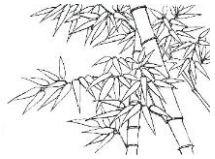
Sustainable Engineering Materials
ISSN: 3105-1960 (Print); ISSN: 3105-1979 (Online)
http://www.sustain-dpl.com/picnews.asp?id=177
DOI: 10.54113/j.suem.2026.000017



ORIGINAL ARTICLE

Predicting compressive strength of fly ash blended sandcrete using machine learning models

Navaratnarajah Sathiparan*

Department of Civil Engineering, Faculty of Engineering, University of Jaffna, Jaffna 40000, SriLanka
*Corresponding Author: Navaratnarajah Sathiparan. Email: sakthi@eng.jfn.ac.lk

Abstract: Cement-based materials, particularly sandcrete, play a crucial role in the construction industry, where demand for sustainable and high-performance materials is increasing. Fly ash, a byproduct of coal combustion, has gained attention as a supplementary cementitious material (SCM) to improve the sustainability of these materials. However, predicting the compressive strength of fly ash-blended mortars, which is essential for ensuring the structural integrity and durability of construction materials, remains a challenge. This study aims to address this gap by leveraging advanced machine learning techniques to predict the compressive strength of fly ash blended sandcrete, using key input variables such as aggregate-to-binder ratio (Agg/B), fly ash-to-binder ratio (FA/B), water-to-binder ratio (W/B), and curing time. While previous research has focused on conventional cement mortars, few studies have integrated these key variables using machine learning models. The goal of this research is to develop a reliable and accurate predictive model for compressive strength, filling a gap in existing literature regarding the influence of multiple input variables on mortar performance. The study employs four machine learning models - Artificial Neural Networks (ANN), K-Nearest Neighbors (KNN), Support Vector Regression (SVR), and Extreme Gradient Boosting (XGB) - to evaluate their performance in predicting compressive strength. The results show that XGB outperforms other models, achieving an R^2 of 0.84 for training data and 0.74 for testing data, along with the lowest prediction errors. These findings demonstrate the potential of machine learning models, particularly XGB, in optimizing mix designs and improving sustainability in construction, offering valuable insights for future material innovations.

Keywords: fly ash, sandcrete, compressive strength, machine learning, extreme gradient boosting

1 Introduction

Sandcrete is a mixture of cement, sand, and water, commonly used in construction for bonding building materials, such as bricks, stones, and tiles. The ratio of cement to sand in the mixture can vary depending on the strength requirements of the application. Mortar serves as a strong adhesive, ensuring the structural integrity of walls, pavements, and other masonry elements [1]. Its applications include bricklaying, plastering, tiling, repair works, and decorative features in landscaping. This mixture is essential for ensuring stability and durability in various construction projects. The need for more cement in sandcrete primarily arises from the requirements for higher strength, durability, and workability. A higher proportion of cement helps achieve greater strength, enabling the mortar to withstand greater

Received: 23 December 2025; Received in revised form: 25 March 2026; Accepted: 8 May 2026

loads and stresses. It also ensures improved durability, protecting the structure from environmental wear, water penetration, and other external factors [2]. Additionally, using more cement provides better workability during application, making it easier to achieve the desired finish. However, this increased cement content raises concerns due to the environmental impacts associated with cement production [3]. Cement production significantly impacts the environment, being a major source of global CO_2 emissions due to its energy-intensive processes and reliance on fossil fuels [4, 5]. The conversion of limestone to lime in kilns releases substantial CO_2, and cement manufacturing depletes natural resources and generates waste. To address these issues, reducing cement use in construction is essential. Supplementary cementitious materials (SCMs) can partially replace cement in mortar and concrete, lowering CO_2 emissions and enhancing sustainability [6]. SCMs, byproducts of industries such as coal combustion and steel manufacturing, improve properties like workability, strength, and durability. Fly ash is particularly notable for its cost-effectiveness, abundance, and ability to recycle industrial waste. Its use not only reduces demand for cement but also improves the durability of concrete, making it a sustainable choice for the construction industry [7, 8].

Predicting compressive strength is vital for cement mortar, as it ensures structural integrity for construction uses. Fly ash-based mortars differ from traditional ones due to fly ash's lower reactivity, making strength prediction crucial for performance [9]. Machine learning models can help engineers optimize the mix of fly ash, cement, and sand to achieve the right balance of strength, workability, and sustainability. This approach allows for customized mortar mixtures that meet specific project needs, improving durability and supporting environmental goals in the construction industry [10]. The incorporation of machine learning models into the study of fly ash-based cement mortar is of great significance for several reasons. Traditional methods of mortar mix design and performance evaluation often fail to effectively model the complex relationships among these variables and their impact on the final properties. Machine learning models are particularly valuable in this context because they can process large datasets and identify intricate patterns and correlations that would be challenging to uncover through conventional analytical methods [11]. These models can predict the performance of fly ash-based cement mortar under different conditions, offering a more accurate and efficient approach to optimizing mix designs. For example, a machine learning model can assess how varying the proportions of fly ash in the mix affects key properties, such as compressive strength, workability, and setting time, while accounting for other factors, including the sand-to-cement ratio, water content, and curing time [12].

The prediction of compressive strength in fly ash incorporated construction materials has been actively studied using machine learning techniques. However, a closer look at the existing literature reveals several areas that still require attention. Many studies focus primarily on conventional concrete [13-22], with fewer applications in other construction materials such as self-compacting concrete [23-26] and pervious concrete [27, 28], especially with fly ash as an active binder. Additionally, while the studies leverage various machine learning models such as artificial neural networks (ANN), support vector regression (SVR), and random forest regression (RF), most are limited to a smaller set of input variables, often overlooking the aggregate-to-binder ratio (Agg/B), fly ash-to-binder ratio (FA/B), water-to-binder ratio (W/B), and curing time, which play significant roles in the overall strength prediction of fly ash-based mortars. Furthermore, there is a limited exploration into the application of more advanced machine learning techniques and the impact of complex interactions among these input variables, which could enhance predictive accuracy. This gap presents an opportunity to explore the synergistic effects of these parameters on mortar strength, providing a more comprehensive and refined prediction model [29].

The novelty of this study lies in its comprehensive approach to predicting the compressive strength of fly ash blended sandcrete, incorporating a wider range of input variables. By utilizing the aggregate-to-binder ratio (Agg/B), fly ash-to-binder ratio (FA/B), water-to-binder ratio (W/B), and curing time as input variables, this research acknowledges the complex interactions among these factors that have been underexplored in previous studies. Furthermore, this study innovates by applying machine learning techniques to a broader dataset, analyzing the intricate relationships between these variables, and exploring their combined effects on compressive strength. Through this approach, the research aims to provide more accurate and reliable predictive models, thereby filling the gap in current literature and offering practical insights for the construction industry. The utilization of advanced machine learning

algorithms for this purpose further adds to the novelty of the research, positioning it as a pioneering work in the field of fly ash-based construction materials.

2.Methodology

The methodology for this study, as shown in **Fig. 1**, aims to predict the compressive strength of fly ash-blended sandcrete using advanced machine learning techniques. The primary goal is to develop a reliable model that accurately forecasts compressive strength based on key input variables, including the Aggregate-to-Binder ratio (Agg/B), Fly Ash-to-Binder ratio (FA/B), Water-to-Binder ratio (W/B), and curing time. To achieve this, a structured approach is employed, incorporating phases of data collection, preprocessing, model development, and evaluation. The dataset for this study was compiled from 34 published literature sources, totaling 915 data points.

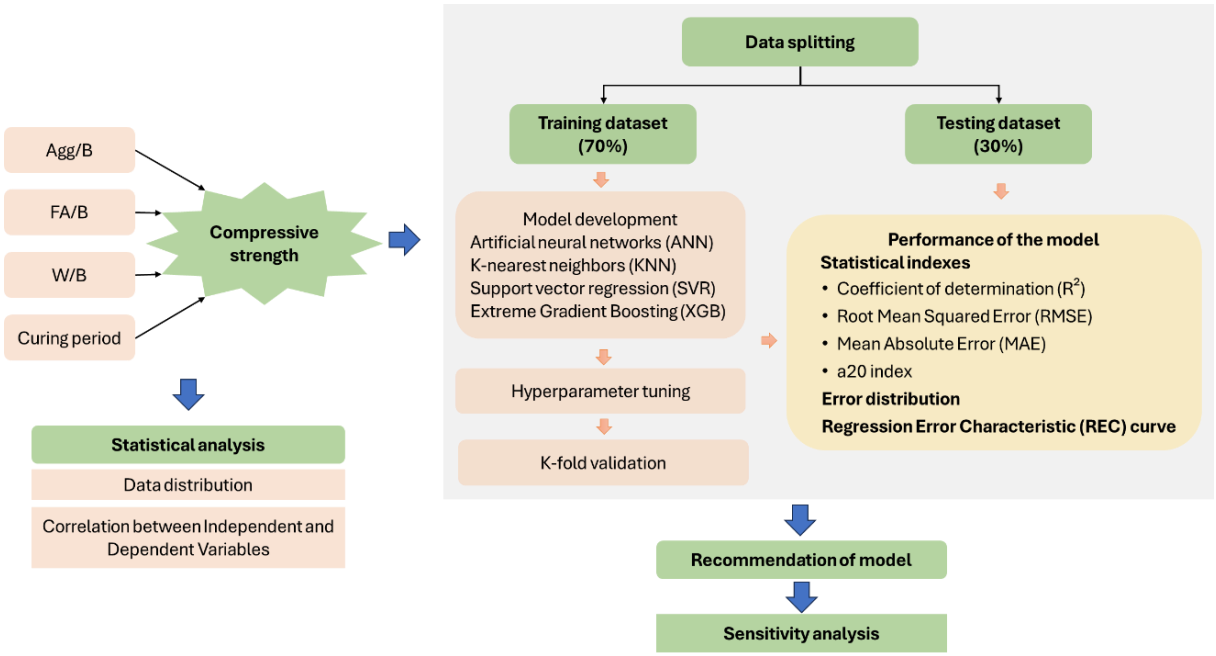


Fig. 1. Flow chart for the methodology used to predict the compressive strength of fly ash blended sandcrete

Various machine learning algorithms were employed, including Artificial Neural Networks (ANN), K-nearest Neighbors (KNN), Support Vector Regression (SVR), and Extreme Gradient Boosting (XGB), to identify the most effective approach for predicting compressive strength. The methodology incorporates best practices in model development, including data splitting, hyperparameter tuning, and model validation, ensuring that the final model provides accurate and reliable predictions. In addition to developing the predictive models, a comprehensive performance evaluation was conducted using several statistical metrics. Furthermore, a sensitivity analysis was performed to assess the impact of variations in input variables on model performance.

2.1 Data collection

For this study, a comprehensive dataset comprising 915 unique data points was compiled to develop machine learning models for predicting the compressive strength of fly ash-blended sandcrete, and is summarized in **Table 1**. The dataset was compiled from 34 peer-reviewed published sources, ensuring a diverse range of experimental conditions and material compositions. The literature review was conducted using prominent academic databases, including Scopus, Web of Science, ScienceDirect, and Google Scholar. The search was conducted using the keywords “fly ash,” “cement mortar,” and “sandcrete,” to identify experimental studies that reported compressive strength values for fly ash-cement mortar or sandcrete. Only studies that reported mix ratios and curing durations were included in the dataset. Articles with duplicated datasets, incomplete data, or inappropriate experimental designs (e.g., the use of additives outside the scope of this study) were excluded.

The dataset focused on four input variables that significantly influence the mechanical performance of cement-based composites. These included the aggregate-to-binder ratio (Agg/B), which ranged from 1.50 to 6.00, the fly ash-to-binder ratio (FA/B), varying between 0 and 0.80, the water-to-binder ratio (W/B), ranging from 0.30 to 0.80, and the curing period (days), spanning from 1 to 360 days. The target variable was the compressive strength (MPa), which ranged from 1.19 MPa to 89.96 MPa across the dataset. The inclusion of diverse studies covering various curing regimes, fly ash contents, and regional material properties helps ensure the generalizability of the predictive models developed in this research.

Table 1. Summary of the dataset collected from published literature

Ref	Agg/B	FA/B	W/B	Curing period (days)	Comp. strength (MPa)	No. of data
Alvarenga and Cordeiro [30]	3.00	0-0.30	0.48	7-120	20.7-40.1	12
Arenas-Piedrahita et al. [31]	3.50	0-0.20	0.60	3-180	24.4-46.4	35
Argiz et al. [32]	3.00	0.1-0.35	0.50	1-90	1.2-9.6	36
Atis et al. [33]	2.00	0-0.40	0.40	1-28	11.2-48	20
Baeza et al. [34]	3.00	0-0.10	0.50	28-90	33.5-41.1	4
Chandra et al. [35]	4.00-6.00	0-0.33	0.63-0.80	7-28	2.7-8.4	15
Cheerarota and Jaturapitakkul [36]	2.75	0-0.30	0.68-0.70	3-90	11.8-55.5	60
Chindaprasirt and Rukzon [37]	2.75	0-0.40	0.50	7-90	33-63.5	12
Fattouh and Abouhalawa [38]	3.00	0-0.30	0.50	3-28	19.9-42.3	12
Hatungimana et al. [39]	3.00	0-0.30	0.50	7-90	29.4-70.3	12
Hsu et al. [40]	2.75	0-0.20	0.49	3-56	19.8-55.1	48
Huang and Zhao [41]	3.00	0-0.30	0.55	3-28	6.6-30	12
Islam and Islam [42]	2.75	0-0.60	0.49	3-90	3.7-38	42
Javed et al [43]	3.00	0-0.15	0.42	14-90	15.2-38.1	6
Jozic et al. [44]	3.00	0-0.40	0.48-0.50	7-300	25.6-70.5	42
Khan et al. [45]	1.50	0-0.25	0.35	1-28	8.1-90	15
Kim et al. [46]	2.45	0-0.20	0.49	3-28	17.7-40.2	12
Madurwar et al. [47]	3.00	0-0.20	0.50	7-28	29.5-42.5	15
Mardani-Aghabaglou et al. [48]	2.75	0-0.10	0.49	7-300	29.6-58.6	10
Nadekar and Nandurkar [49]	5.00	0-0.40	0.60-0.80	3-90	2.1-20.2	24
Nandurkar and Pande [50]	3.00-5.00	0-0.40	0.50-0.80	3-90	1.8-29.6	72
Nguyen et al. [51]	3.00	0-0.30	0.50	3-28	17.8-55.1	9
Ramaswamy et al. [52]	3.00	0-0.50	0.50	3-90	6.3-30.7	36
Ratnaweera [53]	1.67-2.75	0-0.39	0.30-0.50	1-180	8.1-79.2	125
Ravichandran and Balasubramanian [54]	3.00	0-0.70	0.48	7-121	3.4-30	48
Rukzon and Chindaprasirt [55]	2.75	0-0.40	0.50	7-90	32.5-61.9	9
Supit et al. [56]	2.75-4.58	0-0.50	0.40-0.67	7-28	6.1-43	20
Wang et al. [57]	2.62-3.05	0-0.24	0.64-0.75	1-56	3.7-35.9	16
Wang et al. [58]	2.75	0-0.20	0.59-0.68	3-56	22.2-42.7	16
Wibowo and Saidani [59]	3.00	0-0.50	0.40	7-28	11.9-31.2	12
Xu et al. [60]	2.50	0-0.60	0.46	1-180	1.7-58.7	15
Yang et al. [61]	2.99-3.03	0-0.30	0.50	7-91	29.1-71.3	33
Yerramala et al. [62]	3.00	0-0.25	0.50	7-180	10.4-33	24
Yu et al. [63]	3.00	0-0.80	0.50	3-360	5-62.1	36
	1.50-6.00	0-0.80	0.30-0.80	1-360	1.19-89.96	915

2.2 Machine learning models

2.2.1 Artificial Neural Networks (ANN)

Artificial Neural Networks (ANNs) are a class of machine learning models inspired by the structure and functioning of the human brain. These models consist of layers of interconnected nodes, known as neurons, that process information in a hierarchical manner. The architecture of an ANN typically includes three types of layers: the input layer, one or more hidden layers, and the output layer. In the context of this study, the network receives four input values at the input layer, where each input node corresponds to one feature [64]. The input data is then passed through multiple hidden layers. Each

hidden layer consists of neurons (e.g., H-1, H-2, H-3, etc.), where complex transformations and non-linearities are applied to the data to capture the underlying patterns and interactions between features, as shown in **Fig. 2**. The depth of the hidden layers allows the ANN to model complex relationships within the data, improving its predictive capability. Finally, the output layer of the ANN produces the predicted compressive strength of the mortar. The connection between the layers is weighted, and the network is trained using backpropagation to minimize the error between predicted and actual compressive strength values. The training process adjusts the weights to optimize the network's performance in predicting the output. The use of ANNs in this study is particularly advantageous due to their ability to learn from data, adapt to complex patterns, and generalize well to unseen data [65]. To ensure a fair comparison across all models, the ANN architecture was optimized using a systematic grid search. We evaluated hidden layers ranging from 1 to 3 and neurons per layer from 5 to 50. The selected architecture was found to provide the best balance between training stability and generalization, preventing the overfitting observed in deeper configurations.

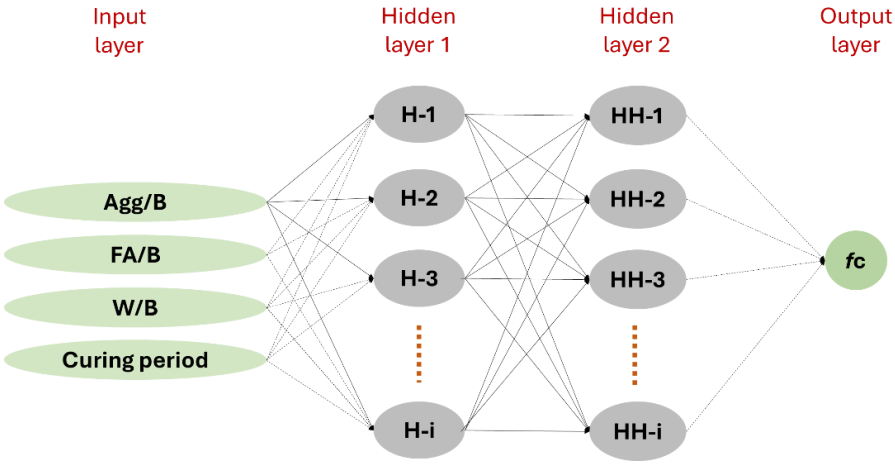


Fig. 2. Schematic diagram of an artificial neural network (ANN) architecture

2.2.2 K-Nearest Neighbors (KNN)

The K-Nearest Neighbors (KNN) algorithm is a non-parametric, instance-based learning method used for regression tasks, where the prediction of a data point is based on the average of its k nearest neighbors in the feature space. In this study, the KNN algorithm is utilized to predict the compressive strength of fly ash-blended sandcrete. KNN works by calculating the distances between the query point and all other points in the training dataset. Typically, the Euclidean distance is used as the metric for this calculation. Once the distances are computed, the algorithm identifies the k closest data points (neighbors). The prediction is then made by averaging the target values (compressive strength) of these k nearest neighbors. The value of k is a critical hyperparameter that determines the model's complexity: a smaller k can lead to overfitting, while a larger k may result in better generalization but with the potential for underfitting [66]. One of the key characteristics of KNN is that it does not require explicit training; instead, the model stores the entire training dataset and makes predictions by comparing the query point to all points in the dataset. As a result, KNN is computationally expensive during inference, particularly with large datasets, as it requires calculating the distance for each query point relative to all training samples [67].

The KNN algorithm is ideal for this study because of its ability to model complex, non-linear relationships between the input variables and the compressive strength of the mortar. The method is simple to implement and does not make assumptions about the underlying data distribution. However, selecting an optimal value for k is crucial, and it is typically done through cross-validation to ensure the best balance between bias and variance.

2.2.3 Support Vector Regression (SVR)

Support Vector Regression (SVR) is a powerful regression method based on the principles of Support Vector Machines (SVM), which are primarily designed for classification tasks. SVR works by mapping the input data into a higher-dimensional feature space using a kernel function, typically the

Radial Basis Function (RBF) kernel in this case. The goal is to find a hyperplane in this higher-dimensional space that maximizes the margin of error within a specified tolerance, often denoted as ε (epsilon) [68]. This margin allows for some deviation from the actual output (i.e., the model can tolerate minor errors within the ε-bound), but the model aims to minimize the overall error outside of this boundary, as shown in **Fig. 3**.

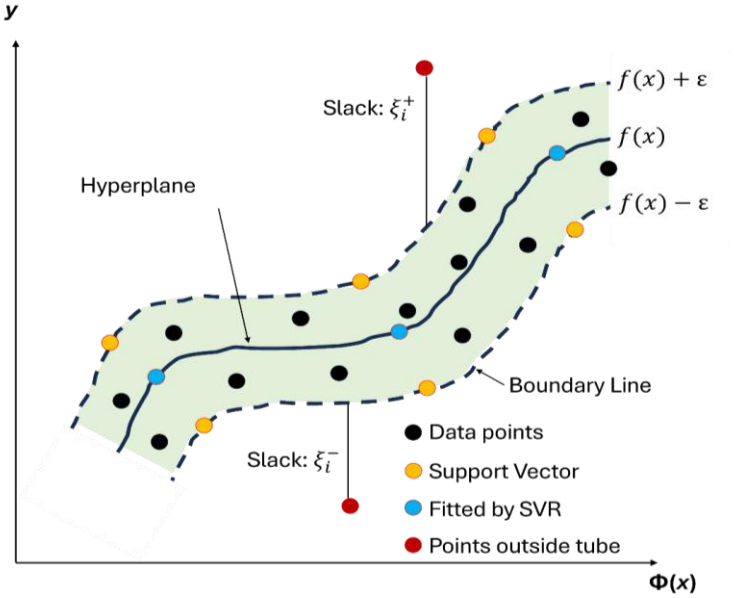


Fig. 3. Schematic diagram of support vector regression (SVR) architecture

SVR focuses on finding the best balance between fitting the data (minimizing error) and maintaining a generalizable model (avoiding overfitting). This is achieved by defining a cost function that penalizes large deviations from the actual values, while simultaneously minimizing overfitting by keeping the model's complexity as low as possible. The cost function is controlled by two hyperparameters: C (which determines the penalty for errors) and ε (which defines the margin of error within which no penalty is applied). One of the primary advantages of SVR is its ability to model non-linear relationships between input features and the target variable through the use of the kernel trick. By applying the kernel function, SVR can effectively capture complex patterns in the data without requiring explicit feature transformation into a higher-dimensional space. This makes SVR particularly well-suited for this study, as the relationship between the input variables and compressive strength may be highly non-linear.

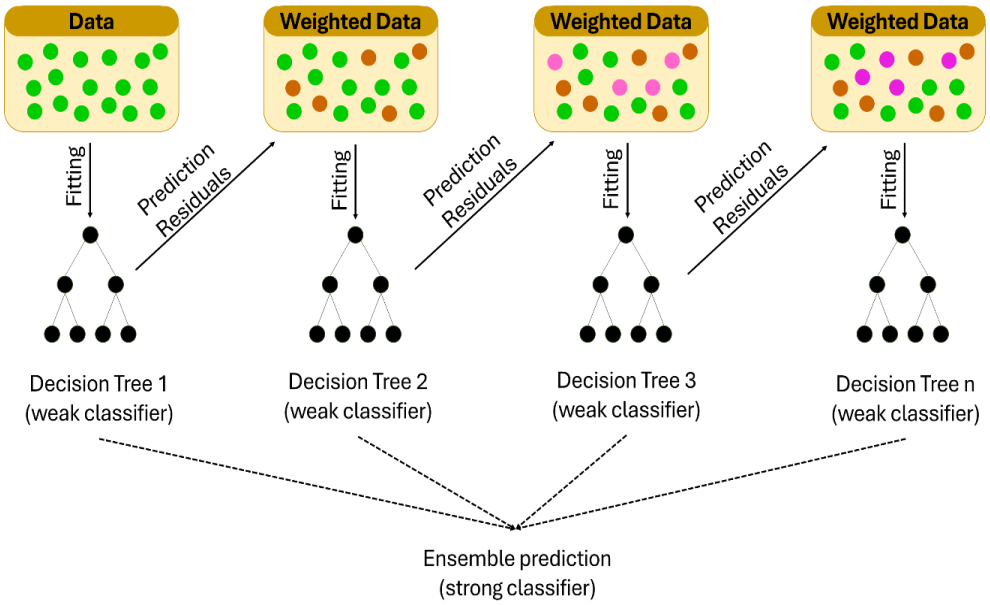


Fig. 4. Schematic diagram of extreme gradient boosting (XGB) architecture

2.2.4 Extreme Gradient Boosting (XGB)

Extreme Gradient Boosting (XGB) is a powerful machine learning technique used for regression tasks that builds an ensemble of decision trees sequentially. It works by correcting the errors of previous trees, with each new tree learning from the residuals (errors) of the existing ensemble. The model optimizes a loss function that balances prediction error and model complexity, using regularization techniques to avoid overfitting. The workflow of XGB begins by applying weights to the data points, as shown in **Fig. 4**. Initially, the data is input into the model, and the first decision tree is created. Following this, the weighted data, where misclassified or difficult-to-predict data points are given higher weights, is used to train the next tree. Each subsequent tree focuses on reducing the error from the previous trees, refining the predictions by learning from the weighted data. This iterative process continues until the ensemble of trees provides a sufficiently accurate prediction [69].

XGB's key features include its ability to capture complex, non-linear relationships between input features and target variables through the use of the gradient boosting technique. The algorithm also utilizes regularization to control the model's complexity and prevent overfitting. Hyperparameters such as the learning rate, maximum tree depth, and number of trees are tuned to achieve optimal model performance. In this study, XGB was employed to predict the compressive strength of fly ash-blended sandcrete due to its effectiveness in handling large datasets, its robustness against overfitting, and its ability to model non-linear interactions. The model's performance can be significantly improved by selecting the correct hyperparameters through cross-validation, ensuring the best fit to the data.

2.3 Performance indicators

In this study, the performance of the machine learning models was evaluated using several well-established performance indicators as follows [70]:

2.3.1 Coefficient of determination (R^2)

The R^2 score measures the proportion of the variance in the target variable (compressive strength) that is predictable from the input features. It is defined mathematically as Eq. (1).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - y_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}_i\right)^2} \tag{1}$$

Where, y_i is the actual value, $\hat{y}_i$ is the predicted value, $\bar{y}$ is the mean of the actual values, n is the number of data points. An R^2 value close to 1 indicates that the model explains most of the variability in the data, while a value close to 0 suggests that the model does not explain much of the variance.

2.3.2 Root mean squared error (RMSE)

The RMSE measures the average magnitude of the errors between predicted and observed values. It is beneficial for quantifying the difference between predicted and actual values on the same scale as the target variable. Lower RMSE values indicate better model performance. The formula is given by Eq. (2).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i\right)^2} \tag{2}$$

2.3.3 Mean absolute error (MAE)

The MAE is the average of the absolute differences between the predicted and actual values, providing a straightforward measure of error that does not consider direction. MAE is a valuable metric when the scale of errors is more important than the squared magnitude of the errors. The formula is given by Eq. (3).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - y_i\right| \tag{3}$$

2.3.4 Scatter index (SI)

The Scatter Index (SI) is used to assess the relative variability of the predictions compared to the actual values. The SI indicates the percentage of the average actual value that the RMSE represents. A smaller SI indicates better predictive accuracy. It is defined as Eq. (4).

$$SI = \frac{RMSE}{\overline{y}} \tag{4}$$

2.3.5 Fraction of data points within 20% error (a20)

The a20 metric indicates the fraction of datapoints where the predicted value falls within 20% of the actual value. A higher a20 value indicates better predictive accuracy. It is calculated as Eq. (5).

$$a_{20} = \frac{1}{n}\sum_{i=1}^{n} 1\left(\left|\frac{y_i - y_i}{y_i}\right| \le 0.20\right) \tag{5}$$

Where: a_{20} is the indicator function that returns 1 if the condition is true (i.e., the prediction is within 20% of the actual value) and 0 otherwise.

2.3.6 Prediction error distribution

The Prediction Error Distribution provides a visual representation of the distribution of errors between predicted and actual values. It is used to identify any systematic bias or skewness in the model's predictions, helping to evaluate whether the model tends to overestimate or underestimate the target variable.

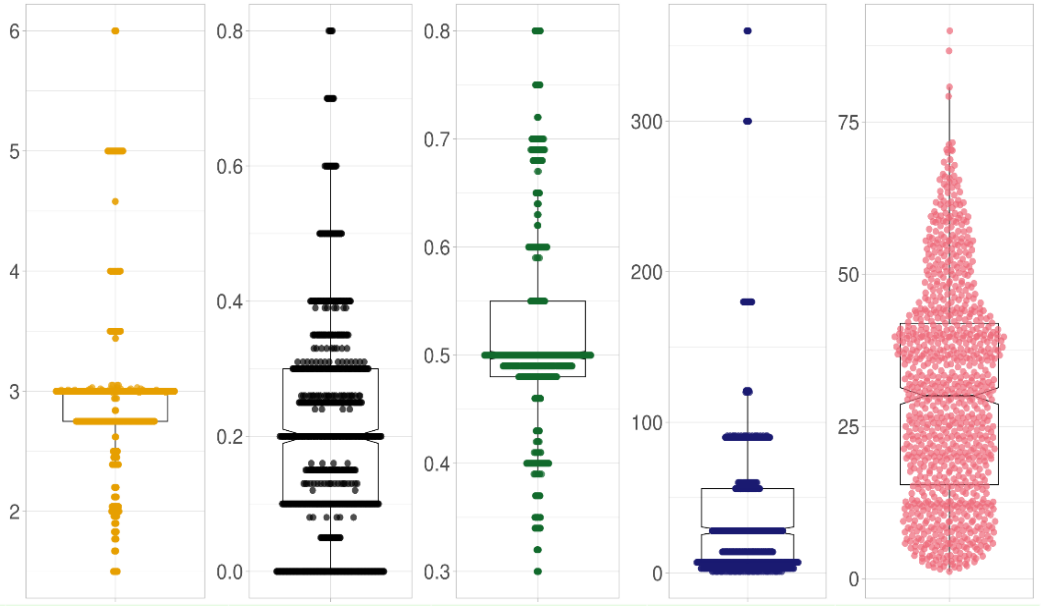
2.3.7 Regression error characteristic (REC) Curves

Regression error characteristic (REC) curves are a graphical tool to visualize the performance of a regression model. These curves plot the model's accuracy as a function of the prediction error margin (ε). The area under the REC curve (AUC) provides a single summary measure of model performance, with higher AUC values indicating better model accuracy over a range of error margins.

3 Results and discussion

3.1 Data distribution and statistical analysis

Fig. 5 shows the data distribution of the dataset used in the research. The Agg/B has a range from 1.50 to 6.00, with an average of 2.96 and a standard deviation of 0.71. The distribution of this variable is positively skewed (skewness = 1.49) with a kurtosis of 3.95, indicating a long right tail and a higher concentration of values around the lower end of the range. This suggests that while most observations tend to have a lower aggregate-to-binder ratio, there are outliers at the higher end. Similarly, FA/B ranges from 0.00 to 0.80, with a mean of 0.21 and a standard deviation of 0.16. The skewness of 0.87 indicates a moderate positive skew, and a kurtosis of 0.98 suggests a relatively flat distribution with fewer extreme values. This suggests that the data is more evenly distributed in this range but slightly favors lower fly ash content. The W/B has a range from 0.30 to 0.80, with an average of 0.52 and a standard deviation of 0.10. It shows moderate positive skewness (0.78) and low kurtosis (0.59), indicating that most of the data is concentrated towards the lower values, with a slight tail extending to the right. Curing time, on the other hand, shows a much wider distribution, with values ranging from 1 day to 360 days, averaging 39.82 days. The high skewness of 2.99 and kurtosis of 11.81 indicate a significant concentration of data in the lower curing period range, with a long tail extending to the higher end. This suggests that while most data points reflect short curing periods, there are instances where extremely long curing times (up to a year) were used.



d	Agg/B	FA/B	W/B	Curing period (days)	Compressive strength (MPa)
Minimum	1.50	0.00	0.30	1	0.75
Maximum	6.00	0.80	0.80	360	77.02
Average	2.96	0.21	0.52	39.82	18.16
Std	0.71	0.16	0.10	54.21	16.83
Skewness	1.49	0.87	0.78	2.99	1.20
Kurtosis	3.95	0.98	0.59	11.81	0.71

Fig. 5. Distribution of the dataset and statistical indices

Finally, the compressive strength of the mortar, which is the target variable, ranges from 0.75 MPa to 77.02 MPa, with an average of 18.16 MPa and a standard deviation of 16.83. The skewness of 1.20 indicates a positively skewed distribution, with a significant number of data points concentrated around the lower compressive strength values, though some outliers exhibit much higher strengths. The kurtosis of 0.71 further supports this by indicating a flatter distribution, with fewer extreme values compared to a normal distribution. Overall, the dataset exhibits significant variability across all variables, with several highly skewed and wide-ranging distributions, indicating that care should be taken in data preprocessing, including the application of transformations or feature engineering, to model these relationships better. The significant variability and positive skewness observed in variables like curing time (skewness: 2.99) and Agg/B (skewness: 1.49) represent the inherent heterogeneity of the data. While this diversity increases the predictive uncertainty, it significantly broadens the model's application range to various fly ash types and mix designs.

The correlation matrix for the variables in this study is presented in **Fig. 6**. The Agg/B shows a moderate negative correlation (−0.44) with compressive strength, indicating that as the aggregate content increases relative to the binder, the mortar's strength tends to decrease. This is consistent with the general understanding that higher aggregate content may weaken the final mixture, as the mortar becomes less cohesive. Additionally, Agg/B has a strong positive correlation (0.71) with FA/B, suggesting that as the fly ash content increases, there is a tendency for the Agg/B ratio also to increase. This relationship suggests that mixes with higher fly ash content may exhibit a corresponding increase in aggregate content, potentially affecting the overall material behavior. This interaction between Agg/B and FA/B should be further explored in the machine learning model, as it may significantly influence how both variables affect compressive strength. On the other hand, FA/B exhibits a moderate negative correlation (−0.3) with compressive strength, indicating that a higher fly ash content tends to reduce the compressive strength of the mortar. This aligns with previous findings that fly ash, while beneficial for

enhancing the sustainability and long-term durability of cementitious materials, can reduce early-age strength, particularly when used in larger quantities. However, the relatively weak correlation suggests that other factors, such as the curing period and the water-to-binder ratio, may also play significant roles in determining strength.

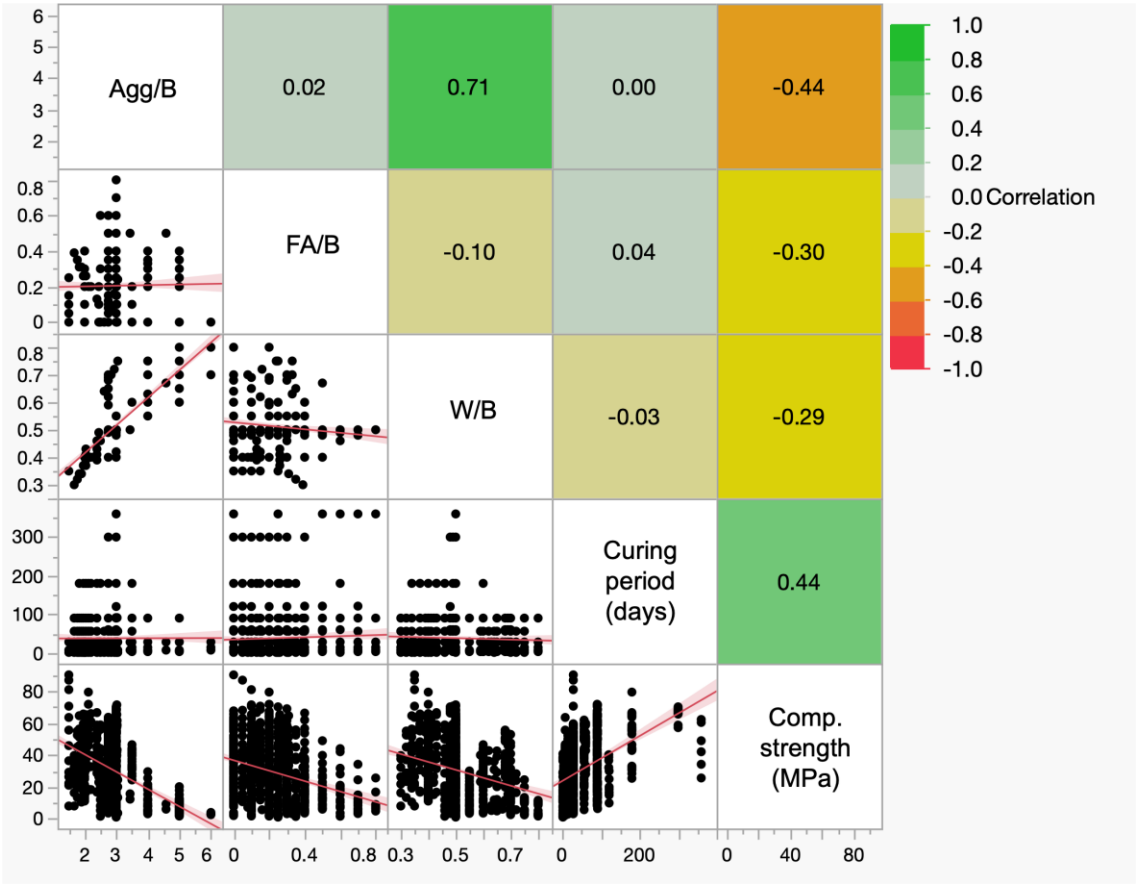


Fig. 6. Correlation matrix for input and output variables

The W/B exhibits a weak negative correlation (−0.29) with compressive strength, indicating that a higher water-to-binder ratio may slightly decrease the compressive strength. This is a common observation in cementitious materials, where excessive water content increases porosity, leading to lower strength. Lastly, the curing period has a positive correlation (r = 0.44) with compressive strength, which aligns with the conventional understanding. Longer curing times generally lead to improved compressive strength, especially in fly ash-blended mortars. The positive correlation between curing time and strength suggests that, beyond just the material composition, the time allowed for proper hydration and pozzolanic reactions in fly ash plays a crucial role in determining the final strength of the mortar. This variable, therefore, holds significant predictive power and should be a central feature in the machine learning model for predicting compressive strength.

3.2 Hyperparameter tuning and K-fold validation

Each model underwent rigorous hyperparameter tuning to ensure optimal performance, which is summarized in **Table 2**. For the ANN, the activation functions chosen were Linear, Gaussian, and TanH, with TanH selected as the optimal choice. This is a commonly used activation function for capturing non-linear relationships in data, as it maps the outputs between -1 and 1, providing non-linearity. The model was configured with two layers, which strikes a balance between complexity and overfitting. The number of nodes per layer was optimized at 4 and 14, ensuring sufficient capacity for learning without excessive complexity. The learning rate was set at 0.05, which provides a moderate pace for weight updates during training, ensuring efficient convergence without overshooting optimal values.

Table 2. Hyperparameter Optimization for Prediction Models

Model	Hyperparameters	Analyzed Values	Selected value
ANN	Activation Functions	[Linear, Gaussian, TanH]	[Linear, Gaussian, TanH]
	Number of layers	[1, 2]	2
	Nodes per layer	[1-16]	[4,14]
	Learning rate	[0.01, 0.025, 0.05, 0.1, 0.5]	0.05
KNN	K	[1-10]	8
SVR	Kernel	[linear, Radial basis function]	Radial basis function
	Cost	[0.01-5]	4.72611
	Gamma	[0.001-0.5]	0.48745
XGB	Max_depth	[1-7]	7
	Subsample	[0.5-1]	0.9701
	Colsample_bytree	[0.5-1]	0.8387
	Min_child_weight	[1-3]	1.2687
	Alpha	[0-0.5]	0.0386
	Lambda	[0-2]	0.0606
	Learning rate	[0.05-0.2]	0.1682
	Iterations	[0-50]	36

For the KNN model, the hyperparameter K (the number of neighbors) was tuned to 8, indicating that the model considers the eight closest neighbors when making predictions. This value suggests a balance between bias and variance, preventing the model from being overly sensitive to noise (which would occur with a smaller K) while avoiding over-smoothing the predictions (which would occur with a larger K). The choice of K optimizes the model's ability to generalize from the data. In the case of SVR, the Radial Basis Function (RBF) kernel was selected, which is highly effective for capturing complex, non-linear relationships between the input variables and the target. The cost parameter was set to 4.72611, balancing the model's tolerance for errors and its focus on minimizing training error, while gamma was set to 0.48745, determining the extent of influence each data point has. These settings allow the SVR model to create a highly flexible and accurate regression surface.

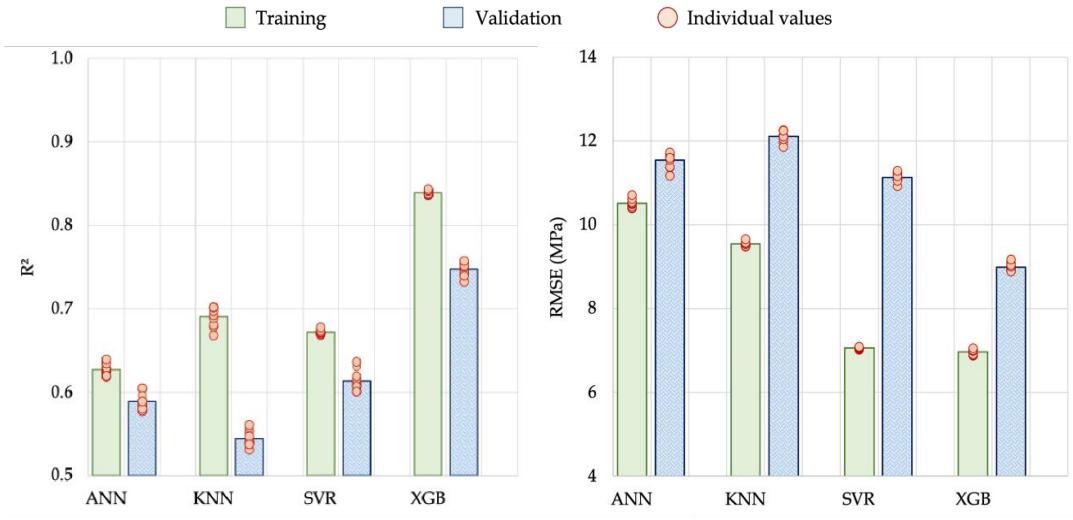
For the XGB model, multiple hyperparameters were optimized to fine-tune the model's performance. The max_depth of 7 allows the model to learn complex patterns without becoming overly deep and prone to overfitting. The subsample parameter of 0.9701 ensures that approximately 97% of the data is used during each iteration, introducing randomness that helps reduce overfitting. The colsample_bytree value of 0.8387 further aids regularization by limiting the number of features used in each tree, thereby enhancing the model's ability to generalize. The min_child_weight was selected as 1.2687, and the alpha value of 0.0386 introduces regularization to prevent overfitting. The learning rate of 0.1682 ensures the model converges steadily, while 36 iterations were selected, providing sufficient rounds of boosting to optimize the model without excessive computation.

The K-fold validation results for the machine learning models used in predicting the compressive strength of fly ash-blended sandcrete are shown in **Fig. 7**. The ANN model demonstrated moderate training accuracy (ranging from 0.62 to 0.64) and slightly lower testing accuracy (ranging from 0.58 to 0.61). This indicates a relatively stable performance on training data but a slight drop in generalization on unseen data. The increase in RMSE from 10.40 to 10.71 MPa during training to 11.17 to 11.73 MPa during testing suggests a small degree of overfitting, where the model performs well on the training data but slightly less effectively on new data.

The KNN model achieved higher training accuracy, ranging from 0.67 to 0.70, indicating a good fit to the training data. However, its testing accuracy was significantly lower (ranging from 0.53 to 0.56), and the RMSE for testing varied between 11.85 and 12.25 MPa. This suggests that the KNN model struggles with generalization and is sensitive to overfitting, performing well on training data but poorly on the test data. The SVR model showed consistent training accuracy (ranging from 0.67 to 0.68) and good generalization to testing data (ranging from 0.60 to 0.64). Its RMSE was lower during training (7.02–7.09 MPa) compared to testing (10.93–11.29 MPa), indicating a solid ability to generalize, although there is still room for improvement.

The XGB model outperformed all other models, with consistently high training accuracy of 0.84 across all folds. The testing accuracy ranged from 0.73 to 0.76, and its RMSE was the lowest, ranging

from 6.87 to 7.06 MPa during training and from 8.88 to 9.17 MPa during testing. This indicates that XGB not only performs well on the training data but also generalizes effectively to new, unseen data. The model's ability to minimize overfitting and achieve high performance on both training and test sets makes it the most robust and reliable model for predicting compressive strength in this study.



(a) R^2 (b) RMSE

Fig. 7. K-fold validation results (a) R^2 and (b) RMSE for different machine learning models

3.3 Performance of the models

Fig. 8 illustrates the predicted versus measured compressive strength of fly ash-blended sandcrete using different machine learning models. The figure illustrates the correlation between the predicted values and the actual measured values for each model, showcasing the accuracy and generalization performance of each. **Table 3** summarizes the performance indicators of different machine learning models for predicting compressive strength. These metrics help assess the models' prediction accuracy and generalization ability.

Table 3. Performance indicator for training and testing datasets for different machine learning models

	ANN		KNN		SVR		XGB	
	Train	Test	Train	Test	Train	Test	Train	Test
R2	0.63	0.59	0.69	0.54	0.67	0.62	0.84	0.74
RMSE	10.48	11.46	9.57	12.07	9.86	11.08	6.95	9.02
MAE	8.11	9.07	7.16	9.13	7.11	8.40	4.72	6.57
SI	0.35	0.37	0.32	0.39	0.33	0.35	0.23	0.29
a20	0.40	0.39	0.51	0.43	0.53	0.46	0.65	0.54

Among the four models, XGB demonstrated the highest predictive accuracy. With an R^2 value of 0.84 on the train set and 0.74 on the test set, XGB outperformed all other models in both training and testing. It also had the lowest RMSE and MAE, at 9.02 and 6.57 MPa, respectively, indicating superior precision in its predictions. Furthermore, the Scatter Index (SI) for XGB was the lowest at 0.29, reflecting its consistency in predicting the compressive strength with minimal variability. Additionally, XGB achieved the highest a20 value of 0.54, meaning that over 54% of its predictions were within a ±20% error margin, which is a favorable result for practical applications.

In comparison, ANN and KNN models displayed moderate performance. ANN showed an R^2 of 0.59 on the test set, with an RMSE of 11.46 MPa, indicating that it struggled to generalize effectively to unseen data. Similarly, KNN, despite having a higher R^2 of 0.69 during training, exhibited a significant drop to 0.54 on the test set, highlighting its tendency to overfit the data. Both models showed higher SI and lower a20 values, suggesting less stable predictions and a greater proportion of predictions falling outside the ±20% error margin.

SVR performed reasonably well, with an R^2 of 0.62 on the test set, and had an RMSE of 11.08 MPa, making it more reliable than ANN and KNN in terms of generalization. It achieved a lower MAE (8.40 MPa) compared to ANN and KNN, and the SI value of 0.35 indicated a relatively consistent performance. The a20 value for SVR (0.46) was also moderate, suggesting that while it performed better than ANN and KNN, it still left room for improvement in prediction accuracy. In conclusion, XGB emerged as the best-performing model for predicting the compressive strength, outperforming ANN, KNN, and SVR in all key performance indicators.

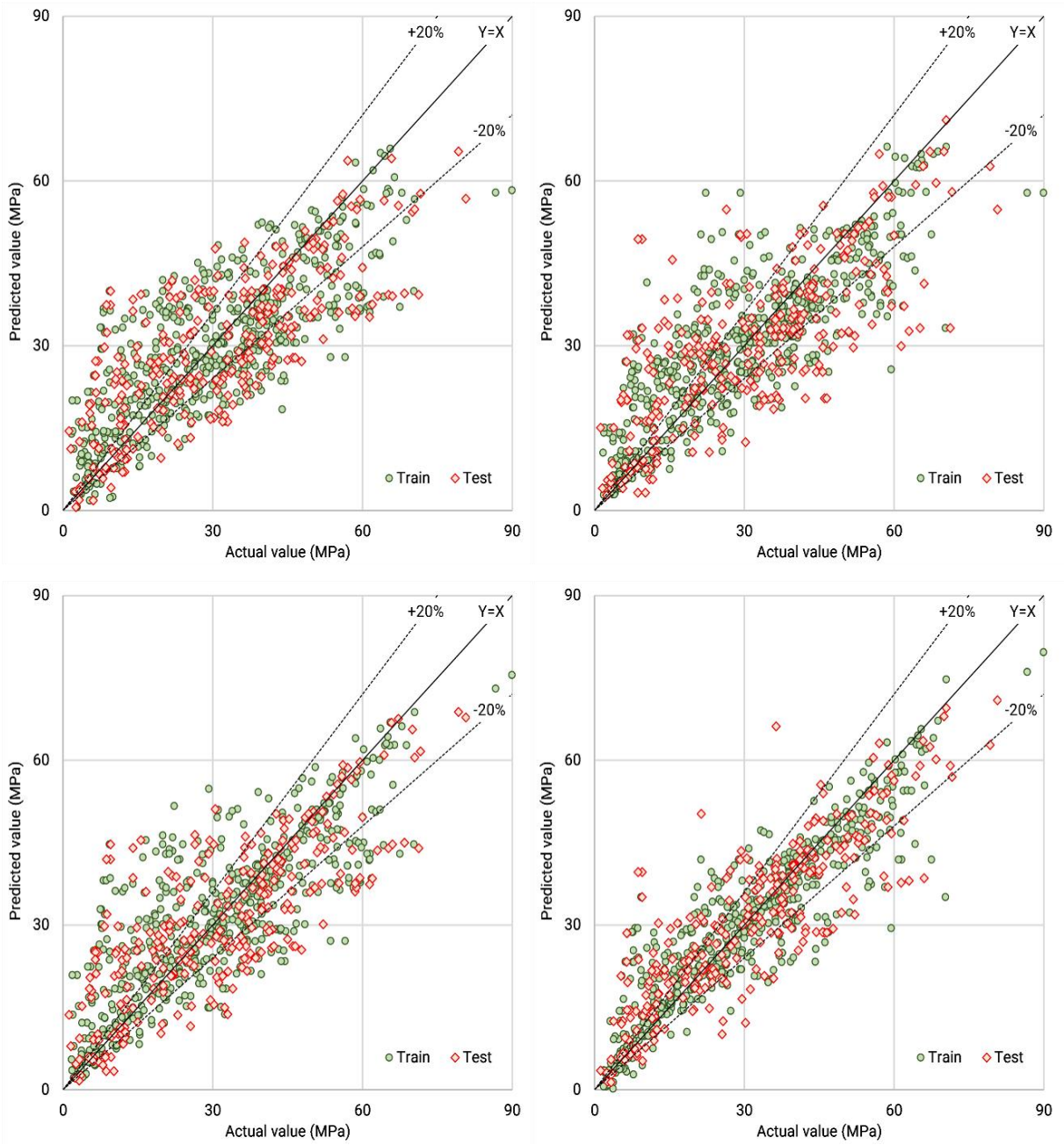


Fig. 8. Predicted vs. actual compressive strength of cement mortar.

3.4 Prediction error of the models

The machine learning models were evaluated based on their prediction error distributions, as shown in **Fig. 9**, for both the training and testing datasets. The training dataset results indicate that ANN performed the most consistently with an average error of -0.03 MPa, accompanied by a narrow 95% confidence interval ranging from -0.40 to 0.35 MPa. The skewness of 0.18 suggests a mild positive

skew, indicating that the model tended to slightly over-predict the compressive strength. KNN exhibited a similar performance, with an average error of -0.07 MPa; however, its skewness of -0.62 indicates a tendency towards under-prediction. SVR, with an average error of 0.16 MPa, showed a slight over-prediction tendency, reflected in its skewness of -0.04. XGB, with an average error of -0.04 MPa and a skewness of -0.41, also performed well, showing a slight under-prediction. The relatively narrow confidence intervals for all models on the training dataset suggest a high degree of certainty in their predictions.

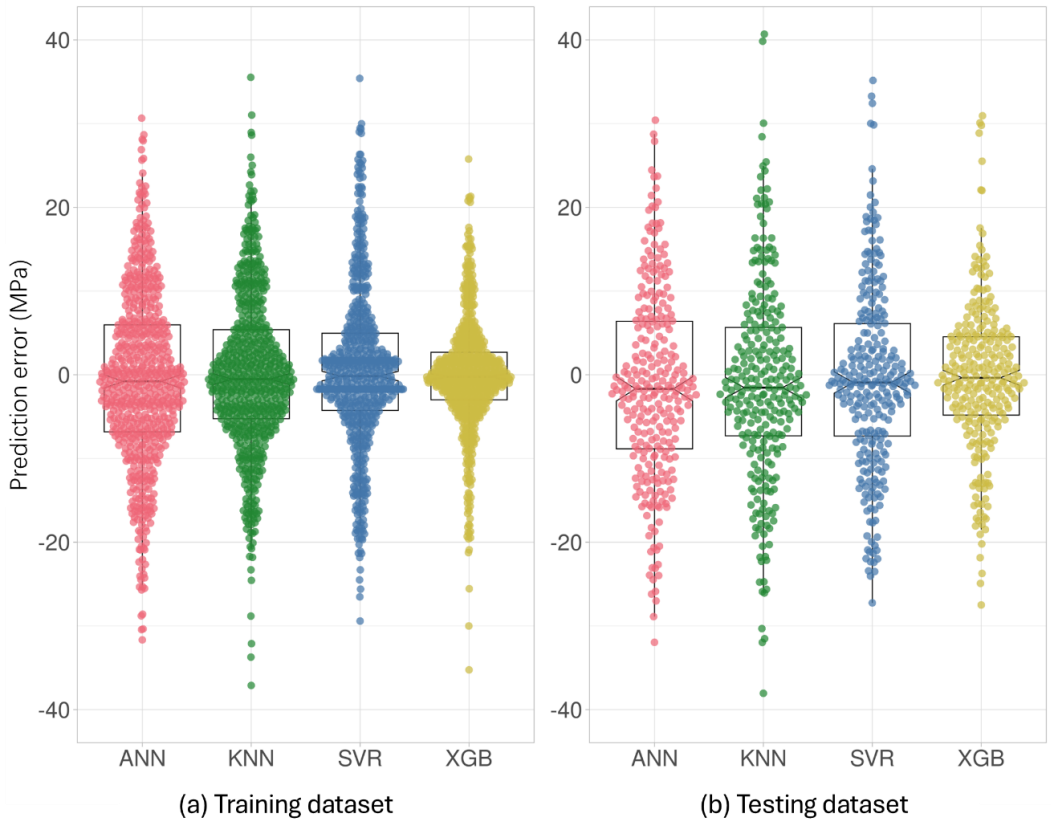


Fig. 9. Prediction error distribution for (a) the training dataset and (b) the testing dataset

When applied to the testing dataset, the models exhibited more variability. ANN had an average error of 0.10 MPa, with a skewness of 0.24, indicating a slight positive bias. However, the confidence interval for ANN was much broader, ranging from -0.66 to 0.86 MPa, suggesting increased uncertainty in its predictions on the testing dataset. KNN had an average error of 0.04 MPa, with a skewness of -0.86, indicating a stronger tendency towards under-prediction. The confidence interval for KNN was also wide, from -0.88 to 0.96 MPa, highlighting the model's uncertainty in the testing phase. SVR showed an average error of 0.46 MPa, with a skewness of -0.01, indicating almost symmetric errors, but its confidence interval was the widest, ranging from -0.27 to 1.19 MPa. Similarly, XGB had an average error of 0.47 MPa, with a skewness of -0.37, and a confidence interval ranging from -0.27 to 1.21 MPa. The broader confidence intervals for these models on the testing dataset suggest higher uncertainty in their generalization ability compared to their performance on the training data. Overall, ANN and XGB demonstrated the best overall performance, with minimal bias and relatively stable predictions across both training and testing datasets. SVR and KNN demonstrated higher variability, particularly on the testing dataset, with wider confidence intervals and larger prediction errors. A limitation of this study is the lack of categorical input variables for specimen geometry and specific testing standards (e.g., ASTM vs. EN). This heterogeneity introduces a 'noise floor' in the predictions. However, the uniform distribution of errors across the REC curves suggests that the XGB model remains robust against these variations and does not exhibit systematic bias toward any single data source. In machine learning training, such skewed distributions can lead to biased models that perform

better on frequent observations but struggle with 'tail' events (e.g., very long curing durations). To mitigate this, the decision-tree-based XGB model was prioritized, as it is inherently more robust to skewed features compared to distance-based models like KNN and SVR, which generally require symmetric distributions for optimal performance.

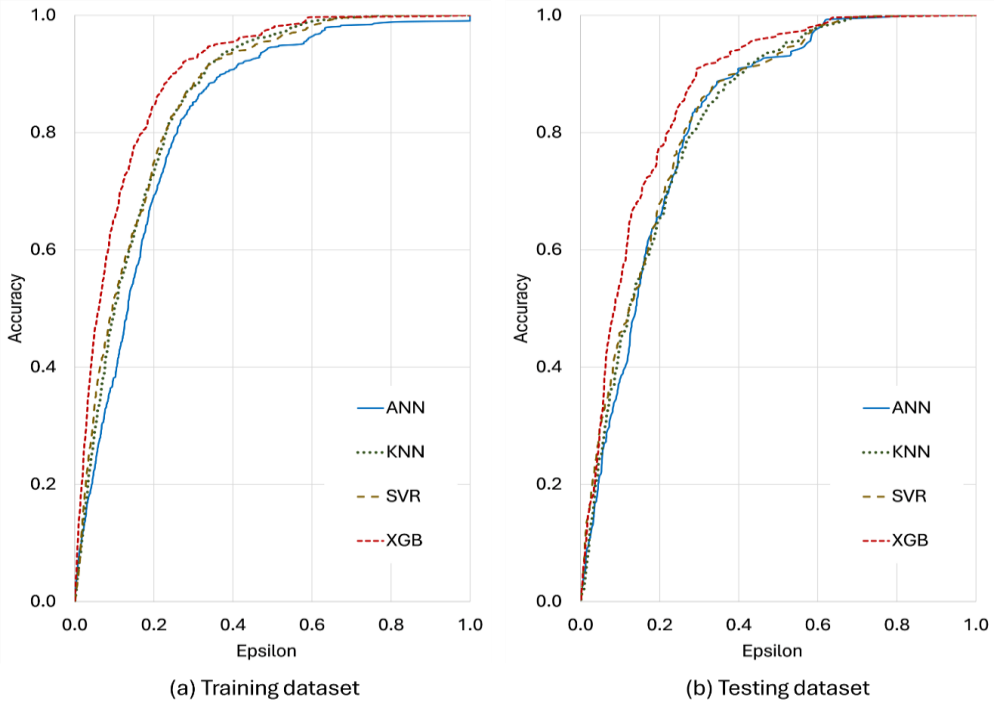


Fig. 10. Regression error characteristic (REC) Curves for (a) training dataset and (b) testing dataset

The REC curves presented in **Fig. 10** show the accuracy of each model as a function of the error margin (epsilon). All models demonstrated improved accuracy as the error margin increased, a typical outcome for predictive models in regression tasks. Among the models, XGB exhibited the best performance, as its REC curve approached a near-perfect accuracy of 1.0 faster than the other models. This suggests that XGB was the most accurate model in predicting compressive strength, particularly when a small error margin was allowed. In comparison, SVR and KNN displayed similar performance across the epsilon values, consistently lagging behind XGB but outperforming ANN. The results suggest that while SVR and KNN perform adequately, they are not as robust as XGB in handling the underlying data complexity and variability. The ANN, despite being a powerful model for many tasks, showed the weakest performance among the four models, indicating that it might not have been as well-suited for this particular prediction task.

When considering the AOC value or the training dataset, XGB outperformed all other models, with an AUC of 0.90, indicating a strong fit to the training data. SVR and KNN both obtained an AUC of 0.86, while ANN recorded the lowest AUC of 0.83. These values corroborate the observations from the REC curves, further supporting the superior performance of XGB. When evaluating the testing dataset, the XGB model maintained its top position with an AUC of 0.87, which reflects its ability to generalize well to new, unseen data. In contrast, SVR, KNN, and ANN all displayed slightly reduced AUC values, with SVR and KNN both achieving an AUC of 0.83 and ANN obtaining 0.82. This demonstrates that XGB not only performed well on the training dataset but also exhibited strong generalization capabilities, which is a critical factor for practical applications where model performance on unseen data is paramount.

Fig. 11 illustrates the number of data points that machine learning models accurately predicted within a 20% error margin. For the training dataset (640 data points), 183 data points were accurately predicted by all models, while none of the models accurately predicted 155 data points. Additionally, 69 data points were not accurately predicted by XGB but were correctly predicted by one or more of the other models. The corresponding values for the ANN, KNN, and SVR models are 226, 158, and

147, respectively. There are 59 data points that the XGB model uniquely predicted, while the corresponding values for the ANN, KNN, and SVR models are 3, 23, and 14, respectively.

For the testing dataset (275 data points), 65 data points were accurately predicted by all models, while none of the models accurately predicted 91 data points. Of the remaining data points, 35 were not accurately predicted by XGB but were accurately predicted by one or more of the other models. The corresponding values for ANN, KNN, and SVR models are 78, 66, and 58, respectively. Furthermore, 25 data points were uniquely predicted by the XGB model, with corresponding values of 3, 8, and 3 for the ANN, KNN, and SVR models, respectively. This suggests that the XGB model demonstrated strong generalization capabilities for new, unseen data. However, some predictions still fell outside the 20% error margin, highlighting the challenges of achieving perfect accuracy in all cases.

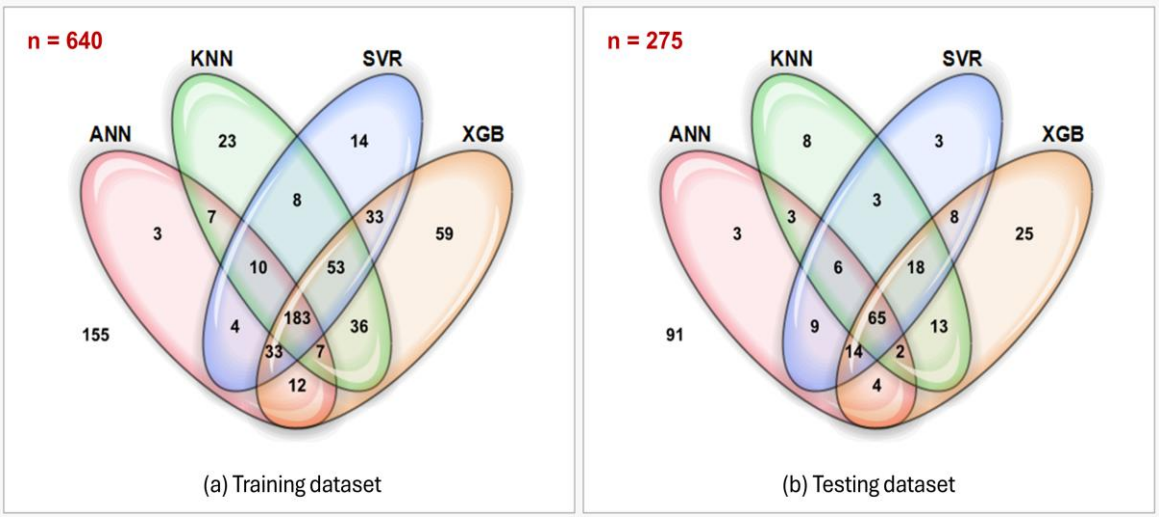


Fig. 11. Venn diagram for prediction accuracy for a) the training dataset and (b) the testing dataset

3.5 Sensitivity analysis

Fig. 12 and **Fig. 13** illustrate the influence of input features on model predictions, which was assessed using SHapley Additive exPlanations (SHAP), providing insights into feature importance and their respective contributions to the model's output. The results from the mean absolute SHAP values and the summary SHAP plots offer a detailed understanding of how each input variable affects the compressive strength predictions.

In terms of feature importance, the curing period stands out as the most influential variable, with a mean absolute SHAP value of 7.707. This suggests that curing time plays a crucial role in predicting the compressive strength of fly ash-blended cement mortar, with longer curing periods leading to higher strength predictions. This finding is consistent with established knowledge in the field of materials science, where extended curing periods facilitate the hydration process, thereby enhancing the material's strength. The Agg/B and the FA/B are both significant contributors to the model's predictive accuracy; however, the impact of FA/B is somewhat lower than that of Agg/B. With a mean absolute SHAP value of 6.617 for Agg/B and 3.530 for FA/B, the Agg/B ratio is shown to have a stronger influence on compressive strength predictions. The similarity in their SHAP values suggests that changes in either of these ratios significantly affect the predicted compressive strength, but Agg/B plays a more dominant role. As both ratios increase, the predicted strength tends to decrease, which aligns with the expected behavior in concrete and mortar mixtures. Higher Agg/B or FA/B ratios may result in a less cohesive matrix, thereby reducing the overall strength of the material. This inverse relationship is also corroborated by the summary SHAP plot, where high values of Agg/B and FA/B are associated with lower SHAP values, reflecting weaker compressive strength predictions. In contrast, W/B exhibits a weaker influence, with a mean absolute SHAP value of 1.158, making it the least significant variable among the four. This suggests that, while changes in the W/B ratio do affect compressive strength predictions, their impact is not as pronounced as that of the other variables. The summary SHAP plot further supports this, showing that the W/B ratio has a positive effect on compressive strength, where

higher W/B values tend to increase the predicted strength, but the overall effect remains less substantial compared to the Agg/B, FA/B, and curing period.

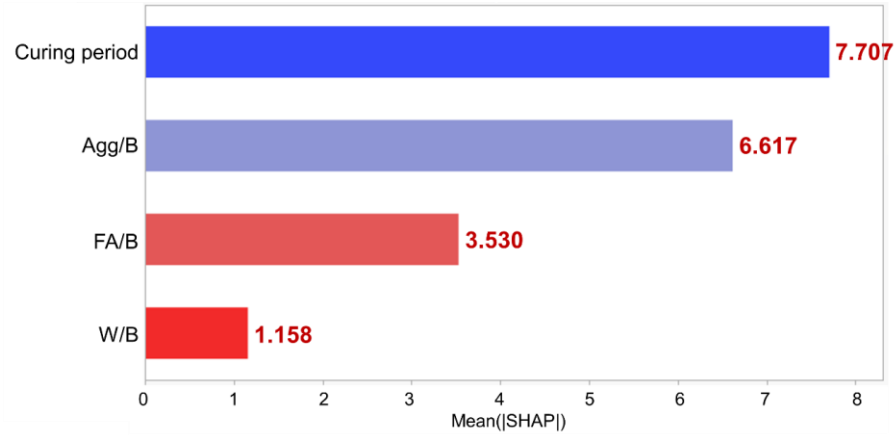


Fig. 12. Mean SHAP value for the input variables

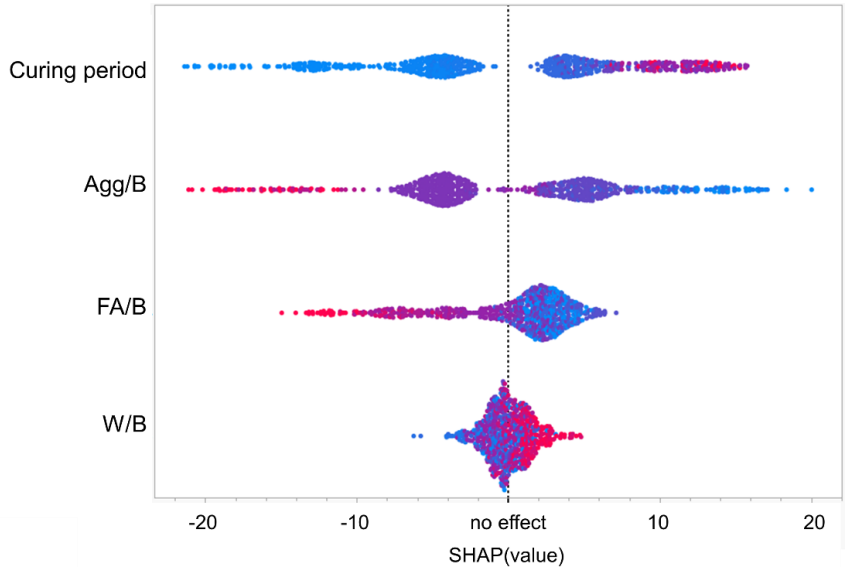


Fig. 13. SHAP summary plot for the input variables

The summary SHAP plot provides a more granular view of how the different input variables affect the model's predictions across a range of values. The plot demonstrates that as the curing period increases, the SHAP value consistently rises, indicating that longer curing periods lead to stronger predictions. For instance, a lower curing period yields a SHAP value of around -20, whereas higher curing periods result in SHAP values of up to +16. This reinforces the idea that curing time is a critical factor in determining compressive strength, as longer curing times allow for better hydration and bonding within the cement matrix. Similarly, the Agg/B ratio shows a negative correlation with compressive strength, with high Agg/B values corresponding to lower SHAP values (around -20) and lower Agg/B values associated with higher SHAP values (up to +20). This suggests that an excess of aggregate relative to binder in the mixture weakens the material, reducing its compressive strength. The same negative trend is observed for the FA/B ratio, with high values yielding SHAP values of around -15, and lower values resulting in positive SHAP values of +8. This inverse relationship suggests that a higher fly ash content in the mix may dilute the reactivity of the cement, potentially reducing the mortar's strength. Finally, the W/B ratio shows a moderate positive relationship with compressive strength, where higher values of W/B correspond to stronger predictions. The SHAP values range from -6 for low W/B values to +5 for higher W/B values, suggesting that increasing the W/B ratio can enhance the compressive strength, albeit to a lesser extent compared to the curing period and the aggregate-to-binder and fly ash-to-binder ratios.

4 Conclusion

This study aimed to predict the compressive strength of fly ash-blended sandcrete using advanced machine learning models, thereby contributing to the optimization of sustainable construction materials. A comprehensive dataset of 915 data points, sourced from 34 peer-reviewed studies, was utilized to evaluate the performance of four machine learning models: ANN, KNN, SVR and XGB.

The findings revealed that the compressive strength of fly ash blended mortar is significantly influenced by curing time, Agg/B, and FA/B, while the W/B has a comparatively lesser impact. Among the models, XGB demonstrated the highest predictive accuracy across all performance metrics. Specifically, XGB achieved an R^2 value of 0.84 for training and 0.74 for testing, indicating its superior ability to explain the variance in compressive strength. Additionally, XGB exhibited the lowest Root RMSE of 6.95 MPa during training and 9.02 MPa during testing, along with the lowest MAE and the highest fraction of predictions within a ±20% error margin (54%). In contrast, while KNN and SVR models showed moderate performance, they exhibited significant variability and overfitting tendencies, particularly in the testing phase. ANN displayed the weakest performance, with lower R^2 values and higher prediction errors. The sensitivity analysis further emphasized the importance of curing time in predicting compressive strength, followed by the Agg/B and FA/B ratios. The results suggest that while higher curing time leads to improved compressive strength, higher Agg/B and FA/B ratios generally result in lower strength predictions.

The results of this study highlight the potential of machine learning, particularly XGB, in accurately predicting the compressive strength of fly ash blended cement mortars. The XGB model's robust performance across diverse datasets positions it as a valuable tool for optimizing mix designs and promoting sustainability in construction practices. This study contributes to the growing body of knowledge on the application of machine learning for construction material optimization, offering a pathway for more efficient and accurate predictions of material performance in real-world scenarios. While factors such as fly ash chemical composition (CaO content), cement type, and curing temperature significantly influence the hydration process, these parameters were not consistently reported across all 34 independent literature sources. To maintain a robust and statistically significant sample size, the study focused on the four most universally reported and influential mix design parameters: Agg/B, FA/B, W/B, and curing time. This approach ensures the model's high applicability to practical scenarios where detailed chemical analysis of constituents may not be readily available. The exclusion of specimen size and fly ash classification (Class F vs. Class C) likely contributes to the observed residual errors in the testing phase. Future studies could enhance predictive precision by incorporating categorical inputs for specimen geometry or the Calcium-to-Silica (C/S) ratio of the binder system as more standardized data becomes available in the literature.

Following further research, the application of even more advanced machine learning techniques could be explored, refining model predictions by incorporating additional variables and more extensive datasets.

- Future studies could investigate the inclusion of additional input variables that may affect the compressive strength of fly ash blended mortars, such as temperature, humidity, or the chemical composition of the fly ash. These factors could further refine the predictive models and improve the accuracy of compressive strength predictions.
- Combining multiple machine learning techniques, such as ensemble models or hybrid algorithms, could improve the predictive power of the models. For example, combining XGB with other models, such as neural networks or deep learning methods, may lead to better generalization and accuracy in predicting compressive strength, particularly in complex material systems.
- Further exploration into the optimization of machine learning models, such as deep learning and reinforcement learning, could offer improvements in prediction accuracy. Deep learning models, in particular, might be able to capture more complex patterns in data, improving the model's ability to predict the compressive strength of mortars with greater precision.
- Future research could focus on developing multi-objective optimization models that not only predict the compressive strength of fly ash-based mortars but also integrate sustainability metrics to enhance their environmental impact. This would enable the simultaneous

optimization of mechanical properties and environmental impact, considering factors such as the carbon footprint and energy consumption associated with various mix designs.

- The generalizability of the current models could be tested using real-world data from construction sites or large-scale production plants. By applying these models to actual construction scenarios, researchers can assess their performance in practical settings and refine them based on real-time data, improving their robustness and reliability.
- While this study primarily focused on fly ash, the use of other SCMs such as slag, silica fume, or natural pozzolans as partial replacements for cement could be explored. Investigating their synergistic effects with fly ash would provide a broader understanding of how multiple SCMs can optimize the performance and sustainability of cementitious materials.
- Conducting studies that evaluate fly ash sourced from different regions and industrial processes could provide valuable insights into how variations in fly ash composition and quality affect the performance of cementitious materials. Such studies would help refine the models to account for regional differences in materials, leading to more adaptable and versatile predictive models.

By pursuing these directions, future research can further improve the predictive accuracy, sustainability, and practical application of machine learning models in the field of construction materials. These advancements will be crucial for developing more efficient, cost-effective, and environmentally friendly construction practices in the years to come.

Funding Statement

The author received no specific funding for this study.

CRediT authorship contribution statement

Sathiparan: Conceptualization, Investigation, Formal analysis, Writing – original draft, Writing – review & editing.

Conflicts of Interest

The author affirms that there are no conflicts of interest to disclose in relation to this study.

Data Availability Statement

All data and models that support the findings of this study are available from the corresponding author upon reasonable request.

References

[1] Sathiparan N: Mesh type seismic retrofitting for masonry structures: critical issues and possible strategies. European Journal of Environmental and Civil Engineering 2015; 19(9): 1136-1154. https://doi.org/10.1080/19648189.2015.1005160.

[2] Nawab MS, Ali T, Qureshi MZ, Zaid O, Ben Kahla N, Sun Y, Anwar N, Ajwad A: A study on improving the performance of cement-based mortar with silica fume, metakaolin, and coconut fibers. Case Studies in Construction Materials 2023; 19: e02480. https://doi.org/10.1016/j.cscm.2023.e02480.

[3] Sathiparan N: A systematic review of corncob ash in construction: Current findings and future directions. Sustainable Materials and Technologies 2025; 43: e01315. https://doi.org/10.1016/j.susmat.2025.e01315.

[4] Sathiparan N, Anburuvel A, Muralitharan M, Isura Kothalawala DA: Sustainable use of coco pith in cement-sand mortar for masonry block production: Mechanical characteristics, durability and environmental benefit. Journal of Cleaner Production 2022; 360: 132243. https://doi.org/10.1016/j.jclepro.2022.132243.

[5] Alli YA, Bamisaye A, Bamidele MO, Etafo NO, Chkirida S, Lawal A, Hammed VO, Akinfenwa AS, Hanson E, Nwakile C et al: Transforming waste to wealth: Harnessing carbon dioxide for sustainable solutions. Results in Surfaces and Interfaces 2024; 17: 100321. https://doi.org/10.1016/j.rsurfi.2024.100321.

[6] Knight KA, Cunningham PR, Miller SA: Optimizing supplementary cementitious material replacement to minimize the environmental impacts of concrete. Cement and Concrete Composites 2023; 139: 105049. https://doi.org/10.1016/j.cemconcomp.2023.105049.

[7] Rumman R, Kamal MR, Bediwy A, Alam MS: Partially burnt wood fly ash characterization and its application in low-carbon mortar and concrete. Construction and Building Materials 2023; 402: 132946. ttps

://doi.org/10.1016/j.conbuildmat.2023.132946.
[8] Sathiparan N, Jaasim JHM, Banujan B: Sustainable production of cement masonry blocks with the combined use of fly ash and quarry waste. Materialia 2022; 26: 101621. https://doi.org/10.1016/j.mtla.2022.101621.
[9] Al-Ahdal A, AbdelRahman B, Galal K: Compressive, shear, and tensile behaviours of concrete masonry: Experimental and numerical study. Construction and Building Materials 2025; 458: 139266. ttps://doi.org/10.1016/j.conbuildmat.2024.139266.
[10] Chen Q, Zhang J, Zhang L, Wang Z, Zhao T, Zhang Y, Wang Z: Compressive strength predictio of high-performance concrete: Integrating multi-ingredient influences and mix proportion insights. Construction and Building Materials 2024; 451: 138791. https://doi.org/10.1016/j.conbuildmat.2024.138791.
[11] Rathnayaka M, Karunasinghe D, Gunasekara C, Wijesundara K, Lokuge W, Law DW: Machine learning approaches to predict compressive strength of fly ash-based geopolymer concrete: A comprehensive review. Construction and Building Materials 2024; 419: 135519. https://doi.org/10.1016/j.conbuildmat.2024.135519.
[12] Sathiparan N: Predicting compressive strength in cement mortar: The impact of fly ash composition through machine learning. Sustainable Chemistry and Pharmacy 2025; 43: 101915. https://doi.org/10.1016/j.scp.2025.101915.
[13] Ahmad A, Farooq F, Niewiadomski P, Ostrowski K, Akbar A, Aslam F, Alyousef R: Prediction of Compressive Strength of Fly Ash Based Concrete Using Individual and Ensemble Algorithm. In: Materials 2021; 14.
[14] Jiang Y, Li H, Zhou Y: Compressive Strength Prediction of Fly Ash Concrete Using Machine Learning Techniques. Buildings 2022; 12(5): 690. https://doi.org/10.3390/buildings12050690.
[15] Kakasor Ismael Jaf D, Abdulrahman AS, Abdulrahman PI, Salih Mohammed A, Kurda R, Ahmed HU, Faraj RH: Effitioned soft computing models to evaluate the impact of silicon dioxide (SiO2) to calcium oxide (CaO) ratio in fly ash on the compressive strength of concrete. Journal of Building Engineering 2023; 74: 106820. https://doi.org/10.1016/j.jobe.2023.106820.
[16] Kakasor Ismael Jaf D, Ismael Abdulrahman P, Salih Mohammed A, Kurda R, Qaidi SMA, Asteris PG: Machine learning techniques and multi-scale models to evaluate the impact of silicon dioxide (SiO2) and calcium oxide (CaO) in fly ash on the compressive strength of green concrete. Construction and Building Materials 2023; 400: 132604. https://doi.org/10.1016/j.conbuildmat.2023.132604.
[17] Khursheed S, Jagan J, Samui P, Kumar S: Compressive strength prediction of fly ash concrete by using machine learning techniques. Innovative Infrastructure Solutions 2021; 6(3): 149. https://doi.org/10.1007/s41062-021-00506-z.
[18] Loureiro AAB, Stefani R: Comparing the performance of machine learning models for predicting the compressive strength of concrete. Discover Civil Engineering 2024; 1(1): 19. https://doi.org/10.1007/s44290024-00022-w.
[19] Omer B, Jaf DKI, Abdalla A, Mohammed AS, Abdulrahman PI, Kurda R: Advanced modeling for predicting compressive strength in fly ash-modified recycled aggregate concrete: XGboost, MEP, MARS, and ANN approaches. Innovative Infrastructure Solutions 2024; 9(3): 61. https://doi.org/10.1007/s41062-024-01365-0.
[20] Onyelowe KC, Jagan J, Kontoni D-PN, Moghal AAB, Onuoha IC, Viswanathan R, Soni DK: Utilization of GEP and ANN for predicting the net-zero compressive strength of fly ash concrete toward carbon neutrality infrastructure regime. International Journal of Low-Carbon Technologies 2023; 18: 902-914. https://doi.org/10.1093/ijlct/ctad081.
[21] Song H, Ahmad A, Farooq F, Ostrowski KA, Maślak M, Czarnecki S, Aslam F: Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms. Construction and Building Materials 2021; 308: 125021. https://doi.org/10.1016/j.conbuildmat.2021.125021.
[22] Wadhawan S, Bassi A, Singh R, Patel M: Prediction of Compressive Strength for Fly Ash-Based Concrete: Critical Comparison of Machine Learning Algorithms. Journal of Soft Computing in Civil Engineering 2023; 7(3): 68-110. https://doi.org/10.22115/scce.2023.353183.1493.
[23] Abu Yaman M, Abd Elaty M, Taman M: Predicting the ingredients of self compacting concrete using artificial neural network. Alexandria Engineering Journal 2017; 56(4): 523-532. https://doi.org/10.1016/j.aej.2017.04.007.
[24] Belalia Douma O, Boukhatem B, Ghrici M, Tagnit-Hamou A: Prediction of properties of self-compacting concrete containing fly ash using artificial neural network. Neural Computing and Applications 2017, 28(1): 707-718. https://doi.org/10.1007/s00521-016-2368-7.
[25] Ismael Jaf DK: Soft Computing and Machine Learning-Based Models to Predict the Slump and Compressive Strength of Self-Compacted Concrete Modified with Fly Ash. In: Sustainability 2023; 15.
[26] Saha P, Debnath P, Thomas P: Prediction of fresh and hardened properties of self-compacting concrete using support vector regression approach. Neural Computing and Applications 2020; 32(12): 7995-8010. https://doi.org/10.1007/s00521-019-04267-w.
[27] Sathiparan N, Jeyananthan P, Subramaniam DN: Predicting compressive strength of pervious concrete with

fly ash: a machine learning approach and analysis of fly ash compositional influence. Multiscale and Multidisciplinary Modeling, Experiments and Design 2024; 7(6): 5651-5671. 10.1007/s41939-024-00551-y.
[28] Sathiparan N, Jeyananthan P, Subramaniam DN: Prediction of compressive strength of fly ash blended pervious concrete: a machine learning approach. International Journal of Pavement Engineering 2023; 24(2): 2287146.
[29] Abdellatief 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 in Engineering 2025; 25: 104542. https://doi.org/10.1016/j.rineng.2025.104542.
[30] Alvarenga KP, Cordeiro GC: Evaluating sugarcane bagasse fly ash as a sustainable cement replacement for enhanced performance. Cleaner Engineering and Technology 2024; 20: 100751. ttps://doi.org/10.1016/j.clet. 2024.100751.
[31] Arenas-Piedrahita JC, Montes-García P, Mendoza-Rangel JM, López Calvo HZ, Valdez-Tamez PL, Martínez-Reyes J: Mechanical and durability properties of mortars prepared with untreated sugarcane bagasse ash and untreated fly ash. Construction and Building Materials 2016; 105: 69-81. https://doi.org/10.1016/j.conbuildmat.2015.12.047.
[32] Argiz C, Menéndez E, Sanjuán MA: Effect of mixes made of coal bottom ash and fly ash on the mechanical strength and porosity of Portland cement. Materiales de Construcción 2013; 63(309): 49-64. https://doi.or g/10.3989/mc.2013.03911.
[33] Atiş CD, Kiliç A, Sevim UK: Strength and shrinkage properties of mortar containing a nonstandard high-calcium fly ash. Cement and Concrete Research 2004; 34(1): 99-102. ttps://doi.org/10.1016/S0008-8846(03) 00247-3.
[34] Baeza F, Payá J, Galao O, Saval JM, Garcés P: Blending of industrial waste from different sources as partial substitution of Portland cement in pastes and mortars. Construction and Building Materials 2014; 66: 645-653. https://doi.org/10.1016/j.conbuildmat.2014.05.089.
[35] Chandra S, Takhelmayum G: Study on flow and compressive strength properties of mortars using fly ash and lime. International Journal of Engineering Science and Innovative Technology 2013; 2(6): 309-315.
[36] Cheerarot R, Jaturapitakkul C: A study of disposed fly ash from landfill to replace Portland cement. Waste Management 2004; 24(7): 701-709. https://doi.org/10.1016/j.wasman.2004.02.003.
[37] Chindaprasirt P, Rukzon S: Strength, porosity and corrosion resistance of ternary blend Portland cement, rice husk ash and fly ash mortar. Construction and Building Materials 2008; 22(8): 1601-1606. https://doi.org/10.1016/j.conbuildmat.2007.06.010.
[38] Fattouh MS, Abouhalawa AS: fly ash as a partial replacement of cement on the compressive strength of cement mortar using north sinai martials. American Journal of Engineering Research 2021; 10(11): 162-167.
[39] Hatungimana D, Taşköprü C, İçhedef M, Saç MM, Yazıcı Ş: Compressive strength, water absorption, water sorptivity and surface radon exhalation rate of silica fume and fly ash based mortar. Journal of Building Engineering 2019; 23: 369-376. https://doi.org/10.1016/j.jobe.2019.01.011.
[40] Hsu S, Chi M, Huang R: Effect of fineness and replacement ratio of ground fly ash on properties of blended cement mortar. Construction and Building Materials 2018; 176: 250-258. https://doi.org/10.1016/j.conbuildmat.2018.05.060.
[41] Huang Q, Zhao L: Correlation between compressive strengths and water absorption of fly ash cement mortar immersed in water. Archives of Civil Engineering 2019; 1(15): 141-152.
[42] Islam MM, Islam MS: Strength behaviour of mortar using fly ash as partial replacement of cement. Concrete Research Letters 2010; 1(3): 98-106.
[43] Javed K, Hassan MI-u, Farooq MA, Sharif MB: Detailed investigation of compressive and bond strength for sustainable brick masonry developed by using various types of bricks and green mortars. Journal of Building Engineering 2024; 84: 108477. https://doi.org/10.1016/j.jobe.2024.108477.
[44] Jozic D, Zelic J, Janjatovic I: Influence of the coarse fly ash on the mechanical properties of the cement mortars. Cermics - Silikaty 2010; 54(2): 144-151.
[45] Khan MI, Abbas YM, Fares G, Alqahtani FK: Flowability and Strength Characteristics of Binary Cementitious Systems Containing Silica Fume, Fly Ash, Metakaolin, and Glass Cullet Powder. Materials 2023; 16(19). https://doi.org/10.3390/ma16196436.
[46] Kim J, Yi C, Zi G: Waste glass sludge as a partial cement replacement in mortar. Construction and Building Materials 2015; 75: 242-246. https://doi.org/10.1016/j.conbuildmat.2014.11.007.
[47] Madurwar KV, Burile AN, Sorte AM: Compressive strength of cement & fly ash mortar - A case study. In: Proceedings of Sustainable Infrastructure Development & Management (SIDM) 2019; Nagpur. Visvesvaraya National Institute of Technology: 3376014.
[48] Mardani-Aghabaglou A, Andiç-Çakir Ö, Ramyar K: Freeze–thaw resistance and transport properties of high-volume fly ash roller compacted concrete designed by maximum density method. Cement and Concrete Composites 2013; 37: 259-266. https://doi.org/10.1016/j.cemconcomp 2013.01.009.
[49] Nandekar PR, Nandurkar BP: Effect of industrial byproducts such as fly ash and rice husk ash on the

performance of cement sand mortar. International Journal of Latest Trends Engineering and Technology 2016; 6(4): 531-536.
[50] Nandurkar BP, Pande AM: Compressive strength of mortars containing fly ash and rice husk ash. International Journal of Civil Engineering and Technology 2018; 9(7): 1012-1020.
[51] Nguyen NTT, Ngo TV, Nguyen KK, Vu VQ, Xia Y, Tran MQ, Dang HT, Matos J, Dang SN: Effects of Fly Ash and Graphene Oxide in Cement Mortar Considering the Local Recycled Material Context. In: Applied Sciences 2024; 14: 6140.
[52] Ramaswamy KP, Siddik M, Nazeer M: Workability ans strength studies on fly ash modified masonry mortars. In: International Conference on Modeling & Simulation in Civil Engineering (ICMSC 2011): 2011; Kerala.
[53] Ratnaweera PR: Influence of fly ash characteristics on compressive strength of mortar. New Jersey Institute of Technology 1991.
[54] Ravichandran Y, Balasubramanian M: Effect of processed fly ash on cement mortar for standard fine and normal sand. Journal of Civil Engineering and Environmental Technology 2014; 1(1): 1-4.
[55] Rukzon S, Chindaprasirt P: Use of disposed waste ash from landfills to replace Portland cement. Waste Management & Research 2009; 27(6): 588-594. https://doi.org/10.1177/0734242X09103189.
[56] Supit SWM, Shaikh FUA, Sarker PK: Effect of ultrafine fly ash on mechanical properties of high volume fly ash mortar. Construction and Building Materials 2014; 51: 278-286. https://doi.org/10.1016/j.conbuildmat.2013.11.002.
[57] Wang H-Y, Wang C-C, Chang S-C, Lin J-C: A study of engineering and electricity properties of cement mortar added with recycled materials and piezoelectric powders. Construction and Building Materials 2016; 113: 297-305. https://doi.org/10.1016/j.conbuildmat.2016.03.056.
[58] Wnag CP, Chuang YC, Lin JX: Evaluation on feasibility of fly ash cement mortar as adhesive of post-installed rebar. Materials Science and Engineering 2019; 615: 012021. https://doi.org/10.1088/1757-899X/615/1/012021.
[59] Wibowo AP, Saidani M: Mortar with fly ash as a partial cement replacement: analysing the compressive strength and heat of hydration. E3S Web of Conferences 2023; 429: 05033. https://doi.org/10.1051/e3sconf/202342905033.
[60] Xu A, Sarkar SL, Nilsson LO: Effect of fly ash on the microstructure of cement mortar. Materials and Structures 1993; 26(7): 414-424. https://doi.org/10.1007/BF02472942.
[61] Yang Y, Takasu K, Suyama H, Ji X, Xu M, Liu Z: Comparative Analysis of Woody Biomass Fly Ash and Class F Fly Ash as Supplementary Cementitious Materials in Mortar. In: Materials 2024; 17: 3723.
[62] Yang Y, Takasu K, Suyama H, Ji X, Xu M, Liu Z: Comparative Analysis of Woody Biomass Fly Ash and Class F Fly Ash as Supplementary Cementitious Materials in Mortar. In: Materials 2024; 17.
[63] Yu J, Wu H-L, Mishra DK, Li G, Leung CKY: Compressive strength and environmental impact of sustainable blended cement with high-dosage Limestone and Calcined Clay (LC2). Journal of Cleaner Production 2021; 278: 123616. https://doi.org/10.1016/j.jclepro.2020.123616.
[64] Sathiparan N, Jeyananthan P, Subramaniam DN: Effect of aggregate size, aggregate to cement ratio and compaction energy on ultrasonic pulse velocity of pervious concrete: prediction by an analytical model and machine learning techniques. Asian Journal of Civil Engineering 2024; 25(1): 495-509. https://doi.org/10.1007/s42107-023-00790-3.
[65] Qayyum Khan A, Ahmad Awan H, Rasul M, Ahmad Siddiqi Z, Pimanmas A: Optimized artificial neural network model for accurate prediction of compressive strength of normal and high strength concrete. Cleaner Materials 2023; 10: 100211. https://doi.org/10.1016/j.clema.2023.100211.
[66] Sathiparan N, Jeyananthan P: Predicting compressive strength of cement-stabilized earth blocks using machine learning models incorporating cement content, ultrasonic pulse velocity, and electrical resistivity. Nondestructive Testing and Evaluation 2024; 39(5): 1045-1069. https://doi.org/10.1080/10589759.2023.2240940.
[67] Halder RK, Uddin MN, Uddin MA, Aryal S, Khraisat A: Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications. Journal of Big Data 2024; 11(1): 113. https://doi.org/10.1186/s40537-024-00973-y.
[68] Sathiparan N: Prediction model for compressive strength of rice husk ash blended sandcrete blocks using a machine learning models. Asian Journal of Civil Engineering 2024; 25(6): 4745-4758. https://doi.org/10.1007/s42107-024-01077-x.
[69] Sathiparan N: Predicting compressive strength of grouted masonry using machine learning models with feature importance analysis. Materials Today Communications 2024; 41: 110487. https://doi.org/10.1016/j.mtcomm.2024.110487.
[70] Sathiparan N, Jeyananthan P, Subramaniam DN: A machine learning approach to predicting pervious concrete properties: a review. Innovative Infrastructure Solutions 2025; 10(2): 55. https://doi.org/10.1007/s41062-024-01829-3.