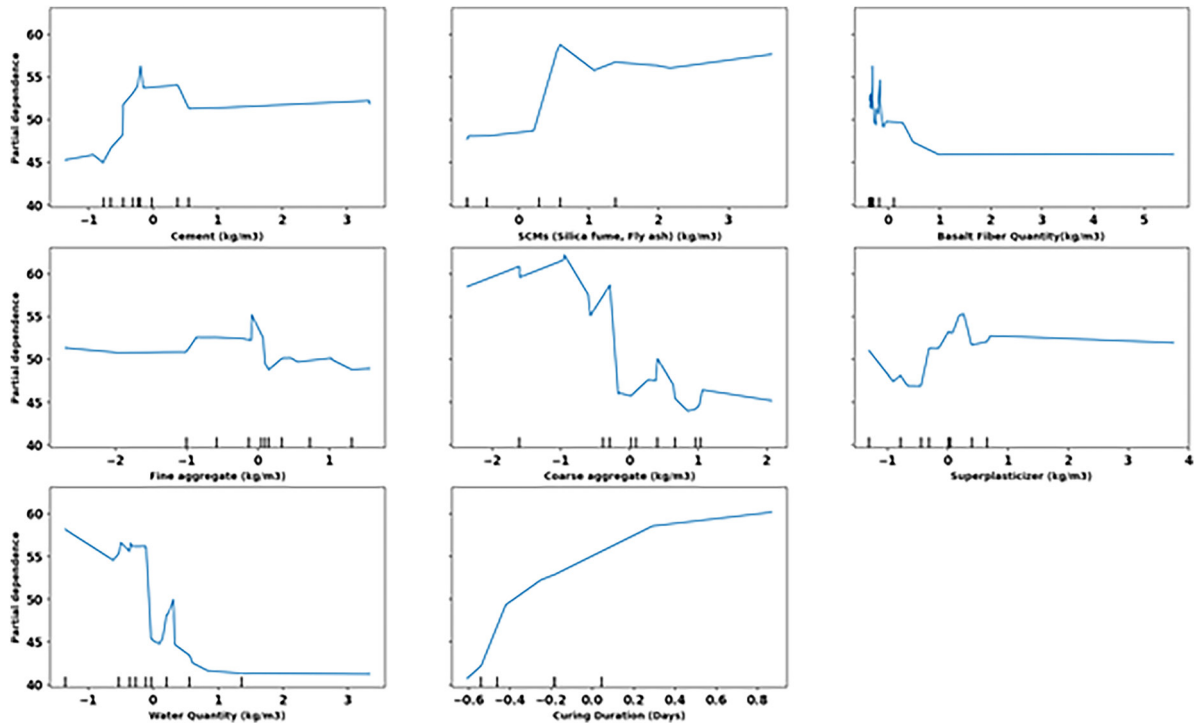


Figure 21: Partial dependence plots for key concrete mix parameters.

4. CONCLUSION

The finding of this study given below:

- This paper proposes a holistic Machine Learning (ML) methodology to predict the 28-day compressive strength of Basalt Fiber Reinforced Concrete (BFRC) using novel ensemble approaches on 450 concrete mix designs.
- Of all models, the GB+LGBM hybrid ensemble was found to be the best performing model with the highest R^2 (0.9445) and the lowest RMSE (5.99 MPa) on the test dataset when compared to other single as well as hybrid models like Cat Boost, Light GBM, GB+CB, and GB+CB+LGBM.
- The SHAP (Shapley Additive Explanations) figure showed that compressive strength was most influenced by the content of coarse aggregate and cement ($\text{SHAP} \approx +8$ and $+7$, respectively). The influence of the water content was negative but significant, which proved the harmfulness of the high water-cement ratios. Basalt fiber, while being an integral component of BFRC, has been demonstrated to have a moderate impact when overloaded, even resulting in strength decrease if it's agglomerated and workability loss.
- Partial Dependence Plots (PDP) showed that the effects of the curing period are significant for strength development, especially above 90 days, and then become constant. SCMs (silica fume/fly ash) have a positive effect on the compressive strength to an optimum level after which the gain in compressive strength it reduces. There is a great influence on the rate of dosage of superplasticizer, mainly in the range of 2.8–4.3 kg/m^3 in order to increase strength due to the improved compaction as well as workability.
- Residuals and prediction error distributions also supported GB+LGBM as the best model, and the fact that the model presented slight overfitting with balanced predicted values and almost all the predicted errors taking place within ± 5 MPa range made it a better choice than all the other regressors.
- Visual interpretation (heatmaps, correlation matrices, box plots, violin plots) further reinforced that cement, coarse aggregates, SCMs, and superplasticizer are the primary contributors of performance, while excess of fine aggregate and water are the offenders.

5. LIMITATIONS

While this research proves the robust predictive capacity of hybrid ensemble machine learning models for predicting the compressive strength of BFRC, some limitations have to be recognized. First, although the data

used in this study is from a large sample, it was limited concerning variation of fiber content, aggregate grading, and curing conditions. These restrictions could affect the applicability of the built models to other geographical areas, sources of materials, and environments of concrete production. Furthermore, the dataset was derived from several studies, which brought in sources of variation regarding testing standards, specimen preparation, and lab environment. These inconsistencies may result in uncertainty of predictions and could compromise the model's robustness. In addition, while SHAP-based interpretability provides useful explanations, it may miss complex multi-feature interactions or non-linear dependencies present in more heterogeneous data. Lastly, the investigation is more predictive modelling-oriented rather than based on multi-objective aspects (like cost, sustainability durability) required in practice for mix design optimization.

6. FUTURE DIRECTIONS

Future studies could overcome these limitations by utilizing larger and more diverse experimental datasets, including region-specific material properties as well as a wide variety of BFRP dosage levels and curing conditions. In this study, there is the possibility that more advanced hyper-parameter tuning techniques (e.g., Bayesian optimization, evolutionary learning, or meta-learning) would be proposed to enhance the stability and generalization of the model, as many previous studies dealing with a hybrid ML framework, as well as with a hyper-tuned ML models to accurately predict concrete properties [125–127]. Moreover, one could consider further inclusion of microstructural, rheological, and environmental input to improve the physical interpretability of predictions. Multi-goal optimization methods can be additionally utilized to compromise between the mechanical property and the economic eco-environment. Last but not least, cross-validation in independent laboratories or validation on a field scale would give more convincing evidence on model applicability for practical engineering use. These innovations will, in turn, contribute to the C&EW and concrete industry materializing more robust, scalable, and sustainable ML-based solutions for mix design.

7. PRACTICAL IMPLICATIONS

The practical implications of the findings in this paper for BFRC design and construction could be substantial. The hybrid machine learning models proposed in this completed research can be adopted by engineers to improve BFRC mix design, that is achieve the target compressive strength and minimize the use of materials. When the dominant factors affecting strength, specifically cement content, coarse aggregate contents, and fiber dosages, are determined, the waste of time taken sampling concrete in the lab to optimize mixes will be effectively decreased. From an economic standpoint, the results of such models can significantly diminish the requirement of costly and time-consuming laboratory tests for investigation purposes, thus reducing overall project cost. Additionally, adopted mix designs will save on both material wastage and cost efficiency, providing sustainable development. From an environmental point of view, the ability to optimize BFRC mixtures may help to reduce the quantity of supplementary cementitious materials required and thereby the global carbon footprint for concrete production. Hence, it is believed that the use of these hybrid models can lead to better economic and sustainable performance of BFRC in practical structures.

The important properties determined in this study, which are cement content, coarse aggregate, and water-cement ratio, will help to optimize the mix design of BFRC. Understanding the relative significance of these features will allow engineers to optimize the material's proportions in relation to the compressive strength and performance of NAC. For example, the findings indicate that an optimal amount of cement between 440 and 600 kg/m³ results in enhanced compressive strength. Also, coarse aggregate content has a strong influence on compressive strength; therefore, appropriate proportioning of aggregates is significant for concrete resistance and strength improvement. The water should also be regulated to avoid the loss of strength, and it is recommended to use a content between 120–160 kg/m³ so as not to have the paste diluted. The knowledge obtained from the machine learning models and the importance of these features present a data-driven approach toward mix design optimization, which is essential for developing sustainable as well as high-performance concrete. Practitioners could apply these results to optimize the mixing ratios, thereby saving materials and optimally promoting the economic and environmental performance of concrete.

8. ACKNOWLEDGMENTS

The authors are thankful to the Deanship of Graduate Studies and Scientific Research at Najran University for funding this work under the Student Research Funding Program grant code (NU/SRP/SERC/14/3391-1).

9. DATA AVAILABILITY

The dataset used in this study has been provided as supplementary data with this manuscript.

10. SUPPLEMENTARY MATERIAL

Dataset_Basalt Fiber Reinforced Concrete

11. BIBLIOGRAPHY

- [1] SONG, Y., XU, R., HONG, J., *et al.*, “Preparation of a novel green lithium slag expansive agent: Enhancing deformation control, crack resistance, and durability of mortar”, *Construction & Building Materials*, v. 495, pp. 143644, 2025. doi: <https://doi.org/10.1016/j.conbuildmat.2025.143644>.
- [2] WEI, X., XIE, S., LUO, A., *et al.*, “Performance improvement of ecologically friendly ultra-high performance concrete with waste tire recycled steel fiber”, *Construction & Building Materials*, v. 495, pp. 143607, 2025. doi: <https://doi.org/10.1016/j.conbuildmat.2025.143607>.
- [3] SOBUZ, H.R., VISINTIN, P., MOHAMED ALI, M.S., *et al.*, “Manufacturing ultra-high performance concrete utilising conventional materials and production methods”, *Construction & Building Materials*, v. 111, pp. 251–261, 2016. doi: <https://doi.org/10.1016/j.conbuildmat.2016.02.102>.
- [4] HABIBUR RAHMAN SOBUZ, M., KHAN, M.H., KAWSARUL ISLAM KABBO, M., *et al.*, “Assessment of mechanical properties with machine learning modeling and durability, and microstructural characteristics of a biochar-cement mortar composite”, *Construction & Building Materials*, v. 411, pp. 134281, 2024. doi: <https://doi.org/10.1016/j.conbuildmat.2023.134281>.
- [5] LEE, H.-P., CHAK, W.-F., TEOW, K.-L., *et al.*, “Investigation of catalyzed biomass thermoelectric concrete with palm oil fuel ash”, In: *Proceedings of the International Conference on Sustainable Civil Engineering Structures and Construction Materials*, pp. 451–463, 2020.
- [6] FU, L.Y., YANG, H., FU, B.Y., *et al.*, “Stress-dependent wave propagation in fractured rocks with nonlinear elastic and hyperelastic deformations”, *Journal of Geophysical Research. Solid Earth*, v. 130, n. 6, pp. e2024JB030889, 2025. doi: <https://doi.org/10.1029/2024JB030889>.
- [7] WANG, K., QIU, Z., HU, H., *et al.*, “Macro-meso mechanical behavior and degradation mechanisms of silty clay subgrade subjected to coupled traffic loading and dry-wet cycles”, *Transportation Geotechnics*, v. 55, pp. 101725, 2025. doi: <https://doi.org/10.1016/j.trgeo.2025.101725>.
- [8] SU, Z., LU, X., YIN, Q., *et al.*, “Internal and surface deformation evolution and rupture characteristics of rock-like models with different compression strengths”, *Experimental Mechanics*, v. 65, pp. 1547–1566, 2025. doi: <https://doi.org/10.1007/s11340-025-01236-4>.
- [9] DATTA, S.D., SOBUZ, M.H.R., MIM, N.J., *et al.*, “Investigation on the effectiveness of using building information modeling (BIM) tools in project management: a case study”, *Revista de la Construcción*, v. 22, n. 2, pp. 306–320, 2023. doi: <https://doi.org/10.7764/RDLC.22.2.306>.
- [10] LEE, H.-P., JUSLI, E., LING, J.H., *et al.*, “Concrete Paving Blocks incorporating palm oil boiler ash and palm oil clinker as substitute concrete materials”, *Journal of Engineering Science and Technology*, v. 18, n. 6, pp. 75–81, 2023.
- [11] AL-KHARABSHEH, B.N., ARBILI, M.M., MAJDI, A., *et al.*, “Basalt fiber reinforced concrete: a compressive review on durability aspects”, *Materials*, v. 16, n. 1, pp. 429, 2023. doi: <https://doi.org/10.3390/ma16010429>. PubMed PMID: 36614766.
- [12] LI, Y., ZHANG, J., HE, Y., *et al.*, “A review on durability of basalt fiber reinforced concrete”, *Composites Science and Technology*, v. 225, pp. 109519, 2022. doi: <https://doi.org/10.1016/j.compscitech.2022.109519>.
- [13] SOBUZ, M.H.R., DATTA, S.D., AKID, A.S.M., “Investigating the combined effect of aggregate size and sulphate attack on producing sustainable recycled aggregate concrete”, *Australian Journal of Civil Engineering*, v. 21, n. 2, pp. 224–239, 2023. doi: <https://doi.org/10.1080/14488353.2022.2088646>.
- [14] UDDIN, M.A., JAMEEL, M., SOBUZ, H.R., *et al.*, “The Effect of Curing Time on Compressive Strength of Composite Cement Concrete”, *Applied Mechanics and Materials*, v. 204-208, pp. 4105–4109, 2012. doi: <https://doi.org/10.4028/www.scientific.net/AMM.204-208.4105>.
- [15] SOBUZ, M.H.R., DATTA, S.D., RAHMAN, M. “Evaluating the properties of demolished aggregate concrete with non-destructive assessment”, In: Arthur, S., Saitoh, M., Pal S. K. (eds), *Advances in civil engineering, lecture notes in civil engineering*, Singapore, Springer Singapore, 2022, pp. 223–233. doi: https://doi.org/10.1007/978-981-16-5547-0_22.
- [16] JOHN, V.J., DHARMAR, B., “Influence of basalt fibers on the mechanical behavior of concrete—A review”, *Structural Concrete*, v. 22, n. 1, pp. 491–502, 2021. doi: <https://doi.org/10.1002/suco.201900086>.

- [17] JAGADEESH, P., RANGAPPA, S.M., SIENGCHIN, S., “Basalt fibers: An environmentally acceptable and sustainable green material for polymer composites”, *Construction & Building Materials*, v. 436, pp. 136834, 2024. doi: <https://doi.org/10.1016/j.conbuildmat.2024.136834>.
- [18] HUANG, H., HUANG, M., ZHANG, W., *et al.*, “Experimental investigation on rehabilitation of corroded RC columns with BSP and HPFL under combined loadings”, *Journal of Structural Engineering*, v. 146, n. 8, pp. 04020157, 2020. doi: [https://doi.org/10.1061/\(ASCE\)ST.1943-541X.0002725](https://doi.org/10.1061/(ASCE)ST.1943-541X.0002725).
- [19] DAS, S., HABIBUR RAHMAN SOBUZ, M., TAM, V.W.Y., *et al.*, “Effects of incorporating hybrid fibres on rheological and mechanical properties of fibre reinforced concrete”, *Construction & Building Materials*, v. 262, pp. 120561, 2020. doi: <https://doi.org/10.1016/j.conbuildmat.2020.120561>.
- [20] SOBUZ, M.H.R., SAHA, A., ANAMIKA, J.F., *et al.*, “Development of self-compacting concrete incorporating rice husk ash with waste galvanized copper wire fiber”, *Buildings (Basel, Switzerland)*, v. 12, n. 7, pp. 1024, 2022. doi: <https://doi.org/10.3390/buildings12071024>.
- [21] MANSOUR, W., OSAMA, B., LI, W., *et al.*, “Numerical Response of Concrete-Filled Steel Tubular (CFST) Columns Externally Strengthened with FRP Composites Subjected to Cyclic Loading”, *International Journal of Concrete Structures and Materials*, v. 18, n. 1, pp. 82, 2024. doi: <https://doi.org/10.1186/s40069-024-00716-6>.
- [22] WANG, X., HE, J., MOSALLAM, A.S., *et al.*, “The effects of fiber length and volume on material properties and crack resistance of basalt fiber reinforced concrete (BFRC)”, *Advances in Materials Science and Engineering*, v. 2019, n. 1, pp. 7520549, 2019. doi: <https://doi.org/10.1155/2019/7520549>.
- [23] AKID, A.S.M., WASIEW, Q.A., SOBUZ, M.H.R., *et al.*, “Flexural behavior of corroded reinforced concrete beam strengthened with jute fiber reinforced polymer”, *Advances in Structural Engineering*, v. 24, n. 7, pp. 1269–1282, 2021. doi: <https://doi.org/10.1177/1369433220974783>.
- [24] HASAN, N., SOBUZ, H.R., AUWALU, A.S., *et al.*, “Investigation into the suitability of kenaf fiber to produce structural concrete”, *Advanced Materials Letters*, v. 6, n. 8, pp. 731–737, 2015. doi: <https://doi.org/10.5185/amlett.2015.5818>.
- [25] RAHMAN SOBUZ, M.H., ALAM, A., JOHN OEHLERS, D., *et al.*, “Experimental and analytical studies of size effects on compressive ductility response of Ultra-High-Performance Fiber-Reinforced concrete”, *Construction & Building Materials*, v. 409, pp. 133864, 2023. doi: <https://doi.org/10.1016/j.conbuildmat.2023.133864>.
- [26] HAMA, S.M., AL-JUMAILY, I.A., DAHAMM, E.K., *et al.*, “Synergistic structural performance of green self-compacting concrete enhanced by nano-clay and waste glass powder”, *Scientific Reports*, v. 15, n. 1, pp. 45241, 2025. doi: <https://doi.org/10.1038/s41598-025-28595-1>. PubMed PMID: 41275037.
- [27] SOBUZ, M.H.R., ALZLFAWI, A., HAKEEM, I.Y., *et al.*, “Machine learning-based modeling to predict and parametrically optimize the compressive strength of nanomaterial concrete composites exposed to elevated temperatures”, *Case Studies in Construction Materials*, v. 23, pp. e05546, 2025. doi: <https://doi.org/10.1016/j.cscm.2025.e05546>.
- [28] SOBUZ, M.H.R., AYON, I.S., HASAN, N.M.S., *et al.*, “Analytical and Finite Element Modeling for Time-Dependent Deflection of Ultra-High-Performance Fiber-Reinforced Concrete Beam”, *Latin American Journal of Solids and Structures*, v. 22, n. 11, pp. e8655–e8655, 2025. doi: <https://doi.org/10.1590/1679-7825/e8655>.
- [29] SOBUZ, M.H.R., KABBO, M.K.I., KHAN, M.H., *et al.* “Predicting the compressive strength of sustainable cement mortar composite incorporating biochar: A machine learning modeling approach”, in Kapoor, R. T., Sillanpää, M. (eds), *Applications of artificial intelligence in removal of emerging contaminants*, Amsterdam, Elsevier, 2026, pp. 39–62. doi: <https://doi.org/10.1016/B978-0-443-26779-6.00007-3>.
- [30] FAYED, S., GHALLA, M., HAMODA, A., *et al.*, “Using steel mesh as internal reinforcement in the concrete supports subjected to partially compressive load”, *Scientific Reports*, v. 15, n. 1, pp. 42258, 2025. doi: <https://doi.org/10.1038/s41598-025-26266-9>. PubMed PMID: 41298725.
- [31] NAWAR, M.E., ALKHARISI, M.K., BAYOUMI, E.L.A., *et al.*, “Shear behavior of reinforced concrete beams with vertical holes externally strengthened with near surface mounted bars”, *Scientific Reports*, v. 15, n. 1, pp. 41746, 2025. doi: <https://doi.org/10.1038/s41598-025-27250-z>. PubMed PMID: 41291037.
- [32] SOBUZ, M.H.R., AHAMED, M.M., RAHMAN, S.M.A., *et al.*, “Experimental and machine learning evaluation of gamma versus non-gamma irradiated polyethylene terephthalate (PET) waste as a replacement of natural sand: Toward plastic waste reduction and sustainable development

- goals”, *Construction & Building Materials*, v. 494, pp. 143446, 2025. doi: <https://doi.org/10.1016/j.conbuildmat.2025.143446>.
- [33] CHEN, D., CHEN, Y., MA, L., *et al.* “A state of review on manufacturing and effectiveness of ultra-high-performance fiber reinforced concrete for long-term integrity of concrete structures”, *Advances in concrete construction*, v. 17, n. 5, pp. 293–310, 2024.
- [34] SOBUZ, H.R., AHMED, E., SUTAN, N.M., *et al.*, “Bending and time-dependent responses of RC beams strengthened with bonded carbon fiber composite laminates”, *Construction & Building Materials*, v. 29, pp. 597–611, 2012. doi: <https://doi.org/10.1016/j.conbuildmat.2011.11.006>.
- [35] HASAN, N.M.S., SHAURDHO, N.M.N., SOBUZ, M.H.R., *et al.*, “Utilization of waste glass cullet as partial substitutions of coarse aggregate to produce eco-friendly concrete: role of metakaolin as cement replacement”, *Sustainability (Basel)*, v. 15, n. 14, pp. 11254, 2023. doi: <https://doi.org/10.3390/su151411254>.
- [36] SOBUZ, M.H.R., JABIN, J.A., ASHRAF, J., *et al.*, “Enhancing Sustainable Concrete Production by Utilizing Fly Ash and Recycled Concrete Aggregate with Experimental Investigation and Machine Learning Modeling”, *Journal of Building Pathology and Rehabilitation*, v. 9, n. 2, pp. 134, 2024. doi: <https://doi.org/10.1007/s41024-024-00474-8>.
- [37] SOBUZ, M.H.R., KABBO, M.K.I., ALAHMARI, T.S., *et al.*, “Microstructural behavior and explainable machine learning aided mechanical strength prediction and optimization of recycled glass-based solid waste concrete”, *Case Studies in Construction Materials*, v. 22, pp. e04305, 2025. doi: <https://doi.org/10.1016/j.cscm.2025.e04305>.
- [38] ALAHMARI, T.S., ASHRAF, J., SOBUZ, M.H.R., *et al.*, “Predicting the compressive strength of fiber-reinforced self-consolidating concrete using a hybrid machine learning approach”, *Innovative Infrastructure Solutions*, v. 9, n. 11, pp. 446, 2024. doi: <https://doi.org/10.1007/s41062-024-01751-8>.
- [39] SHI, C., WU, Z., XIAO, J., *et al.*, “A review on ultra high performance concrete: Part I. Raw materials and mixture design”, *Construction & Building Materials*, v. 101, pp. 741–751, 2015. doi: <https://doi.org/10.1016/j.conbuildmat.2015.10.088>.
- [40] RAJ, S., KUMAR, V.R., KUMAR, B.B., *et al.*, “Basalt: Structural insight as a construction material”, *Sadhana*, v. 42, n. 1, pp. 75–84, 2017. doi: <https://doi.org/10.1007/s12046-016-0573-9>.
- [41] KOGISO, T., HIRSCHMANN, M.M., FROST, D.J., “High-pressure partial melting of garnet pyroxenite: possible mafic lithologies in the source of ocean island basalts”, *Earth and Planetary Science Letters*, v. 216, n. 4, pp. 603–617, 2003. doi: [https://doi.org/10.1016/S0012-821X\(03\)00538-7](https://doi.org/10.1016/S0012-821X(03)00538-7).
- [42] AHMED, E., SOBUZ, H.R., “Sobuz, H. R., “Flexural and time-dependent performance of palm shell aggregate concrete beam””, *KSCE Journal of Civil Engineering*, v. 15, n. 5, pp. 859–865, 2011. doi: <https://doi.org/10.1007/s12205-011-1148-2>.
- [43] YAN, L., CHOUW, N., “A comparative study of steel reinforced concrete and flax fibre reinforced polymer tube confined coconut fibre reinforced concrete beams”, *Journal of Reinforced Plastics and Composites*, v. 32, n. 16, pp. 1155–1164, 2013. doi: <https://doi.org/10.1177/0731684413487092>.
- [44] JAMSHAI, H., MILITKÝ, J., MISHRA, R., *et al.* “Basalt fibers and their composites”, *Novelties in Fibrous Material Science*, v. 2, 2017.
- [45] MILITKÝ, J., MISHRA, R., JAMSHAI, H. “Basalt fibers”, In: Bunsell, A. R. (ed), *Handbook of properties of textile and technical fibres*, Sawston, Woodhead Publishing, 2018, pp. 805-840. doi: <https://doi.org/10.1016/B978-0-08-101272-7.00020-1>.
- [46] ASADI, A., BAAIJ, F., MAINKA, H., *et al.*, “Basalt fibers as a sustainable and cost-effective alternative to glass fibers in sheet molding compound (SMC)”, *Composites. Part B, Engineering*, v. 123, pp. 210–218, 2017. doi: <https://doi.org/10.1016/j.compositesb.2017.05.017>.
- [47] MOROVA, N., “Investigation of usability of basalt fibers in hot mix asphalt concrete”, *Construction & Building Materials*, v. 47, pp. 175–180, 2013. doi: <https://doi.org/10.1016/j.conbuildmat.2013.04.048>.
- [48] ISLAM, M.S., HAMDAN, S., HASSAN, A., *et al.*, “The chemical modification of tropical wood polymer composites”, *Journal of Composite Materials*, v. 48, n. 7, pp. 783–789, 2014. doi: <https://doi.org/10.1177/0021998313477894>.
- [49] UDDIN, M.A., JAMEEL, M., SOBUZ, H.R., *et al.*, “Experimental study on strength gaining characteristics of concrete using Portland Composite Cement”, *KSCE Journal of Civil Engineering*, v. 17, n. 4, pp. 789–796, 2013. doi: <https://doi.org/10.1007/s12205-013-0236-x>.

- [50] LAM, N.N., VAN HUNG, L., “Mechanical and shrinkage behavior of basalt fiber reinforced ultra-high-performance concrete”, *GEOMATE Journal*, v. 20, n. 78, pp. 28–35, 2021. doi: <https://doi.org/10.21660/2021.78.86151>.
- [51] HIGH, C., SELIEM, H.M., EL-SAFETY, A., *et al.*, “Use of basalt fibers for concrete structures”, *Construction & Building Materials*, v. 96, pp. 37–46, 2015. doi: <https://doi.org/10.1016/j.conbuildmat.2015.07.138>.
- [52] DHAND, V., MITTAL, G., RHEE, K.Y., *et al.*, “A short review on basalt fiber reinforced polymer composites”, *Composites. Part B, Engineering*, v. 73, pp. 166–180, 2015. doi: <https://doi.org/10.1016/j.compositesb.2014.12.011>.
- [53] HUANG, H., LI, M., ZHANG, W., *et al.*, “Seismic behavior of a friction-type artificial plastic hinge for the precast beam–column connection”, *Archives of Civil and Mechanical Engineering*, v. 22, n. 4, pp. 201, 2022. doi: <https://doi.org/10.1007/s43452-022-00526-1>.
- [54] HUANG, H., LI, M., YUAN, Y., *et al.*, “Theoretical analysis on the lateral drift of precast concrete frame with replaceable artificial controllable plastic hinges”, *Journal of Building Engineering*, v. 62, pp. 105386, 2022. doi: <https://doi.org/10.1016/j.jobbe.2022.105386>.
- [55] JABIN, J.A., KHONDOKER, M.T.H., SOBUZ, M.H.R., *et al.*, “High-temperature effect on the mechanical behavior of recycled fiber-reinforced concrete containing volcanic pumice powder: An experimental assessment combined with machine learning (ML)-based prediction”, *Construction & Building Materials*, v. 418, pp. 135362, 2024. doi: <https://doi.org/10.1016/j.conbuildmat.2024.135362>.
- [56] ALAHMARI, T.S., KABBO, M.K.I., SOBUZ, M.H.R., *et al.*, “Experimental assessment and hybrid machine learning-based feature importance analysis with the optimization of compressive strength of waste glass powder-modified concrete”, *Materials Today. Communications*, v. 44, pp. 112081, 2025. doi: <https://doi.org/10.1016/j.mtcomm.2025.112081>.
- [57] KHARUN, M., KOROTEEV, D. “Effect of basalt fibres on the parameters of fracture mechanics of MB modifier based high-strength concrete”, In: *VI International Scientific Conference “Integration, Partnership and Innovation in Construction Science and Education*, Moscow, 2018. doi: <https://doi.org/10.1051/mateconf/201825102003>.
- [58] YANG, L., XIE, H., FANG, S., *et al.*, “Experimental study on mechanical properties and damage mechanism of basalt fiber reinforced concrete under uniaxial compression”, *Structures (Oxford)*, v. 31, pp. 330–340, 2021. doi: <https://doi.org/10.1016/j.istruc.2021.01.071>.
- [59] BIRADAR, S.V., DILEEP, M.S., GOWRI, T.V., “Studies of concrete mechanical properties with basalt fibers”, *IOP Conference Series. Materials Science and Engineering*, v. 1006, n. 1, pp. 012031, 2020. doi: <https://doi.org/10.1088/1757-899X/1006/1/012031>.
- [60] KATKHUDA, H., SHATARAT, N., “Improving the mechanical properties of recycled concrete aggregate using chopped basalt fibers and acid treatment”, *Construction & Building Materials*, v. 140, pp. 328–335, 2017. doi: <https://doi.org/10.1016/j.conbuildmat.2017.02.128>.
- [61] AITCIN, P., “The durability characteristics of high performance concrete: a review”, *Cement and Concrete Composites*, v. 25, n. 4-5, pp. 409–420, 2003. doi: [https://doi.org/10.1016/S0958-9465\(02\)00081-1](https://doi.org/10.1016/S0958-9465(02)00081-1).
- [62] YANG, L.T., GUO, M. *High-performance computing: paradigm and infrastructure*, Hoboken, John Wiley & Sons, 2006.
- [63] AYUB, T., SHAFIQ, N., NURUDDIN, M.F., “Mechanical properties of high-performance concrete reinforced with basalt fibers”, *Procedia Engineering*, v. 77, pp. 131–139, 2014. doi: <https://doi.org/10.1016/j.proeng.2014.07.029>.
- [64] MOHAMMADI MOHAGHEGH, A., SILFWERBRAND, J., ÅRSKOG, V., “Shear behavior of high-performance basalt fiber concrete—Part I: Laboratory shear tests on beams with macro fibers and bars”, *Structural Concrete*, v. 19, n. 1, pp. 246–254, 2018. doi: <https://doi.org/10.1002/suco.201700208>.
- [65] KHARUN, M., AL ARAZA, H.A., HEMATIBAHAR, M., *et al.* “Experimental study on the effect of chopped basalt fiber on the mechanical properties of high-performance concrete”, *AIP Conference Proceedings*, v. 2559, n. 1, 2022. doi: <https://doi.org/10.1063/5.0099042>.
- [66] XU, L., WANG, X., WANG, Z., *et al.*, “Hybrid quantum genetic algorithm for structural damage identification”, *Computer Methods in Applied Mechanics and Engineering*, v. 438, pp. 117866, 2025. doi: <https://doi.org/10.1016/j.cma.2025.117866>.
- [67] ZHOU, Z., TU, H., YANG, J., *et al.*, “Prediction and Analysis of Slurry Pressure in the Upper Cutting Face of a Shield Based on a GA-APSO-RF Ensemble Model”, *International Journal of Geomechanics*, v. 25, n. 12, pp. 04025292, 2025. doi: <https://doi.org/10.1061/IJGNAI.GMENG-11572>.

- [68] SOBUZ, M.H.R., RAHMAN, M.M., AAYAZ, R., *et al.*, “Combined influence of modified recycled concrete aggregate and metakaolin on high-strength concrete production: Experimental assessment and machine learning quantifications with advanced SHAP and PDP analyses”, *Construction & Building Materials*, v. 461, pp. 139897, 2025. doi: <https://doi.org/10.1016/j.conbuildmat.2025.139897>.
- [69] SOBUZ, M.H.R., ADITTO, F.S., DATTA, S.D., *et al.*, “High-strength self-compacting concrete production incorporating supplementary cementitious materials: experimental evaluations and machine learning modelling”, *International Journal of Concrete Structures and Materials*, v. 18, n. 1, pp. 67, 2024. doi: <https://doi.org/10.1186/s40069-024-00707-7>.
- [70] ALDREES, A., SIDDIQ, B., AL-RASHED, W.S., *et al.*, “Leveraging machine learning to model salinity and water flux for improved insights into forward osmosis membrane bioreactors”, *Journal of Water Process Engineering*, v. 68, pp. 106585, 2024. doi: <https://doi.org/10.1016/j.jwpe.2024.106585>.
- [71] LONG, X., MAO, M., SU, T., *et al.*, “Machine learning method to predict dynamic compressive response of concrete-like material at high strain rates”, *Defence Technology*, v. 23, pp. 100–111, 2023. doi: <https://doi.org/10.1016/j.dt.2022.02.003>.
- [72] MOMSHAD, A.M., ISLAM, M.H., DATTA, S.D., *et al.*, “Assessing the engineering properties and environmental impact with explainable machine learning analysis of sustainable concrete utilizing waste banana leaf ash as a partial cement replacement”, *Cleaner Engineering and Technology*, v. 24, pp. 100886, 2025. doi: <https://doi.org/10.1016/j.clet.2025.100886>.
- [73] SOBUZ, M.H.R., SALEH, M.A., SAMIUN, M., *et al.*, “AI-driven modeling for the optimization of concrete strength for low-cost business production in the USA construction industry”, *Engineering, Technology and Applied Science Research*, v. 15, n. 1, pp. 20529–20537, 2025. doi: <https://doi.org/10.48084/etasr.9733>.
- [74] SOBUZ, M.H.R., KHATUN, M., KABBO, M.K.I., *et al.*, “An explainable machine learning model for encompassing the mechanical strength of polymer-modified concrete”, *Asian Journal of Civil Engineering*, v. 26, n. 2, pp. 931–954, 2024. doi: <https://doi.org/10.1007/s42107-024-01230-6>.
- [75] MISHRA, M., SOBUZ, M.H.R., KABBO, M.K.I., *et al.*, “Experimental assessment and data-driven hybrid machine learning quantification with parametric optimization of compressive strength of ceramic waste concrete”, *Case Studies in Construction Materials*, v. 23, pp. e05043, 2025. doi: <https://doi.org/10.1016/j.cscm.2025.e05043>.
- [76] SU, M., ZHONG, Q., PENG, H., *et al.*, “Selected machine learning approaches for predicting the interfacial bond strength between FRPs and concrete”, *Construction & Building Materials*, v. 270, pp. 121456, 2021. doi: <https://doi.org/10.1016/j.conbuildmat.2020.121456>.
- [77] NGUYEN, K.T., NGUYEN, Q.D., LE, T.A., *et al.*, “Analyzing the compressive strength of green fly ash based geopolymer concrete using experiment and machine learning approaches”, *Construction & Building Materials*, v. 247, pp. 118581, 2020. doi: <https://doi.org/10.1016/j.conbuildmat.2020.118581>.
- [78] ULLAH, H.S., KHUSHNOOD, R.A., FAROOQ, F., *et al.*, “Prediction of compressive strength of sustainable foam concrete using individual and ensemble machine learning approaches”, *Materials (Basel)*, v. 15, n. 9, pp. 3166, 2022. doi: <https://doi.org/10.3390/ma15093166>. PubMed PMID: 35591498.
- [79] LIU, Y., “High-performance concrete strength prediction based on machine learning”, *Computational Intelligence and Neuroscience*, v. 2022, n. 1, pp. 5802217, 2022. doi: <https://doi.org/10.1155/2022/5802217>. PubMed PMID: 35669631.
- [80] REZA KASHYZADEH, K., AMIRI, N., GHORBANI, S., *et al.*, “Prediction of concrete compressive strength using a back-propagation neural network optimized by a genetic algorithm and response surface analysis considering the appearance of aggregates and curing conditions”, *Buildings (Basel, Switzerland)*, v. 12, n. 4, pp. 438, 2022. doi: <https://doi.org/10.3390/buildings12040438>.
- [81] ONYELOWE, K.C., EBID, A.M., HANANDEH, S., *et al.*, “Modeling the compressive strength behavior of concrete reinforced with basalt fiber”, *Scientific Reports*, v. 15, n. 1, pp. 11493, 2025. doi: <https://doi.org/10.1038/s41598-025-96343-6>. PubMed PMID: 40181176.
- [82] PAL, A., AHMED, K.S., HOSSAIN, F.Z., *et al.*, “Machine learning models for predicting compressive strength of fiber-reinforced concrete containing waste rubber and recycled aggregate”, *Journal of Cleaner Production*, v. 423, pp. 138673, 2023. doi: <https://doi.org/10.1016/j.jclepro.2023.138673>.
- [83] NAJM, O.F., MOHAMED, O.A., ALZARD, M.H., “The efficiency of statistical and artificial neural network techniques on evaluating and predicting compressive strength of sustainable SCC with basalt fibers”, *Materials Today: Proceedings*, ahead of print, 2023. doi: <https://doi.org/10.1016/j.matpr.2023.03.185>.

- [84] ALZLFAWI, A., SOBUZ, M.H.R., KAWSARUL, M., *et al.*, “Prediction and parametric modeling of compressive strength of waste marble dust concrete through machine learning and experimental analysis”, *Scientific Reports*, v. 15, n. 1, pp. 42982, 2025. doi: <https://doi.org/10.1038/s41598-025-26978-y>. PubMed PMID: 41330987.
- [85] AL-RASHED, W.S., “Assessing public perception and parametric analysis of municipal solid waste management solutions in Tabuk City, Saudi Arabia”, *Scientific Reports*, v. 15, n. 1, pp. 11304, 2025. doi: <https://doi.org/10.1038/s41598-025-94525-w>. PubMed PMID: 40175483.
- [86] KABBO, M.K.I., SOBUZ, M.H.R., ADITTO, F.S., *et al.*, “Experimental assessment and machine learning quantification of structural eco-cellular lightweight concrete incorporating waste marble powder and silica fume”, *Journal of Building Engineering*, v. 105, pp. 112557, 2025. doi: <https://doi.org/10.1016/j.jobbe.2025.112557>.
- [87] UDDIN, M.A., SOBUZ, M.H.R., KABBO, M.K.I., *et al.*, “Predicting the mechanical performance of industrial waste incorporated sustainable concrete using hybrid machine learning modeling and parametric analyses”, *Scientific Reports*, v. 15, n. 1, pp. 26330, 2025. doi: <https://doi.org/10.1038/s41598-025-11601-x>. PubMed PMID: 40685423.
- [88] RAHMAN SOBUZ, M.H., CHOWDHURY TILAK, M.K., RAHMAN, S.M.A., *et al.*, “Comparative performance evaluation of quartz and snail shell powders modified concrete: Mechanical, machine learning, and microstructural assessments”, *Journal of Materials Research and Technology*, v. 36, pp. 7766–7782, 2025. doi: <https://doi.org/10.1016/j.jmrt.2025.04.296>.
- [89] SAFAYET, M.A., ABDULLAH, A.A., ISLAM, M.F., *et al.*, “Combined influence of eggshell powder and nylon fiber on self-compacting concrete production: experimental assessment and machine learning quantifications”, *Materials Research Express*, v. 12, n. 2, pp. 025003, 2025. doi: <https://doi.org/10.1088/2053-1591/adb0a6>.
- [90] YANG, G., ZHAO, H., HU, Z., *et al.*, “Prediction of restrained stress for UHPC: Considering relationship between long-term and in-situ creep”, *Construction & Building Materials*, v. 484, pp. 141722, 2025. doi: <https://doi.org/10.1016/j.conbuildmat.2025.141722>.
- [91] ZHENG, Y., LI, J., LI, J., “Determination of Initial Void Ratio of Remolded Clay Using Bayesian Data Fusion Approach”, *Geotechnical Testing Journal*, v. 48, n. 3, pp. 346–363, 2025. doi: <https://doi.org/10.1520/GTJ20240111>.
- [92] QIAN, K., CUI, L., DENG, X., *et al.*, “Experimental and Analytical Study of BFRP Bar Reinforced UHPC Beams Under Static and Impact Loading”, *International Journal of Impact Engineering*, v. 206, pp. 105456, 2025. doi: <https://doi.org/10.1016/j.ijimpeng.2025.105456>.
- [93] YANG, Y., CAO, J., LI, Y., *et al.*, “Effect of electric field on water-film between concrete and formwork: Atomic-scale mechanisms and technology indicators”, *Construction & Building Materials*, v. 498, pp. 143979, 2025. doi: <https://doi.org/10.1016/j.conbuildmat.2025.143979>.
- [94] HAO, X.-K., ZHANG, H.-Y., DENG, T., *et al.*, “Experimental and theoretical investigation of low-shrinkage alkali-activated materials permanent formwork reinforced concrete prisms under axial load”, *Construction & Building Materials*, v. 500, pp. 144156, 2025. doi: <https://doi.org/10.1016/j.conbuildmat.2025.144156>.
- [95] ZHANG, W., YANG, X., LIN, J., *et al.*, “Experimental and numerical study on the torsional behavior of rectangular hollow reinforced concrete columns strengthened By CFRP”, *Structures (Oxford)*, v. 70, pp. 107690, 2024. doi: <https://doi.org/10.1016/j.istruc.2024.107690>.
- [96] SUN, L., WANG, X., ZHANG, C., “Three-dimensional high fidelity mesoscale rapid modelling algorithm for concrete”, *Structures (Oxford)*, v. 70, pp. 107561, 2024. doi: <https://doi.org/10.1016/j.istruc.2024.107561>.
- [97] YANG, H., LI, S., SHU, J., *et al.*, “Data-driven high-resolution total focus imaging from array ultrasonic time-domain signals of reinforced concrete material”, *Construction & Building Materials*, v. 492, pp. 143048, 2025. doi: <https://doi.org/10.1016/j.conbuildmat.2025.143048>.
- [98] ZHOU, Z., YANG, J., SONG, J., *et al.*, “Research on shield parameter prediction method based on multivariate variable weight combination model”, *Tunnelling and Underground Space Technology*, v. 168, pp. 107189, 2026. doi: <https://doi.org/10.1016/j.tust.2025.107189>.
- [99] SUN, S., JIANG, Z., LI, L., *et al.*, “DEM analysis on rock-breaking impact effect of shield disc cutter in typical soft and hard composite strata”, *International Journal for Numerical and Analytical Methods in Geomechanics*, v. 49, n. 10, pp. 2414–2431, 2025. doi: <https://doi.org/10.1002/nag.3991>.

- [100] SOBUZ, M.H.R., KABBO, M.K.I., ALZLFAWI, A., *et al.*, “A supervised machine learning modeling with parametric optimization of the compressive strength of graphene-nano-engineered concrete”, *Matéria (Rio de Janeiro)*, v. 30, pp. e20240989, 2025. doi: <https://doi.org/10.1590/1517-7076-rmat-2024-0989>.
- [101] MALASHIN, I., TYNCHENKO, V., GANTIMUROV, A., *et al.*, “Boosting-Based Machine Learning Applications in Polymer Science: A Review”, *Polymers*, v. 17, n. 4, pp. 499, 2025. doi: <https://doi.org/10.3390/polym17040499>. PubMed PMID: 40006161.
- [102] YE, J., CHOW, J.-H., CHEN, J., *et al.* “Stochastic gradient boosted distributed decision trees”, In *Proceedings of the 18th ACM conference on Information and knowledge management*, Hong Kong, 2–6 November 2009.
- [103] DEV, V.A., EDEN, M.R., “Formation lithology classification using scalable gradient boosted decision trees”, *Computers & Chemical Engineering*, v. 128, pp. 392–404, 2019. doi: <https://doi.org/10.1016/j.compchemeng.2019.06.001>.
- [104] NOVIANDY, T. R., IDROES, G. M., HARDI, I., “Integrating explainable artificial intelligence and light gradient boosting machine for glioma grading”, *Informatics and Health*, v. 2, pp. 1–8, 2025. doi: <https://doi.org/10.1016/j.infoh.2024.12.001>
- [105] OSMAN, M., HE, J., MOKBAL, F.M.M., *et al.*, “ML-LGBM: A machine learning model based on light gradient boosting machine for the detection of version number attacks in RPL-based networks”, *IEEE Access : Practical Innovations, Open Solutions*, v. 9, pp. 83654–83665, 2021. doi: <https://doi.org/10.1109/ACCESS.2021.3087175>.
- [106] CHU, Z., YU, J., HAMDULLA, A., “LPG-model: A novel model for throughput prediction in stream processing, using a light gradient boosting machine, incremental principal component analysis, and deep gated recurrent unit network”, *Information Sciences*, v. 535, pp. 107–129, 2020. doi: <https://doi.org/10.1016/j.ins.2020.05.042>.
- [107] NAGASSOU, M., MWANGI, R.W., NYARIGE, E., “A hybrid ensemble learning approach utilizing light gradient boosting machine and category boosting model for lifestyle-based prediction of type-II diabetes mellitus”, *Journal of Data Analysis and Information Processing*, v. 11, n. 4, pp. 480–511, 2023. doi: <https://doi.org/10.4236/jdaip.2023.114025>.
- [108] CHA, G.-W., MOON, H.-J., KIM, Y.-C., “Comparison of random forest and gradient boosting machine models for predicting demolition waste based on small datasets and categorical variables”, *International Journal of Environmental Research and Public Health*, v. 18, n. 16, pp. 8530, 2021. doi: <https://doi.org/10.3390/ijerph18168530>. PubMed PMID: 34444277.
- [109] JAVAID, N., AKBAR, M., ALDEGHEISH, A., *et al.*, “Employing a machine learning boosting classifiers based stacking ensemble model for detecting non technical losses in smart grids”, *IEEE Access : Practical Innovations, Open Solutions*, v. 10, pp. 121886–121899, 2022. doi: <https://doi.org/10.1109/ACCESS.2022.3222883>.
- [110] SIBINDI, R., MWANGI, R.W., WAITITU, A.G., “A boosting ensemble learning based hybrid light gradient boosting machine and extreme gradient boosting model for predicting house prices”, *Engineering Reports*, v. 5, n. 4, pp. e12599, 2023. doi: <https://doi.org/10.1002/eng2.12599>.
- [111] ZHANG, X., DENG, Q., JIA, M., *et al.*, “Integrating a hybrid data processing strategy into an optimized light gradient boosting machine for photovoltaic power forecasting”, *Neural Computing & Applications*, v. 37, n. 3, pp. 1633–1650, 2025. doi: <https://doi.org/10.1007/s00521-024-10691-4>.
- [112] CHAKRABORTY, D., ELHEGAZY, H., ELZARKA, H., *et al.*, “A novel construction cost prediction model using hybrid natural and light gradient boosting”, *Advanced Engineering Informatics*, v. 46, pp. 101201, 2020. doi: <https://doi.org/10.1016/j.aei.2020.101201>.
- [113] EKANAYAKE, I., MEDDAGE, D., RATHNAYAKE, U., “A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP)”, *Case Studies in Construction Materials*, v. 16, pp. e01059, 2022. doi: <https://doi.org/10.1016/j.cscm.2022.e01059>.
- [114] ULLAH, I., LIU, K., YAMAMOTO, T., *et al.*, “Modeling of machine learning with SHAP approach for electric vehicle charging station choice behavior prediction”, *Travel Behaviour & Society*, v. 31, pp. 78–92, 2023. doi: <https://doi.org/10.1016/j.tbs.2022.11.006>.
- [115] AHMED, U., JIANGBIN, Z., ALMOGREN, A., *et al.*, “Hybrid bagging and boosting with SHAP based feature selection for enhanced predictive modeling in intrusion detection systems”, *Scientific Reports*, v. 14, n. 1, pp. 30532, 2024. doi: <https://doi.org/10.1038/s41598-024-81151-1>. PubMed PMID: 39690165.

- [116] MOLNAR, C., *et al.* “Relating the partial dependence plot and permutation feature importance to the data generating process”, In: World Conference on Explainable Artificial Intelligence, Lisbon, 26–28 July 2023. doi: https://doi.org/10.1007/978-3-031-44064-9_24.
- [117] GNAT, S., “Determining real estate features influence on prices with partial dependence plots: a case study in Szczecin, Poland”, *Real Estate Management and Valuation*, v. 32, n. 4, pp. 105–116, 2024. doi: <https://doi.org/10.2478/remav-2024-0039>.
- [118] ALHARTHI, A., ALFARES, A., Sha’aban, Y.A., *et al.*, “Improving the interpretability of ANN-based predictions of lattice constants in aliovalently doped perovskites using partial dependence plots”, *Crystals*, v. 15, n. 6, pp. 538, 2025. doi: <https://doi.org/10.3390/cryst15060538>.
- [119] TAŞAN, M., TAŞAN, S., DEMIR, Y., “Estimation and uncertainty analysis of groundwater quality parameters in a coastal aquifer under seawater intrusion: a comparative study of deep learning and classic machine learning methods”, *Environmental Science and Pollution Research International*, v. 30, n. 2, pp. 2866–2890, 2023. doi: <https://doi.org/10.1007/s11356-022-22375-4>. PubMed PMID: 35941499.
- [120] SINGH, M., CHOUDHARY, P., BEDI, A.K., *et al.*, “Compressive strength estimation of waste marble powder incorporated concrete using regression modelling”, *Coatings*, v. 13, n. 1, pp. 66, 2022. doi: <https://doi.org/10.3390/coatings13010066>.
- [121] MILLER, S.A., JOHN, V.M., PACCA, S.A., *et al.*, “Carbon dioxide reduction potential in the global cement industry by 2050”, *Cement and Concrete Research*, v. 114, pp. 115–124, 2018. doi: <https://doi.org/10.1016/j.cemconres.2017.08.026>.
- [122] KASHEM, A., KARIM, R., MALO, S.C., *et al.*, “Hybrid data-driven approaches to predicting the compressive strength of ultra-high-performance concrete using SHAP and PDP analyses”, *Case Studies in Construction Materials*, v. 20, pp. e02991, 2024. doi: <https://doi.org/10.1016/j.cscm.2024.e02991>.
- [123] SHAFIGHFARD, T., KAZEMI, F., ASGARKHANI, N., *et al.*, “Machine-learning methods for estimating compressive strength of high-performance alkali-activated concrete”, *Engineering Applications of Artificial Intelligence*, v. 136, pp. 109053, 2024. doi: <https://doi.org/10.1016/j.engappai.2024.109053>.
- [124] SOBUZ, M.H.R., KABBO, M.K.I., KHATUN, M., *et al.*, “Microstructural assessment and supervised machine learning-aided modeling to explore the potential of quartz powder as an alternate binding material in concrete”, *Case Studies in Construction Materials*, v. 22, pp. e04568, 2025. doi: <https://doi.org/10.1016/j.cscm.2025.e04568>.
- [125] ALYAMI, M., ONYELOWE, K., ALATEAH, A.H., *et al.*, “Innovative hybrid machine learning models for estimating the compressive strength of copper mine tailings concrete”, *Case Studies in Construction Materials*, v. 21, pp. e03869, 2024. doi: <https://doi.org/10.1016/j.cscm.2024.e03869>.
- [126] ULLAH, I., JAVED, M.F., ALABDULJABBAR, H., *et al.*, “Estimating the compressive and tensile strength of basalt fibre reinforced concrete using advanced hybrid machine learning models”, *Structures (Oxford)*, v. 71, pp. 108138, 2025. doi: <https://doi.org/10.1016/j.istruc.2024.108138>.
- [127] ALYAMI, M., ULLAH, I., ALATEAH, A.H., *et al.*, “Machine learning models for predicting the compressive strength of cement-based mortar materials: Hyper tuning and optimization”, *Structures (Oxford)*, v. 71, pp. 107931, 2025. doi: <https://doi.org/10.1016/j.istruc.2024.107931>.