



OPEN Compressive strength prediction of carbonated recycled aggregate concrete using regression based machine learning models

Hints G. Gebremariam✉, Shifferaw Taye & Abrham Gebre Tarekgn

Carbonation treatment of recycled concrete aggregates (RCA) enhances aggregate performance while enabling CO₂ sequestration, thereby supporting sustainable construction practices. Despite the growing application of machine learning (ML) in concrete research, predictive modeling of concrete incorporating carbonated RCA remains insufficiently explored. This study investigates regression based ML approaches for predicting the compressive strength of concrete containing carbonated recycled coarse aggregates. A total of 108 experimentally obtained data points were compiled and statistically validated prior to model development. The input variables included water to cement ratio, cement to coarse aggregate ratio, cement to fine aggregate ratio, water absorption and crushing value of natural and recycled coarse aggregates, parent concrete strength, degree of carbonation, and replacement ratio of natural coarse aggregates with recycled aggregates. Composite indices were introduced to simplify predictive formulations and to evaluate the influence of multicollinearity. Six ML algorithms namely, multilinear regression (MLR), ridge regression, polynomial MLR, decision tree, random forest, and LightGBM were developed and optimized using 5-fold cross-validation. Model performance was evaluated using mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R²), supported by residual and prediction analyses. The decision tree and LightGBM models achieved the highest predictive accuracy on the test dataset, yielding RMSE of 1.59 MPa, MAE of 1.17 MPa, and R² of 0.991. The random forest model closely followed, with RMSE of 1.64 MPa, MAE of 1.25 MPa, and R² of 0.991. Polynomial MLR demonstrated moderate performance, with RMSE of 2.23 MPa, MAE of 1.89 MPa, and R² of 0.983. In contrast, conventional MLR and ridge regression exhibited lower accuracy, with RMSE values of 3.19 and 3.08 MPa, MAE of 2.65 and 2.56 MPa, and R² values of 0.965 and 0.967, respectively. The tree based ensemble models provided the most accurate and stable predictions, maintaining a maximum error range of approximately ±13%. Although the decision tree and LightGBM models achieved marginally lower prediction errors, the random forest model was identified as the most reliable overall due to its ability to capture physically meaningful relationships, as evidenced by SHAP analysis, together with its high predictive accuracy and robustness. Composite indexing had a minimal effect on model performance. Sensitivity and SHAP analyses identified the mix proportioning index as the most influential parameter, followed by degree of carbonation and aggregate performance. Overall, the results demonstrate that ML based models, particularly tree based ensembles, effectively capture the experimental behavior of carbonated RCA concrete and enable efficient compressive strength prediction.

Keywords Classical regression model, Machine learning, Carbonation, Performance metrics, Compressive strength

Concrete and other materials recycling is increasingly recognized as a vital strategy in sustainable construction. This is aimed at reducing the environmental burden of construction industry while conserving natural materials^{1,2}. RCAs, obtained from crushed and processed waste concrete offers significant environmental and economic benefits. These benefits include reduced landfill disposal, lower extraction of natural resources, and cost savings^{1,3-5}. However, RCA typically exhibits inferior qualities due to the residual cement paste adhered

School of Civil and Environmental Engineering, College of Technology and Built Environment, Addis Ababa University, Addis Ababa, Ethiopia. ✉email: hints.g.gebrezgiher@adu.edu.et

to the aggregate surface. These characteristics often result in reduced mechanical performance and durability compared to concrete made with natural aggregates, limiting the broader application of RCA in structural projects^{6–9}.

Carbonation treatment of RCA has emerged as an effective technique to address the above mentioned limitations. In the carbonation treatment process, carbon dioxide (CO₂) is introduced under controlled conditions to react with calcium hydroxide (Ca(OH)₂) and calcium silicate hydrate phases (C-S-H), forming calcium carbonate (CaCO₃) that fills pores. Then this carbonate polymorph densifies interfacial transition zones, and heals micro cracks. This chemical transformation improves aggregate quality. It can also enhance density, reduce water absorption, and increasing strength. Moreover, carbonation treatment provides the dual benefit of improving concrete performance while contributing to carbon sequestration and storage^{10–13}.

While the benefits of carbonated recycled aggregate concrete (ARAC) are well established, traditional laboratory testing of its mechanical properties is labor intensive, costly, and time consuming¹⁴. These constraints hinder rapid mix optimization and large scale adoption. Predictive modeling offers a practical alternative by enabling accurate property estimation based on mix proportions, material properties, and other parameters, thereby reducing the need for exhaustive experimental testing^{15–19}. Classical regression techniques, such as MLR, Ridge regression and polynomial MLR, have been widely applied in predicting concrete strengths due to their interpretability and ability to quantify linear and simple nonlinear relationships^{20–22}. However, these methods performs less when dealing with complex, nonlinear interactions.

Recent advances in tree and ensemble based ML models offer powerful tools for overcoming these limitations. Models such as random forest, decision tree, and LightGBM can capture nonlinearities, handle high dimensional data, and automatically identify important input features^{23–25}. Optimization frameworks such as k-fold cross validation enable systematic hyper parameter tuning and Shapley Additive explanations (SHAP) for interpretability²⁶.

Despite the growing body of research on predictive modeling for conventional concrete and RAC, there is a notable gap. To the best of the authors' knowledge, no study to date has developed predictive models considering parameters like carbonation treatment and parent concrete strength effects on the characteristics of RCAs and ARACs. Additionally, effects of multicollinearity and correlation between data sets on the predictive capacity of regression based models are not explored. Furthermore, systematic comparisons between classical regression based ML models and new ensemble based ML approaches for carbonated RCA remain unexplored.

The present study aims to fill these gaps by developing and evaluating predictive models for the compressive strength of carbonated RCA concrete. Composite indices for correlated data will be developed and a comparison of the ML data outputs using the individual dataset and aggregated datasets will be explored. Both classical regression based ML methods (MLR, ridge and Polynomial MLR) and advanced ML techniques (random forest, decision tree, and LightGBM) are employed. Model performance is compared in terms of predictive accuracy, robustness, interpretability, and computational efficiency. The research further seeks to identify the most influential parameters governing carbonated RAC performance.

Methodology

Source of data

The data used for strength predictions in this study were obtained from an experimental program reported in a previous work by the present authors²⁷. In that study, 108 concrete samples were tested for compressive strength. The datasets were categorized into five groups: (i) mix design proportioning of the target concrete, (ii) physical and mechanical properties of the coarse aggregates, (iii) parent concrete strength effects on RCA carbonation degree, (iv) quantified CO₂ stored by carbonation (expressed as percent by mass of mortar attached to RCAs), and (v) percentage replacement of natural aggregates with RCAs. The eight individual data sets from the 5 categories were, water to cement ratio (W/C), cement to Coarse aggregate ratio (C/C_a), cement to fine aggregate ratio (C/F_a), water absorption (W_a) of natural and recycled coarse aggregates in percent, crushing value (C_v) of natural and recycled coarse aggregates in percent, parent concrete strength (P_{cs}) of recycled aggregates, degree of carbonation (D_c) of recycled aggregates, and Replacement ratio (R_r) of natural coarse aggregates by recycled coarse aggregates. The descriptive statistics of these data sets and the output variable which is compressive strength (CS) are shown in Table 1.

Variable	Mean	St. Dev	Minimum	Maximum	Skewness	Kurtosis
C/C _a	0.3855	0.1251	0.2610	0.5100	-0.00	-2.04
C/F _a	0.7510	0.4179	0.3350	1.1670	-0.00	-2.04
W/C	0.4150	0.1557	0.2600	0.5700	0.00	-2.04
W _a (%)	3.629	1.096	2.250	5.910	0.82	-0.28
C _v (%)	22.952	4.073	18.010	30.710	0.76	-0.71
R _r (%)	57.89	38.83	1.00	100.00	0.05	-1.76
D _c (%)	4.141	3.500	0.100	11.930	1.06	0.25
P _{cs} (Mpa)	26.57	23.19	0.10	67.00	0.91	-0.58
CS (Mpa)	39.38	17.98	18.07	70.42	0.26	-1.63

Table 1. Descriptive statistics of the predictor variables and target variable.

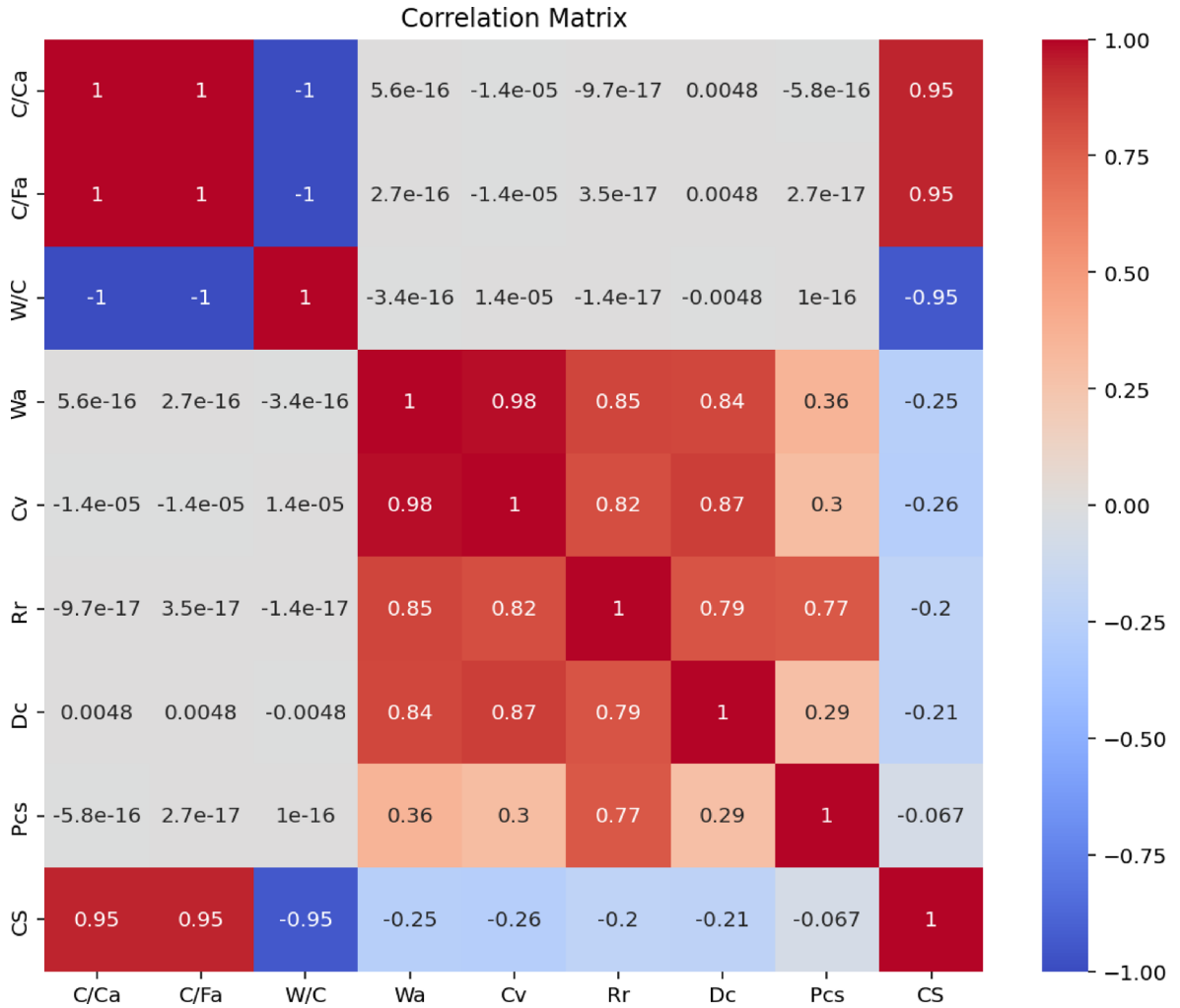


Fig. 1. Pearson’s correlation matrix between predictor variables before aggregation.

Term	Constant	C/C _a	C/C _a	W/C	W _a (%)	C _v (%)	R _r (%)	D _c (%)	P _{cs} (Mpa)	CS (Mpa)
VIF	0.00E+00	inf	inf	inf	4.46E+03	4.32E+02	4.35E+04	3.60E+03	1.31E+04	3.37E+01

Table 2. Multicollinearity between predictors before aggregation.

Data validity and assumption checks prior to regression based machine learning modeling

Prior to modeling, a series of statistical validity checks were conducted to ensure that the predictor variables were accurately represented in the models and that the fundamental assumptions of multiple regression and data processing were reasonably met^{28,29}. The initial step involved was examining the pairwise correlations among the predictor variables as shown in Fig. 1. It was observed that C/C_a, W/C, and C/F_a exhibited high intercorrelation to each other. In addition, a strong correlation was found between W_a, C_v, R_r, and D_c suggesting a high degree of correlation in the information provided by these predictors.

The presence of such high correlations raised a substantial concern regarding multicollinearity, which was further confirmed through Variance Inflation Factor (VIF) analysis as shown in Table 2³⁰. Extremely large VIF values (Table 2) were seen in the data sets, indicating multicollinearity and may lead to unstable coefficient estimates, inflate standard errors, and difficulties in determining the relative importance of individual predictors.

To address this issue, a combined index approach was implemented as a countermeasure^{30,31}. This involved aggregating highly correlated variables into composite indices, thereby retaining the essential information they contribute while reducing the interdependence between predictors. In this study, two composite predictors were developed to address multicollinearity among highly correlated variables, while retaining the essential engineering significance of each constituent parameter. The composites were formulated based on essential physical and mechanical characteristics of aggregates (Performance Index) and the strength governing proportions of the mix (proportioning Index) in a compact and statistically sound form using multiplicative, divisive, and exponential

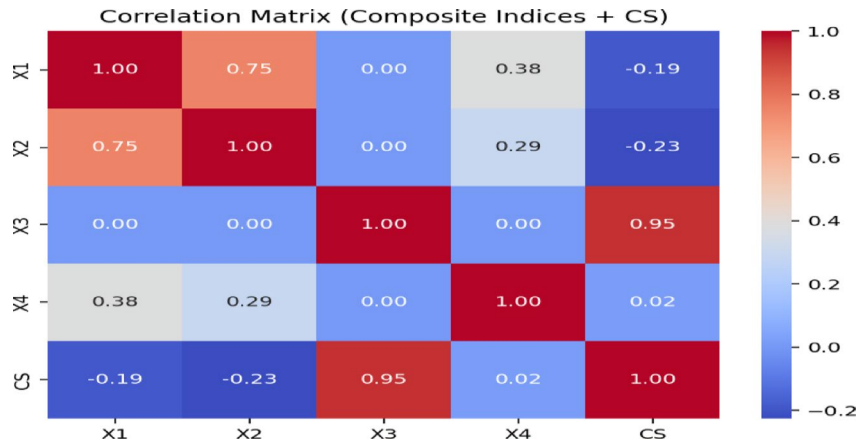


Fig. 2. Pearson’s correlation matrix after aggregation.

Feature	X ₁	X ₂	X ₃	X ₄
VIF	2.4911	2.3209	1.0001	1.1732

Table 3. VIF after aggregation.

relationships informed by domain knowledge and statistical optimization³¹. The composite indices were created for the proportioning variables (W/C , C/C_a and C/F_a), and aggregate performance indicators (W_a , C_v and P_{cs}). The new data set after aggregation were then became, replacement ratio (x_1), degree of carbonation (x_2), proportioning index (x_3) and performance index (x_4).

To perform the aggregation, domain knowledge was used to derive the proportioning index and performance indices (see Eq. 1 to 4). A statistical optimization was then applied to ensure that the correlation between predictors was reduced to below 0.8 and the VIF to below 5 (see Fig. 2; Table 3). This step was necessary to minimize instability and overfitting caused by multicollinearity. Based on these criteria, each variable was assigned an exponent of a_1 to a_9 . Having these exponents, the new indices were fed into a machine learning model, which was used to learn and determine the optimized exponents:

$$x_1 = R_r^{a_1} \tag{1}$$

$$x_2 = D_c^{a_2} \tag{2}$$

$$x_3 = \frac{(C/C_a)^{a_3} \times (C/F_a)^{a_4}}{(W/C)^{a_5}} \tag{3}$$

$$x_4 = \left(\frac{P_{cs}^{a_6}}{(W_a^{a_7} \times C_v^{a_8})} \right)^{a_9} \tag{4}$$

Where, x_1 is the replacement ratio =, x_2 is the degree of carbonation =, x_3 is the proportioning index and x_4 is the performance index of the aggregates.

After optimization using ML the values of the exponents were obtained as follows:

$a_1 = 1.911$, $a_2 = 0.103$, $a_3 = 1.165$, $a_4 = 0.801$, $a_5 = 1.481$, $a_6 = 1.053$, $a_7 = 1.752$, $a_8 = 0.986$, $a_9 = 0.659$. All correlation values and VIFs fell below their respective limits, as shown in Table 3; Fig. 2.

The proportioning index (x_3) reflects the balance between binder content, coarse and fine aggregate mass, and water dosage in the recycled aggregate concrete mix. The product of the cement to coarse aggregate and cement to fine aggregate ratios represents the overall binder richness of the mix relative to the aggregate mass. Dividing by the water to cement ratio adjusts the index to account for water content, a critical determinant of workability and strength development. Physically, a higher proportion index value corresponds to mixes with greater binder content, generally associated with higher strength and reduced porosity. The index therefore encapsulates the combined influence of mix composition parameters in a single value, reducing collinearity between the original ratios while retaining their collective predictive power.

The performance index (x_4) integrated three critical aggregate characteristics into a single measure of aggregate quality. Parent concrete strength, in the numerator, reflects the intrinsic load bearing capacity of the original aggregate source. The denominator combines water absorption, representing pore structure, and aggregate crushing value, representing resistance to mechanical breakdown. The physical interpretation of this composite is straightforward where higher parent strength increases the index, while higher porosity or lower

mechanical durability reduces it. Thus, the index expresses a net measure of aggregate performance. Statistically, combining these collinear parameters into one variable mitigates redundancy in regression modelling, ensuring VIF remain within acceptable limits while preserving the behavioral influence of each individual property. The exponents in each predictor are determined based on the statistical optimization that the data fulfils all the statistical checks (see Table 3; Fig. 2). To prevent the results from the composite indexes from becoming zero or infinite, small insignificant values were assumed instead of using zero for the natural aggregate concrete. For the replacement ratio, a value of 1% was used instead of 0%, while for the degree of carbonation and parent concrete effects, a value of 0.1% was assigned. These substitutions were necessary because several of the results in the composite indexes or individual results involved terms in which these variables appear in multiplication or division; using zero would lead to multiplication or division by zero errors and consequently produce zero or infinite index values. The selected small values are sufficiently low to preserve the physical meaning of “no replacement” or “no measurable effect,” while simultaneously providing numerical stability in the calculations. They act only as mathematical placeholders and do not influence the actual material behavior or alter the interpretation of natural aggregate concrete.

Following the construction of composite indices for highly correlated predictors, two separate modeling pathways were pursued. The aim of this stage was to determine whether the presence of intercorrelation among predictors influences the predictive performance of regression based ML models. In the first pathway, models were trained using the original dataset containing the individual predictors. In the second pathway, models were trained using the composite dataset, in which correlated predictors had been aggregated into indices. Both modeling pathways were implemented under identical conditions where the same data preprocessing steps, including training and testing partitioning, and algorithm settings were applied to ensure fair comparison. Model performance was assessed using standard regression evaluation metrics. Comparing the predictive outcomes of the two models, we can see if intercorrelation among predictors, has a significant effect on the models’ accuracy and stability.

Figure 3 shows the frequency distribution of the aggregated predictors and the target compressive strength values. The compressive strength exhibits a bimodal distribution, corresponding to normal strength concrete and high strength concrete.

Principles of the classical statistical models and ensemble based algorithms

In numerical modeling, selecting the right regression approach depends not only on the nature of the data but also on the underlying assumptions about how features influence the target³². This section explores a spectrum of regression methods from classical statistical models to advanced machine learning architectures. Each technique is briefly outlined, focusing on its internal logic, typical use cases, and how it adapts to different data structures. Three classical regression based algorithm and three tree and ensemble learning based algorithms were used in this study and these were, MLR, ridge regression, ploy MLR, random forest, decision tree, and LightGBM. Abrams-law baseline was used to show the value added by ML models beyond classical relations. Abrams law states that the strength of concrete is inversely related to the water to cement ratio and depends only on this ratio. As the water to cement ratio increases, the compressive strength decreases³³.

Classical MLR models

A regression model provides a mathematical framework for describing how one variable of interest changes in relation to one or more influencing factors. If the model incorporates only a single predictor, it is termed a simple regression. But, when it includes two or more predictors, it becomes a multiple regression model. Such models are widely used to quantify the degree and nature of association between explanatory variables and the response, with the aim of revealing patterns and dependencies that might otherwise remain hidden³².

In a simple regression scenario, the model seeks a single straight line that best represents the relationship between the dependent and independent variable across all observations. Multiple regression extends this principle to higher dimensions that, instead of fitting a line, it estimates a hyperplane that best represents the combined effect of several predictors. While the inclusion of multiple variables increases analytical complexity, it also allows for richer interpretations, including the possibility that the influence of one predictor may vary depending on the level of another. These dependencies may manifest as linear trends, curvature, or conditional effects³⁴.

The general structure of MLR model can be expressed as:

$$y = b_0 + \sum_{j=1}^n b_j x_j \quad (5)$$

b_0 is the intercept, and b_j are the partial regression coefficients reflecting the marginal contribution of each predictor. The coefficients are typically estimated using the least squares method, which minimizes the sum of squared differences between observed and predicted values³⁴. Ridge regression is a type of regularized linear regression that helps prevent overfitting by adding a penalty on the size of the model coefficients.

Polynomial regression builds on the MLR framework by introducing higher order and interaction terms, enabling the model to capture curved relationships without discarding the linear estimation approach. This enhancement increases the model’s flexibility in fitting data that exhibit nonlinear patterns. However, the inclusion of many polynomial terms also increases the likelihood of overfitting, particularly when the number of predictors or polynomial degree is high. Nevertheless, for datasets with relatively smooth curvature, polynomial MLR can yield highly accurate and interpretable fits. The polynomial MLR equation of degree two is expressed as:

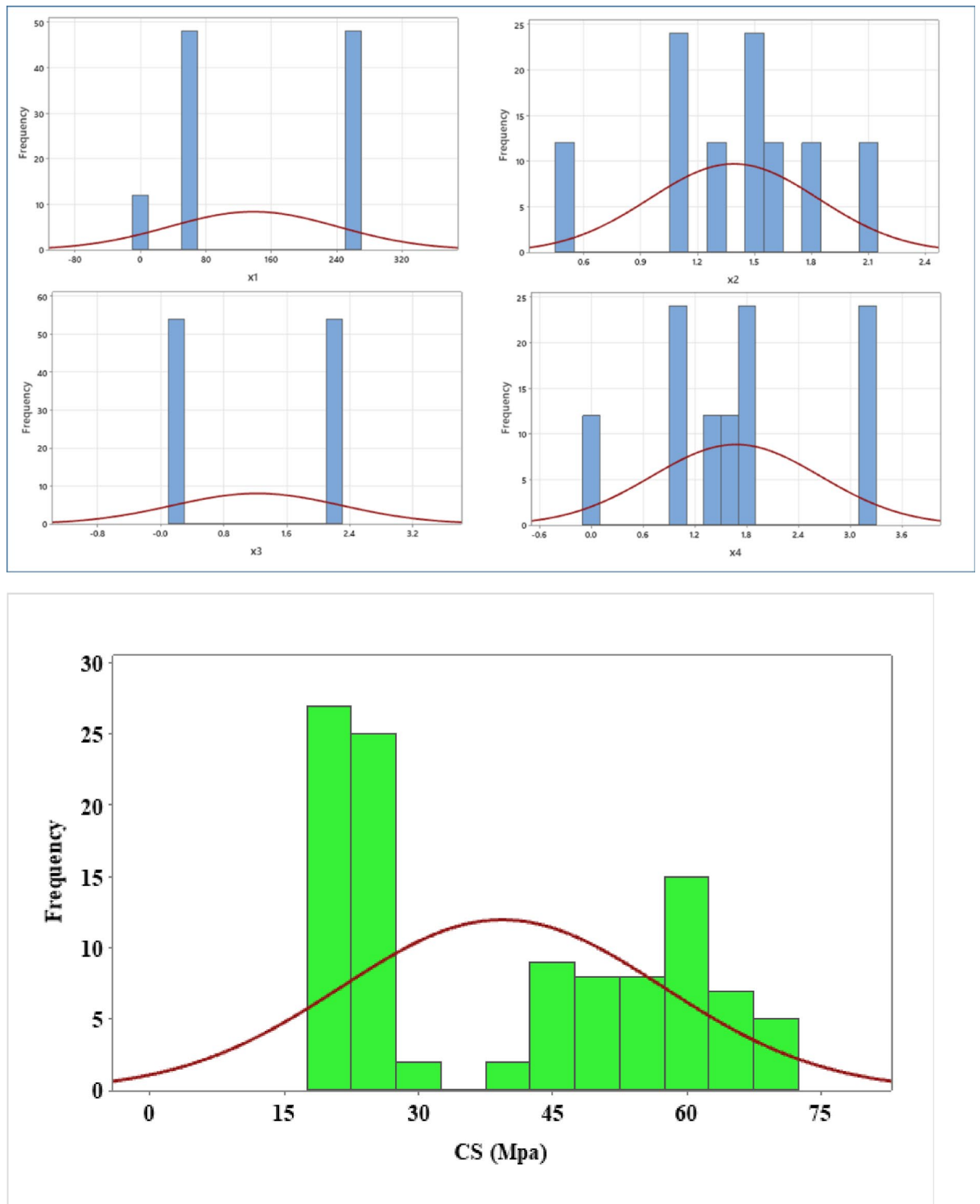


Fig. 3. Frequency distribution of features after aggregation.

$$y = b_0 + \sum_{j=1}^n b_j x_j + \sum_{j=1}^n b_{jj} x_j^2 + \sum_{1 \leq i < j \leq n} b_{ij} x_i x_j \quad (6)$$

Where b_{jj} denotes coefficients for the squared terms of each predictor, and b_{ij} corresponds to coefficients for interaction terms between predictors.

Tree and ensemble based learning algorithms

(a) Decision tree

The algorithm works by recursively splitting the dataset based on feature values that provides the best discriminatory power, resulting in a tree like structure³⁵. Each internal node represents a decision based on the input feature, and the leaves represent the specific classification outcome or value of the tested attribute. The tree is constructed using a greedy approach, selecting the best feature to split the data at each step³⁶. Decision trees partition the feature space into regions using recursive binary splits. For input x , prediction is the average of target values in the corresponding region:

$$y(x) = \sum_{n=1}^n c_n 1(x \in R_n) \quad (7)$$

Where, R_n are regions (leaf nodes), C_n is the mean response in R_n , and $1(x \in R_n)$ is the indicator function. Splits are chosen to minimize impurity measures.

(b) Random forest

Random forest, an ensemble learning technique, builds multiple decision trees and aggregates their results to improve predictive performance and reduce overfitting^{35,37}. Each tree in the forest is trained on a random subset of the data, and predictions of each are combined through majority voting to produce the final classification. The Random forest model was used to compare its performance to the decision tree, with expectations of improved accuracy and reduced variance. This method is often preferred in situations where robustness and generalization are important³⁵. Random forest's average predictions are generated from many decision trees built on bootstrapped data samples and random feature subsets^{38,39}:

$$y(x) = \frac{1}{T} \sum_{t=1}^T y_t(x) \quad (8)$$

Where, $y_t(x)$ is the prediction of the t^{th} tree and T is the total number of trees. This reduces variance and overfitting compared to a single decision tree^{39,40}.

(c) LightGBM

LightGBM is a gradient boosting framework that builds decision trees sequentially, each one correcting errors of the previous. It uses advanced optimizations such as leaf wise tree growth, gradient based one side sampling, and exclusive feature bundling, making it highly efficient for large scale tabular data⁴¹. LightGBM achieves better performance on many tabular tasks, handling complex nonlinear relationships with minimal manual feature engineering. Interpretability is moderate, with tools like SHAP often needed to explain predictions. Overfitting is possible, but well controlled with parameters like tree depth, learning rate, and regularization. It is highly scalable and computationally efficient, though essentially a black box model. The model results are found by adding the prediction of all the ensemble trees (Eq. 9), where $f_n(x)$ indicates the prediction from the n^{th} individual tree^{40,41}.

$$y(x) = \sum_{n=1}^n f_n(x) \quad (9)$$

Cross validation strategy and hyperparameter tuning

Model evaluation employed fivefold cross validation to ensure robust performance assessment and minimize bias due to data partitioning. Model hyperparameters were tuned using grid search with cross validation. To ensure full reproducibility, all procedures used fixed random seeds (split seed=42, kfold seed=42, search seed=42) and search spaces as shown in Table 4.

Quantitative performance metrics

To provide a comprehensive evaluation of predictive accuracy and robustness, multiple statistical metrics were computed for both training and testing sets:

R^2 was applied to assess explanatory power. RMSE used to measure prediction error in absolute terms. While MAE used to determine average magnitude of error. The expected values for the indice R^2 should ideally be 1, indicating good prediction level, while RMSE, and MAE should ideally be 0, indicating less errors in the predictions. In general, a predictive model is ideal if its performance indices are close to or strictly at these values⁴². The equations for calculating these indices are presented in Eqs. (10–12) as follows^{42,43}.

$$R^2 = 1 - \frac{\sum_{n=1}^n (y_i - \hat{y}_i)^2}{\sum_{n=1}^n (y_i - \bar{y})^2} \quad (10)$$

Model	Hyperparameter	Search space
Decision tree	Max depth	None, 3, 5, 10, 20
	Min samples split	2, 5, 10, 20
	Min samples leaf	1, 2, 4, 8
Random forest	n_estimators	100, 200, 400
	Max depth	None, 5, 10, 20
	Min samples split	2, 5, 10
	Min samples leaf	1, 2, 4
LightGBM	n_estimators	100, 200, 400
	num_leaves	15, 31, 63
	Max depth	-1, 5, 10, 20
	Learning rate	0.001, 0.01, 0.05, 0.1
	Min child samples	5, 10, 20, 50
Ridge Regression	Alpha	50 logarithmically spaced values between 1×10^{-6} and 1000
Polynomial MLR (deg=2)	Polynomial degree	Fixed at 2
MLR	Parameters	Default fit (no tuning)

Table 4. Search spaces and hyperparameters. Optimal configurations used for each model were MLR: default fit Polynomial MLR (deg = 2): poly_degree = 2 Ridge: alpha = 0.09102981779915227 Decision tree: max_depth = None, min_samples_split = 2, min_samples_leaf = 1 Random forest: n_estimators = 400, max_depth = 20, min_samples_split = 2, min_samples_leaf = 1 LightGBM: n_estimators = 400, num_leaves = 15, min_child_samples = 10, max_depth = -1, learning_rate = 0.1

Where, y_i are observed value, \hat{y} are predicted value and \bar{y} are mean of observed value.

$$RMSE = \sqrt{\frac{1}{n} \sum_{n=1}^n (y_i - \hat{y})^2} \quad (11)$$

$$MAE = \frac{1}{n} \sum_{n=1}^n |y_i - \hat{y}| \quad (12)$$

Visual performance metrics

When evaluating models, visual checks are extremely useful because they make complex statistics easier to understand at a glance⁴². To complement the numerical evaluation of model performance, visual diagnostic techniques were applied. These techniques were, scatter plots, residual plots, histograms of residuals, and quantile-quantile (Q-Q) plots. Scatter plots, in which predicted values were plotted against the observed data, provided an overview of how closely the model output followed the actual trend. Ideally, points cluster around the line, and noticeable departures from this line highlight areas of reduced reliability. Residual plots offered further insight by presenting the distribution of prediction errors. Systematic patterns, such as consistent underestimation or overestimation in certain ranges or increasing error spread with larger values, were examined to identify potential weaknesses in the models.

Histograms of the residuals were used to evaluate the overall error distribution. A distribution centered near zero with a roughly symmetric shape indicates unbiased predictions, while skewed or shifted distributions suggest bias. In addition, Q-Q plots were generated to compare the residuals with a theoretical normal distribution for the test dataset. Alignment of points along the diagonal indicated that the normality assumption was reasonable, whereas systematic deviations reflected non-normal error structures.

Feature importance analysis

To investigate the sensitivity of machine learning models and to interpret their behavior at both broad and granular levels, the SHAP framework, grounded in cooperative game theory principles, was applied⁴⁴. This method was used to evaluate the relative contribution of each input variable to the model's predictions. As a sophisticated tool within explainable artificial intelligence, SHAP provides a transparent view of how input features interact and influence the output. By quantifying feature impacts, it highlights which variables play the most significant roles in shaping predictions and clarifies the direction and magnitude of their effects on the model's results^{42,45}.

Pipeline and scientific libraries

The experimental pipeline was implemented in Python 3.10 with a notebook of Spider version 6, using Scikit-learn, LightGBM, and supporting scientific libraries.

Result and discussion

Data correlation effects on model performance

The outputs from the different machine learning modeling approaches showed that the predictions using the individual datasets and composite indices on the MLR based models have shown some differences. But in the tree based models there have not shown tangible differences. Figure 4 shows the effects of data type on the residual compressive strength outputs of different ML models. The residual are scattered following the same line on the tree based models for both datasets but some deviations are seen on MLR, Ridge regression and Poly MLR. It is shown that the individual data performed well on these models.

The quantitative matrices comparisons between models trained on individual predictors and those using composite indices are shown on Tables 5 and 6. As a baseline model, Abrams' law yielded R^2 , MAE, and RMSE values of 0.668, 5.774, and 9.849, respectively. These results are considerably lower in accuracy compared to the ML models. With the composite predictors, MLR, Ridge regression and Poly MLR achieved fair accuracy with test R^2 of 0.965, 0.967 and 0.983 respectively. Similarly with the individual predictors they achieved a test R^2 values of 0.975, 0.975 and 0.991 respectively. Decision tree, random forest and LightGBM models all achieved very high predictive performance (test $R^2 > 0.991$) under both datasets, with minimal train-test differences. Overall, the results indicate that predictor intercorrelation is a minimal concern for classical linear models, where it may contribute to overfitting and unstable parameter estimates. For more advanced models, particularly ensemble tree methods, multicollinearity had no effect, as they inherently manage redundancy among predictors. This suggests that while composite indexing can improve the stability of simpler models such as MLR, modern ML algorithms can reliably predict concrete compressive strength even in the presence of highly correlated predictors. Nonetheless, the composite indices continue to demonstrate substantial value by greatly simplifying model variables. Equations (15) and (16) reveal that prediction equations from classical models, especially polynomial MLR models are often long. Using composite indices, as shown in Eqs. (13) and

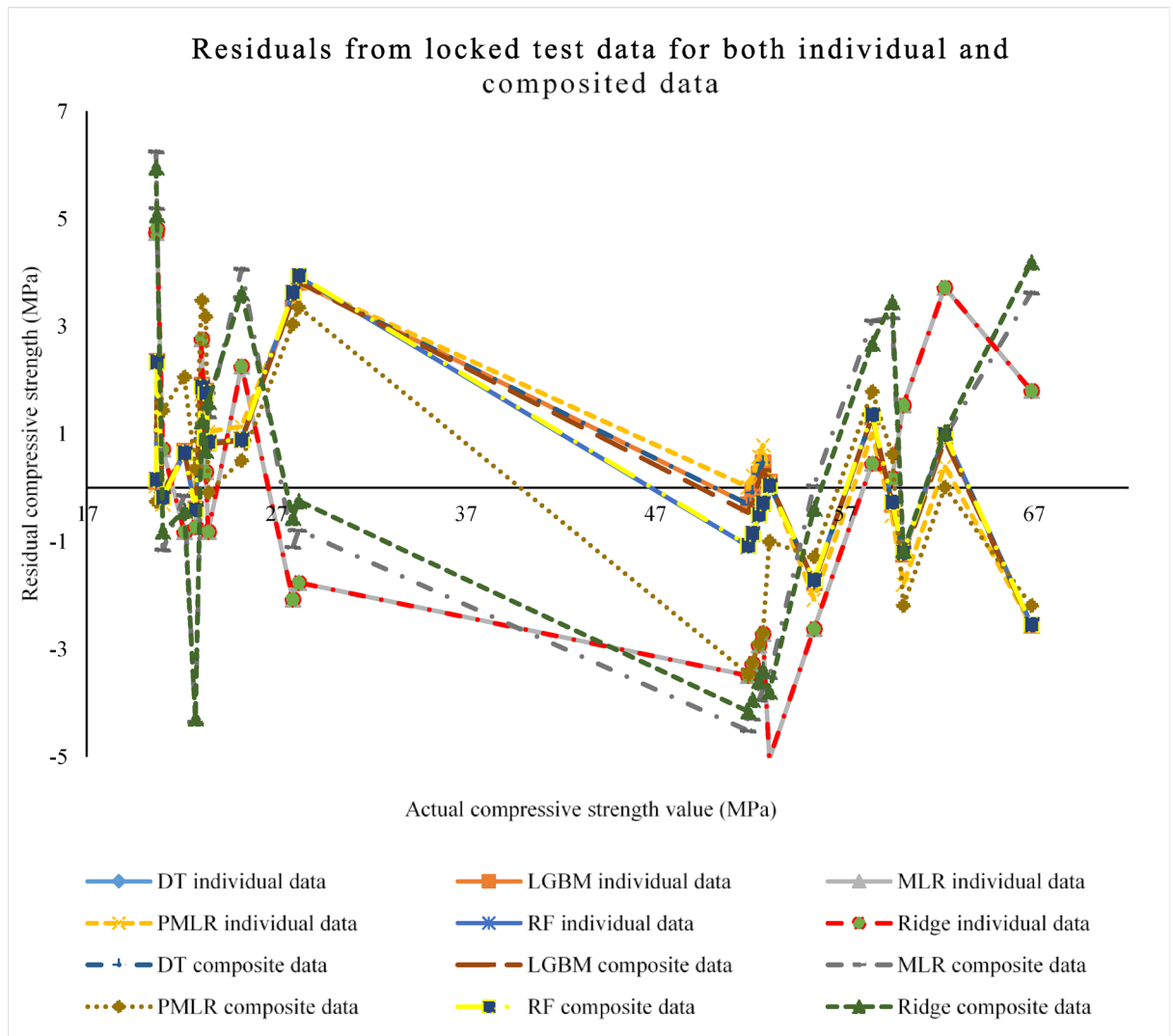


Fig. 4. Model output comparisons for locked datasets.

Final test results (Aggregated composite data)													
Model	Training						Testing						
	MAE	RMSE	R2	95th percentile error	99th percentile error	Max. error	MAE	RMSE	R2	95th percentile error	99th percentile error	Max. error	
Decision tree	0.7276	0.9032	0.9975	1.7733	2.1496	2.3283	1.1732	1.5914	0.9913	3.4575	3.7499	3.8150	
LightGBM	0.7247	0.9041	0.9975	1.8207	2.1463	2.2699	1.1683	1.5941	0.9913	3.4623	3.7551	3.8202	
Random forest	0.7349	0.9115	0.9975	1.7680	2.0993	2.3577	1.2493	1.6363	0.9908	3.5792	3.8788	3.9439	
Polynomial MLR (deg = 2)	1.0225	1.2883	0.9949	2.3979	2.9620	3.6508	1.8854	2.2251	0.9830	3.4555	3.4822	3.4879	
Ridge	3.3034	3.9182	0.9531	7.1056	7.8891	8.6654	2.5568	3.0824	0.9675	5.0352	5.7710	5.9561	
MLR	3.3152	3.9067	0.9534	6.8441	7.5090	8.5545	2.6536	3.1926	0.9651	5.1591	6.0281	6.2501	

Table 5. Quantitative performance indicators of models using aggregated data.

(14), has reduce this complexity. The indices condensed multiple variables into single, coherent metrics, making the models far more manageable and actionable.

Where, DT is decision tree, PMLR is polynomial MLR, LGBM is lightGBM, RF is random forest.

$$MLR CS_{Agg.} = 48.150 - 0.00048x_1 - 26.935x_2 + 9.525x_3 + 13.573x_4 \quad (13)$$

Where MLR CS_{Agg.}, is the compressive strength prediction equation derived from multilinear regression based machine learning and aggregated composite predictors.

$$\begin{aligned} poly MLR CS_{Agg.} = & 6435.953 + 1.417x_1 - 12271.462x_2 + 1.2448x_3 \\ & - 5030.874x_4 - 0.000144x_1^2 - 0.294x_1x_2 - 0.000018x_1x_3 - 0.110x_1x_4 \\ & + 5275.472x_2^2 - 11.267x_2x_3 + 4889.829x_2x_4 + 4.975x_3^2 + 4.125x_3x_4 - 174.564x_4^2 \end{aligned} \quad (14)$$

Where, poly MLR CS_{Agg.}, is the compressive strength prediction equation derived from polynomial multilinear regression based machine learning and aggregated composite predictors.

$$\begin{aligned} MLR CS_{Indv.} = & 374.051 + 10.277 \frac{C}{C_a} + 34.339 \frac{C}{F_a} - 12.794 \frac{W}{C} \\ & - 91.591W_a - 8.310C_v + 8.076R_r - 25.664D_c - 7.424P_{cs} \end{aligned} \quad (15)$$

Where MLR CS_{Indv.}, is the compressive strength prediction equation derived from multilinear regression based machine learning with individual predictor data sets.

$$\begin{aligned} Poly MLR CS_{Indv.} = & 792224.051 - 170.639 \frac{C}{C_a} - 570.167 \frac{C}{F_a} + 212.442 \frac{W}{C} - 1425.529W_a \\ & - 187.573C_v - 1523.173R_r - 529.283D_c - 1756.424P_{cs} - 131.562 \left(\frac{C}{C_a}\right)^2 - 347.949 \frac{C}{C_a} \frac{C}{F_a} \\ & + 11.081 \frac{C}{C_a} \frac{W}{C} - 138.582 \frac{C}{C_a} W_a - 28.542 \frac{C}{C_a} C_v - 337.069 \frac{C}{C_a} R_r - 67.744 \frac{C}{C_a} D_c \\ & - 359.610 \frac{C}{C_a} P_{cs} - 856.391 \left(\frac{C}{F_a}\right)^2 - 77.075 \frac{C}{F_a} \frac{W}{C} + 302.591 \frac{C}{F_a} W_a + 5.373 \frac{C}{F_a} C_v \\ & - 308.182 \frac{C}{F_a} R_r + 57.918 \frac{C}{F_a} D_c - 258.222 \frac{C}{F_a} P_{cs} + 176.327 \left(\frac{W}{C}\right)^2 - 1103.230 \frac{W}{C} W_a \\ & - 132.331 \frac{W}{C} C_v - 943.503 \frac{W}{C} R_r - 389.336 \frac{W}{C} D_c - 1124.185 \frac{W}{C} P_{cs} - 5794.084W_a^2 \\ & - 23828.927W_aC_v + 9449.029W_aR_r + 425.264W_aD_c + 2064.348W_aP_{cs} + 581.360C_v^2 \\ & + 1165.162C_vR_r - 6957.866C_vD_c - 2097.720C_vP_{cs} - 555.217R_r^2 + 3997.994R_rD_c \\ & + 355.644R_rP_{cs} + 444.873D_c^2 - 7790.791D_cP_{cs} + 526.695P_{cs}^2 \end{aligned} \quad (16)$$

Where, poly MLR CS_{Indv.} is the compressive strength prediction equation derived from polynomial multilinear regression based machine learning with individual predictor data sets.

Model performance comparisons

The series of plots comparing actual versus predicted compressive strength values for different models are shown on Fig. 5 revealing distinct performance patterns across the methods. MLR, Ridge and Polynomial MLR (Fig. 5, a, b and f) has shown reasonably good agreement between predicted and actual values. The points are largely clustered around the ideal 45° line, indicating that these models capture the general trend of the data. The polynomial MLR improves alignment for some values, suggesting that introducing a nonlinear term helps capture some curvature in the relationship. However, deviations at some values indicate some residual bias that these classical regression models may not fully capture^{46–49}.

Decision tree, random forest and LightGBM (Fig. 5, c, d and e) has demonstrated improved prediction accuracy over classical regression models. Decision tree models already provide a better fit compared to MLR for nonlinear patterns, while Random Forest further enhance predictive performance by averaging multiple trees, reducing overfitting, and capturing complex interactions. The clustering of points near the ideal line in (Fig. 5, c, d and e) confirms their robustness across test datasets. These models effectively manage the nonlinearity and heteroscedasticity in the data^{48–50}.

The figural representations of the quantitative performance metrics (R² and MAE) and visual performance metrics (Q-Q plots and residual histograms) of the regression based ML models are shown on Figs. 6, 7 and 8, as well as all quantitative performance metrics scores of these models are shown on Table 5 for aggregated datasets and Table 6 for individual data sets. These results provides valuable insights into the predictive performance of the various ML models on the given dataset. The evaluation of the qualitative performance metrics R², MAE and RMSE for training and testing datasets, as well as an analysis of the visual performance metrics like the distribution of the models' residuals, revealed a clear hierarchy of model effectiveness, with nonlinear, tree based models outperforming traditional linear and polynomial regression methods. As shown in the performance metrics, the decision tree, random forest and LightGBM models demonstrated superior predictive accuracy and stability. These models achieved the best qualitative metrics values, indicating that they were most effective at capturing the underlying patterns in the data and minimizing prediction error. This performance is consistent with the nature of these models, which are well suited to handle complex, nonlinear relationships and intricate

Final test results (Individual data)													
Model	Training						Testing						
	MAE	RMSE	R2	95th percentile error	99th percentile error	Max_error	MAE	RMSE	R2	95th percentile error	99th percentile error	Max_error	
Decision Tree	0.7276	0.9032	0.9975	1.7733	2.1495	2.3283	1.1732	1.5914	0.9913	3.4574	3.7499	3.8150	
LightGBM	0.7276	0.9032	0.9975	1.7733	2.1495	2.3283	1.1733	1.5914	0.9913	3.4578	3.7503	3.8154	
Polynomial MLR (deg = 2)	0.7439	0.9505	0.9972	1.9631	2.1576	2.3507	1.2056	1.6245	0.9910	3.4574	3.7499	3.8150	
Random Forest	0.7364	0.9117	0.9975	1.7680	2.0993	2.3577	1.2497	1.6322	0.9909	3.5530	3.8513	3.9164	
Ridge	2.6039	3.1980	0.9688	5.8326	6.3379	6.5610	2.2520	2.6890	0.9752	4.8088	4.9973	5.0464	
MLR	2.6039	3.1980	0.9688	5.8326	6.3379	6.5611	2.2520	2.6890	0.9752	4.8087	4.9973	5.0465	

Table 6. Quantitative performance indicators of models using individual data.

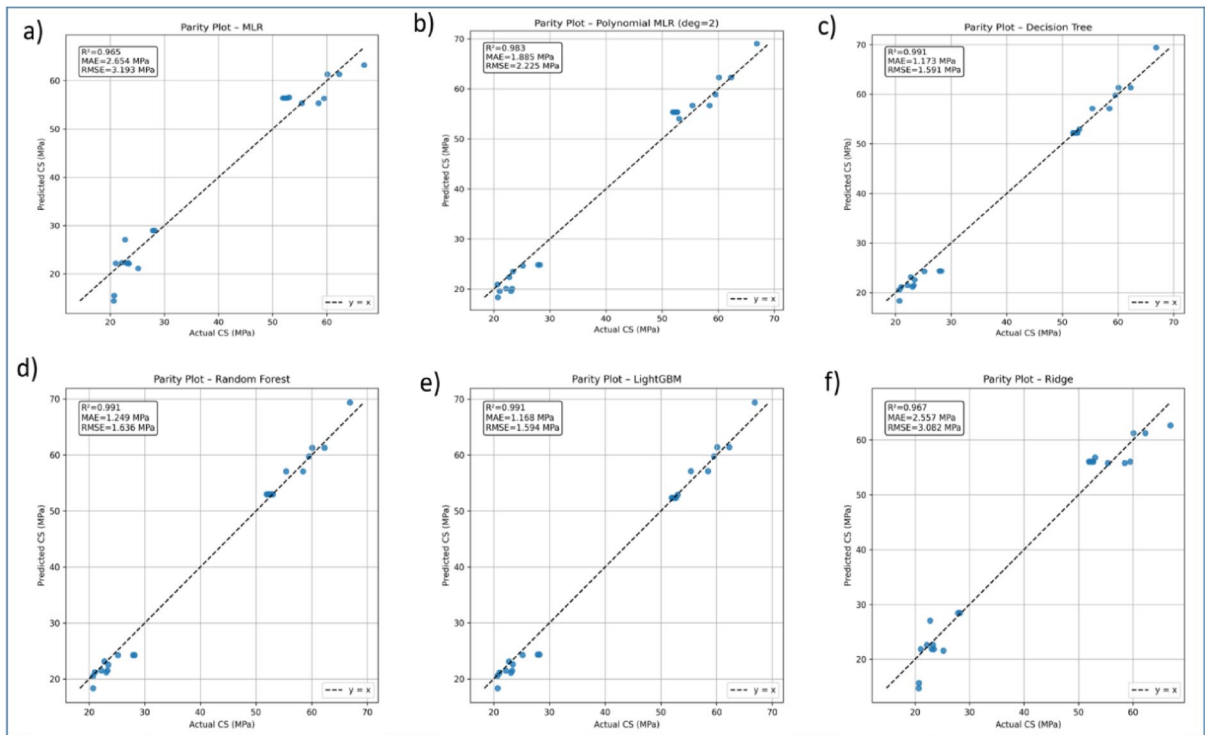


Fig. 5. Relationship between actual experimental compressive strength and predicted values in multiple models.

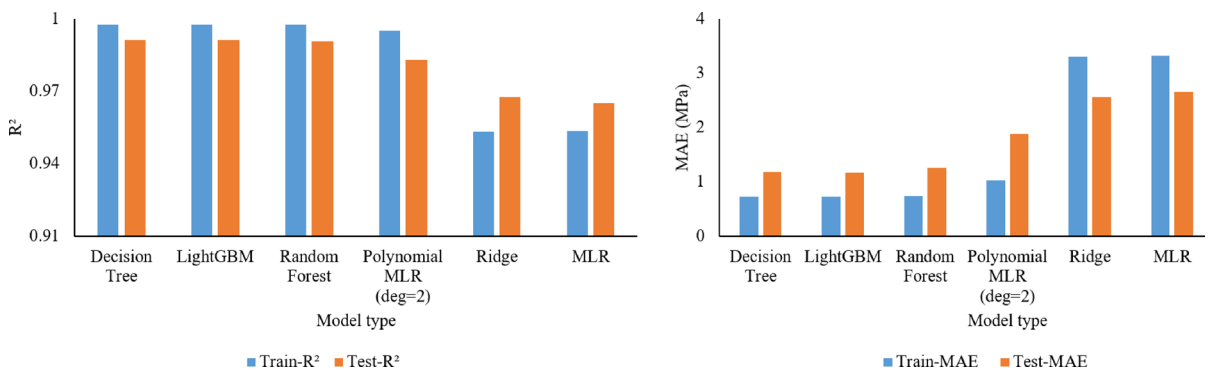


Fig. 6. Models performance comparisons using R² and MAE.

feature interactions that linear models cannot. Many studies have compared different models for better performances based on their data types^{18,51,52}.

Further evidence of the models' performance can be found by examining the distribution of their residuals using Q-Q plots and residual histograms. The histograms visually represent the frequency of different residual values, while the Q-Q plots compare the quantiles of the residuals to the quantiles of a theoretical normal distribution. For a well fitting model, both plots should demonstrate a bell shaped curve and a close alignment with the diagonal line, respectively⁵³.

The residual histograms for the decision tree, random forest and LightGBM models show a fair distribution that is close to a bell shaped curve, with residuals centered around zero. This is a strong indication that these models are fairly capturing the systematic relationships in the data. The corresponding Q-Q plots for these models show that their residuals are close to a normal distribution.

Conversely, the MLR and ridge regression models yielded the weakest performance, with the lowest test R² (0.965) and the highest MAE (2.65). This indicates that the relationship between the predictors and the target variable is likely not linear, and that a simple linear model is unable to model the data as accurate as tree based models. The residual histogram for the Linear MLR models is noticeably less centered and shows a wider spread, indicating less prediction capability compared to other models. An improvement in performance observed in the Polynomial MLR model which achieved a better qualitative performance than the standard MLR further

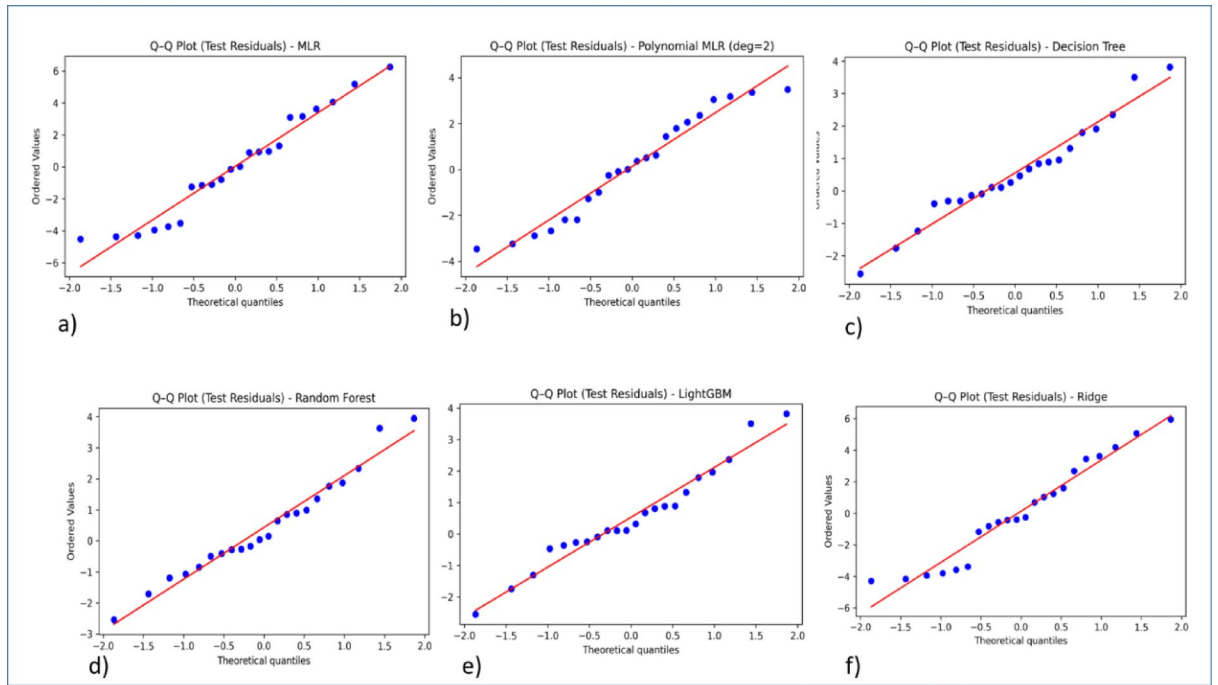


Fig. 7. Q-Q plot of the test residuals of the models.

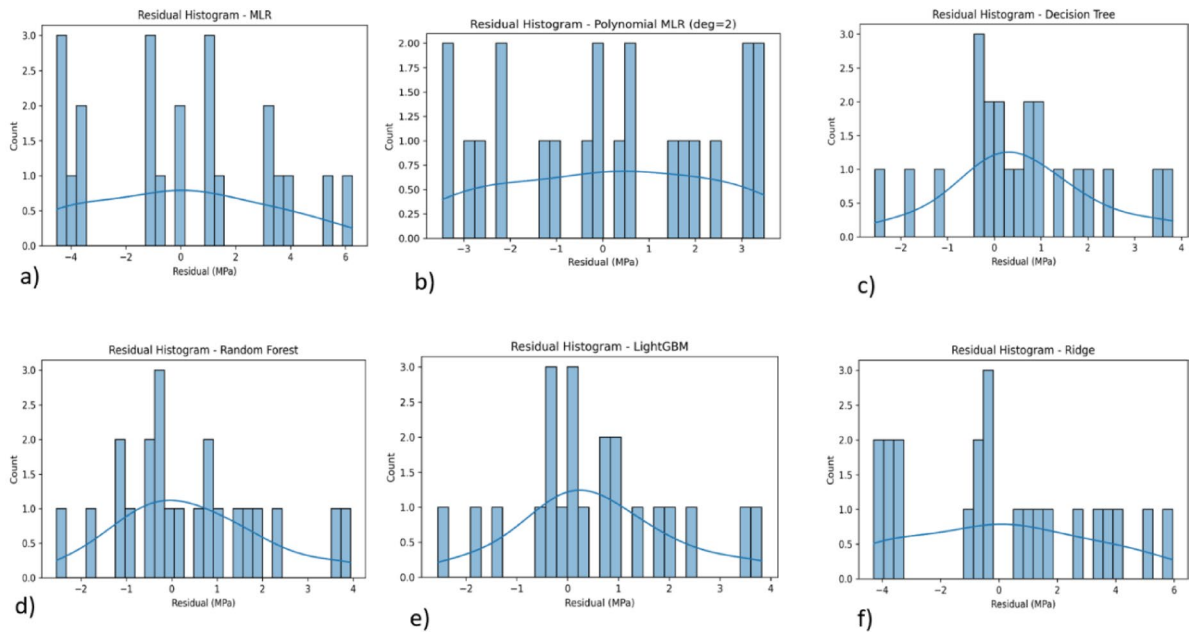


Fig. 8. Histogram of residual of the prediction models.

supports the conclusion that nonlinear relationships are a key characteristic of this dataset. However, its residual histogram, while better than the standard MLR, is still not as well formed as the tree based models’ histograms, suggesting that even with polynomial terms, a linear based model struggles to fully capture the data’s structure.

Analysis of average residuals for each model also showed that the poly MLR, Decision Tree, Random Forest and LightGBM models all demonstrated strong performance on the testing data, outperforming the standard linear MLR models.

The simple multilinear models exhibited an average locked test residual values of not less than 2.56 Mpa on the individual data and 2.65 Mpa on composite data. The scatter plots further confirm this performance, as the predictions from the MLR models show a wide deviation of up to 6.3 MPa, falling within a $\pm 30\%$ maximum

error range from the actual experimental data. In contrast, the poly MLR model, which added complexity, achieved a lower testing average residual of 1.21 Mpa on the individual data and 1.88 Mpa on the composite data, showing its superior ability to generalize to new, unseen data. This model's better ability to generalize is further supported by its scatter plot analysis, which shows a significantly more accurate prediction with only a $\pm 15\%$ maximum error range.

Similarly, the decision tree, random forest and LightGBM models also performed well on the test set. The decision tree, random forest and LightGBM models testing average residuals were 1.17 Mpa, 1.25 MPa and 1.17 MPa, respectively, on both the individual and composited datasets. The scatter plots for Decision Tree, Random Forest and LightGBM models reinforce their accuracy, showing a tight error range of $\pm 13\%$ from the experimental data.

Sensitivity analysis

This section examines the SHAP summary plots of the individual predictors (C/C_a , C/F_a , W/C , W_a , C_v , D_c , P_{cs} and R_r) as can be seen on Fig. 9 and the composite induced predictors (x_1 , x_2 , x_3 and x_4) as can be seen on Fig. 10, to evaluate the influence of predictor variables on the ML model for concrete compressive strength⁵⁴. SHAP values on the horizontal axis indicate each feature's contribution to model predictions were positive values enhance the output, and negative values reduce it based on the predictor behavior. The color scale represents feature magnitude, allowing identification of whether high or low feature values drive positive or negative contributions¹⁴.

The analysis shows that all predictors impact the model, though to varying degrees. The most influential variables are the cement and water related ratios (W/C , C/C_a and C/F_a), forming a composite index x_3 . High C/C_a and C/F_a ratios combined with low W/C strongly increase predicted compressive strength, consistent with standard mix design principles.

The degree of carbonation ranks second in importance. Its effect looks complex and nonlinear, with both high and low values producing substantial positive and negative contributions. This was due to the incomplete carbonation of the high parent concrete aggregates. The percentage of the carbonation byproducts in low strength parent concrete aggregates were higher compared to aggregates sourced from high strength parent concrete. But conversely, even though the amount of carbonation by products were less in the high strength parent concrete aggregates, the resultant compressive strength of the concrete produced from these aggregates was higher compared to the aggregates produced from low strength parent concrete. This behavior introduced the nonlinear behavior in the SHAP.

Composite index x_4 , encompassing (W_a and C_v along with P_{cs}), shows moderate influence. Individually these factors were impactful. However, collectively they exert less effect than the mix proportioning index variables and carbonation. The replacement ratio of natural with recycled aggregates (R_r , composite index x_1) has the least

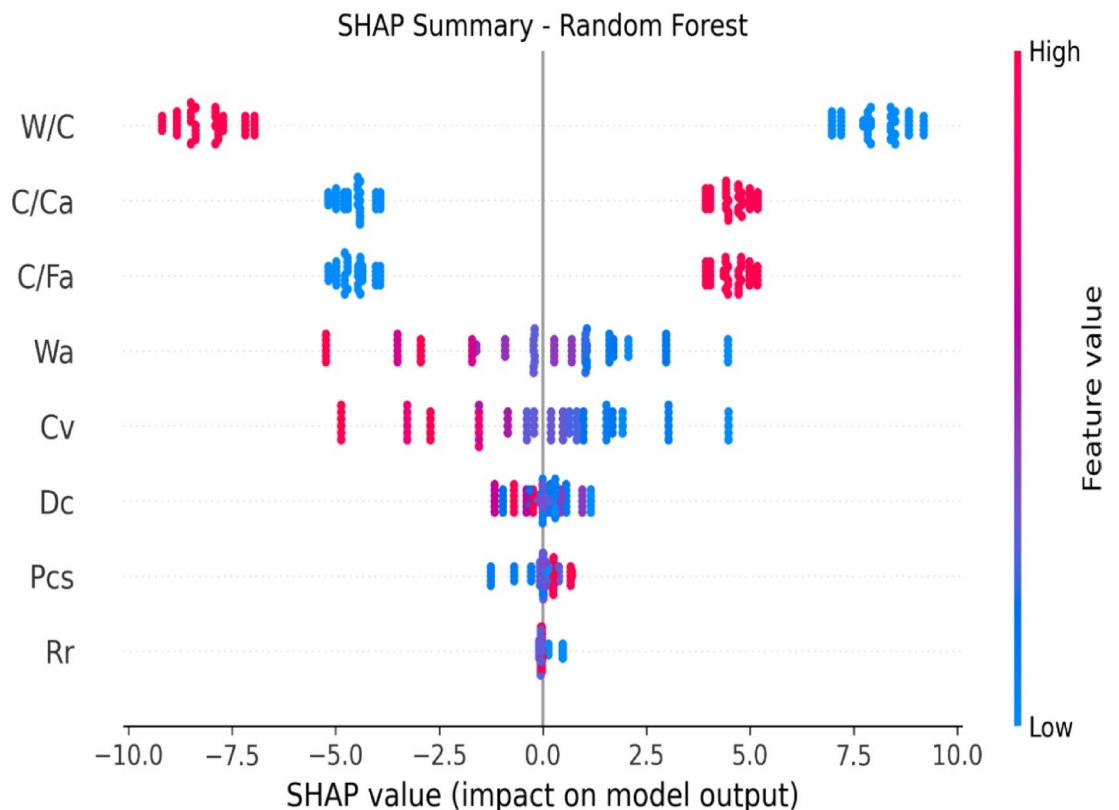


Fig. 9. Random forest's SHAP analysis report of the individual predictors.

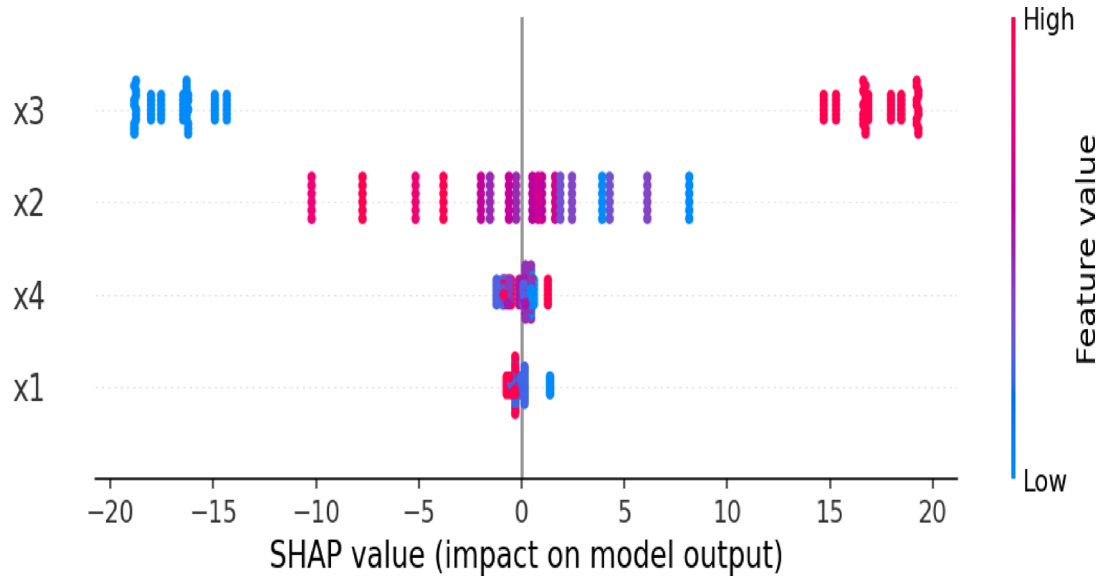


Fig. 10. Random forest’s SHAP analysis report of the aggregated predictors.

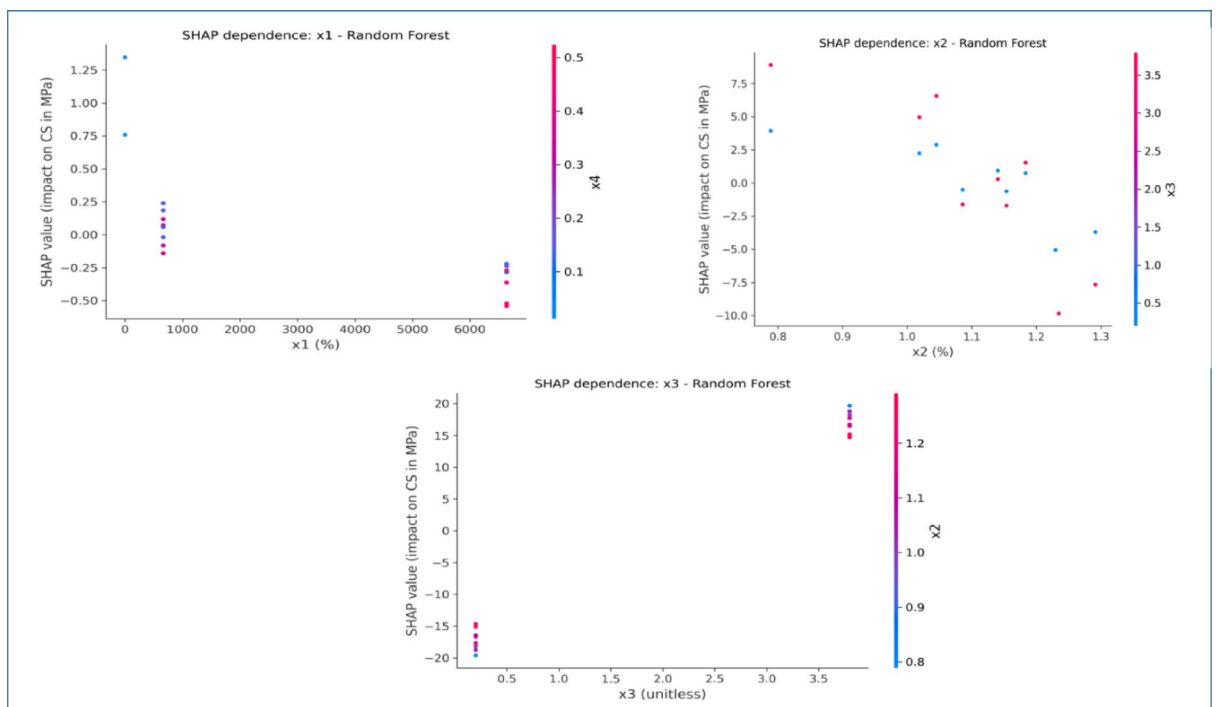


Fig. 11. Sample SHAP dependence plots between aggregated variables.

impact, with SHAP values clustered near zero. In summary, proportioning index elements were the primary drivers of model predictions, followed by carbonation, while recycled aggregate characteristics play a secondary role. This insight can guide optimized concrete mix design for enhanced performance. A study by Singh et al.⁵⁵ on predicting the flexural strength of recycled aggregate based concrete using an integrated approach combining machine learning models and global sensitivity analysis demonstrated that ML techniques can effectively estimate key mechanical properties. Based on the findings of their work, ML models can also be used to predict the compressive strength of RAC, and the sensitivity of each predictor can be evaluated in a manner similar to the analysis presented in this study .

Additionally, SHAP dependence plots were used to examine the marginal effects of the key input variables on the random forest predictions of compressive strength. Figure 11 shows the dependence plots between variables x_1 , x_2 and x_3 . The variable x_1 showed SHAP values clustered near zero across its entire range, indicating that it

contributed minimally to the model's predictions. In contrast, x_2 exhibited both positive and negative SHAP values (approximately -10 to $+9$ MPa), demonstrating a meaningful and nonlinear influence on compressive strength. Higher values of x_2 generally corresponded to positive SHAP contributions, while lower values were associated with reduced predicted strength. Among the evaluated predictors, x_3 displayed the strongest and most consistent effect, with SHAP values ranging from approximately -20 MPa at low x_3 levels to $+20$ MPa at higher levels. This indicates that increasing x_3 substantially increases the predicted compressive strength.

Table 7 shows an example of converting the individual predictors to composite indices then to predictions. It is clearly shown that the optimized indices predicted the outputs well.

The ML models developed in this study were trained and validated using datasets that fall within the parameter ranges summarized in Table 8. These ranges define the domain in which the models can make predictions. Inputs that fall outside these intervals may lead to extrapolation, which may reduce prediction accuracy.

Table 9 show the global SHAP values of each model. SHAP values quantify the contribution of each feature to the model's prediction and are expressed in the same units as the model output. In this study, the model predicts compressive strength (MPa), so a SHAP value represents how many Mpa a feature increases or decreases the predicted strength. For example, a mean absolute SHAP value of approximately 17 for feature x_3 (see Table 9) indicates that this variable changes the prediction by about 17 MPa on average, whereas a value around 3 for x_2 reflects a smaller influence of roughly 3 MPa. Importantly, Polynomial regression models can exhibit very large SHAP magnitudes due to the presence of squared and interaction terms, which increase the internal numerical scale of the model. These inflated SHAP values do not imply disproportionately higher importance; instead, they reflect the model's expanded feature space and sensitivity to unscaled polynomial terms. Therefore, comparisons of absolute SHAP magnitudes should be made within the same model, while cross model comparisons should rely primarily on relative rankings rather than numerical size.

Conclusion and recommendation

This study demonstrated that carbonation treatment of RCAs can be effectively integrated with ML to support sustainable construction practices. The investigation showed that, unlike conventional testing methods that are expensive and time consuming, regression based ML models provide an efficient and accurate alternative. For the first time, predictive modeling was applied specifically to concretes containing carbonated RCAs, filling a gap in existing research. Using 108 experimental tests of individual datasets that are later composited to indices representing mix proportioning, aggregate performance, carbonation degree, and replacement percentage, six ML algorithms were evaluated for their predictive ability. Based on the outputs of the study the authors concluded and recommended the following:

- ML models predicted the compressive strength of carbonated RAC with good accuracy.
- Intercorrelation and multicollinearity within the datasets have no noticeable effect on the performance of the ML models, particularly tree based and ensemble based models.
- Comparing classical regression based and tree based models on the same dataset, it was found that tree based models performed better.
- Numerical prediction equations developed using polynomial MLR performed well after the ensemble based models and may be applied in practical compressive strength predictions.
- Although the decision tree and LightGBM models achieved marginally lower prediction errors, the random forest model was identified as the most reliable overall due to its ability to capture physically meaningful relationships, as evidenced by SHAP analysis, together with its high predictive accuracy and robustness.
- Variations in fine aggregate type and properties, variations in cement type, addition of super plasticizers and coarse aggregate size effect were not considered in this study. Therefore, future studies can work on compressive strength prediction of carbonated RCA concrete considering these properties together with deep learning models and sufficient datasets.
- Modeling of Tensile strength, modulus of elasticity and other durability properties of carbonated RCA concrete can be another area for future research.
- The datasets used in this study are limited and cover extreme strength ranges, low to normal strength (17 to 29 MPa) and normal to high strength (40 to 72 MPa). This limitation may have contributed to the relatively high R^2 values produced by the models, which should therefore be interpreted with caution. Although the dataset includes 18 mix combinations, the limited variation in water to cement ratios, with only two values across all mixtures, likely reduced the overall variability, simplifying the prediction task and allowing the models to achieve high goodness of fit scores. Even when additional parameters such as carbonation degree, replacement ratio, and parent concrete strength are considered, the restricted range of water to cement ratios may have constrained the dataset's diversity, reflecting the narrow experimental conditions. Including additional strength values both within and beyond the current ranges, as well as a wider range of water to cement ratios, could improve model accuracy, reduce overfitting, and enhance robustness by capturing a broader spectrum of concrete mix behavior.

Individual predictors										Actual CS (Mpa)	Composite predictors					Predicted CS (MPa)						
C/C _a	C/F _a	W/C	W/C _a	W _a (%)	C _v (%)	R _r (%)	D _c (%)	P _σ (Mpa)	P _σ (Mpa)	Actual CS (Mpa)	X ₁	X ₂	X ₃	X ₄	MLR	Ridge	Poly MLR	DT	RF	LightGBM		
0.51	1.167	0.26	2.76	20.09	30	1.2	20.1	60.1	60.1	60.1	585.042	1.307	2.636	0.525	61.350	61.256	62.283	61.335	61.297	61.403		
0.261	0.335	0.57	3.35	21.82	30	2.227	7.4	23.28	23.28	23.28	585.042	3.241	0.088	0.406	22.307	22.589	20.098	21.478	21.517	21.484		
0.261	0.335	0.57	2.25	18.011	1	0.1	0.1	28.19	28.19	28.19	1.000	0.034	0.088	0.226	28.981	28.439	24.835	24.375	24.246	24.370		
0.51	1.167	0.26	3.94	24.95	100	4	67	52.13	52.13	52.13	585.042	7.659	2.636	0.549	56.425	56.064	55.371	52.220	52.976	52.373		
0.261	0.335	0.57	2.76	20.09	30	1.2	20.1	22.74	22.74	22.74	585.042	1.307	0.088	0.525	27.097	27.027	22.376	23.134	23.150	23.098		
0.261	0.335	0.57	3.94	24.95	100	4	67	23.07	23.07	23.07	585.042	7.659	0.088	0.549	22.171	21.835	19.582	21.158	21.196	21.109		
0.261	0.335	0.57	3.03	21.18	30	3.579	7.4	25.17	25.17	25.17	585.042	6.505	0.088	0.422	21.103	21.551	24.661	24.278	24.276	24.288		
0.51	1.167	0.26	3.03	21.18	30	3.579	7.4	55.39	55.39	55.39	585.042	6.505	2.636	0.422	55.356	55.780	56.669	57.148	57.100	57.136		
0.261	0.335	0.57	3.85	22.01	100	5.12	67	23.42	23.42	23.42	585.042	11.006	0.088	0.573	22.098	21.826	23.510	22.580	22.571	22.612		
0.261	0.335	0.57	3.35	21.82	30	2.227	7.4	22.16	22.16	22.16	585.042	3.241	0.088	0.406	22.307	22.589	20.098	21.478	21.517	21.484		
0.51	1.167	0.26	3.94	24.95	100	4	67	51.91	51.91	51.91	585.042	7.659	2.636	0.549	56.425	56.064	55.371	52.220	52.976	52.373		
0.51	1.167	0.26	3.94	24.95	100	4	67	52.48	52.48	52.48	585.042	7.659	2.636	0.549	56.425	56.064	55.371	52.220	52.976	52.373		
0.261	0.335	0.57	2.25	18.011	1	0.1	0.1	27.88	27.88	27.88	1.000	0.034	0.088	0.226	28.981	28.439	24.835	24.375	24.246	24.370		
0.51	1.167	0.26	3.94	24.95	100	4	67	52.69	52.69	52.69	585.042	7.659	2.636	0.549	56.425	56.064	55.371	52.220	52.976	52.373		
0.51	1.167	0.26	3.35	21.82	30	2.227	7.4	53.03	53.03	53.03	585.042	3.241	2.636	0.406	56.560	56.818	54.025	52.922	52.986	52.916		
0.261	0.335	0.57	3.94	24.95	100	4	67	21.02	21.02	21.02	585.042	7.659	0.088	0.549	22.171	21.835	19.582	21.158	21.196	21.109		
0.51	1.167	0.26	2.25	18.01	1	0.1	0.1	66.86	66.86	66.86	1.000	0.034	2.636	0.226	63.234	62.668	69.046	69.414	69.399	69.413		
0.51	1.167	0.26	2.76	20.09	30	1.2	20.1	62.29	62.29	62.29	585.042	1.307	2.636	0.525	61.350	61.256	62.283	61.335	61.297	61.403		
0.51	1.167	0.26	3.85	22.01	100	5.12	67	59.51	59.51	59.51	585.042	11.006	2.636	0.573	56.351	56.056	58.890	59.820	59.778	59.778		
0.51	1.167	0.26	3.03	21.18	30	3.579	7.4	58.46	58.46	58.46	585.042	6.505	2.636	0.422	55.356	55.780	56.669	57.148	57.100	57.136		
0.261	0.335	0.57	5.91	30.71	100	7.426	25	20.7	20.7	20.7	585.042	19.001	0.088	0.386	15.507	15.625	18.346	18.346	18.360	18.332		
0.261	0.335	0.57	4.84	28.59	100	11.93	25	20.66	20.66	20.66	585.042	38.117	0.088	0.418	14.410	14.704	20.917	20.546	20.505	20.550		

Table 7. Example that converts Raw inputs to an index and then to a prediction.

C/C _a	C/F _a	W/C	W _a (%)	C _v (%)	R _r (%)	D _c (%)	P _{cs} (Mpa)
0.261–0.51	0.335–1.167	0.26–0.58	2.25–5.91	18–31	1–100	0.1–12	25–67

Table 8. The ML models prediction ranges.

Models	Feature	SHAP global mean
MLR	x ₁	1.446
MLR	x ₂	2.920
MLR	x ₃	17.127
MLR	x ₄	2.002
Polynomial MLR (deg=2)	x ₁	4264.070
Polynomial MLR (deg=2)	x ₂	1330.491
Polynomial MLR (deg=2)	x ₃	2.238
Polynomial MLR (deg=2)	x ₄	742.073
Polynomial MLR (deg=2)	x ₁ ²	3101.106
Polynomial MLR (deg=2)	x ₁ x ₂	1096.625
Polynomial MLR (deg=2)	x ₁ x ₃	0.147
Polynomial MLR (deg=2)	x ₁ x ₄	102.858
Polynomial MLR (deg=2)	x ₂ ²	1232.142
Polynomial MLR (deg=2)	x ₂ x ₃	22.456
Polynomial MLR (deg=2)	x ₂ x ₄	771.937
Polynomial MLR (deg=2)	x ₃ ²	35.755
Polynomial MLR (deg=2)	x ₃ x ₄	1.984
Polynomial MLR (deg=2)	x ₄ ²	14.255
Ridge	x ₁	1.766
Ridge	x ₂	2.567
Ridge	x ₃	17.115
Ridge	x ₄	1.912
Decision tree	x ₁	0.187
Decision tree	x ₂	4.003
Decision tree	x ₃	17.064
Decision tree	x ₄	0.697
Random forest	x ₁	0.303
Random forest	x ₂	3.838
Random forest	x ₃	17.079
Random forest	x ₄	0.592
LightGBM	x ₁	0.775
LightGBM	x ₂	3.532
LightGBM	x ₃	17.417
LightGBM	x ₄	1.138

Table 9. Mean global SHAP of models.

Data availability

The data used and analyzed in this study are available from the corresponding author upon reasonable request.

Received: 3 October 2025; Accepted: 9 January 2026

Published online: 20 January 2026

References

- Li, S., Wang, B., Gao, Y. & Yu, Y. A compressive strength prediction model for precast recycled aggregate concrete based on machine learning. *AIP Adv.* **15**, 075133. <https://doi.org/10.1063/5.0280234> (2025).
- Sinkhonde, D., Bezabih, T., Mirindi, D., Mashava, D. & Mirindi, F. Ensemble machine learning algorithms for efficient prediction of compressive strength of concrete containing tyre rubber and brick powder. *Clean. Waste Syst.* **10** (2025), 100236. <https://doi.org/10.1016/j.clwas.2025.100236> (2025).
- Lin, L., Xu, J., Ying, W., Yu, Y. & Zhou, L. Post fire compressive mechanical behaviors of concrete incorporating coarse and fine recycled aggregates. *Constr. Build. Mater.* **461**, 139948. <https://doi.org/10.1016/j.conbuildmat.2025.139948> (2025).

4. Yu, Y., Xu, J., Su, J., Xu, L. & Luo, Y. Investigating specimen size and shape effects on compressive mechanical behaviors of recycled aggregate concrete using discrete element mesoscale Modeling. *Constr. Build. Mater.* **438**, 137196. <https://doi.org/10.1016/j.conbuildmat.2024.137196> (2024).
5. Wang, D. et al. Mechanical performance of recycled aggregate concrete in green civil engineering: review. *Case Stud. Constr. Mater.* **19**, e02384. <https://doi.org/10.1016/j.cscm.2023.e02384> (2023).
6. Arora, S., Singh, B. & Bhardwaj, B. Strength performance of recycled aggregate concretes containing mineral admixtures and their performance prediction through various modeling techniques. *J. Build. Eng.* **24**, 100741. <https://doi.org/10.1016/j.jobbe.2019.100741> (2019).
7. Deng, F. et al. Compressive strength prediction of recycled concrete based on deep learning. *Constr. Build. Mater.* **175**, 562–569. <https://doi.org/10.1016/j.conbuildmat.2018.04.169> (2018).
8. Duan, J., Asteris, P. G., Nguyen, H., Bui, X. N. & Moayedi, H. A novel artificial intelligence technique to predict compressive strength of recycled aggregate concrete using ICA-XGBoost model. *Eng. Comput.* **37**, 3329–3346. <https://doi.org/10.1007/s00366-020-01003-0> (2021).
9. Zhang, J., Huang, Y., Aslani, F., Ma, G. & Nener, B. A hybrid intelligent system for designing optimal proportions of recycled aggregate concrete. *J. Clean. Prod.* **273**, 122922. <https://doi.org/10.1016/j.jclepro.2020.122922> (2020).
10. Peiris, D. et al. Impact of treatment methods on recycled concrete aggregate performance: a comprehensive review. *Environ. Sci. Pollut. Res.* **32**, 14405–14438. <https://doi.org/10.1007/s11356-025-36497-y> (2025).
11. Fortunato, L. R., Parsekian, G. A. & Junior, A. N. Developing an easy to build laboratory chamber for CO₂ experiments. *Matéria (Rio J)* <https://doi.org/10.1590/1517-7076-RMAT-2023-0078> (2023).
12. Bui, H., Delattre, F. & Levacher, D. Experimental methods to evaluate the carbonation degree in Concrete-State of the Art review. *Appl. Sci.* **13** (4), 2533. <https://doi.org/10.3390/app13042533> (2023).
13. Liu, K. et al. Carbonation of recycled aggregate and its effect on properties of recycled aggregate concrete: A review. *Mater. Express.* **11** (9), 1439–145214. <https://doi.org/10.1166/mex.2021.2045> (2021).
14. Kashem, A. et al. Hybrid data-driven approaches to predicting the compressive strength of ultra-high-performance concrete using SHAP and PDP analyses, case stud. *Constr. Mater.* **20**, e02991. <https://doi.org/10.1016/j.cscm.2024.e02991> (2024).
15. Alyami, M. et al. Estimating compressive strength of concrete containing rice husk Ash using interpretable machine learning-based models. *Case Stud. Constr. Mater.* **20**, e02901. <https://doi.org/10.1016/j.cscm.2024.e02901> (2024).
16. Phillip, S. & Nidhi, M. Performance comparison of artificial neural network and random forest models for predicting the compressive strength of fibre – reinforced GGBS – based geopolymer concrete composites. *Mater. Circ. Econ.* **6** (34), 1–18. <https://doi.org/10.1007/s42824-024-00128-7> (2024).
17. Tipu, R. K., Arora, R. & Kumar, K. Machine learning-based prediction of concrete strength properties with coconut shell as partial aggregate replacement: A sustainable approach in construction engineering. *Asian J. Civ. Eng.* **25** (3), 2979–2992. <https://doi.org/10.1007/s42107-023-00957-y> (2024).
18. Tipu, R. K., Batra, V., Suman, Panchal, V. R., Pandya, K. S. & Patel, G. A. Optimizing compressive strength in sustainable concrete: a machine learning approach with iron waste integration. *Asian J. Civ. Eng.* **25**, 4487–4512. <https://doi.org/10.1007/s42107-024-01061-5> (2024b).
19. Lamba, P. et al. Repurposing plastic waste: experimental study and predictive analysis using machine learning in bricks. *J. Mol. Struct.* **1317**, 139158. <https://doi.org/10.1016/j.molstruc.2024.139158> (2024).
20. Rusu, A., Bărbuță, M. & Sabina, S. Multiple linear regression model to predict compressive strength of concrete with silica fume and metallic fibers. In: Moldovan, L., Gligor, A. (eds) *The 17th International Conference Interdisciplinarity in Engineering. Inter-ENG 2023. Lecture Notes in Networks and Systems*, 926. Springer, Cham. (2024). https://doi.org/10.1007/978-3-031-54664-8_23
21. Chore, H. & Shelke, N. L. Prediction of compressive strength of concrete using multiple regression model. *Struct. Eng. Mech.* <https://doi.org/10.12989/sem.2013.45.6.837> (2013).
22. Aamir, H., Aamir, K. & Javed, M. F. Linear and Non-Linear regression analysis on the prediction of compressive strength of sodium hydroxide Pre-Treated crumb rubber concrete. *Eng. Proc.* **44** (1), 5. <https://doi.org/10.3390/engproc20230404005> (2023).
23. Yuan, Y., Yang, M., Shang, X., Xiong, Y. & Zhang, Y. Predicting the compressive strength of UHPC with coarse aggregates in the context of machine learning. *Case Stud. Constr. Mater.* **19**, e02627. <https://doi.org/10.1016/j.cscm.2023.e02627> (2023).
24. Abdellatif, M., Murali, G. & Dixit, S. Leveraging machine learning to evaluate the effect of Raw materials on the compressive strength of ultra-high-performance concrete. *Results Eng.* **25**, 104542. <https://doi.org/10.1016/j.rineng.2025.104542> (2025).
25. Abuodeh, O. R., Abdalla, J. A. & Hawileh, R. A. Assessment of compressive strength of Ultra-high performance concrete using deep machine learning techniques. *Appl. Comput.* **95**, 106552. <https://doi.org/10.1016/j.asoc.2020.106552> (2020).
26. Salman, K. M. et al. Explainable AutoML models for predicting the strength of high-performance concrete using Optuna, SHAP and ensemble learning. *Front. Mater.* <https://doi.org/10.3389/fmats.2025.1542655> (2025).
27. Gebremariam, H. G., Taye, S. & Tarekegn, A. G. CO₂ Uptake in recycled concrete aggregates: influence of parent concrete strength and preprocessing carbonation. unpublished manuscript. (2025).
28. Alahakoon, Y. et al. Prediction of alkali–silica reaction expansion of concrete using explainable machine learning methods. *Discov. Appl. Sci.* **7**, 407. <https://doi.org/10.1007/s42452-025-06880-y> (2025).
29. Meddage, D. P. P., Mohotti, D., Wijesooriya, K., Lee, C. K. & Kwok, K. C. S. Interpolating wind pressure time-histories around a tall building - a deep learning-based approach. *J. Wind Eng. Ind. Aerodyn.* **256**, 105968. <https://doi.org/10.1016/j.jweia.2024.105968> (2025).
30. O'brien, R. M. A caution regarding rules of thumb for variance inflation factors. *Qual. Quant.* **41**, 673–690. <https://doi.org/10.1007/s11335-006-9018-6> (2007).
31. Singh, S. G. & Kumar, S. V. Dealing with multicollinearity problem in analysis of side friction characteristics under urban heterogeneous traffic conditions. *Arab. J. Sci. Eng.* **46**, 10739–10755. <https://doi.org/10.1007/s13369-020-05213-y> (2021).
32. Loureiro, A. A. B. & Stefani, R. Comparing the performance of machine learning models for predicting the compressive strength of concrete. *Discov. Civ. Eng.* **1**, 19. <https://doi.org/10.1007/s44290-024-00022-w> (2024).
33. Mondal, A. & Bhanja, S. Augmentation of Abrams law for fly Ash concrete. *Mater. Today Proc.* **65** (2), 644–650. <https://doi.org/10.1016/j.matpr.2022.03.204> (2022).
34. Sah, A. K. & Hong, Y. M. Performance comparison of machine learning models for concrete compressive strength prediction. *Mater. (Basel)*. **7** (9), 2075. <https://doi.org/10.3390/ma17092075> (2024).
35. Gayathri, R., Rani, S. U., Čepová, L., Rajesh, M. & Kalita, K. A comparative analysis of machine learning models in prediction of mortar compressive strength. *Processes* **10** (7), 1387. <https://doi.org/10.3390/pr10071387> (2022).
36. Biau, G. & Scornet, E. A random forest guided tour. *TEST* **25**, 197–227. <https://doi.org/10.1007/s11749-016-0481-7> (2016).
37. Padula, W. V. Introduction to decision tree modeling, in Bishai, D., Brenzel L., Padula, W. (eds), *Handbook of Applied Health Economics in Vaccines, Handbooks in Health Economic Evaluation* (Oxford Academic), accessed 16 Aug. 2025. (2023). <https://doi.org/10.1093/oso/9780192896087.003.0020>
38. James, G., Witten, D., Hastie, T. & Tibshirani, R. *An Introduction To Statistical Learning with Applications in R. Part of the Book Series: Springer Texts in Statistics* (Springer, 2021). <https://doi.org/10.1007/978-1-0716-1418-1>
39. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32. <https://doi.org/10.1023/A:1010933404324> (2001).
40. Elshaarawy, M. K., Zelenáková, M. & Armanuos, A. M. Hydraulic performance modeling of inclined double cutoff walls beneath hydraulic structures using optimized ensemble machine learning. *Sci. Rep.* **15**, 27592. <https://doi.org/10.1038/s41598-025-10990-3> (2025).

41. Hajihosseini, M., Maghsoudi, A. & Ghezlbash, R. A novel scheme for mapping of MVT-Type Pb-Zn prospectivity: LightGBM, a highly efficient gradient boosting decision tree machine learning algorithm. *Nat. Resour. Res.* **32**, 2417–2438. <https://doi.org/10.1007/s11053-023-10249-6> (2023).
42. Elshaarawy, M. K., Alsaadawi, M. M. & Hamed, A. K. Machine learning and interactive GUI for concrete compressive strength prediction. *Sci. Rep.* **14**(1), 16694 (2024).
43. Harith, I. K. et al. Prediction of high-performance concrete strength using machine learning with hierarchical regression. *Multiscale Multidiscip Model. Exp. Des.* **7**, 4911–4922. <https://doi.org/10.1007/s41939-024-00467-7> (2024).
44. Das, P. & Kashem, A. Hybrid machine learning approach to prediction of the compressive and flexural strengths of UHPC and parametric analysis with Shapley additive explanations. *Case stud. Constr. Mater.* **20**, e02723. <https://doi.org/10.1016/j.cscm.2023.e02723> (2024).
45. Elhishi, S., Elashry, A. M. & El-Metwally, S. Unboxing machine learning models for concrete strength prediction using XAI. *Sci. Rep.* **13**, 19892. <https://doi.org/10.1038/s41598-023-47169-7> (2023).
46. Arunvivek, G. K., Anandaraj, S., Kumar, P., Pratap, B. & Sembeta, R. Y. Compressive strength modelling of cenosphere and copper slag-based geopolymer concrete using deep learning model. *Sci. Rep.* **15**, 27849 (2025).
47. Albostami, A. S., Al-Hamd, R. K. S. & Al-Matwari, A. A. Data-Driven predictive modeling of steel slag concrete strength for sustainable construction. *Buildings* **14** (8), 2476. <https://doi.org/10.3390/buildings14082476> (2024).
48. Pratap, B., Kumar, P., Shubham, K. & Chaudhary, N. Soft computing-based investigation of mechanical properties of concrete using ready-mix concrete waste water as partial replacement of mixing portable water. *Asian J. Civil Eng.* **25** (2), 1255–1266. <https://doi.org/10.1007/s42107-023-00841-9> (2024).
49. Kumar, P., Pratap, B., Sharma, S. & Kumar, I. Compressive strength prediction of fly Ash and blast furnace slag-based geopolymer concrete using convolutional neural network. *Asian J. Civil Eng.* **25** (2), 1561–1569. <https://doi.org/10.1007/s42107-023-00861-5> (2024).
50. Gogineni, A., Panday, I. K., Kumar, P. & Paswan, R. K. Predicting compressive strength of concrete with fly Ash and admixture using xgboost: a comparative study of machine learning algorithms. *Asian J. Civil Eng.* **25** (1), 685–698. <https://doi.org/10.1007/s42107-023-00804-0> (2024).
51. Anwar, M. K., Qurashi, M. A., Zhu, X., Shah, R., Siddiq, S. A. & M.U A comparative performance analysis of machine learning models for compressive strength prediction in fly ash-based geopolymers concrete using reference data, case stud. *Constr. Mater.* **22**, e04207. <https://doi.org/10.1016/j.cscm.2025.e04207> (2025).
52. Dhengare, S., Waghe, U., Yenurkar, G. & Shyamala, A. A comprehensive model for concrete strength prediction using advanced learning techniques. *Discov Appl. Sci.* **7**, 551. <https://doi.org/10.1007/s42452-025-07095-x> (2025).
53. Cho, S. J., De Boeck, P., Naveiras, M. & Ervin, H. Level-specific residuals and diagnostic measures, plots, and tests for random effects selection in multilevel and mixed models. *Behav. Res.* **54**, 2178–2220. <https://doi.org/10.3758/s13428-021-01709-z> (2022).
54. Haque, M. A., Chen, B., Kashem, A., Qureshi, T. & Ahmed, A. A. M. hybrid intelligence models for compressive strength prediction of MPC composites and parametric analysis with SHAP algorithm. *Mater. Today Commun.* **35**, 105547 (2023).
55. Singh, R., Tipu, R. K., Mir, A. A. & Patel, M. Predictive modelling of flexural strength in recycled Aggregate-Based concrete: A comprehensive approach with machine learning and global sensitivity analysis. *Iran. J. Sci. Technol. Trans. Civ. Eng.* **49**, 1089–1114. <https://doi.org/10.1007/s40996-024-01502-w> (2025).

Acknowledgements

College of technology and built environment (CTBE), Addis Ababa University (AAU) is honored for infrastructural support during the study.

Author contributions

Gebremariam H.G. conceptualization, data analysis, wrote the initial draft; Taye S. methodology development, visualization, and supervision, Tarekn A.G. conceptualization, revision and supervision. All authors reviewed and approved the final manuscript and contributed to discussion.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.G.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026