

STRENGTH PREDICTION MODELS FOR CONCRETE INCORPORATING FLY ASH AS A WASTE MATERIAL

M. Kumar ^a, V. Kumar^{b,*}, A. Priyadarshee^b, A. Kumar Rahul^b, R. Kumar ^c, Shweta Kumari ^d

^a Department of Civil Engineering, Government Polytechnic Muzaffarpur, Naya Tola, Muzaffarpur, Bihar, India-842001,
E-mail: madan123kumar@gmail.com

^b Department of Civil Engineering, Muzaffarpur Institute of Technology, Muzaffarpur, Laxmi Chowk, Muzaffarpur, Bihar, India-842003, E-mails: * kumar2vijay@gmail.com; i.akashpriyadarsheel@gmail.com; atulcivil.iitbhu@gmail.com

^c Department of Electronics and Communication Engineering, Muzaffarpur Institute of Technology, Muzaffarpur, Laxmi Chowk, Muzaffarpur, Bihar, India-842003, E-mail: ampuravi@gmail.com

^d Department of Information Technology, Muzaffarpur Institute of Technology, Muzaffarpur, Laxmi Chowk, Muzaffarpur, Bihar, India-842003, E-mail: shweta@mitmuzaffarpur.org

Received: 27.01.2026 / Accepted: 17.02.2026 / Revised: 30.03.2026 / Available online: 31.05.2026

DOI: [10.2478/jaes-2026-0007](https://doi.org/10.2478/jaes-2026-0007)

KEY WORDS: waste material, SCM, fly ash, machine learning, strength prediction, sustainable concrete.

ABSTRACT:

A comparative machine learning-based methodology was adopted to predict the compressive strength of fly ash concrete using Multiple Linear Regression (MLR), Support Vector Regression (SVR), AdaBoost Regressor (ABR), Random Forest (RF), and Extreme Gradient Boosting models (XG). A dataset of 498 mix designs collected from published literature was used, considering cement, fine and coarse aggregate, fly ash content, water content, water-cement ratio, and curing period as input parameters. Model performance was evaluated using mean absolute error, root mean square error, and coefficient of determination. The Extreme Gradient Boosting model showed the best predictive capability ($R^2 = 0.881$; RMSE = 5.65 MPa). Sensitivity analysis identified curing period, cement content, and water content as the most influential variables. The results demonstrate reliable strength prediction and enable model comparison to support data-driven mix optimization for sustainable fly ash concrete (FAC).

1. INTRODUCTION

The growing population is increasing the demand for human needs, driving up industrial and agricultural activities. Such activities expand the production of waste materials. Poor management of such waste material threatens all living organisms and the environment. Utilizing such waste material is one of the most effective ways of handling it. Waste materials like fly ash (FA), rice husk ash (RHA), tire crumbles (TC), municipal solid waste (MSW), etc., are becoming popular in the construction of different structures (Lee et al. 2003; Priyadarshee et al. 2021). Waste materials like FA, pond ash (PA), etc., are generally not very aggressive. However, the large production of such materials creates enormous problems. The bulk utilization of such waste material is possible in construction. Another issue of a growing population is a boom in the infrastructural development to fulfill the requirements of the people. To support the infrastructural development, materials like cement, steel, etc., must be produced. Cement production alone accounts for 5-8% of global CO₂ emissions (Benhelal et al. 2013). Cement production has significant environmental impacts across all stages of its manufacturing process (Benhelal et al. 2013; MacLaren and White 2003). Each year, approximately 4 billion

tons of Portland cement (PC) are produced, with the manufacturing of one ton of cement releasing roughly 0.8 tons of CO₂ into the atmosphere (Hanifa et al. 2023; Ren et al. 2023). This substantial carbon dioxide emission poses a critical threat to environmental sustainability.

To address these challenges, researchers have increasingly focused on developing Supplementary Cementing Materials (SCMs) as a sustainable alternative. These materials mitigate environmental issues and substantial waste disposal (Khatib and Wild 1998). Studies have demonstrated that incorporating SCMs in the construction industry effectively diverts waste from landfills while offering environmental and economic advantages (Ondova et al. 2012; Pitroda et al. 2012). SCMs can be derived from various industrial and agricultural by-products, further supporting sustainable construction practices. Fly ash (FA) is a by-product generated during coal combustion in power plants. It is a pozzolanic by-product generated from thermal power plants. Once regarded as waste, it is recognized for its economic benefits and ability to enhance the performance of fresh and hardened concrete. Due to its cementitious characteristics, FA is an adequate partial replacement for cement in concrete mixtures. Its

* Corresponding author. Vijay Kumar, Assistant Professor, Department of Civil Engineering, MIT Muzaffarpur, Laxmi Chowk, Muzaffarpur, Bihar, India-842003, e-mail: kumar2vijay@gmail.com.

widespread application as a Supplementary Cementing Material (SCM) in the construction industry improves material properties and contributes significantly to environmental sustainability by reducing reliance on traditional cement.

Many researchers have investigated the impact of adding FA on the strength of concrete. Lam et al. (2000), Thomas (2007), Harwalker and Awati (2014), and others have shown that adding FA suppresses the early strength of concrete but improves its long-term strength (Singh et al. 2024). They have also found that adding FA increases the durability of the concrete. The content of FA generally varies between 15% to 30% by weight of cement. For structures like foundations and dams, a higher content of FA is used (Xu and Shi 2018). Mardani-Aghabaglou and Ramyar 2013; Liu and Paresuel-Moreno 2014; Amran et al. 2020; and others have shown that partial replacement of cement with FA decreases strength due to an increase in the water-binder ratio. Researchers like Barbuta et al. 2017; Sadrmomtazi et al. 2018; and Kumar and Prasad 2019 etc. have investigated the impact of admixtures like silica fume, glass fiber, and polyester fibers along with FA on the performance of concrete (Antoni et al. 2015; Pitroda et al. 2012). Similarly, other researchers like Jena and Panda 2018, Iqbal et al. 2017; Bellum et al. 2020, etc., have investigated the impact of FA on the properties of concrete tensile strength, flexural strength, modulus of elasticity, impact resistance, etc. (Behnood et al. 2017). They found that adding the FA as a replacement gives satisfactory results with concrete.

Different research works are done to develop mathematical models to predict the compressive strength of concrete. Such models help engineers during various stages, from design to construction (Lee et al. 2003). Traditionally, statistical methods were used to develop such a mathematical model (Singh et al. 2024). Nowadays, machine learning is becoming a widespread technique in developing such models. Machine learning is popular in various fields of Civil engineering (Priyadarshee et al. 2021; Kumar et al. 2023; Singh et al. 2024; Priyadarshee et al. 2025). Artificial neural networks (ANN), Deep neural networks (DNN), Multiple linear regression (MLR), Support vector machine (SVM), etc., are examples of different machine learning techniques that can be utilized in the development of strength prediction models of concrete (Khambra and Shukla 2023) researchers like Nguyen et al. Dao et al. (2019) and Ling et al. (2019) have used DNN and ANN techniques to predict the strength of geopolymers concrete (Priyadarshee et al. 2021). Similarly, Choudhary (2019), Al-Shamiri et al. (2020), Chopra et al. (2018), Drousseau et al. (2019), etc. have developed prediction models based upon ANN, sequential feature selection (SFS), Decision tree (DT), Random Forest (RF), etc. (Asteris and Mokos 2020; Y. Song et al. 2022a). Some of the work, like Naseri et al. (2020), Ling et al. (2021), Azmi-pour et al. (2020), etc., is done to predict the strength of fly-ash-based concrete based on ANN, SVM, and other machine learning techniques (Asteris and Mokos 2020; Feng et al. 2020). Considering the different studies, it is found that there are limited comparative studies, which are done on small or homogeneous datasets. Also, there is limited study on the quantification of the relative influence of mix

parameters on CS, which is essential for practical mix design optimization. Based on these limitations mentioned in the present study, a comprehensive side-by-side comparison of five ML algorithms MLR, SVR, ABR, RF, and XGBR using a large, diverse dataset of 498 FAC mixed compiled from published literature is done. The specific contributions are:

- Development of predictive models under identical data preprocessing and evaluation protocols, enabling a fair comparison of algorithm performance.
- Application of hyperparameter tuning to ensemble models (ABR, XGBR) to maximize prediction accuracy.
- Sensitivity analysis to identify and rank the most influential mix parameters affecting CS.
- Integration of predictive modeling with practical engineering insights for sustainable concrete mix optimization.

By combining model benchmarking with feature importance evaluation, this study not only identifies the most accurate algorithm for FAC strength prediction but also provides actionable guidance for practitioners. It enables the bridging the gap between data-driven modelling. Also, the provide the solution of real-world application in sustainable concrete design.

2. METHODOLOGY

As mentioned above, the main objective of the present study is to compare the performance of ML models developed using different techniques to predict the CS of FA concrete. Figure 1 illustrate the methodology adopted in the present study. This indicates that the first data collection was done from different literature. After this, the data was preprocessed. Other models were developed, and their training and testing were done. Based on the data from training and testing, performance matrices were prepared. Based on the performance metrics, a comparison of the models was done. In the following subsection, details of each stage are presented.

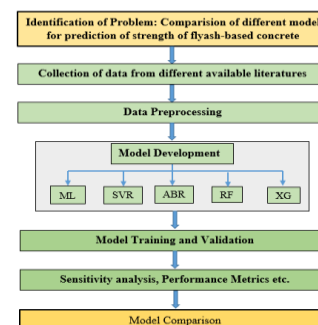


Figure 1. Methodology adopted in the present study

2.1 Data Collection

In the present study, 498 data sets on the CS of fly ash content (FC) were gathered from literature surveys of past research papers (Antoni et al. 2015; Behnood et al. 2017; Benhelal et al. 2013; Bildirci 2019; Krishnan 2016; Kim et al. 2018; Kumar and Mehta 2004; Ling et al. 2013; Mukarjee 2012; Kumar et al. 2006; Ondova et al. 2012; Paudel et al. 2023; Pitroda et al. 2012; Rashad 2015; Sakir et al. 2020; Salami et al. 2021; Sandhu and Siddique 2017; Shaikh 2016; Siriruang et al. 2016; Solanki and Pitroda 2013; Song et al.

2021; Song et al. 2022b; Teixeira et al. 2019; Vairagade et al. 2015; Wadhawan et al. 2023). For the development of models, CS was taken as the dependent variable, and fly ash content (FC), CA content (CAC), FAG content (FAGC), superplasticizer content (SC), water content (WC), water cement ratio (WCR), and CP were taken as independent variables. So, data related to independent and dependent variables were taken from the literature. The statistical summary of the adopted dataset used in this study is depicted in Table 1. This table provides the minimum and maximum values, standard deviations (SD), and average values for input and output variables. The proposed model should be applied only within the specified parameter ranges to ensure accurate predictions. The statistical analysis reveals that the dataset used encompasses a wide range of ingredient values, and the SD indicates the extent of data distribution around the mean. As per Table 2, all the datasets that have been gathered have a CS range of 6.27–79.99 MPa. The proposed model can estimate the CS of the samples. It is applicable only within the specified range, as mentioned in statical dataset. The table demonstrates that the variables in the database span a wide and representative range of values. This ensures the robustness and comprehensiveness of the collected data. The diversity in input parameters enhances the reliability of the database. It also allows the proposed ML models to capture complex relationships between variables. As a result, the models can provide accurate and consistent predictions of CS within the considered range.

Ingre dient	Ce me nt (kg/ m ³)	Fly Ash (kg/ m ³)	Wat er (kg/ m ³)	SP (kg/ m ³)	CA (kg/ m ³)	FA G (kg/ m ³)	CP (day s)	w/c -	CS (MP a)
count	498	498	498	498	498	498	498	498	498
mean	329.73	64.77	178.41	3.25	101.70	741.93	45.48	0.36	35.20
SD	91.18	68.78	23.87	5.35	127.56	109.81	60.26	0.25	15.30
min	136.10	0.00	126.70	0.00	533.00	291.00	1.00	0.00	6.27
25%	275.00	0.00	159.00	0.00	966.00	688.90	7.00	0.00	23.11
50%	317.45	48.00	184.00	0.00	100.35	763.00	28.00	0.40	35.04
75%	379.00	124.78	192.00	7.00	108.35	808.00	56.00	0.58	43.93
max	540.00	300.00	254.00	28.20	125.60	905.40	365.00	0.70	79.99

Table 2. Summary of descriptive statistics for the input and output parameters used in the study

2.2 Data Preprocessing

After the data collection, preprocessing was done. Data preprocessing is a crucial step as it directly impacts the performance and accuracy of the model. It generally involves data cleaning, transformation, feature selection and extraction, splitting data, etc. During data collection, missing data and duplicity were taken care of carefully. To avoid the issue raised due to the scale sensitivity of models. All the variables were normalized in the range of [0,1]. From Table 2, it can be observed that the variables are on different scales and ranges. So, in such conditions, 'Data Normalization,' also termed as the rescaling process, become important. The normalized data (X_n) expression can be written as below.

$$X_n = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

Where: X is the variable's original value, and $X_{\min}$ and $X_{\max}$ are each input variable's minimum and maximum values, respectively.

Further, 498 data points were divided randomly into two groups at this stage. In the first group, 70% of the total data, 349 data points, were retained for training. The remaining 30%, 149 data points, were kept in the second group. These data were utilized for testing models.

2.3 Model Development

As mentioned above, five different techniques were used to develop various models. The descriptions of the other methods used in the present study are summarized below.

2.3.1 Multivariable linear Regression (MLR): MLR is a statistical technique. It is used to model the relationship between a target variable (i.e., dependent variable) and two or more input predictors (i.e., independent variables). This is achieved by fitting a linear equation to the observed data while assessing the degree of correlation between the variables. The model is typically developed using the least squares approach. This method minimizes the sum of squared differences between the observed and predicted values. The MLR technique evaluates how multiple independent variables collectively influence the dependent variable and produces a linear relationship. According to Neter et al. (1996), the methodology assumes that each observation is independent, the variance of residuals (errors) is constant across all levels of the independent variables (homoscedasticity), and the independent variables are not highly correlated with each other (i.e., there is no multicollinearity). The general form of the MLR equation is expressed as:

$$Y = A_0 + A_1X_1 + A_2X_2 + \dots + A_nX_n + \epsilon \quad (2)$$

Where: Y is the target variable; $X_1, X_2, X_3, \dots, X_n$ are input predictor variables; A_0 is intercept; $A_1, A_2, \dots, A_n$ are the coefficients of independent variables, and ϵ is the error.

2.3.2 Support Vector Regression (SVR): It is a robust supervised machine learning algorithm. It is widely applied to regression and classification tasks. Renowned for its ability to handle nonlinear relationships, small datasets, and high-dimensional feature spaces, SVR is particularly effective for predictive modeling in fields like civil engineering. Unlike traditional regression techniques, SVR maps input data into a higher-dimensional feature space using a nonlinear kernel function, such as the radial basis function (RBF) or polynomial kernel. This transformation enables the algorithm to identify an optimal hyperplane that minimizes prediction error while maintaining a maximum tolerance margin, ensuring both accuracy and generalization. The core objective of the SVR is to find the following function.

$$f(x) = (w, \phi(x)) + b \quad (3)$$

It predicts the target variable y while ensuring deviations from the observed data points (x_i, y_i) remain within a predefined margin ϵ . This is achieved by solving the following optimization problem:

$$\sum_{w, b, \xi, \zeta^*}^{\min} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \zeta_i^*) \quad (4)$$

Subjected to:

$$y_i - (w, \phi(x_i)) - b \leq \epsilon + \xi_i \quad (5)$$

$$\langle w, \phi(x_i) \rangle + b - y_i \leq \epsilon + \xi_i^* \quad (6)$$

$$\xi_i, \xi_i^* \geq 0 \quad (7)$$

Where, $\|w\|^2$ ensures a flat regression function, ξ_i and ξ_i^* are slack variables allowing tolerance for outliers, and C is a regularization parameter controlling the trade-off between margin width and prediction error. By leveraging its kernel-based transformation and optimization framework, SVR can effectively model complex nonlinear relationships between input features and the target variable. It then precious for predicting the construction materials' CS of the desired dataset.

2.3.3 Adaptive Boosting Regression (ABR): Adaptive Boosting (AdaBoost) is an ensemble learning method that sequentially combines weak learners, such as shallow decision trees, to create a strong predictive model. Iteratively adjusting sample weights emphasizes correcting errors from prior iterations, improving overall accuracy. Particularly effective in capturing non-linear relationships, AdaBoost is well-suited for modelling complex behaviours like the CS of FA concrete, which depends on mixed design parameters such as cement content, water-to-cement ratio, and fly ash proportion. Its structure enhances predictive performance while reducing overfitting, ensuring reliable and generalizable predictions under diverse conditions. AdaBoost is a valuable tool for advancing material property modelling. AdaBoost follows a systematic approach to iteratively improve the performance of weak learners through weighted adjustments and ensemble predictions. At each iteration, the weighted error (ϵ_t) of the weak learner is calculated based on the misclassification of samples, given by the expression:

$$\epsilon_t = \frac{\sum_{i=1}^N w_i^{(t)} \cdot \mathbb{I}(h_t(x_i) \neq y_i)}{\sum_{i=1}^N w_i^{(t)}} \quad (8)$$

Where: $w_i^{(t)}$ represents the weight of sample i , $h_t(x_i)$ is the weak learner's prediction, and $\mathbb{I}$ is an indicator function. The contribution of each weak learner to the final model is quantified by its weight (α_t), determined using:

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right) \quad (9)$$

This ensures learners with lower errors have higher weights, making them more influential in the ensemble. To focus on correcting errors, the weights of the training samples are updated after each iteration as:

$$w_i^{(t+1)} = w_i^{(t)} \cdot \exp(\alpha_t \cdot \mathbb{I}(h_t(x_i) \neq y_i)) \quad (10)$$

The updated weights are then normalized to ensure proper scaling for subsequent iterations. The final prediction of the AdaBoost model for regression tasks is obtained as a weighted sum of predictions from all the weak learners:

$$H(x) = \sum_{t=1}^T \alpha_t \cdot h_t(x) \quad (11)$$

This formulation enables AdaBoost to effectively capture the non-linear relationships between input parameters, such as the mix design variables in fly ash concrete and the target variable, compressive strength. By iteratively minimizing the exponential loss function:

$$L = \sum_{i=1}^N \exp(-y_i \cdot H(x_i)) \quad (12)$$

AdaBoost enhances prediction accuracy and ensures generalization. These mathematical foundations underline the algorithm's ability to model the complex interactions in fly ash concrete behaviour, making it a robust tool for analysing and predicting material properties. Including such expressions in your study demonstrates the algorithm's systematic approach and highlights its suitability for regression tasks in material science research.

2.3.4 Random Forest (RF): RF regression is a robust supervised machine learning technique. This method is widely used for predicting the CS of materials like FA concrete. This method is used for its high performance in both regression and classification tasks; RF leverages an ensemble learning approach to enhance predictive accuracy. Using this method, it reduces the overfitting of the dataset. The algorithm constructs multiple decision trees during training; each built on a randomly bootstrapped subset of the dataset and aggregates their outputs—using averaging for regression tasks or majority voting for classification. This ensemble structure makes RF highly effective for modeling complex, nonlinear relationships in large-scale problems. The core mechanism of RF involves creating diverse decision trees by randomly selecting subsets of features for each tree split, ensuring reduced correlation among trees and greater model generalization. For regression tasks, the final prediction is obtained by averaging the predictions from all individual trees:

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T h_t(x) \quad (13)$$

Where: T is the total number of decision trees, $h_t(x)$ is the prediction of the t -th tree, and $\hat{y}$ is the overall prediction. An additional advantage of this method is its ability to provide feature importance metrics, enabling insights into which mix design parameters such as, FA content, water-to-cement ratio (W/C ratio), or curing duration are the most influential factors for determining CS of the dataset. Combining individual decision trees' strengths while mitigating their weaknesses, Random Forest regression delivers reliable and interpretable results, making it an invaluable tool for complex predictive modelling in civil engineering research.

2.3.5 Extreme gradient boosting (XGB): This method is a highly efficient and scalable supervised ML algorithm. This method is widely used to predict the materials CS, like FAC. Built on the principles of gradient boosting, XGBoost is an ensemble method that constructs multiple decision trees in sequence, where each tree aims to correct the errors of its predecessor. Optimized for speed and performance, XGBoost is particularly effective for handling large datasets and capturing complex, nonlinear relationships, making it well-suited for regression tasks. The strength of XGBoost lies in its advanced optimization techniques and system-level enhancements, including parallel and distributed computing. It used a novel tree-learning algorithm designed for sparse data, and regularization to prevent overfitting. For regression tasks, XGBoost minimizes a regularized objective function that combines the prediction error and the complexity of the model:

$$L(y) = \sum_{k=1}^K (y_i - y_j)^2 + \sum_{k=1}^K \lambda (J_k)$$

Where: $l(y_i, \hat{y}_i)$ is a loss function measuring the difference between observed (y_i) and predicted $(\hat{y}_i)$ values, $\Omega(f_k) = \frac{1}{2} \lambda \|w_k\|^2$ represents the regularization term, and Θ encompasses the model parameters. Each tree's prediction is added to refine the overall model output:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(x_i) \quad (15)$$

Where: $f_t(x_i)$ is the prediction of the t-h tree, the scalability of XGBoost is further enhanced by features like weighted quantile sketch for tree splitting and hardware optimizations, allowing it to handle large and high-dimensional datasets efficiently. In the context of FAC, XGBoost's ability to model intricate relationships between mixed design parameters—such as fly ash content, water-to-cement ratio, and curing conditions—and compressive strength makes it a powerful tool for predictive modelling. Its regularization techniques and advanced optimization ensure high accuracy and robustness, positioning XGBoost as a competitive alternative to other machine learning methods like Random Forest.

2.4 Model Training and Testing

The dataset was prepared for model development by defining the architecture, including key parameters such as the number of layers, units, and activation functions. It was then split into training, validation, and testing subsets to ensure reliable evaluation of model performance. For the ML models MLR, and SVR, default parameters were used during initialization. For the XGBoost model, the objective function was set to "reg:squarederror" to minimize the squared error during training. Each model was trained using the designated training dataset, with the validation set used to fine-tune the models and assess their generalizability. After completing the training phase, the models were evaluated on the test dataset to measure prediction errors and examine correlations between predicted and observed values. Once testing was complete, the entire dataset was utilized to generate predictions for the target variable. To maintain model performance over time, periodic updates were performed by retraining on newly available data or by optimizing hyperparameters. Coding and implementation were carried out using Google Colab notebooks, providing an efficient platform for computational tasks. Graphical analysis and data visualization were performed using OriginLab and Microsoft Excel, ensuring clear representation of results and trends.

2.5 Performance Metrics

Moreover, the output of the ML model was evaluated using several performance indices. These include mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), mean absolute percentage error (MAPE), coefficient of correlation (R), and coefficient of determination (R^2). The mathematical expressions of these equations are presented in Eq. (8) and Eq. 13, respectively.

$$MAE = \frac{1}{n} \sum_{i=1}^n |EV_i - PV_i| \quad (16)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (EV_i - PV_i)^2 \quad (17)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (PV_i - EV_i)^2}{n}} \quad (18)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{EV_i - PV_i}{EV_i} \right| \quad (19)$$

$$R = \frac{\sum_{i=1}^n (EV_i - EV_{mean})(PV_i - PV_{mean})}{\sqrt{\sum_{i=1}^n (EV_i - EV_{mean})^2 \sum_{i=1}^n (PV_i - PV_{mean})^2}} \quad (20)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (EV_i - PV_i)^2}{\sum_{i=1}^n (EV_i - EV_{mean})^2} \quad (21)$$

Where:

PV_i is the predicted value of the output variable;
 EV_i is the measured value of the output variable;
 n is the number of data used in the analysis;
 PV_{mean} is the average model output value; and
 EV_{mean} is the average experimental value.

For the CS prediction problem, the accuracy and reliability of a ML model are often evaluated using key performance metrics. An R^2 value closer to 1 indicates a strong correlation between the predicted and actual values, signifying a more accurate model. Meanwhile, metrics such as MSE and RMSE quantify the level of discrepancy between predicted and observed results, with smaller values reflecting improved prediction accuracy. MAE further provides insight into the average magnitude of prediction errors, with lower MAE values indicating better model performance. Additionally, MAPE evaluates the relative error as a percentage of the actual values, helping to assess the model's performance across datasets (Song et al., 2021; Li et al., 2023).

Hyperparameters, which are set prior to the training process, play a critical role in controlling the learning process and the behaviour of an ML model. These parameters influence various aspects of the model, including regularization techniques to prevent overfitting, the search space for optimization, and the overall performance of the algorithm. Optimizing hyperparameters is essential for minimizing the loss function and achieving better predictive accuracy with fewer errors.

In this study, hyperparameter tuning was applied to the XGBoost and AdaBoost regressor models to enhance their performance. A systematic approach was adopted to explore a larger search space, enabling the identification of optimal parameter combinations. This tuning process improved the model outputs, ensuring better prediction accuracy and reliability.

3. RESULTS AND DISCUSSION

The results obtained from the analysis are presented in the following section. It begins with an exploration of the dataset's statistical characteristics through correlation matrices, histograms, boxplots, and scatter plots, providing insight into the relationships between input variables and CS. This is followed by a performance comparison of five selected ML algorithms considered in this study as mentioned previously namely; MLR, SVR, ABR, RF, and XGBR using both testing and validation datasets. The section further includes a sensitivity analysis to identify the most influential mix parameters and an evaluation of

each model using multiple performance metrics, enabling a comprehensive assessment of prediction accuracy, robustness, and practical applicability.

3.1 Correlation matrix

The correlation matrix (Figure 3) provides an overview of the pairwise Pearson correlation coefficients between all input variables and the compressive strength (CS) of FAC. From Figure 3, it can be observed that CC and CP show the strongest positive correlations with CS, confirming their well-established role in strength development. Longer curing promotes continued hydration and pozzolanic reactions of fly ash, leading to higher long-term strength. WC exhibits a moderate negative correlation with CS, consistent with the detrimental effect of higher water-cement ratios on concrete strength. SP content shows a weak positive correlation with CS, suggesting its influence via improved workability and reduced required water content. CA displays the lowest correlation with CS, indicating that within the study range, its variation does not strongly affect compressive strength. No variable shows a correlation coefficient close to zero with CS, suggesting all variables contribute to some extent to the prediction model. Importantly, intercorrelations among input variables such as between WC and WCR highlight the risk of multicollinearity in traditional regression models. This reinforces the choice of ensemble ML methods like RF and XGBR, which are less sensitive to multicollinearity.

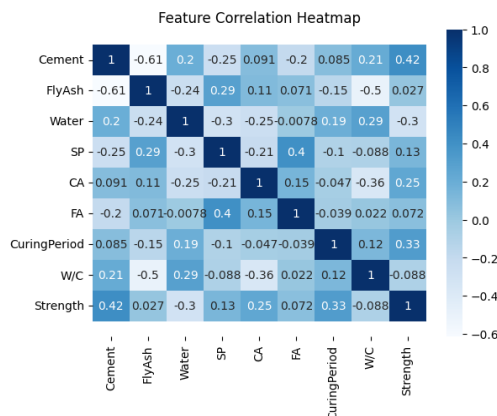


Figure 3. Heat map showing Pearson's correlation coefficients among the input variables considered in the current study

3.2 Histogram, Boxplot of the Dataset

A graphical representation of the distribution of a set of discrete data is presented by a histogram plot (Tipu et al. 2022; Ly et al. 2021; Wadhawan et al. 2023). This kind of bar chart displays the density or frequency of various values or ranges of values within the data. It serves as an effective tool for understanding data distribution. It identifies patterns and trends in the dataset. The ranges of values are displayed on the X-axis, while the frequency or density of the data is plotted on the Y-axis. The distributions of various datasets can be compared with it. Additionally, it can assist in determining the type of distribution. The histograms presented in Figures 4 and 5 illustrate the distribution of normalized values of all input variables and CS. The distributions are generally well spread across the 0-1 normalized range, indicating that the dataset covers a broad spectrum of mix designs, which enhances model robustness.

CS distribution shown in Figure 4 shows a right-skewed pattern, with a higher frequency of samples in the 20-45 MPa range. CC

content distribution shown in Figure 5 is moderately concentrated with fewer high-cement mixes. While FC exhibits a widespread, with a noticeable cluster near zero, reflecting some mixes without FA addition. The boxplot analysis (Figure 6) compares predicted versus actual CS values. While the medians are close, a slight upward shift in the predicted median suggests a minor overestimation bias in the models, particularly in higher strength ranges. The interquartile range (IQR) for predicted CS is slightly broader, indicating marginally higher variability in predictions compared to actual values. Outliers are present in both datasets, primarily associated with mixes having extreme curing times or unusual binder proportions, which may warrant targeted data augmentation in future work.

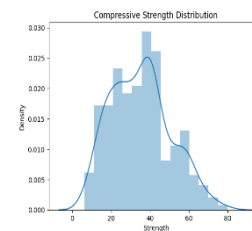


Figure 4. Histogram plot of compressive strength of FA based concrete

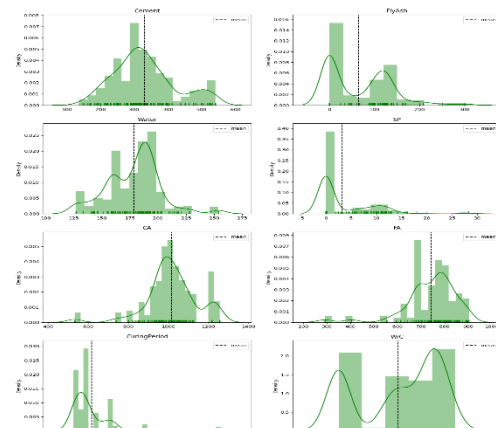


Figure 5. Combined histogram of input variables

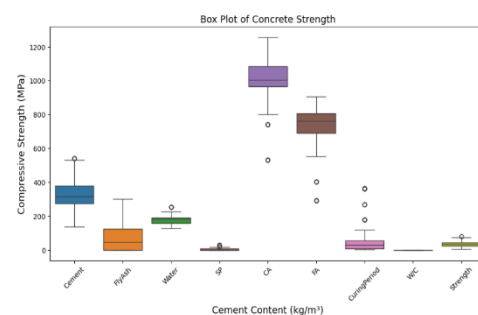


Figure 6. Boxplot of concrete strength of fly ash concrete

3.3 Scatter plots

Scatter plots shown in Figure 7 visualize the relationship between each input variable and CS. CP displays a clear upward trend with CS, confirming its dominant influence on strength development, especially beyond 28 days where fly ash's pozzolanic activity becomes more pronounced. CC shows a strong linear relationship with CS, although the slope tapers off at higher cement contents, suggesting diminishing returns. WC exhibits a negative slope,

consistent with WCR theory. FC presents a mixed pattern: moderate contents (15-30% of cement mass) are associated with competitive or even higher CS values at later ages, whereas very high contents tend to reduce early strength. CA show weaker trends, reinforcing their lower feature importance ranking in the sensitivity analysis. The scatter plots also confirm that the dataset is sufficiently diverse, with no extreme clustering that might bias the models toward a narrow mix composition range.

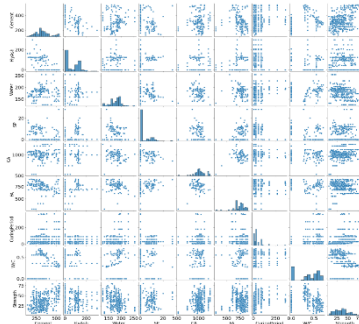


Figure 7. Scatter plots of Input variables and Compressive strength

3.4 Performance of models during testing

The performance of all five ML models considered in this study; as, MLR, SVR, ABR, RF, and XGBR was evaluated using the testing dataset comprising 149 samples (30% of the total data). The scatter plots of predicted versus actual CS for each model are presented in Figure 8. It can be noted that XGBR predictions are highly concentrated along the 1:1 reference line, indicating strong agreement with actual CS values and minimal bias. ABR also shows high predictive accuracy but with slightly more dispersion than XGBR in higher CS ranges. RF predictions cluster reasonably well around the ideal line but exhibit minor underestimation at higher strengths. SVR demonstrates a wider spread, particularly in mid-range CS values (25–45 MPa), leading to reduced R^2 . MLR shows the weakest performance, with a pronounced scatter and systematic underprediction for higher CS values. Quantitatively, the R^2 values obtained during testing (Table 15) confirm that XGBR achieved the highest correlation ($R^2 = 0.881$) and lowest RMSE (5.65 MPa), followed by ABR and RF. The MLR model performed the worst, with an R^2 of 0.620 and RMSE of 9.93 MPa, indicating its limited ability to capture the nonlinear relationships between concrete mix of various parameters considered and CS of FAC.

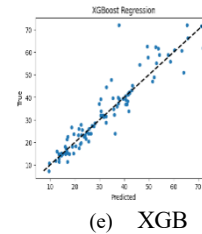
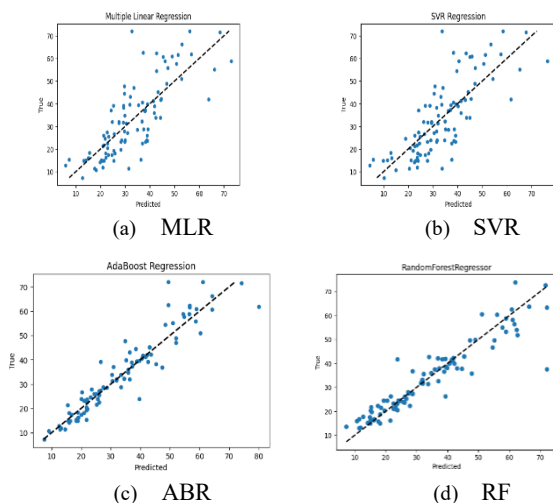
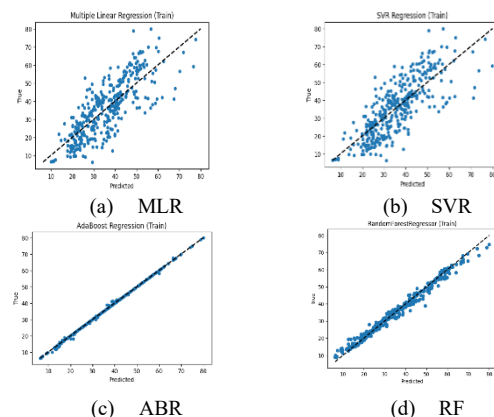


Figure 8. Performance of the models during testing (a) MLR, (b) SVR, (c) AdaBoost Regressor, (d) Random Forest and (e) XGBoost models, respectively

3.5 Performance of models during validation

Validation performance was assessed using the validation subset from the training data (20% of the training set), ensuring that model hyperparameters were optimized without overfitting. Figure 9 illustrates the predicted versus actual CS values during validation. The validation metrics indicate that XGBR maintained its superior performance, with $R^2 \approx 0.88$ and minimal deviation from test results, confirming strong generalization capability. ABR also retained high accuracy with only a marginal increase in RMSE compared to testing, highlighting its robustness to data partitioning. RF showed stable but slightly lower performance in validation, suggesting sensitivity to the random feature selection inherent to the method. SVR displayed noticeable variance between training and validation results, indicating potential sensitivity to kernel parameter selection. MLR again lagged significantly, reinforcing its inadequacy for modelling the complex nonlinear behaviour of FAC strength. The similarity between validation and test set performances for XGBR and ABR suggests that these models are the most reliable for practical applications, as they do not suffer from significant overfitting or underfitting.

Figure 10 compares the RMSE values for the training and testing phases across all ML models considered. The results highlight that XGBR and ABR not only achieved the lowest RMSE in testing but also maintained minimal differences between training and testing RMSE values, indicating strong generalization and absence of significant overfitting. RF also showed relatively balanced train–test RMSE, though slightly higher than the top two models. In contrast, SVR and MLR exhibited larger train–test RMSE gaps, with SVR showing higher testing errors, suggesting model instability and sensitivity to parameter settings, and MLR underperforming consistently due to its inability to capture nonlinear relationships. This analysis further confirms that ensemble-based models, particularly XGBR and ABR, deliver the most stable and accurate predictions for FAC compressive strength.



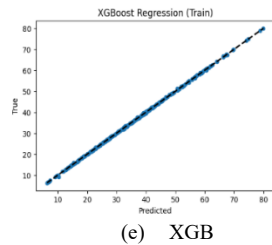


Figure 9. Performance of the models during training (a) MLR, (b) SVR, (c) AdaBoost Regressor, (d) Random Forest and (e) XGBoost models, respectively

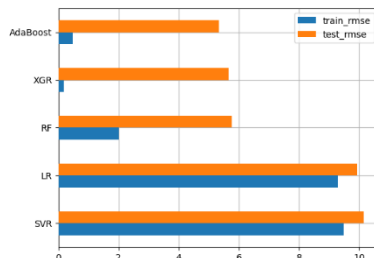


Figure 10. Train and test of RMSE for all ML models

3.6 Sensitivity analysis

For determining the relative importance of input variables sensitivity analysis was conducted for predicting the compressive strength of FAC. Feature importance was calculated using the permutation importance method for RF and XGBR models, which measures the increase in prediction error when the values of a given feature are randomly shuffled. Figure 11 shows the results obtained from the sensitivity analysis. From this it can be noted that CP and CC emerged as the most influential factors, together accounting for over 45% of the model's predictive power. WC ranked third, highlighting the strong influence of water–cement ratio on concrete strength development. FA and SP had moderate influence, with FA's importance aligning with its role as a supplementary cementitious material affecting long-term strength. FAG and CA showed relatively lower importance, indicating that within the tested ranges, their variations had less direct impact on CS compared to binder-related variables. These findings have direct practical implications: optimizing curing duration, cement content, and water proportion could yield the most significant improvements in FAC strength, while changes in aggregate proportions would have comparatively minor effects.

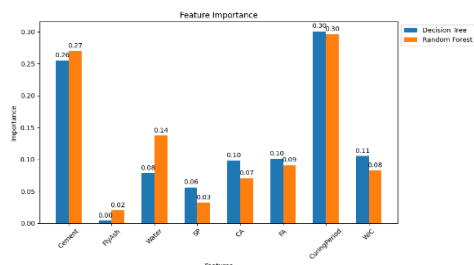


Figure 11. Relative importance of different parameters affecting CS of concrete

3.7 Evaluation metrics

The predictive accuracy of each model was assessed using MAE, MSE, RMSE, and the R^2 , as summarized in Table 12. Among all the algorithms considered in this study, the XGBR model achieved the best overall performance, with an MAE of 3.43

MPa, MSE of 31.96 MPa², RMSE of 5.65 MPa, and R^2 of 0.881. Which indicating its superior capability to capture the complex, nonlinear dependencies between input variables and compressive strength. The ABR model ranked second, with performance metrics very close to XGBR, confirming its suitability for accurate FAC prediction. RF exhibited slightly higher errors (RMSE = 5.97 MPa) but remained competitive, making it a viable option when model interpretability and feature importance are priorities. In contrast, SVR and MLR recorded noticeably higher errors, with RMSE values of 10.15 MPa and 9.93 MPa, respectively, highlighting their limited capacity to handle the nonlinear relationships present in the dataset. The visual comparisons in Figures 13 to 15 support these findings. Figure 13 shows that XGBR and ABR produce the lowest MAE values, indicating minimal average deviation from actual strength values, with RF following closely, whereas SVR and MLR exhibit much larger errors. Figure 14 reveals a similar trend for MSE, with XGBR and ABR effectively minimizing large deviations, RF performing moderately, and SVR and MLR producing substantially higher values. Figure 15 confirms these results in terms of RMSE, where XGBR and ABR again lead, RF occupies the middle ground, and MLR demonstrates the weakest fit to the data. The consistency of these trends across all metrics confirms that ensemble-based approaches, particularly XGBR and ABR, provide the most accurate and robust predictions for FAC compressive strength, followed by RF, while SVR and MLR are less suitable for this application.

Algorithm	MAE (MPa)	MSE (MPa ²)	R ²	RMSE (MPa)
MLR	7.61	98.62	0.620	9.93
SVR (Tuning)	7.76	102.94	0.601	10.15
ABR	3.72	31.69	0.880	5.63
RF	3.76	35.69	0.860	5.97
XGBoost (tuning)	3.43	31.96	0.881	5.65

Table 12. Evaluation metrics for different ML algorithms considered in present study

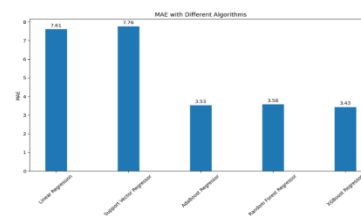


Figure 13. Comparative representation MAE of all adopted ML models in this study

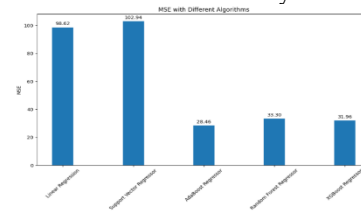


Figure 14. Comparative representation of MSE of all adopted ML models in this study

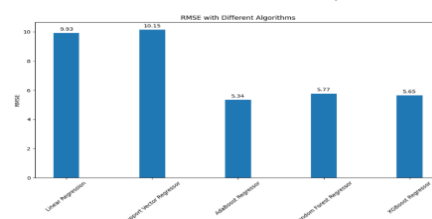


Figure15. Comparative representation RMSE of all the adopted ML models in this study

4. CONCLUSION

This study presented a comprehensive comparative evaluation of five machine learning algorithms; Multiple Linear Regression (MLR), Support Vector Regression (SVR), AdaBoost Regressor (ABR), Random Forest Regressor (RF), and Extreme Gradient Boosting Regressor (XGBR) for predicting the CS of fly ash-based concrete (FAC) using a large, diverse secondary dataset of 498 experimental records collected from published literature. Unlike prior studies that focused on a single algorithm or limited datasets, this work systematically compared multiple modelling approaches under identical conditions and quantified feature importance to provide actionable guidance for mix design optimization. Key findings and contributions are as follows:

- XGBR demonstrated the highest predictive accuracy, achieving an R^2 of 0.881, MAE of 3.43 MPa, and RMSE of 5.65 MPa after hyperparameter tuning, outperforming all other models. This represents an improvement of ~42% in R^2 over the conventional MLR approach.
- Ensemble learning techniques (XGBR, ABR, RF) consistently outperformed both linear regression (MLR) and kernel-based methods (SVR), highlighting their superiority for modelling the nonlinear, multivariate relationships inherent in FAC mix parameters.
- Sensitivity analysis revealed that curing period, cement content, and water content are the most influential parameters for predicting FAC strength, together accounting for over 60% of the total predictive contribution. This insight provides a direct basis for practical mix design adjustments to maximize strength.
- Validation results closely matched testing performance for XGBR and ABR, confirming their robust generalization capability and suitability for real-world predictive applications.
- By integrating feature importance ranking with high-accuracy ML predictions, the study bridges the gap between data-driven modelling and practical engineering decision-making for sustainable concrete production.

The results suggest that practitioners seeking accurate and reliable FAC strength predictions should prioritize XGBR or ABR models, particularly when datasets are moderately sized but feature a wide variability in mix parameters. Additionally, resource allocation in mix optimization should focus on controlling curing period, cement dosage, and water proportion, as these have the most significant impact on performance. Future studies should extend this analysis by incorporating fresh experimental data to further validate the models, exploring deep learning architectures for potential incremental gains in accuracy, and developing user-friendly prediction tools for practitioners in the concrete industry. Furthermore, integrating cost and environmental impact data alongside strength prediction could enable multi-objective optimization for sustainable infrastructure development.

References

Antoni, Lucky Chandra, and Djwantoro Hardjito. 2015. "The Impact of Using Fly Ash, Silica Fume and Calcium Carbonate on the Workability and Compressive Strength of Mortar." In *Procedia Engineering*, doi:10.1016/j.proeng.2015.11.132.

Asteris, Panagiotis G., and Vaseilios G. Mokos. 2020. "Concrete Compressive Strength Using Artificial Neural Networks." *Neural Computing and Applications* 32(15): 11807–26. doi:10.1007/S00521-019-04663-2.

Behnood, Ali, Venous Behnood, Mahsa Modiri Gharehveran, and Kursat Esat Alyamac. 2017. "Prediction of the Compressive Strength of Normal and High-Performance Concretes Using M5P Model Tree Algorithm." *Construction and Building Materials* 142: 199–207.

doi:10.1016/j.conbuildmat.2017.03.061.

Benhelal, Emad, Gholamreza Zahedi, Ezzatollah Shamsaei, and Alireza Bahadori. 2013. "Global Strategies and Potentials to Curb CO₂ Emissions in Cement Industry." *Journal of Cleaner Production*. doi:10.1016/j.jclepro.2012.10.049.

Bildirici, Melike E. 2019. "Cement Production, Environmental Pollution, and Economic Growth: Evidence from China and USA." *Clean Technologies and Environmental Policy*. doi:10.1007/s10098-019-01667-3.

Feng, De Cheng, Zhen Tao Liu, Xiao Dan Wang, Yin Chen, Jia Qi Chang, Dong Fang Wei, and Zhong Ming Jiang. 2020. "Machine Learning-Based Compressive Strength Prediction for Concrete: An Adaptive Boosting Approach." *Construction and Building Materials* 230. doi:10.1016/j.conbuildmat.2019.117000.

Hanifa, Mohd, R. Agarwal, U. Sharma, P. C. Thapliyal, and L. P. Singh. 2023. "A Review on CO₂capture and Sequestration in the Construction Industry: Emerging Approaches and Commercialised Technologies." *Journal of CO₂ Utilization* 67(November 2022): 102292. doi:10.1016/j.jcou.2022.102292.

Khambra, Geetanjli, and Prashant Shukla. 2023. "Novel Machine Learning Applications on Fly Ash Based Concrete: An Overview." *Materials Today: Proceedings*. doi:10.1016/j.matpr.2021.07.262.

Khatib, J. M., and S. Wild. 1998. "Sulphate Resistance of Metakaolin Mortar." *Cement and Concrete Research* 28(1): 83–92. doi:10.1016/S0008-8846(97)00210-X.

Kim, Taehwan, Jeffrey M. Davis, M. Tyler Ley, Shinhyu Kang, and Pouya Amrollahi. 2018. "Fly Ash Particle Characterization for Predicting Concrete Compressive Strength." *Construction and Building Materials* 165: 560–71. doi:10.1016/J.CONBUILDMAT.2018.01.059.

Kumar Mehta, P. 2004. "High-Performance, High-Volume Fly Ash Concrete for Sustainable Development." *International Workshop on Sustainable Development and Concrete Technology*.

Kumar Tipu, Rupesh, V. R. Panchal, and K. S. Pandya. 2022. "An Ensemble Approach to Improve BPNN Model Precision for Predicting Compressive Strength of High-Performance Concrete." *Structures*. doi:10.1016/j.istruc.2022.09.046.

Kumar V., Priyadarshree A., Chandra S., Jindal A. and Rana D. (2023). Behavioral study of raft reinforced with geogrid and geocell through experiments and neural models. *Civil Engineering Infrastructure Journal*, 56(2), 321–332. 10.22059/cej.2023.343951.1847

Lee, C. Y., H. K. Lee, and K. M. Lee. 2003. "Strength and Microstructural Characteristics of Chemically Activated Fly Ash-Cement Systems." *Cement and Concrete Research*. doi:10.1016/S0008-8846(02)00973-0.

Ling, X. H., S. Setunge, and I. Patnaikuni. 2013. "Effect of Different Concentrations of Lime Water on Mechanical Properties of High Volume Fly Ash Concrete." *From Materials to Structures: Advancement Through Innovation - Proceedings of the 22nd Australasian Conference on the Mechanics of Structures and Materials, ACMSM 2012* (May 2016): 1135–41. doi:10.1201/b15320-202.

Ly, Hai-Bang, Thuy-Anh Nguyen, Hai-Van Thi Mai, and Van Quan Tran. 2021. "Development of Deep Neural Network Model to Predict the Compressive Strength of Rubber Concrete." *Construction and Building Materials*. doi:10.1016/j.conbuildmat.2021.124081.

MacLaren, Douglas C., and Mary Anne White. 2003. "Cement: Its Chemistry and Properties." *Journal of Chemical Education*. doi:10.1021/ed080p623.

Mukarjee, Saswatha. 2012. "Study on the Physical and Mechanical Property of Ordinary Portland Cement and Fly Ash Paste." *International Journal of Civil and Structural Engineering* 2(3). doi:10.6088/ijcsesr.00202030003.

Naveen Kumar, C., Kiran V. John, Jagadish Vengala, and R. V. Ranganath. 2006. "Self-Compacting Concrete with Fly Ash and

Metakaolin.” *Indian Concrete Journal*.

Ondova, M., N. Stevulova, and A. Estokova. 2012. “The Study of the Properties of Fly Ash Based Concrete Composites with Various Chemical Admixtures.” In *Procedia Engineering*, doi:10.1016/j.proeng.2012.07.582.

Paudel, Satish, Anil Pudasaini, Rajesh Kumar Shrestha, and Ekta Kharel. 2023. “Compressive Strength of Concrete Material Using Machine Learning Techniques.” *Cleaner Engineering and Technology* 15: 100661. doi:10.1016/J.CLET.2023.100661.

Pitroda, Jayeshkumar, L B Zala, and F S Umrigar. 2012. “Experimental Investigations on Partial Replacement of Cement with Fly Ash in Design Mix Concrete.” *International Journal of Advanced Engineering Technology E* 3(4).

Priyadarshree, Akash, Sunayana Chandra, and Vikas Kumar. 2021. “Performance of Grass Ash with Mix of Black Cotton Soil and Lime.” *Innovative Infrastructure Solutions*. doi:10.1007/s41062-021-00518-9.

Priyadarshree A., Chandra S., Kumar V., Rana D., and Singh N. (2025). Modelling the behavior of urban solid waste ash using artificial neural networks. *Asian Journal of Civil Engineering*, <https://doi.org/10.1007/s42107-025-01472-y>

Rashad, Alaa M. 2015. “A Brief on High-Volume Class F Fly Ash as Cement Replacement – A Guide for Civil Engineer.” *International Journal of Sustainable Built Environment* 4(2). doi:10.1016/j.ijsbe.2015.10.002.

Ren, Ming, Teng Ma, Chen Fang, Xiaorui Liu, Chaoyi Guo, Silu Zhang, Ziqiao Zhou, et al. 2023. “Negative Emission Technology Is Key to Decarbonizing China’s Cement Industry.” *Applied Energy* 329: 120254. doi:10.1016/J.APENERGY.2022.120254.

Sakir, Shamir, Sudharshan N. Raman, Md Safiuddin, A. B.M. Amrul Kaish, and Azrul A. Mutalib. 2020. “Utilization of By-Products and Wastes as Supplementary Cementitious Materials in Structural Mortar for Sustainable Construction.” *Sustainability (Switzerland)*. doi:10.3390/su12093888.

Salami, Babatunde Abiodun, Teslim Olayiwola, Tajudeen A. Oyehan, and Ishaq A. Raji. 2021. “Data-Driven Model for Ternary-Blend Concrete Compressive Strength Prediction Using Machine Learning Approach.” *Construction and Building Materials* 301. doi:10.1016/J.CONBUILDMAT.2021.124152.

Sandhu, Ravinder Kaur, and Rafat Siddique. 2017. “Influence of Rice Husk Ash (RHA) on the Properties of Self-Compacting Concrete: A Review.” *Construction and Building Materials*. doi:10.1016/j.conbuildmat.2017.07.165.

Shaikh, Faiz Uddin Ahmed. 2016. “Mechanical and Durability Properties of Fly Ash Geopolymer Concrete Containing Recycled Coarse Aggregates.” *International Journal of Sustainable Built Environment* 5(2). doi:10.1016/j.ijsbe.2016.05.009.

Singh, Neelam, Akash Priyadarshree, Sunayana Chandra, and Vikas Kumar. 2024. “Development of Prediction Models for Strength Behavior of MSWIA Mixed with Fiber and Cement.” *Multiscale and Multidisciplinary Modeling, Experiments and Design* 8(1): 11. doi:10.1007/s41939-024-00595-0.

Siriruang, Chaichan, Pisanu Toochinda, Parnthep Julnipitawong, and Somnuk Tangtermsirikul. 2016. “CO₂ Capture Using Fly Ash from Coal Fired Power Plant and Applications of CO₂-Captured Fly Ash as a Mineral Admixture for Concrete.” *Journal of Environmental Management* 170. doi:10.1016/j.jenvman.2016.01.010.

Solanki, Jayraj Vinodsinh, and Jayeshkumar Pitroda. 2013. “Investigation of Low Cost Concrete Using Industrial Waste as Supplementary Cementitious Materials.” *International Journal of Engineering Science and Innovative Technology (IJESIT)*.

Song, Hongwei, Ayaz Ahmad, Furqan Farooq, Krzysztof Adam Ostrowski, Mariusz Maślak, Sławomir Czarnecki, and Fahid Aslam. 2021. “Predicting the Compressive Strength of Concrete with Fly Ash Admixture Using Machine Learning Algorithms.” *Construction and Building Materials*. doi:10.1016/j.conbuildmat.2021.125021.

Song, Yang, Jun Zhao, Krzysztof Adam Ostrowski, Muhammad Faisal Javed, Ayaz Ahmad, Muhammad Ijaz Khan, Fahid Aslam, and Roman Kinasz. 2022a. “Prediction of Compressive Strength of Fly-Ash-Based Concrete Using Ensemble and Non-Ensemble Supervised Machine-Learning Approaches.” *Applied Sciences (Switzerland)* 12(1). doi:10.3390/app12010361.

Song, Yang, Jun Zhao, Krzysztof Adam Ostrowski, Muhammad Faisal Javed, Ayaz Ahmad, Muhammad Ijaz Khan, Fahid Aslam, and Roman Kinasz. 2022b. “Prediction of Compressive Strength of Fly-Ash-Based Concrete Using Ensemble and Non-Ensemble Supervised Machine-Learning Approaches.” *Applied Sciences (Switzerland)*. doi:10.3390/app12010361.

Teixeira, E. R., A. Camões, F. G. Branco, J. B. Aguiar, and R. Figueiro. 2019. “Recycling of Biomass and Coal Fly Ash as Cement Replacement Material and Its Effect on Hydration and Carbonation of Concrete.” *Waste Management* 94: 39–48. doi:10.1016/j.wasman.2019.05.044.

Vairagade, V. S., D. K. Parbat, and S. A. Dhale. 2015. “Fly Ash as Sustainable Material for Green Concrete - A State of Art.” *International Journal of research in Engineering, Science and Technology (IJRESTs)* 1(2): 17–24.

Wadhawan, Sifti, Akshita Bassi, Rajwinder Singh, and Mahesh Patel. 2023. “Prediction of Compressive Strength for Fly Ash-Based Concrete: Critical Comparison of Machine Learning Algorithms.” *Journal of Soft Computing in Civil Engineering* 7(3): 68–110. doi:10.22115/SCCE.2023.353183.1493.