Research Article

Limei Fu* and Feng Xu

Establishing strength prediction models for low-carbon rubberized cementitious mortar using advanced AI tools

https://doi.org/10.1515/rams-2025-0135
received February 26, 2025; accepted July 14, 2025

Abstract: Rubberized cementitious composites have emerged as a sustainable alternative in the construction sector by promoting circular economy principles. However, their reduced compressive strength (CS) due to the inclusion of rubber remains a significant barrier to widespread adoption. While several experimental studies exist, there is a clear gap in utilizing data-driven strategies to efficiently predict and optimize the strength performance of such materials. This research addresses this gap by evaluating the predictability of machine learning approaches for evaluating the CS of rubberized mortar (RM) incorporating supplementary cementitious materials. Among the tested algorithms, including bagging, gradient boosting, and AdaBoost, the bagging model achieved the highest accuracy (R^2 = 0.975). SHapley Additive exPlanations analysis further identified cement and sand content as the most influential variables affecting CS. The findings were integrated into a graphical user interface for practical, real-time strength estimation. This tool can support engineers and material designers in developing sustainable RM mixes with improved performance and reduced reliance on extensive laboratory testing.

Keywords: rubberized mortar, compressive strength, supplementary cementitious materials

1 Introduction

Several natural resources, including sand, aggregates, and cement, are consumed during the production of concrete [1]. These construction ingredients are manufactured from non-renewable natural resources. The continual consumption of these natural resources is unsustainable. Despite the fact that the supply of these natural sources is vast, and a scarcity of supplies might not be immediately apparent, it is imperative to identify a sustainable and environmentally friendly development path. Research on sustainable and green construction materials has been initiated, involving novel types of materials to address the aforementioned concerns [2–4]. Rubber concrete has the potential to reduce black pollution, which is a result of the accumulation of refuse rubber [5,6]. Since the inception of the twenty-first century, the most severe pollution observed has been black pollution. White pollution is distinguished from black pollution. Black pollution is caused by refuse from rubber products that are difficult to dispose of, whereas the plastic industry produces white pollution. The modernization of rubber inventions has resulted in a substantial quantity of refuse rubber, including automobile tires. There is no effective method of treating waste rubber, as its decomposition is complicated. Regardless of whether this waste is recycled, this will result in secondary pollution of nature [7]. The absence of efficient treatment results in the collection of refuse rubber, which poses a significant environmental issue. Rubber is employed in a significant number of industrial products, such as those utilized in the electronic and automobile sectors [8]. The most significant challenges linked with refuse rubber are its chemical stability and challenging degradation [9]. At present, the most prevalent treatment method is the recycling of waste rubber [10,11].

Concrete is a major building material that has a diverse array of uses in a variety of sectors [12,13]. Rubberized cementitious composites have contributed to the development of a green built environment in recent years. Rubberized concrete is manufactured by grinding waste tires into rubber powder or coarse particles to replace a portion of the sand and/or aggregates in traditional concrete [14,15]. The prevalent use of rubberized composites provides a way to recycle higher volumes of waste tires and decrease the use of natural resources (sand and aggregates). It also generates socio-economic advantages [16,17]. The mechanical strength of

* **Corresponding author: Limei Fu,** Accounting College, Hainan Vocational University of Science and Technology, Haikou, Hainan, 571126, China, e-mail: 20213125400013@hainnu.edu.cn
Feng Xu: Accounting College, Hainan Vocational University of Science and Technology, Haikou, Hainan, 571126, China

rubberized cementitious composites needs to be investigated for structural applications. This is one of the critical performance criteria that must be investigated in order to implement rubberized cementitious composites in the construction sector. Prior research has shown that rubberized cementitious composites can be employed in non-structural members, such as pre-cast bricks/blocks for partition walls, due to their lower strength compared to conventional mortar or concrete [18]. Hence, rubberized cementitious composites are irreplaceable compared to conventional cementitious composites with reduced strength issues. Therefore, further investigations are required to increase the strength properties of rubberized cement-based materials using sustainable approaches.

Studies have been carried out using cement alternatives using supplementary cementitious materials (SCMs), such as a variety of industrial and agricultural residues [19–22]. In the presence of a pore solution, the process of cement hydration can occur either pozzolanically or hydraulically, contingent upon the presence of SCMs [23,24]. The integration of industrial waste such as marble and glass powders and silica fume as SCMs in cement-based materials is an appropriate method for enhancing the sustainability of construction [25–28]. This method has the potential to reduce the necessity for cement, reduce CO_2 discharges, and efficiently manage waste disposal concerns. Additionally, the material strength was enhanced as a consequence of the utilization of these residues as SCMs [29,30]. Despite the observation of the individual use of marble and glass powders and silica fume as SCMs in cement-based materials, there has been limited exploration into their combined effect on cement concrete or mortar that incorporates rubber as a fractional sand substitute [31]. However, further investigations are needed to explore the feasibility of these SCMs in rubberized mortar (RM).

Thorough experimentation and human effort are necessary for traditional experimental approaches to test the strength of building materials [32]. The studied period is lengthy and affected by several uncertain considerations, such as test settings and controlling measures [33]. Thus, it is required to build a reasonable and effective method to examine the strength of rubberized cementitious composites. Recently, several researchers have shown interest in the evaluation of material characteristics using machine learning (ML) methods [34–36]. ML techniques are often categorized as artificial intelligence-based algorithms that assess data to enable the decision process [37,38]. It is an artificial intelligence technique that has the capacity to understand and predict data [39]. ML methods might be employed to predict the RM attributes that incorporate various mix proportions [40]. Ensemble ML techniques are recognized as mostly reliable methods, as these techniques combine several single basic learners to reduce the weaknesses of the individual methods. The technique's generalization capability is improved by diverse perspectives by typical ensemble learning techniques, which include the bagging (Bg), boosting, and stacking procedures. These techniques exhibit better performance related to empirical techniques and individual approaches [41,42]. Hence, it is necessary to study how ML models accurately predict the strength of RM incorporating SCMs and which material components most influence this strength performance. Addressing this question is essential because it offers a pathway to reduce dependence on labor-intensive experimental procedures, while also enabling rapid mix design optimization.

In order to fill this gap, the compressive strength (CS) of RM was assessed in this research using ML methods, which were obtained from the results of laboratory tests. The study's objective was accomplished through the application of gradient boosting (GB), AdaBoost (AB), and Bg ML techniques. The approximation precision of each model was assessed by comparing the projected results with the real test outcomes and implementing statistical tests. It takes significant time, effort, and money to conduct experiments because of all the steps involved in the procedure, including obtaining the necessary ingredients, casting the samples, curing them, and testing them. Furthermore, conventional testing methods present challenges in evaluating the combined influences of input parameters on the CS of RM. SHapley Additive exPlanations (SHAP) analysis was implemented in this investigation to investigate the association between the CS of RM and the input parameters. In this investigation, an experimental data sample was used to assess the efficacy of ML strategies in calculating the CS of RM. The relative influence of raw constituents, taken as inputs, was also determined by using SHAP analysis that can aid in optimizing mix proportions. Finally, based on the optimum ML model, a graphical user interface (GUI) was built to evaluate the RM CS, thereby eliminating further lab tests.

2 Research methods

2.1 Data description

A widespread data sample of 408 points on the CS of RM was generated from experimental procedures in the current study, as described in a recent study [31]. The raw materials for RM were cement, sand, water, superplasticizer, silica

fume, glass powder, and marble powder. Rubber particles similar to sand size were used in proportions of 5–25 wt% as sand substitutes. Three distinct SCMs, including marble powder, silica fume, and glass powder, were used as cement substitutes in proportions of 5–25 wt% individually and 5–15 wt% in combinations. In all batches, the superplasticizer dosage was maintained at 1.25 wt% by cement, the water-to-binder ratio was maintained at 0.28, and the cement-to-sand ratio was maintained at 1:1.5. The aforementioned information pertains to mix proportions that were implemented in experimental tests, which generated the data sample necessary for the modeling in the current study. The 50 mm cubes were cast using a motorized mixer to combine the constituents of RM in order to assess their CS. After casting, the specimens were stored in molds for 24 h prior to being demolded and water cured for 28 days [31]. The CS tests were performed following ASTM C109/C109M-20 [43].

ML models were constructed and refined through the pre-processing of the dataset, which is a critical stage. Processing missing data, detection of outliers, encoding, and splitting data are among the tasks that are frequently employed in data preprocessing [44]. No outliers and missing data were detected in the dataset. Additionally, the composition or type of constituents utilized in each blend was consistent. This research did not use any multivariate outlier detection approach; however, the absence of outliers was guaranteed using a comprehensive data pretreatment strategy. A method of univariate outlier detection was performed on each input prior to the training of the ML models. This involved the detection and elimination of data elements that were beyond the pre-established acceptable limits. Additionally, the database was subjected to a comprehensive statistical and visual analysis to detect and solve any probable outliers.

The six input parameters utilized for modeling using ML techniques were the quantities of raw constituents, which included cement ($kg{\cdot}m^{-3}$), sand ($kg{\cdot}m^{-3}$), rubber ($kg{\cdot}m^{-3}$), silica fume ($kg{\cdot}m^{-3}$), marble powder ($kg{\cdot}m^{-3}$), and glass powder ($kg{\cdot}m^{-3}$). The quantities of water and superplasticizers were constant throughout the dataset. The CS in MPa of samples was employed as an output in the modeling process. In order for the ML methods to generate precise predictions, it has been recommended that the ratio of the number of data points to the number of input parameters should be at least 5 [45,46]. Using a data sample of 408 points and six inputs, the ratio is 81.6, which is substantially higher than in this research. Consequently, the dataset was deemed appropriate for ML modeling.

Table 1 summarizes the statistical data for both inputs and outputs, and Figure 1 illustrates their frequency distribution. Statistical measures are employed to provide the minimum and maximum values for parameters. The statistical distribution metrics of mode, mean, standard deviation, kurtosis, median, sample variance, standard error, and range are also included in Table 1. The occurrence of more values near or equal to zero resulted in the median and mode values for silica fume, glass, and marble powder being zero. The database contains values close to zero for silica fume, glass, and marble powder, which can be ascribed to the fact that these constituents are present in minimal quantities in the RM mixtures. The data dispersal for cement is depicted in Figure 1(a), which suggests that it has two peak values. The diminished values suggest that it has been replaced by other materials, such as rubber. Figure 1(b) contains the dispersal of sand data with values ranging from 920 to 1,240 $kg{\cdot}m^{-3}$. The pattern of rubber values shows five distinct values, as illustrated in Figure 1(c). The dispersal of silica fume, glass, and marble powder

Table 1: Descriptive statistics of inputs and outputs

Parameter	Cement ($kg{\cdot}m^{-3}$)	Sand ($kg{\cdot}m^{-3}$)	Rubber ($kg{\cdot}m^{-3}$)	Silica fume ($kg{\cdot}m^{-3}$)	Marble powder ($kg{\cdot}m^{-3}$)	Glass powder ($kg{\cdot}m^{-3}$)	CS (MPa)
Mean	688.5	1054.8	184.6	45.5	45.5	45.5	34.0
Standard error	3.5	4.4	4.4	2.9	2.9	2.9	0.3
Median	701.3	1054.0	186.0	0.0	0.0	0.0	34.5
Mode	742.5	1178.0	62.0	0.0	0.0	0.0	36.8
Standard deviation	70.9	89.2	88.9	58.7	58.7	58.7	5.7
Sample variance	5029.8	7955.9	7903.4	3443.6	3443.6	3443.6	32.3
Kurtosis	−1.0	−1.3	−1.3	0.3	0.3	0.3	−0.6
Range	247.5	320.0	310.0	206.3	206.3	206.3	28.2
Minimum	577.5	920.0	0.0	0.0	0.0	0.0	20.3
Maximum	825.0	1240.0	310.0	206.3	206.3	206.3	48.5
Sum	280912.5	430350.0	75330.0	18562.5	18562.5	18562.5	13870.0

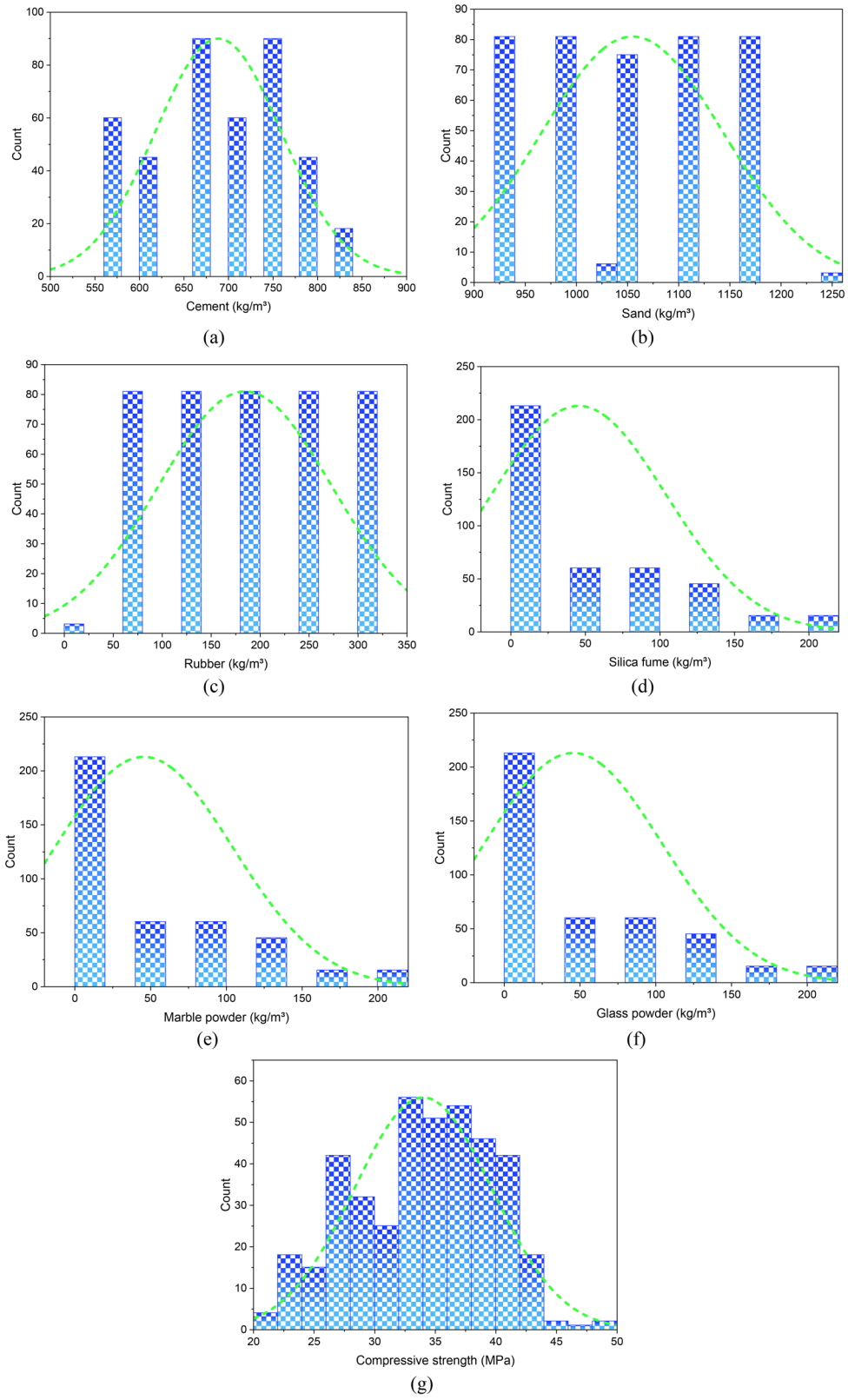


Figure 1: Visual dispersal of the parameters in the dataset: (a) cement, (b) sand, (c) rubber, (d) silica fume, (e) marble powder, (f) glass powder, and (g) CS.

is similar, as depicted in Figure 1(d)–(f), due to their comparable replacement ratios in RM mixes. The dataset's prevalence of CS is illustrated in Figure 1(g), with the majority of results falling within the 25–45 MPa range.

2.2 Development of models

A wide range of inputs is considered vital for ML algorithms to attain the required results [47]. It is advised that the data set utilized for modeling contains variables that fluctuate to optimize an ML model's performance. The use of a stable value or one with minor fluctuations may lead to inaccurate results [48,49]. The experimental database was used to construct the CS estimation models for RM. The modeling stage employed six inputs, as specified in Table 1, and one output. The research's goals were accomplished by using the Python code and Spyder tool (version 5.5.1) in the Anaconda Navigator software. The predictive ML models were constructed using three ensemble methods: GB, AB, and Bg. These techniques were chosen because of their proven capability in handling nonlinear interactions and minimizing overfitting in small-to-medium-sized datasets typical of material science applications [50,51]. These models offer enhanced prediction accuracy through their ability to combine multiple weak learners, making them particularly suitable for complex materials like RM containing SCMs. Furthermore, their compatibility with model interpretation tools such as SHAP enables transparent analysis of feature contributions, supporting the dual goals of predictive accuracy and practical insight for material design. Despite the fact that each model employs distinct methods to address regression issues, their combined use enhances the robustness of our investigation and provides a thorough understanding of the CS of RM. Additionally, the algorithms were configured to generate results exclusively for untrained data (testing set). The ratio of training data and testing data was set at 70:30, as recommended in prior similar studies on distinct materials [52,53]. The research strategy employed is illuminated in Figure 2.

A set of parameters known as hyperparameters can be used to modify the learning process of ML models. The hyperparameter values are necessary for the training of the model in supervised ML (classification and regression) [54]. The user-adjusted trial-and-error technique and the default configuration of the ML package can both be employed to ascertain an appropriate hyperparameter value. Trial and error are time-consuming and exhaustive processes for the operator to ascertain the optimal values

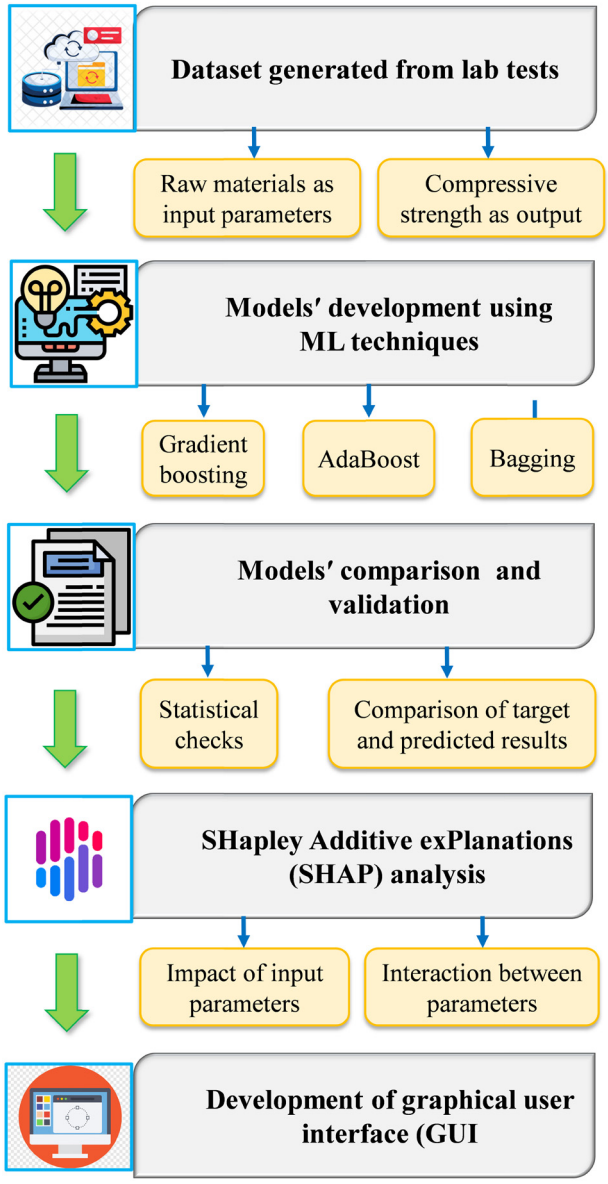


Figure 2: Flow diagram of the steps involved in the present study.

for the hyperparameters [55]. Using a hyperparameter tailoring and optimization procedure, users can save energy and time by choosing appropriate hyperparameter values for a selected ML algorithm [56]. The hyperparameter's optimal value is of the utmost significance in any ML model, as it enables the simultaneous verification of the function's lowest loss and highest accuracy [57]. Although there are alternative methods for optimizing hyperparameters, this investigation implemented grid search optimization strategies in conjunction with 10-fold cross-validation. Table 2 contains the hyperparameters' optimal values and ranges that were established during the modeling process.

2.3 Models' evaluation criteria

The valuation of the built ML prediction models was conducted using statistical performance measures, including the objective function (OBJ), mean absolute error (MAE), coefficient of determination (R^2), scatter index (SI), root mean square error (RMSE), and mean absolute percentage error (MAPE). The R^2 value of the expected result reflects the accurateness of the utilized techniques. The R^2 value quantifies the scale of deviation among the predictive models and the actual values. A lower numerical number near zero denotes a greater scale of deviation, while a greater value near one shows a reduced deviation. The reduced error metrics from the statistical checks denote the better precision of the generated ML model [49]. The assessment of ML models was performed utilizing Eqs. (1)–(5) [49,50,58]:

$$\text{RMSE} = \sqrt{\sum \frac{(P_i - E_i)^2}{n}}, \tag{1}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |P_i - E_i|, \tag{2}$$

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n} \frac{|P_i - E_i|}{E_i}, \tag{3}$$

$$\text{SI} = \frac{\text{RMSE}}{y'}, \tag{4}$$

$$\text{OBJ} = \left(\frac{\text{MAE} + \text{RMSE}}{R^2 + 1}\right), \tag{5}$$

where n represents data point numbers, P_i represents the predicted result, E_i represents the target values, and y' representsthe mean of predicted results.

2.4 SHAP analysis strategy

SHAP evaluation, established by Lundberg and Lee [59], is an extensive strategy for comprehending ML-based modeling. The Shapley value quantifies the comparative influence and contribution of input features on the results. The method closely mirrors parametric assessment, in which other parameters are held constant while one parameter is modified to measure its influence on the target characteristic [60]. Jupyter Notebook (version 7.2.2) was used for SHAP analysis, and Python code was established. This portion analyzes the significance of each component in the CS results, interpreting the impact of input features on the CS of RM. A SHAP outcome map is created to visualize feature significance, facilitating comprehension of the features affecting CS and their relative importance in the prediction modeling.

3 Results and analysis

3.1 GB model

Figure 3 illustrates the CS of RM modified with blends of SCMs, ascertained through the GB method. Figure 3(a) demonstrates the correlation between the target and estimated CS. The GB model accurately predicted the results, exhibiting only a minor disparity between the target and anticipated CS. An R^2 score of 0.963 signifies a robust correlation between real and predicted results. Figure 3(b) shows the spreading of the target (orange), predicted (blue), and absolute error (green) values for the GB model. An utmost disparity of 3.60 MPa was observed, with an average of 0.86 MPa. Furthermore, the absolute errors were examined proportionately, indicating that for 31% of the results, the error was below 0.5 MPa; for 36% results, the error ranged from 0.5 to 1.0 MPa; for 28% results, the error ranged from 1.0 to 2.0 MPa, and for 6% results, the error exceeded 2 MPa. The divergence data distribution indicated that the CS of RM can be precisely predicted using a GB model.

Table 2: Hyperparameter chosen for ML models

Parameter	GB		AB		Bg	
	Normal range	Optimum value	Normal range	Optimum value	Normal range	Optimum value
n-Estimators	10–200	70	10–200	110	10–200	100
Maximum depth	1–5	3	—	—	—	—
Learning rate	0.01–0.50	0.30	0.01–0.50	0.70	—	—
Maximum features	0.8–1.0	1	—	—	0.8–1.0	1
Minimum sample split	2–10	4	—	—	—	—
Minimum sample leaf	1–4	2	—	—	—	—

3.2 AB model

Figure 4 shows the findings of the AB method for the CS of RM modified using combinations of SCMs. Figure 4(a) represents the correlation between the target and anticipated CS. The AB model precisely forecasted the outcomes, demonstrating only a little deviation between the objective and expected CS. An R^2 score of 0.971 signifies a strong connection between experimental and model outcomes and a slightly higher accuracy than the GB model. Figure 4(b) depicts the spreading of the target (orange), predicted (blue), and absolute error (green) values for the AB model. The highest difference of 2.81 MPa was recorded, with a mean of 0.75 MPa. Additionally, the absolute errors were analyzed proportionately, revealing that 37% of the results had an error below 0.5 MPa; 39% of the results exhibited an error between 0.5 and 1.0 MPa; 20% of the results had an error between 1.0 and 2.0 MPa; and 4% of the results had an error beyond 2 MPa. The divergence data distribution demonstrated more absolute errors close to zero than the GB for the CS of RM, indicating the superior precision of the AB model.

3.3 Bagging model

Figure 5 shows the results of the Bg model about the CS of RM modified with various combinations of SCMs. Figure 5(a) illustrates the association between the desired and expected CS. The Bg model more accurately anticipated the results, exhibiting the least divergence between the real and test CS. An R^2 score of 0.975 signifies a robust correlation between experimental and predicted

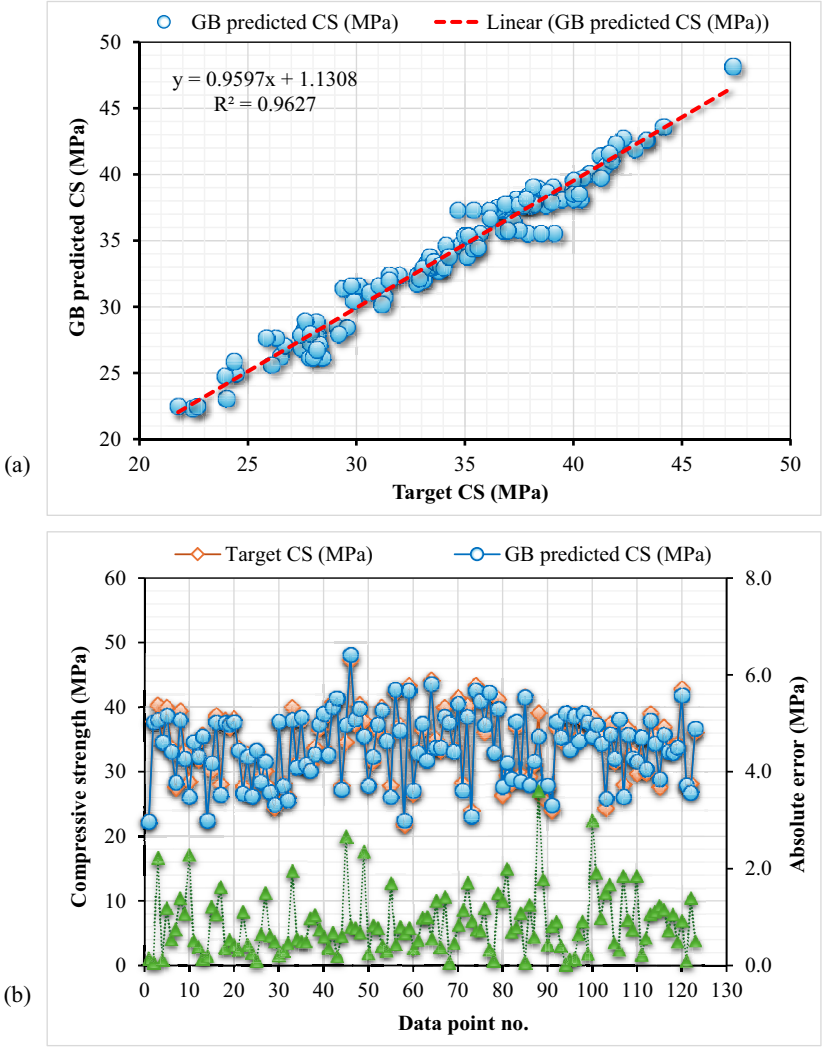
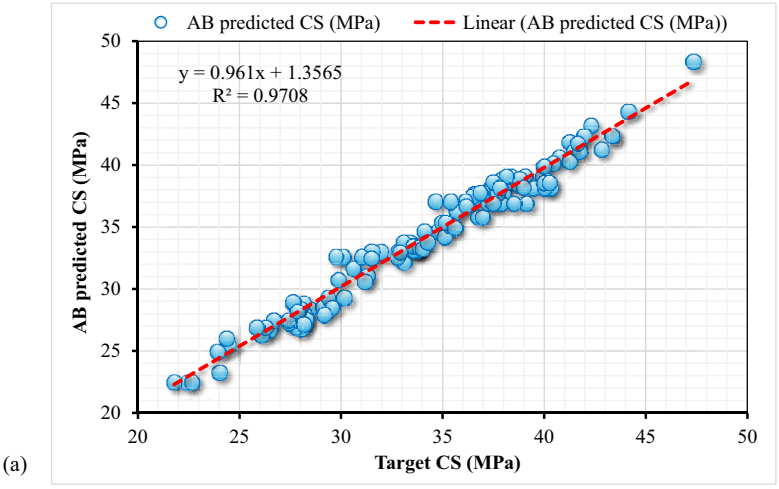
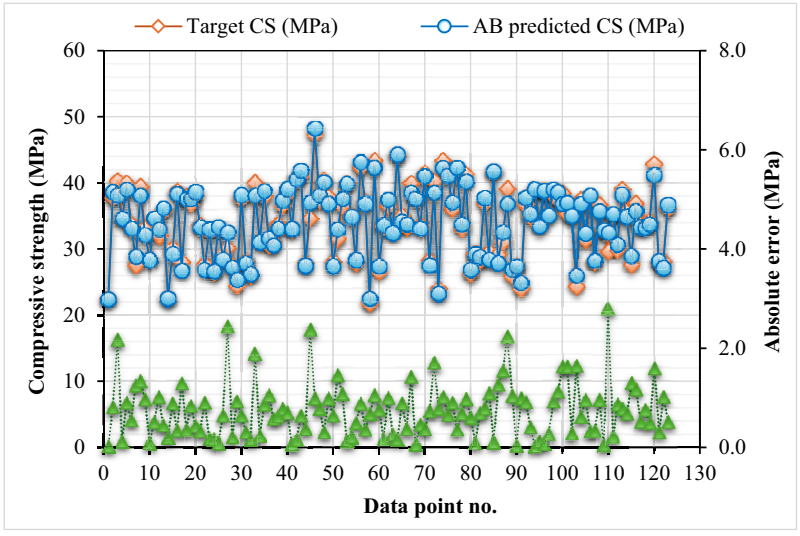


Figure 3: GB model: (a) association between target and predicted CS; and (b) spreading of target, predicted, and absolute error values.



(a)



(b)

Figure 4: AB model: (a) association between target and predicted CS; (b) spreading of target, predicted, and absolute error values.

outcomes, exhibiting marginally superior accuracy compared to the AB and GB. Figure 5(b) illustrates the spreading of the target (orange), predicted (blue), and absolute error (green) for the Bg method. The greatest difference of 2.96 MPa was observed, with a mean value of 0.66 MPa. Additionally, the error values were examined proportionately, indicating that 50% of the results had an error below 0.5 MPa; 31% had an error between 0.5 and 1.0 MPa; 17% had an error between 1.0 and 2.0 MPa; and 2% had an error beyond 2 MPa. The divergence data distribution exhibited a greater number of absolute errors near zero compared to AB and GB models for the CS of RM, signifying the enhanced accuracy of the Bg model. It can be inferred from the findings that all ML models employed in the present research (GB, AB, and Bg) indicated their potential for forecasting the CS of RM with higher precision.

3.4 Comparison and validation of models

Using the statistical checks for the established ML models, Table 3 shows the outcomes of MAE, OBJ, RMSE, SI, MAPE, and R^2. The GB method had an MAE of 0.864 MPa, while AB and Bg had an MAE of 0.746 and 0.658 MPa, respectively. The GB, AB, and Bg approaches were found to be MAPEs of 2.60, 2.20, and 2.00%, respectively. Furthermore, the RMSE for GB was 1.075 MPa, for AB, it was 0.927 MPa, and for the Bg model, it was 0.857 MPa. The models that were used have OBJ values of 0.998 MPa for GB, 0.849 MPa for AB, and 0.767 MPa for Bg. The most accurate model (Bg) had the lowest SI (0.025) compared to GB (0.032) and AB (0.027). Therefore, Bg was found to be the most accurate with respect to higher R^2 and lower MAE, MAPE, RMSE, OBJ, and SI scores.

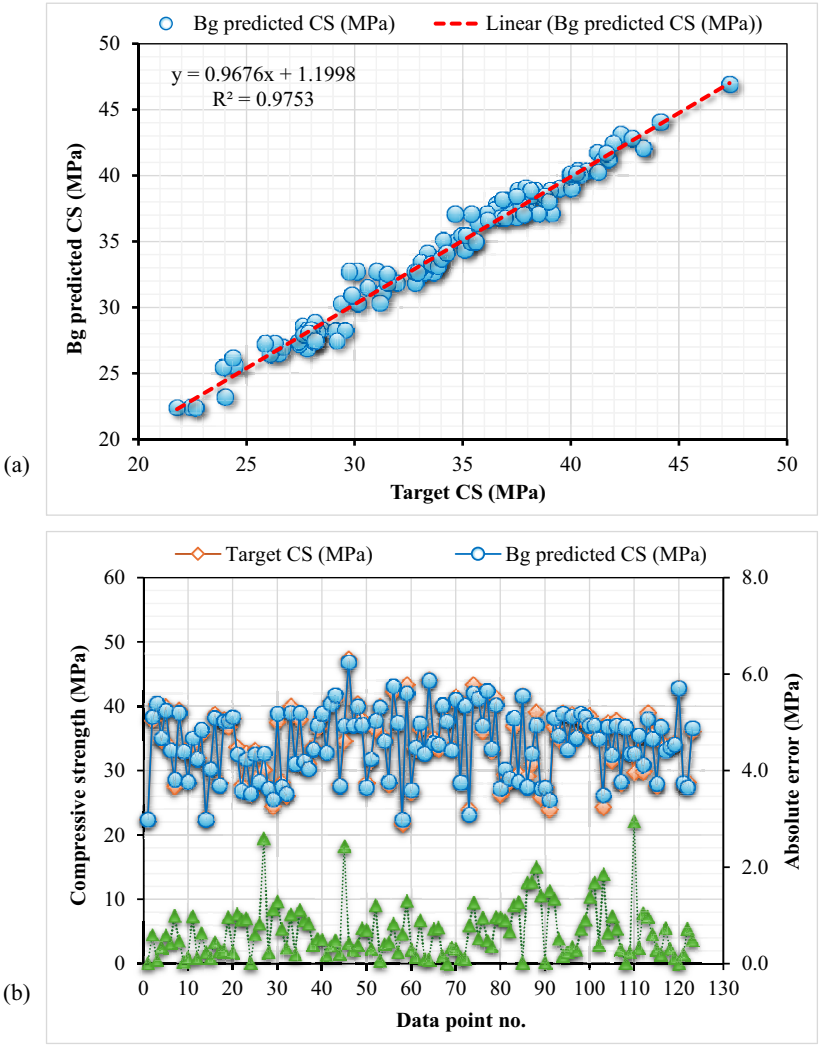


Figure 5: Bg model: (a) association between target and predicted CS; (b) spreading of target, predicted, and absolute error values.

3.5 GUI for the CS of RM

The developed Bg model is employed because of its superior accuracy in generating the GUI. The GUI elements, including layout, icons, and recommendations, can be customized by the ML model in accordance with user feedback, preferences, and behavior. The objective of this modification is to enhance personalization and usability. This methodology enables the interface to adapt and evolve in real time, thereby becoming more attentive to the user's preferences and requirements. This, in turn, leads to increased efficiency and, in the long run, increases user satisfaction and engagement. The integration of ML into GUI design is a significant advancement that enables the creation of more user-friendly, intelligent, and agile applications [61]. Initially, an ML model, *i.e.*, Bg, was implemented and archived for future processing. The GUI was subsequently established by utilizing the stored model mentioned earlier, which enabled it to capitalize on the predictive abilities of the trained model. The model was seamlessly incorporated with the GUI by utilizing the extensive libraries of Python, which allowed for real-time connections and dynamic information that responded to

Table 3: Statistical measures for validation of models

Parameter	GB	AB	Bg
R^2	0.963	0.971	0.975
MAE (MPa)	0.864	0.746	0.658
RMSE (MPa)	1.075	0.927	0.857
MAPE (%)	2.60	2.20	2.00
SI	0.032	0.027	0.025
OBJ (MPa)	0.988	0.849	0.767

inputs. This method ensures a smooth and efficient shift from the model's training to the GUI's deployment, resulting in a user interface that is both intuitive and responsive. Figure 6 illustrates the GUI screenshot that was created using the Bg method to assess the CS of RM, taking into account all input parameters. By altering the value of any input parameter, the GUI provides a straightforward method for examining the anticipated output. However, the developed GUI has certain limitations, including using the same number and type of input parameters with similar units used in the current study. Using different input parameters or altering the units, the GUI will not yield accurate strength values.

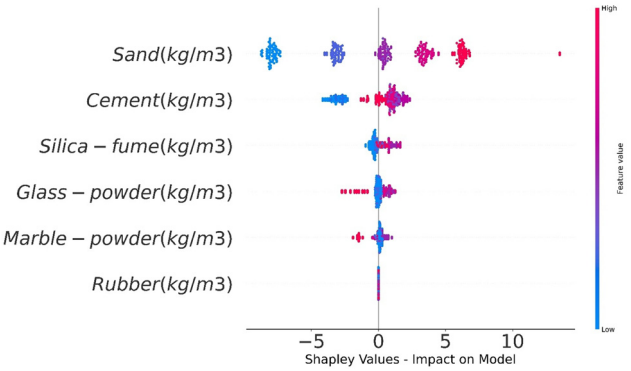


Figure 7: Shapley values implying impact on the outcome.

3.6 SHAP analysis results

The purpose of the SHAP study was to identify how different input parameters affected the CS of RM modified with the blends of SCMs. The extensive dataset used the SHAP tree explainer to yield a detailed interpretation of both the aggregate impact of attributes and localized SHAP explanations. Figure 7 shows the results of the SHAP assessment that was run on the inputs that had an impact on the outcome. On the horizontal axis, the SHAP score indicates the influence of each variable, and the image shows the estimation of distinct input parameters throughout a scale of colors. The sand quantity showed a significant positive correlation and was found to be the principal factor influencing the CS of RM. This suggests that with increased sand quantity, the CS is higher, and replacing sand with rubber causes a reduction in the CS of RM. The results show a stronger positive association between the CS of RM and the amount of cement in the mixture. The impact pattern of silica fume, glass, and marble powder was observed to be both positive and negative, indicating their use in optimum limits in RM.

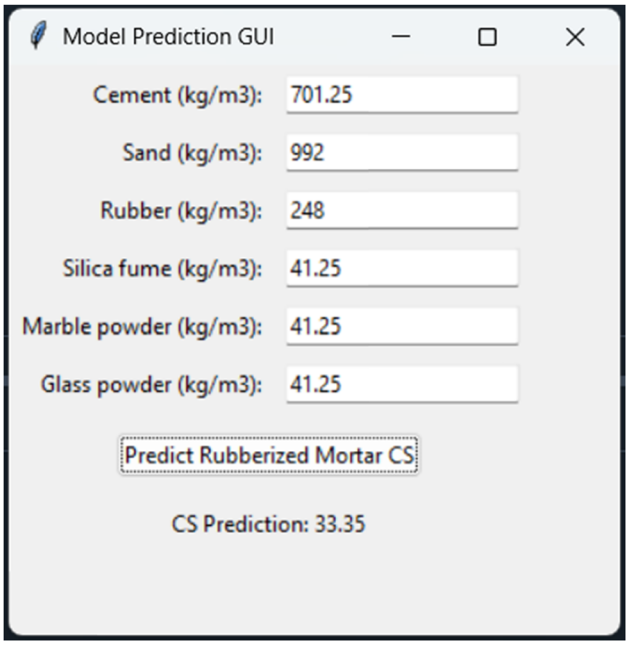


Figure 6: GUI developed for the CS of RM.

The SHAP assessment tool also yields a significant understanding of feature interactions, clearly exhibiting the collective influence of multiple factors on predictability. Employing a unified method is important for recognizing complex relationships inside ML models, as these might not be clearly identifiable through correlation investigation alone. The color variation is used to show the scale of dependency of one variable on another for each specific parameter. In Figure 8, every point is displayed in the shade of red to imply a greater value of the primary dependent parameter, whereas a shade of blue represents a minimal value of the primary dependent parameter. The figure displays the correlation among the values of a specific parameter and their associated SHAP scores. Figure 8(a) shows the affecting pattern and reliance on the cement quantity. The plot depicts a strong, favorable link between the cement quantity and the CS of RM. The optimum level of cement was noted to be about 700 $kg \cdot m^{-3}$, as replacement with SCMs can lead to enhanced strength of RM. Furthermore, feature dependency indicated that the cement quantity primarily interacts with the quantity of sand. The figure demonstrates that higher cement and sand quantities in RM mixes resulted in increased SHAP scores, which are beneficial for enhanced CS. The effect of sand quantity in Figure 8(b) shows that keeping the sand quantity higher and the rubber lower is favorable for the strength of RM. The interdependence pattern of silica fume, glass powder, and marble powder in Figure

8(c)–(e), respectively, exhibits similar patterns with optimum values around 100–150 kg·m^{-3}. The influence of rubber quantity remains ambiguous (Figure 8(f)). This is because the statistical distribution shows that the rubber content changes do not correspond to consistent or significant variations in CS, as evidenced by a low standard deviation in SHAP values for rubber. In contrast, features such as sand, silica fume, marble, and glass powder demonstrate greater variability and a stronger influence on CS predictions. Also, the data structure indicates limited nonlinearity or coupling between rubber and cement, *i.e.*, in the experimental combinations, increases or decreases in rubber were not consistently paired with meaningful changes in cement content that would result in interactive effects. These might be possible reasons for the negligible interaction of rubber during the SHAP analysis. The SHAP study indicates that utilizing silica fume, glass, and marble powder at optimal levels as cement substitutes is advantageous for increasing the CS of RM. The outcomes of the current research were gained from the type of input and

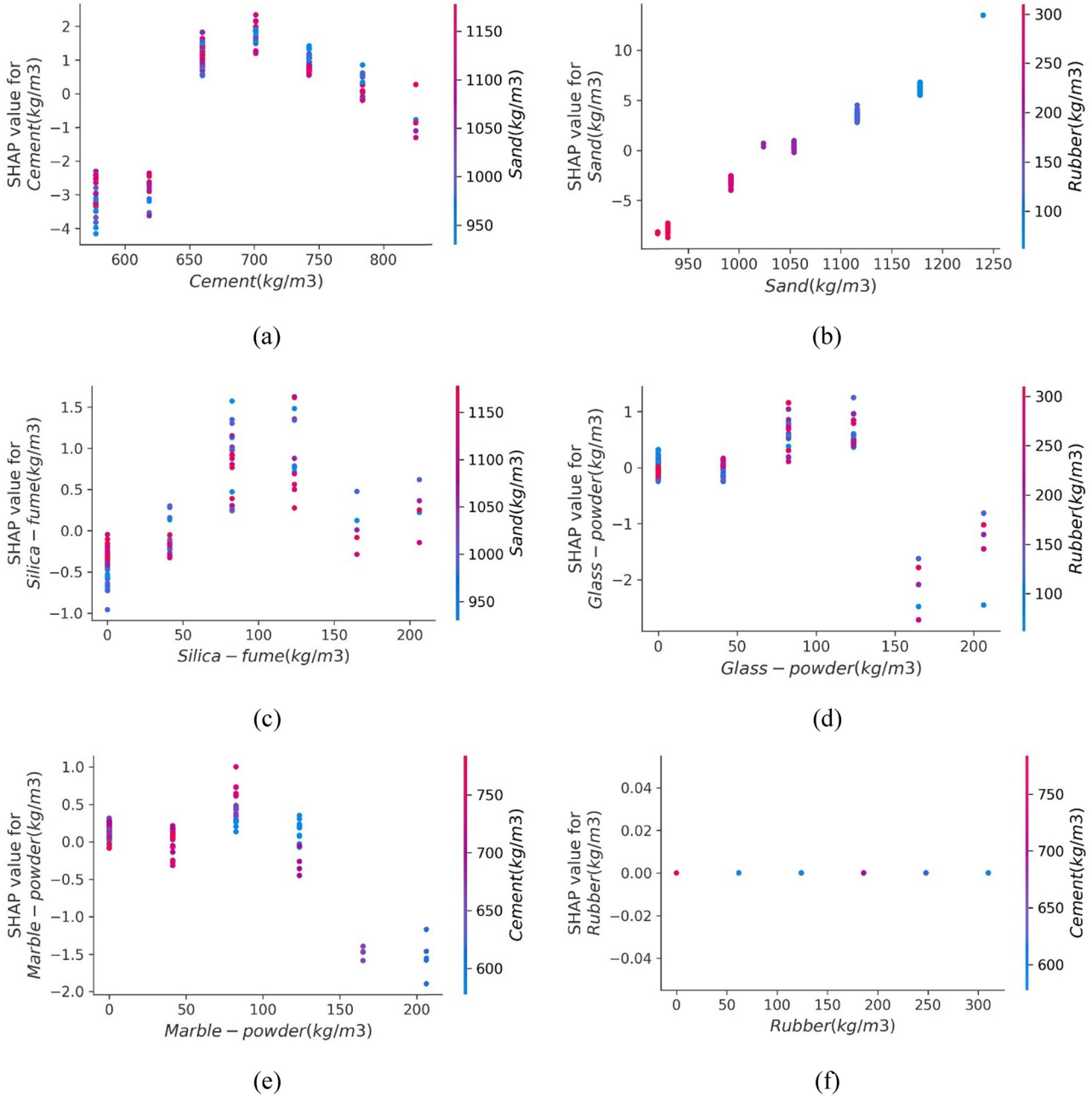


Figure 8: Interdependencies of input features: (a) cement, (b) sand, (c) silica fume, (d) glass powder, (e) marble powder, and (f) rubber.

Table 4: Comparison of the findings with the literature

Ref.	ML algorithm used	Property studied	Material investigated	Best ML model observed
Present study	GB, AB, and Bg	CS	RM	Bg
[50]	Decision tree, AB, extreme GB, multilayer perceptron neuron network, random forest, support vector machine, and Bg	CS	High-performance concrete	Random forest and Bg
[65]	Decision tree, random forest, Bg, and gene expression programming	CS	Geopolymer concrete	Bg
[51]	Decision tree, GB, and Bg	CS	Recycled aggregate concrete	Bg
[66]	Decision tree, AB, and Bg	CS	Geopolymer concrete	Bg
[36]	Bg, gene expression programming, and artificial neural network	Split-tensile strength	Recycled aggregate concrete	Bg
[62]	Decision tree, AB, Bg, and gene expression programming	CS	High-performance concrete	Bg
[63]	Artificial neural network, decision tree, Bg, and gene expression programming	CS	Fly ash-based concrete	Bg
[64]	Bg, decision tree, and gene expression programming	CS	Fly ash-based concrete	Bg

database size used in the SHAP study. Increasing the input range in the dataset size may facilitate more accurate correlations, which is suggested for future research.

4 Discussion

The findings of this study demonstrate the effectiveness of ML techniques in accurately predicting the CS of RM. The high precision of the models, as indicated by R^2 values of 0.975, 0.971, and 0.963 for the Bg, AB, and GB models, respectively, confirms their reliability. These findings emphasize the capability of ML models to facilitate the optimization of RM mixtures by providing reliable strength predictions without the need for extensive laboratory experimentation. Previous studies have demonstrated enhanced accuracy of Bg methods, as listed in Table 4, in forecasting the characteristics of cementitious composites [36,50,51,62–64]. The Bg strategy enhances accuracy by reducing variance by training several models using bootstrapped samples, thereby introducing diversity and mitigating overfitting. The superior accuracy of the Bg model might be ascribed to its robustness, scalability through parallelization, and the application of error-smoothing averaging [49].

A key contribution of this study is the development of a GUI based on the best-performing Bg model. This GUI enables real-time adaptation by allowing users to input material composition parameters and obtain instant predictions of RM strength. The adaptability and efficiency of this tool improve user engagement and enhance the practical applicability of ML-based predictions in engineering practices. However, care must be taken in ensuring that the input values align with the parameters used during model training to maintain accuracy. The predictive models and GUI developed in this study are based on a dataset comprising specific material types, including a single type of ordinary Portland cement and defined ranges of rubber and SCM characteristics. Variations in the physical and chemical properties of these ingredients, such as cement fineness, chemical composition, sand gradation, or rubber texture, may affect the accuracy of predictions. Therefore, users should exercise caution when applying the model beyond the scope of the training dataset. Similarly, the same input parameters with comparable units utilized in the present investigation need to be used to obtain outcomes from the GUI. The GUI will not make accurate strength predictions when alternative input parameters or units are utilized. Also, the SHAP analysis provided valuable insights into the contributions of different input parameters toward CS and suggests that

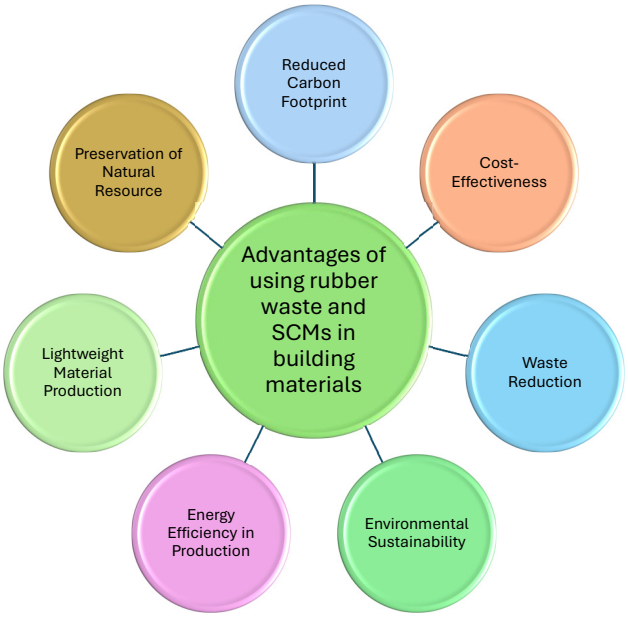


Figure 9: Advantages of using water rubber and SCMs from industrial waste powders in building materials.

strategic mix designs, incorporating these SCMs at appropriate proportions, can lead to enhanced RM performance.

The study also aligns with the broader objective of sustainable construction materials development. The utilization of rubber waste and SCMs such as marble powder, silica fume, and glass powder contributes to waste reduction and promotes eco-friendly construction practices. These SCMs, when used as partial cement replacements, not only enhance the mechanical performance of RM but also reduce the carbon footprint linked with the production of cement. Moreover, the adoption of RM associates with circular economy principles by repurposing waste rubber particles in construction applications. This practice mitigates environmental pollution caused by discarded rubber. Figure 9 shows the benefits related to using waste materials in construction.

5 Conclusions

This study used experimental results observed for the CS of RM to study the suitability of ML methods, including AB, GB, and bagging (Bg). Hyperparameters were refined to obtain the best model outcomes. A GUI was generated for the model with the highest precision. Moreover, SHAP analysis was performed to study the influence of the interaction of input parameters on model results. The main findings of the research are as follows:

- The ML techniques used for the CS prediction of RM exhibited good agreement with the experimental outcomes and can be used for accurate prediction. The R^2 values were 0.975, 0.971, and 0.963 for the Bg, AB, and GB models, respectively, implying their precision accuracy as the values are closer to 1.
- Statistical performance tests conducted for the ML models further confirmed their predictability capabilities. For example, the MAPEs were 2.00, 2.20, and 2.60% for the Bg, AB, and GB, respectively, showing their precisions for estimating the CS of RM.
- The GUI developed based on the Bg model can be used for quick predictions using input values for the input parameters confined to this study, thereby enabling real-time adaptation and personalization based on user inputs, enhancing usability, efficiency, and engagement through dynamic updates and predictive capabilities.
- SHAP analysis results showed that sand and cement quantities in the RM mix were crucial parameters with a more positive impact on the CS of RM, followed by silica fume, glass powder, and marble powder with both positive and negative correlations.
- The SHAP analysis also revealed significant feature interactions, illustrating that the optimal cement quantity was around 700 $kg{\cdot}m^{-3}$ in RM mixes. Higher cement and sand contents result in increased SHAP values, positively influencing strength. Additionally, silica fume, glass powder, and marble powder exhibit optimal values around 100–150 $kg{\cdot}m^{-3}$, enhancing RM strength when used as cement substitutes.

It can be concluded that using ML algorithms like GB, AB, and Bg can precisely evaluate the CS of RM, and modern tools like GUI can help in intelligent calculations of the outcomes with varying input values. However, care needs to be taken while using the GUI regarding input parameter types and units used during modeling.

Acknowledgments: The authors acknowledge the support of Philosophy and Social Sciences Planning Project of Hainan Province, China (Grant No. HNSK(YB)22-23).

Funding information: This work was supported by the Philosophy and Social Sciences Planning Project of Hainan Province, China (Grant No. HNSK(YB)22-23).

Author contributions: L.F.: conceptualization, methodology, software, funding acquisition, supervision, and writing-original draft. F.X.: formal analysis, project administration, resources, visualization, validation, writing, reviewing, and editing. All authors have accepted

responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: The authors state no conflict of interest.

Data availability statement: The dataset generated and/or analyzed during the current study is available from the corresponding author upon reasonable request.

References

[1] Bu, C., D. Zhu, L. Liu, X. Lu, Y. Sun, L. Yu, et al. Research progress on rubber concrete properties: A review. *Journal of Rubber Research*, Vol. 25, 2022, pp. 105–125.

[2] Ahmed, W., G. Lu, S. T. Ng, and G. Liu. Innovative valorization of solid waste materials for production of sustainable low-carbon pavement: A systematic review and scientometric analysis. *Case Studies in Construction Materials*, Vol. 22, 2025, id. e04541.

[3] Ahmad, W., S. J. McCormack, and A. Byrne. Biocomposites for sustainable construction: A review of material properties, applications, research gaps, and contribution to circular economy. *Journal of Building Engineering*, Vol. 105, 2025, id. 112525.

[4] Bouchelil, L., S. B. S. Jafar, and M. Khanzadeh Moradllo. Evaluating the performance of internally cured limestone calcined clay concrete mixtures. *Journal of Sustainable Cement-Based Materials*, Vol. 14, 2025, pp. 198–208.

[5] Thomas, B. S. and R. C. Gupta. Long term behaviour of cement concrete containing discarded tire rubber. *Journal of Cleaner Production*, Vol. 102, 2015, pp. 78–87.

[6] Qaidi, S. M. A., Y. Z. Dinkha, J. H. Haido, M. H. Ali, and B. A. Tayeh. Engineering properties of sustainable green concrete incorporating eco-friendly aggregate of crumb rubber: A review. *Journal of Cleaner Production*, Vol. 324, 2021, id. 129251.

[7] Xiao, Z., A. Pramanik, A. K. Basak, C. Prakash, and S. Shankar. Material recovery and recycling of waste tyres-A review. *Cleaner Materials*, Vol. 5, 2022, id. 100115.

[8] Agarwal, S., M. Tyagi, and R. K. Garg. Conception of circular economy obstacles in context of supply chain: A case of rubber industry. *International Journal of Productivity and Performance Management*, Vol. 72, 2023, pp. 1111–1153.

[9] Soares, F. A. and A. Steinbüchel. Natural rubber degradation products: Fine chemicals and reuse of rubber waste. *European Polymer Journal*, Vol. 165, 2022, id. 111001.

[10] Formela, K. Sustainable development of waste tires recycling technologies – recent advances, challenges and future trends. *Advanced Industrial and Engineering Polymer Research*, Vol. 4, 2021, pp. 209–222.

[11] Surehali, S., A. Singh, and K. P. Biligiri. A state-of-the-art review on recycling rubber in concrete: Sustainability aspects, specialty mixtures, and treatment methods. *Developments in the Built Environment*, Vol. 14, 2023, id. 100171.

[12] Khan, M., J. Lao, M. R. Ahmad, M.-F. Kai, and J.-G. Dai. The role of calcium aluminate cement in developing an efficient ultra-high performance concrete resistant to explosive spalling under high temperatures. *Construction and Building Materials*, Vol. 384, 2023, id. 131469.

[13] Kurniati, E. O., F. Pederson, and H.-J. Kim. Application of steel slags, ferronickel slags, and copper mining waste as construction materials: A review. *Resources, Conservation and Recycling*, Vol. 198, 2023, id. 107175.

[14] Ren, F., J. Mo, Q. Wang, and J. C. M. Ho. Crumb rubber as partial replacement for fine aggregate in concrete: An overview. *Construction and Building Materials*, Vol. 343, 2022, id. 128049.

[15] Xu, Y., J. Wang, P. Zhang, J. Guo, and S. Hu. Enhanced effect and mechanism of colloidal nano-SiO_2 modified rubber concrete. *Construction and Building Materials*, Vol. 378, 2023, id. 131203.

[16] Tanhadoust, A., S. A. A. Emadi, S. Nasrollahpour, F. Dabbaghi, and M. L. Nehdi. Optimal design of sustainable recycled rubber-filled concrete using life cycle assessment and multi-objective optimization. *Construction and Building Materials*, Vol. 402, 2023, id. 132878.

[17] Chen, A., X. Han, M. Chen, X. Wang, Z. Wang, and T. Guo. Mechanical and stress-strain behavior of basalt fiber reinforced rubberized recycled coarse aggregate concrete. *Construction and Building Materials*, Vol. 260, 2020, id. 119888.

[18] Mei, J., G. Xu, W. Ahmad, K. Khan, M. N. Amin, F. Aslam, et al. Promoting sustainable materials using recycled rubber in concrete: A review. *Journal of Cleaner Production*, Vol. 373, 2022, id. 133927.

[19] Barham, W. S., B. Albiss, and O. Latayfeh. Influence of magnetic field treated water on the compressive strength and bond strength of concrete containing silica fume. *Journal of Building Engineering*, Vol. 33, 2021, id. 101544.

[20] Ashish, D. K. Concrete made with waste marble powder and supplementary cementitious material for sustainable development. *Journal of Cleaner Production*, Vol. 211, 2019, pp. 716–729.

[21] Khan, M. and C. McNally. A holistic review on the contribution of civil engineers for driving sustainable concrete construction in the built environment. *Developments in the Built Environment*, Vol. 16, 2023, id. 100273.

[22] Jamil, S., M. Idrees, A. Akbar, and W. Ahmed. Investigating the mechanical and durability properties of carbonated recycled aggregate concrete and its performance with SCMs. *Buildings*, Vol. 15, 2025, id. 201.

[23] Shanmugasundaram, N. and S. Praveenkumar. Influence of supplementary cementitious materials, curing conditions and mixing ratios on fresh and mechanical properties of engineered cementitious composites – A review. *Construction and Building Materials*, Vol. 309, 2021, id. 125038.

[24] Gupta, S. and S. Chaudhary. State of the art review on Supplementary Cementitious Materials in India – I: An overview of legal perspective, governing organizations, and development patterns. *Journal of Cleaner Production*, Vol. 261, 2020, id. 121203.

[25] Sandanayake, M., Y. Bouras, R. Haigh, and Z. Vrcelj. Current sustainable trends of using waste materials in concrete—a decade review. *Sustainability*, Vol. 12, 2020, id. 9622.

[26] Singh, G. V. P. B. and K. V. L. Subramaniam. Production and characterization of low-energy Portland composite cement from post-industrial waste. *Journal of Cleaner Production*, Vol. 239, 2019, id. 118024.

[27] Khan, M., A. Rehman, and M. Ali. Efficiency of silica-fume content in plain and natural fiber reinforced concrete for concrete road. *Construction and Building Materials*, Vol. 244, 2020, id. 118382.

[28] Khan, M., M. Cao, A. Hussain, and S. H. Chu. Effect of silica-fume content on performance of $CaCO_3$ whisker and basalt fiber at

matrix interface in cement-based composites. *Construction and Building Materials*, Vol. 300, 2021, id. 124046.
[29] Ashish, D. K. Feasibility of waste marble powder in concrete as partial substitution of cement and sand amalgam for sustainable growth. *Journal of Building Engineering*, Vol. 15, 2018, pp. 236–242.
[30] Juenger, M. C. G., R. Snellings, and S. A. Bernal. Supplementary cementitious materials: New sources, characterization, and performance insights. *Cement and Concrete Research*, Vol. 122, 2019, pp. 257–273.
[31] Amin, M. N., R.-U.-D. Nassar, K. Khan, S. Ul Arifeen, M. Khan, and M. T. Qadir. Integrating testing and modeling methods to examine the feasibility of blended waste materials for the compressive strength of rubberized mortar. *Reviews on Advanced Materials Science*, Vol. 63, 2024, id. 20240081.
[32] Ahmad, W., V. S. S. C. S. Veeraghantla, and A. Byrne. Advancing sustainable concrete using biochar: Experimental and modelling study for mechanical strength evaluation. *Sustainability*, Vol. 17, 2025, id. 2516.
[33] Alwi Assaggaf, R., S. Uthman Al-Dulaijan, M. Maslehuddin, O. S. Baghabra Al-Amoudi, S. Ahmad, and M. Ibrahim. Effect of different treatments of crumb rubber on the durability characteristics of rubberized concrete. *Construction and Building Materials*, Vol. 318, 2022, id. 126030.
[34] Feng, D.-C., Z.-T. Liu, X.-D. Wang, Y. Chen, J.-Q. Chang, D.-F. Wei, et al. Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach. *Construction and Building Materials*, Vol. 230, 2020, id. 117000.
[35] Nazar, S., J. Yang, X.-E. Wang, K. Khan, M. N. Amin, M. F. Javed, et al. Estimation of strength, rheological parameters, and impact of raw constituents of alkali-activated mortar using machine learning and SHapely Additive exPlanations (SHAP). *Construction and Building Materials*, Vol. 377, 2023, id. 131014.
[36] Zhu, Y., A. Ahmad, W. Ahmad, N. I. Vatin, A. M. Mohamed, and D. Fathi. Predicting the splitting tensile strength of recycled aggregate concrete using individual and ensemble machine learning approaches. *Crystals*, Vol. 12, 2022, id. 569.
[37] Wang, Q., A. Hussain, M. U. Farooqi, and A. F. Deifalla. Artificial intelligence-based estimation of ultra-high-strength concrete's flexural property. *Case Studies in Construction Materials*, Vol. 17, 2022, id. e01243.
[38] Chou, J.-S. and A.-D. Pham. Enhanced artificial intelligence for ensemble approach to predicting high performance concrete compressive strength. *Construction and Building Materials*, Vol. 49, 2013, pp. 554–563.
[39] Ebid, A. and A. Deifalla. Using artificial intelligence techniques to predict punching shear capacity of lightweight concrete slabs. *Materials*, Vol. 15, 2022, id. 2732.
[40] Li, P., M. Ali Khan, A. M. Galal, H. Hassan Awan, A. Zafar, M. Faisal Javed, et al. Sustainable use of chemically modified tyre rubber in concrete: Machine learning based novel predictive model. *Chemical Physics Letters*, Vol. 793, 2022, id. 139478.
[41] Sun, Y., G. Li, N. Zhang, Q. Chang, J. Xu, and J. Zhang. Development of ensemble learning models to evaluate the strength of coal-grout materials. *International Journal of Mining Science and Technology*, Vol. 31, 2021, pp. 153–162.
[42] Arifuzzaman, M., H. J. Qureshi, A. F. Al Fuhaid, F. Alanazi, M. F. Javed, and S. M. Eldin. Novel ensemble modelling for prediction of fundamental properties of bitumen incorporating plastic waste. *Journal of Materials Research and Technology*, Vol. 24, 2023, pp. 3334–3351.
[43] American Society for Testing and Materials. Committee, C.o.C. *Standard test method for compressive strength of hydraulic cement mortars (using 2-in. or [50-mm] cube specimens)*, ASTM International, West Conshohocken, USA, 2013.
[44] Taffese, W. Z. and L. Espinosa-Leal. A machine learning method for predicting the chloride migration coefficient of concrete. *Construction and Building Materials*, Vol. 348, 2022, id. 128566.
[45] Tsanas, A. and A. Xifara. Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and Buildings*, Vol. 49, 2012, pp. 560–567.
[46] Chen, Z., M. N. Amin, B. Iftikhar, W. Ahmad, F. Althoey, and F. Alsharari. Predictive modelling for the acid resistance of cement-based composites modified with eggshell and glass waste for sustainable and resilient building materials. *Journal of Building Engineering*, Vol. 76, 2023, id. 107325.
[47] Sufian, M., S. Ullah, K. A. Ostrowski, A. Ahmad, A. Zia, K. Śliwa-Wieczorek, et al. An experimental and empirical study on the use of waste marble powder in construction material. *Materials*, Vol. 14, 2021, id. 3829.
[48] Khan, M., J. Lao, and J.-G. Dai. Comparative study of advanced computational techniques for estimating the compressive strength of UHPC. *Journal of Asian Concrete Federation*, Vol. 8, 2022, pp. 51–68.
[49] Cao, Q., X. Yuan, M. Nasir Amin, W. Ahmad, F. Althoey, and F. Alsharari. A soft-computing-based modeling approach for predicting acid resistance of waste-derived cementitious composites. *Construction and Building Materials*, Vol. 407, 2023, id. 133540.
[50] Farooq, F., W. Ahmed, A. Akbar, F. Aslam, and R. Alyousef. Predictive modeling for sustainable high-performance concrete from industrial wastes: A comparison and optimization of models using ensemble learners. *Journal of Cleaner Production*, Vol. 292, 2021, id. 126032.
[51] Khan, K., W. Ahmad, M. N. Amin, F. Aslam, A. Ahmad, and M. A. Al-Faiad. Comparison of prediction models based on machine learning for the compressive strength estimation of recycled aggregate concrete. *Materials*, Vol. 15, 2022, id. 3430.
[52] Chen, Z. Application of machine learning boosting and bagging methods to predict compressive and flexural strength of marble cement mortar. *Materials Today Communications*, Vol. 39, 2024, id. 108600.
[53] Chen, Z., B. Iftikhar, A. Ahmad, Y. Dodo, M. A. Abuhussain, F. Althoey, et al. Strength evaluation of eco-friendly waste-derived self-compacting concrete via interpretable genetic-based machine learning models. *Materials Today Communications*, Vol. 37, 2023, id. 107356.
[54] Yang, L. and A. Shami. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, Vol. 415, 2020, pp. 295–316.
[55] Singh, S., S. K. Patro, and S. K. Parhi. Evolutionary optimization of machine learning algorithm hyperparameters for strength prediction of high-performance concrete. *Asian Journal of Civil Engineering*, Vol. 24, No. 8, 2023, pp. 3121–3143.
[56] Thornton, C., F. Hutter, H. H. Hoos, and K. Leyton-Brown. Auto-WEKA: combined selection and hyperparameter optimization of classification algorithms. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, Chicago, Illinois, USA*, 2013, pp. 847-855.
[57] Yeh, I. C. and L.-C. Lien. Knowledge discovery of concrete material using Genetic Operation Trees. *Expert Systems with Applications*, Vol. 36, 2009, pp. 5807–5812.

[58] Aslam, F., F. Farooq, M. N. Amin, K. Khan, A. Waheed, A. Akbar, et al. Applications of gene expression programming for estimating compressive strength of high-strength concrete. *Advances in Civil Engineering*, Vol. 2020, 2020, id. 8850535.
[59] Lundberg, S. M. and S.-I. Lee. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 2017, Vol. 30.
[60] Abdulalim Alabdullah, A., M. Iqbal, M. Zahid, K. Khan, M. Nasir Amin, and F. E. Jalal. Prediction of rapid chloride penetration resistance of metakaolin based high strength concrete using light GBM and XGBoost models by incorporating SHAP analysis. *Construction and Building Materials*, Vol. 345, 2022, id. 128296.
[61] Amin, M. N., A. Ahmad, K. Khan, and M. T. Qadir. Precision assessment of the machine learning tools for the strength optimization of environmental-friendly lightweight foam concrete. *Journal of Environmental Management*, Vol. 373, 2025, id. 123462.
[62] Ahmad, W., A. Ahmad, K. A. Ostrowski, F. Aslam, P. Joyklad, and P. Zajdel. Application of advanced machine learning approaches to predict the compressive strength of concrete containing supplementary cementitious materials. *Materials*, Vol. 14, 2021, id. 5762.
[63] Song, H., A. Ahmad, F. Farooq, K. A. Ostrowski, M. Maślak, S. Czarnecki, et al. Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms. *Construction and Building Materials*, Vol. 308, 2021, id. 125021.
[64] Ahmad, A., F. Farooq, P. Niewiadomski, K. Ostrowski, A. Akbar, F. Aslam, et al. Prediction of compressive strength of fly ash based concrete using individual and ensemble algorithm. *Materials*, Vol. 14, 2021, id. 794.
[65] Zou, Y., C. Zheng, A. M. Alzahrani, W. Ahmad, A. Ahmad, A. M. Mohamed, et al. Evaluation of artificial intelligence methods to estimate the compressive strength of geopolymers. *Gels*, Vol. 8, 2022, id. 271.
[66] Ahmad, A., W. Ahmad, F. Aslam, and P. Joyklad. Compressive strength prediction of fly ash-based geopolymer concrete via advanced machine learning techniques. *Case Studies in Construction Materials*, Vol. 16, 2022, id. e00840.