Research on Engineering Structures & Materials
www.jresm.org
Research Group

*Research Article*

# Mechanical performance and explainable machine learning analysis of manufactured sand high-performance concrete incorporating mineral admixtures

Basavalingappa [1,a], M.S. Shobha [1,b], Poornima Hulipalled [2,c], Veerabhadrappa Algur [*,3,d]

[1]*Dept. of Civil Eng., Rao Bahadur Y Mahabaleswarappa Engineering College, Ballari, Visvesvaraya Technological University, Belagavi, Karnataka, India*
[2]*Dept. of Computer Applications, Kishkinda University, Ballari, Karnataka, India*
[3]*Dept. of Mechanical Eng., Rao Bahadur Y Mahabaleswarappa Engineering College, Ballari, Visvesvaraya Technological University, Belagavi, Karnataka, India*

**Article Info**

***Article History:***

Received 27 Mar 2026

Accepted 30 Apr 2026

***Keywords:***

Manufactured sand;
High-performance concrete;
Mineral admixtures;
Machine learning;
XGBoost;
Sustainable construction

**Abstract**

This study investigates the mechanical performance of manufactured sand–based high-performance concrete (MSHPC) incorporating supplementary cementitious materials through an integrated experimental and data-driven approach. A total of 180 concrete mixes were prepared by varying manufactured sand replacement (0–100%), mineral admixture type and dosage (fly ash, silica fume, and metakaolin at 0–30%), and water–binder ratio (0.30–0.40). The results indicate that the optimal mix consisting of 60% manufactured sand and 10% metakaolin at a W/B ratio of 0.30 achieved the highest performance, at 28days compressive strength increasing from 68.79 MPa to 91.72MPa, (33.33% improvement), along with significant enhancements in split tensile and flexural strengths. The improved performance is attributed to enhanced particle packing, pore refinement, and interfacial transition zone densification. Mechanical performance decreased with increasing water–binder ratio, confirming the role of matrix densification in strength development. To enable predictive modeling, multiple machine learning models were developed, among which XGBoost demonstrated superior performance with $R^2$ values exceeding 0.99. SHAP-based explainable AI analysis revealed that curing age (for compressive strength), water–binder ratio, and manufactured sand content are the dominant governing parameters, while silica fume and metakaolin exhibited strong positive contributions toward strength enhancement. The proposed framework provides a data-driven and interpretable approach for optimizing sustainable high-performance concrete, reducing dependence on natural sand, and enabling efficient utilization of industrial by-products.



## 1. Introduction

High-performance concrete (HPC) is an advanced class of cementitious material engineered to achieve superior mechanical strength, durability and workability, particularly under aggressive environmental and structural loading conditions. In recent years, the development of HPC has increasingly emphasized sustainability, driven by the urgent need to reduce the environmental footprint associated with cement production and natural resource extraction [1, 2] It is estimated that the manufacture of one metric ton of ordinary Portland cement (OPC) releases approximately one metric ton of carbon dioxide, making the cement industry a major contributor to global greenhouse gas emissions [3] In parallel, excessive extraction of river sand for concrete production

*Corresponding author: veereshalgur@gmail.com
[a]orcid.org/0000-0001-8999-0263; [b]orcid.org/0009-0001-1059-9606; [c]orcid.org/0000-0003-4680-8120; [d]orcid.org/0000-0001-6094-6197
DOI: http://dx.doi.org/10.17515/resm2026-1589me0327rs

has resulted in severe ecological degradation, including riverbank erosion, groundwater depletion and loss of aquatic habitats. Satellite-based studies across major river basins have reported a substantial increase in sand mining activities, highlighting the urgent need for alternative fine aggregates [4].

The incorporation of manufactured sand (M-sand) as a replacement for natural river sand has emerged as a viable and sustainable alternative for concrete production. Owing to its controlled particle size distribution, angular morphology and lower impurity content, M-sand enhances particle packing, reduces void content and improves interfacial transition zone (ITZ) characteristics, thereby contributing to improved mechanical performance and durability of HPC [5–8]. However, the mechanical behavior of M-sand–based concrete is strongly influenced by its interaction with supplementary cementitious materials (SCMs) and the W/B ratio, necessitating systematic optimization of mix proportions.

Mineral admixtures such as fly ash (FA), silica fume (SF), and metakaolin (MK) are widely used as partial replacements for cement to enhance both performance and sustainability. These pozzolanic materials refine the pore structure, promote the formation of calcium silicate hydrate (C–S–H) and calcium–aluminosilicate hydrate (C–A–S–H) gels, and densify the cement matrix, resulting in improved compressive, tensile, and flexural strength, as well as enhanced long-term durability [9–14] Binary and ternary blended systems incorporating OPC with SCMs significantly reduce permeability and improve resistance to chemical and mechanical degradation. Among these materials, metakaolin exhibits superior pozzolanic reactivity and micro-filler effects, particularly in high-performance and high-strength concrete systems [15–19].

Despite the growing body of research on HPC incorporating SCMs and M-sand, most studies rely on empirical or regression-based approaches, which are often inadequate for capturing the complex nonlinear interactions among material composition, curing conditions and mechanical performance. Recently, data science and ML techniques have gained considerable attention in civil engineering for their ability to model nonlinear relationships and provide accurate prediction of concrete strength and durability characteristics [20–25]. Ensemble-based models such as Random Forest (RF), Gradient Boosting (GB) and Extreme XGBoost have demonstrated superior predictive capability compared to conventional regression and artificial neural network (ANN) models in modeling sustainable and high-performance concretes [26–38]. However, several critical gaps remain. Most previous studies focus primarily on predicting compressive strength, while limited attention has been given to the simultaneous prediction of compressive, split tensile and flexural strengths within a unified framework. Moreover, existing investigations rarely integrate explainable artificial intelligence (XAI) to establish a physics-consistent interpretation of model predictions. In particular, the combined influence of curing kinetics, water–binder ratio, mineral admixture chemistry, and M-sand replacement on strength evolution has not been systematically quantified. This gap is addressed in the present study using SHAP-based feature attribution. Furthermore, the role of M-sand in optimizing multi-source high-performance concrete has not been explored through an integrated ML–driven optimization and interpretability framework.

Unlike conventional studies that primarily focus on empirical prediction or algorithmic comparison, the present work provides an experimental and explainable machine learning framework to interpret the microstructure–strength relationship in manufactured sand high-performance concrete. The study reveals a packing-density-controlled strength threshold associated with manufactured sand replacement. It establishes the role of binder chemistry and pozzolanic reactions in governing matrix densification and ITZ refinement. In addition, explainable machine learning is used to quantify the combined influence of curing kinetics, water–binder ratio, mineral admixture chemistry, and aggregate morphology on strength evolution. This combined experimental and data-driven approach with interpretable insights provides deeper material understanding while enabling reliable prediction and optimization of multi-property performance in sustainable high-performance concrete. Accordingly, a comprehensive experimental and machine learning framework is developed to investigate and optimize the mechanical performance of MSHPC. A large experimental dataset is generated by systematically varying manufactured sand replacement level, SCM type and dosage, water–binder ratio, and curing age. Multiple machine

learning models, including Artificial Neural Network (ANN), Random Forest (RF), Adaptive Boosting (AdaBoost), Gradient Boosting (GB), and Extreme Gradient Boosting (XGBoost), are trained and evaluated in terms of predictive accuracy, robustness, and generalization. Furthermore, SHapley Additive exPlanations (SHAP)-based explainable AI is employed to quantify feature influence and establish a data-driven linkage governing strength evolution. The proposed framework enables comprehensive prediction framework for multiple properties and interpretation of compressive, split tensile, and flexural strengths, providing a reliable and interpretable tool for sustainable concrete mix optimization while reducing experimental effort and supporting environmentally responsible construction practices.

## 2. Experimental Program

### 2.1 Materials Used

Ordinary Portland Cement (OPC) of 43 grade, conforming to IS 8112:1989, was used as the primary binder in all concrete mixes. The supplementary cementitious materials (SCMs) employed in this study included fly ash (FA), silica fume (SF), and metakaolin (MK). Fly ash was procured from a local thermal power plant and complied with the requirements of IS 3812 (Part 1) for use as a pozzolanic material in concrete. Silica fume was obtained from a commercial supplier and conformed to ASTM C1240, while metakaolin was sourced from a certified manufacturer and complies with the requirements of ASTM C618 (Class N) for natural pozzolanic materials, ensuring its suitability for use in high-performance concrete applications. The physical and chemical properties of the cement and SCMs are presented in Table 1. The material properties reported in Table 1 are based on supplier-provided test certificates and standard characterization data corresponding to the specific batches used in this study. Fig. 1 presents the raw materials used in the experimental program, including OPC, coarse aggregate, natural sand, manufactured sand, fly ash, silica fume, metakaolin, and the high-range water-reducing admixture (superplasticizer SP-430). These materials were selected to ensure uniformity in mix preparation and to achieve the required workability, strength, and durability characteristics of the concrete. Natural river sand was used as the reference fine aggregate, while manufactured sand (M-sand) obtained from a nearby quarry-crushing facility served as the alternative fine aggregate. Both fine aggregates were tested and classified in accordance with IS 383:2016, and their physical properties, including specific gravity, fineness modulus, bulk density, and grading zone, are presented in Table 2. Crushed granite coarse aggregate with a nominal maximum size of 20 mm was used in all mixes and conformed to the specifications of IS 383:2016.

Table 1. Physical and chemical properties of cementitious materials

| Property | Cement | Fly Ash (FA) | Metakaolin (MK) | Silica Fume (SF) |
|---|---|---|---|---|
| Specific gravity | 3.08 | 2.17 | 2.61 | 2.21 |
| Mean particle size (μm) | 8–210 | 20–25 | 2.53 | 1.14 |
| Specific surface area ($cm^2/g$) | 2,947 | 3,988 | 17,100 | 22,000 |
| Color | Grey | Tan to dark grey | Ivory to cream | Light to dark grey |
| Chemical Composition (%) | | | | |
| Oxide | FA | MK | SF | IS 3812 Requirement |
| $SiO_2$ | 59.16 | 62.3 | 91.40 | ($SiO_2$ + $Al_2O_3$ + $Fe_2O_3$) ≥ 70 |
| $Al_2O_3$ | 30.64 | 43.18 | 1.30 | — |
| $Fe_2O_3$ | 4.70 | 0.60 | 1.00 | — |
| CaO | 2.85 | 1.06 | 1.60 | — |
| MgO | 0.36 | 0.61 | 0.90 | ≤ 5.0 |
| Loss on ignition | 0.21 | 0.70 | 2.70 | ≤ 6.0 |

Potable water, free from deleterious substances, was used for mixing and curing of all specimens. A high-range water-reducing admixture (superplasticizer) based on sulfonated naphthalene formaldehyde (SNF), commercially designated as SP-430, was incorporated to achieve the required workability at low water–binder ratios. The dosage of the superplasticizer was maintained at 0.8% by weight of the total binder content for all concrete mixes.

Table 2. Physical properties of fine and coarse aggregates

| Property | Natural Sand | Manufactured Sand (M-sand) | Coarse Aggregate |
|---|---|---|---|
| Specific gravity | 2.50 | 2.56 | 2.70 |
| Fineness modulus | 2.76 | 2.87 | 6.71 |
| Bulk density (kN/m$^3$) | | | |
| – Loose | 14.70 | 15.12 | 13.29 |
| – Compacted | 16.18 | 16.58 | 15.12 |
| Grading zone | Zone II | Zone II | — |


Fig. 1. Raw materials used in the experimental program

## 2.2 Mix Proportioning

In this experiment, mix proportions were assessed using an absolute volume approach. Ingredients were measured by volume and then converted to weight to ensure a uniform mixture. All ingredients were dry-mixed initially, with mineral admixtures in this phase. Water and a superplasticizer were then introduced to achieve a homogeneous concrete mix. The Aggregate binder ratio was kept constant i.e. (A/B=2) for all mixes. OPC was partially replaced by SCMs, MK, FA and SF, at replacement levels of 0%, 10%, 20% and 30% by mass of the total binder. Three W/B ratios (0.30, 0.35 and 0.40) were selected to evaluate the influence of water content on workability and mechanical performance under high-performance concrete conditions. Natural river sand was replaced with M-sand at proportions of 0%, 20%, 40%, 60%, 80% and 100% by mass of fine aggregate. To achieve the desired workability at low W/B ratios, a high-range water-reducing admixture (superplasticizer, SP-430) was incorporated at a constant dosage of 0.8% by weight of the total binder content. A total of 360 distinct concrete mixes were prepared by systematically combining the levels of SCM replacement, W/B ratio, and M-sand substitution.

The first letter in the mix designation indicates a percentage of manufactured sand, i.e. M0 and M20, M40, M60, M80 and M100 indicating that 0%,20%,40%,60%,80% and 100% of manufactured sand. The second letter indicates the percentage of mineral admixtures like MA0 indicating that without mineral admixtures and F10, F20 and F30 indicate that fly ash content is 10%, 20%, and 30%, S10, S20 and S30 indicate that Silica fume content is 10%, 20%, and 30% and M10, M20 and M30 indicate that Metakaolin content is 10%, 20%, and 30%. The last alphabet indicates water binder ratios, i.e. A=0.30, B=0.35, C=0.40, for example M0MA0A indicates a plain high-performance concrete mix 0% of M-Sand i.e. 100% natural sand and 0% mineral admixture i.e. 100% cement for 0.3 W/B. The nomenclature and compositions of various mixes are detailed in Table 3, maintaining consistent design principles across different W/B ratios to ensure a comprehensive assessment.

With absolute volume method quantity of materials for plain mix i.e. M0MA0A, cement=719.88kg/m$^3$, natural-sand =575.90kg/m$^3$, coarse aggregate =863.86kg/m$^3$ with W/B ratio=0.3, Another mix, i.e.M20M10A cement=676.13kg/m$^3$, metakaolin=75.13kg/m$^3$, natural-sand=480.80kg/m$^3$, M-Sand=120.20kg/m$^3$, coarse aggregate =901.50kg/m$^3$ with W/B ratio=0.3. The observed slump values for the plain concrete mixes M0MA0A, M0MA0B, and M0MA0C were 62 mm, 85 mm, and 105 mm, respectively, obtained using a constant superplasticizer dosage of 0.8% by weight of total binder content.

Table 3 Mix proportions of MSHPC mix for W/B= 0.30

| Sl No | Mix Description | W/B | A/B | Metakaolin (%) | Cement (%) | M- Sand (%) | Natural Sand (%) |
|---|---|---|---|---|---|---|---|
| 1 | M0MA0A | 0.3 | 2 | 0 | 100 | 0 | 100 |
| 2 | M20MA0A | 0.3 | 2 | 0 | 100 | 20 | 80 |
| 3 | M40MA0A | 0.3 | 2 | 0 | 100 | 40 | 60 |
| 4 | M60MA0A | 0.3 | 2 | 0 | 100 | 60 | 40 |
| 5 | M80MA0A | 0.3 | 2 | 0 | 100 | 80 | 20 |
| 6 | M100MA0A | 0.3 | 2 | 0 | 100 | 100 | 0 |
| 7 | M0M10A | 0.3 | 2 | 10 | 90 | 0 | 100 |
| 8 | M20M10A | 0.3 | 2 | 10 | 90 | 20 | 80 |
| 9 | M40M10A | 0.3 | 2 | 10 | 90 | 40 | 60 |
| 10 | M60M10A | 0.3 | 2 | 10 | 90 | 60 | 40 |
| 11 | M80M10A | 0.3 | 2 | 10 | 90 | 80 | 20 |
| 12 | M100M10A | 0.3 | 2 | 10 | 90 | 100 | 0 |
| 13 | M0M20A | 0.3 | 2 | 20 | 80 | 0 | 100 |
| 14 | M20M20A | 0.3 | 2 | 20 | 80 | 20 | 80 |
| 15 | M40M20A | 0.3 | 2 | 20 | 80 | 40 | 60 |
| 16 | M60M20A | 0.3 | 2 | 20 | 80 | 60 | 40 |
| 17 | M80M20A | 0.3 | 2 | 20 | 80 | 80 | 20 |
| 18 | M100M20A | 0.3 | 2 | 20 | 80 | 100 | 0 |
| 19 | M0M30A | 0.3 | 2 | 30 | 70 | 0 | 100 |
| 20 | M20M30A | 0.3 | 2 | 30 | 70 | 20 | 80 |
| 21 | M40M30A | 0.3 | 2 | 30 | 70 | 40 | 60 |
| 22 | M60M30A | 0.3 | 2 | 30 | 70 | 60 | 40 |
| 23 | M80M30A | 0.3 | 2 | 30 | 70 | 80 | 20 |
| 24 | M100M30A | 0.3 | 2 | 30 | 70 | 100 | 0 |

**Similarly designs of constituents are used for other W/B ratios W/B= 0.35 (B) and 0.40 (C) for Metakaolin, *Remaining designs of constituents are used for other W/B ratios W/B= 0.30 (A), 0.35 (B) and 0.40 (C) Silica fume and Fly Ash.*

## 2.3 Experimental Methods

Concrete cube specimens of size 100 × 100 × 100 mm were cast to determine compressive strength. Cylindrical specimens of 150 mm diameter and 300 mm height were prepared to evaluate split tensile strength, and beam specimens of 500 × 100 × 100 mm were cast to measure flexural strength. After demolding, all specimens were cured under standard water-curing conditions and tested at both 7 and 28 days.

All mechanical tests were conducted in accordance with relevant Indian Standard (IS) specifications. Compressive strength tests were carried out using a 2000 kN compression testing machine in accordance with IS 516:2018 at a constant loading rate of 14 N/mm$^2$/min. Split tensile strength tests were conducted using the same 2000 kN compression testing machine as per IS 5816:1999, with a loading rate in the range of 1.2–2.4 N/mm$^2$/min. Flexural strength tests were also performed using the 1000 kN universal testing machine (UTM) in accordance with IS 516:2018 at a constant loading rate of 0.7 N/mm$^2$/min. For each mix combination, three specimens were tested, and the reported values represent the average of three measurements. The standard deviation among replicate specimens was generally less than 2% of the mean values; therefore, error bars were omitted from the plots for clarity.

## 2.4 Experimental Program and Data Preparation

The experimental dataset used in this study was obtained from a systematic laboratory investigation on MSHPC. The input variables considered for ML modeling included the M-sand replacement level, type and dosage of mineral admixture (SF, MK and FA), W/B ratio and curing age. The output variables comprised compressive strength, split tensile strength and flexural strength. The dataset used in this study was constructed based on experimentally measured values for each mechanical property. The datasets for compressive strength, split tensile strength, and flexural strength each consist of 360 data instances corresponding to both 7-day and 28-day curing periods. Each mechanical property was modeled independently using its corresponding dataset to ensure accurate prediction of individual strength characteristics. No artificial data generation, duplication, or imputation was performed, and only experimentally measured values were used for model development. Prior to model development, statistical analysis of the input features was conducted to examine data distribution, skewness and inter-feature correlation. Pearson correlation analysis and variance inflation factor (VIF) were employed to assess potential multicollinearity among the input variables. Outliers were identified using the interquartile range (IQR) method and were retained due to their physical relevance in representing extreme but realistic mix design conditions.

Categorical variables, specifically the type of mineral admixture (FA, SF, MK), were converted into numerical form using One-Hot Encoding prior to model training. This approach ensures that each admixture type is represented as an independent binary feature (FA, SF, MK) without imposing any ordinal relationship between categories. Subsequently, feature scaling was performed using Min–Max normalization to map all numerical input variables into the [0,1] range, thereby improving numerical stability and convergence of the learning algorithms. The One-Hot Encoded categorical variables were not subjected to scaling, as they are already represented in binary form. For model development and validation, the dataset was randomly divided into training (288 samples, 80%) and testing (72 samples, 20%) subsets. The training dataset was used for model learning and hyperparameter optimization, while the testing dataset was exclusively used to evaluate the generalization capability and predictive performance of the developed models.

## 2.5 Models Used for ML

In this study, separate machine learning models were developed independently for each mechanical property, namely compressive strength, split tensile strength, and flexural strength. Each model was trained using its corresponding dataset to ensure accurate prediction of individual properties. Therefore, the implemented models follow a single-output regression approach rather than a multi-output framework. For the Artificial Neural Network (ANN), the output layer consists of a single neuron corresponding to the respective target variable. To develop a robust and comparative predictive framework, six ML models were implemented: ANN, Nonlinear Regression (NLR), RF, Adaptive Boosting (AdaB), GB and XGBoost. These models were selected based on their demonstrated effectiveness in capturing nonlinear relationships and complex feature interactions in cementitious and sustainable concrete systems, as reported in recent literature.

### *2.5.1 Artificial Neural Network*

The ANN model was implemented using a multilayer perceptron (MLP) architecture comprising an input layer, one or more hidden layers, and an output layer. Nonlinear activation functions were employed in the hidden layers to capture complex interactions among the mix design parameters. Although ANN models have been widely applied for predicting concrete strength, their performance is sensitive to network topology, learning rate and training configuration, necessitating careful hyperparameter tuning to achieve stable and reliable predictions.

### *2.5.2 Nonlinear Regression*

Nonlinear regression was adopted as a baseline model to establish the relationship between the input variables and the mechanical properties of MSHPC. Polynomial terms and feature interaction terms were incorporated to improve fitting capability. While computationally efficient and

interpretable, NLR exhibits limited capacity to model highly complex and nonlinear feature interactions when compared with ensemble-based learning methods.

#### *2.5.3 Random Forest*

RF is an ensemble learning technique based on the construction of multiple decision trees using bootstrap sampling and random feature selection. The final prediction is obtained by averaging the outputs of the individual trees. This approach enhances generalization performance by reducing variance and mitigating overfitting, making RF well-suited for modeling nonlinear and high-dimensional relationships in concrete property prediction.

#### *2.5.4 Adaptive Boosting*

AdaB is a boosting-based ensemble algorithm that sequentially trains weak learners by assigning greater weights to previously mis predicted samples. The final prediction is generated through a weighted aggregation of the individual learners. While AdaB can significantly improve prediction accuracy, it may exhibit sensitivity to noise and outliers, particularly in experimental datasets with extreme mix design conditions.

#### *2.5.5 Gradient Boosting*

GB constructs predictive models in a stage-wise manner by iteratively minimizing a specified loss function using gradient descent principles. Each successive model focuses on correcting the residual errors of the preceding ensemble, enabling effective learning of complex nonlinear patterns and variable interactions. GB has demonstrated strong performance in modeling mechanical properties of sustainable and high-performance concretes in recent studies.

#### *2.5.6 Extreme Gradient Boosting*

XGBoost is an advanced and optimized implementation of GB that incorporates L1 and L2 regularization, parallel processing and efficient tree pruning strategies. These features improve computational efficiency, enhance predictive accuracy and reduce the risk of overfitting. Owing to its robustness and scalability, XGBoost has gained widespread adoption in civil engineering and materials informatics applications for data-driven performance prediction and optimization.

### 2.6 Model Training, Hyperparameter Selection and Performance Evaluation

All ML models were trained using the same dataset to ensure a fair and unbiased comparison. To enhance robustness and minimize sampling bias, a 5-fold cross-validation strategy was employed during the training phase. The dataset was randomly partitioned into five mutually exclusive subsets, with four folds used for training and one-fold for validation in each iteration. The reported performance metrics represent the average values across all folds, thereby ensuring model stability and reliable generalization capability.

Hyperparameter optimization was performed through iterative, cross-validation–guided tuning, with the objective of minimizing the root mean squared error on the validation folds. The final hyperparameter configurations for each model are summarized in Table 4. To ensure reproducibility, a fixed random state (random_state = 42) was used for dataset splitting, cross-validation, and model training.

Table 4. Hyperparameter configuration of implemented machine learning models

| Model | Hyperparameters (exact values) |
|---|---|
| ANN | Hidden layers = 2; Neurons per layer = 20 (20–20), Activation = ReLU; Learning rate = 0.001; Optimizer = Adam; Epochs = 800 |
| NLR | Polynomial degree = 2; Interaction terms = included |
| RF | Number of trees = 200; Max depth = 20; Min samples/leaf = 2 |
| AdaB | Estimators = 100; Learning rate = 0.1; Base tree depth = 2 |
| GB | Boosting stages = 200; Learning rate = 0.05; Max depth = 4; Subsample = 0.8 |
| XGBoost | Trees = 300; Learning rate = 0.05; Max depth = 6; Subsample = 0.8; Col sample = 0.8; L1 = 0.0; L2 = 1.0 |

Model performance was evaluated using the coefficient of determination ($R^2$), mean absolute error (MAE), mean squared error (MSE) and root mean squared error (RMSE) for both training and testing datasets. This multi-metric evaluation framework ensures a comprehensive assessment of prediction accuracy, robustness and generalization capability for estimating the mechanical properties of MSHPC.

Model robustness was further evaluated using prediction interval analysis. The uncertainty associated with model prediction uncertainty was quantified using a 90% confidence interval computed through a bootstrapping approach with 1000 resamples. This analysis provides insight into the reliability of strength predictions for previously unseen mix designs and supports the practical deployment of the proposed framework in engineering decision-making and concrete mix optimization.

### 2.7 Model Explainability and Feature Importance

To enhance the interpretability of the proposed ML framework, feature importance analysis was conducted using SHAP. This approach quantifies the contribution of each input variable to the model output by computing marginal contributions across all possible feature coalitions. Global SHAP summary plots were generated to identify the dominant parameters influencing compressive, split tensile and flexural strength. To enhance model interpretability, SHAP-based analysis was employed, including global feature importance, summary plots, and local force plots, to understand the influence of input variables on predicted mechanical properties. This explainability framework enhances model transparency and supports engineering interpretation of the data-driven predictions.

## 3. Results and Discussion

This section presents the experimental performance of MSHPC and establishes the mechanistic basis for the subsequent ML modeling and optimization framework. The influence of W/B ratio, SCM type and dosage and M-sand replacement level on compressive, split tensile and flexural strength is systematically analyzed.

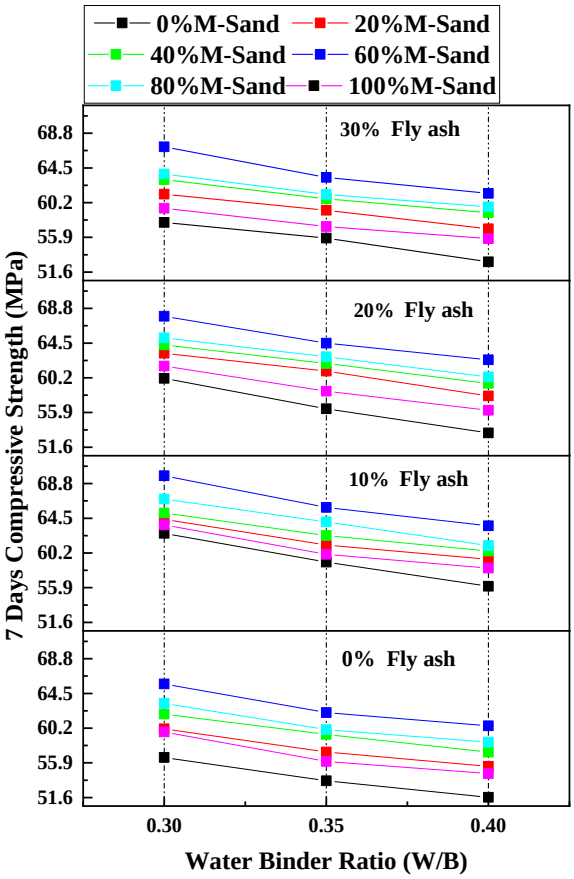
### 3.1 Experimental Evaluation of Mechanical Properties

#### *3.1.1 Effect of Water–Binder Ratio (FA vs. SF vs. MK)*

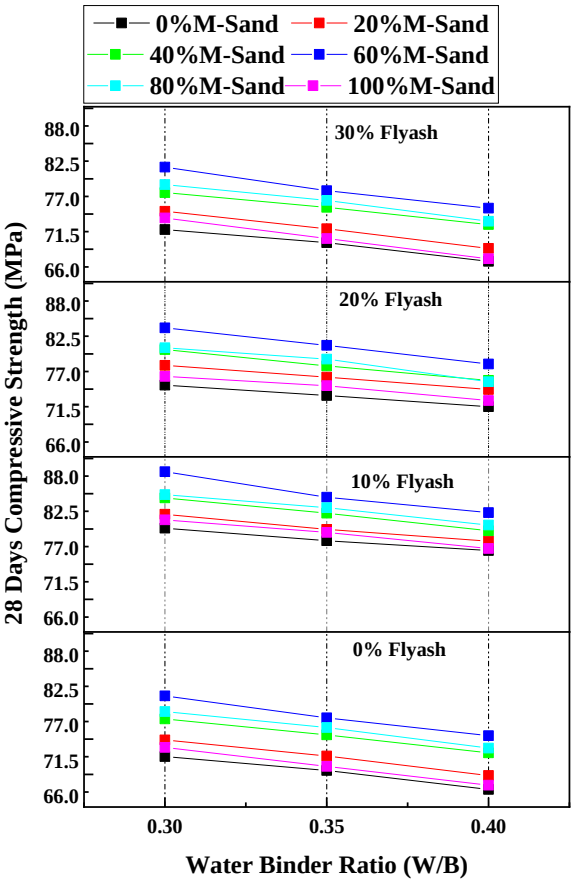
The combined influence of the W/B ratio, manufactured sand (M-sand) replacement level, and SCM type on the 7-day and 28-day compressive strength of MSHPC is illustrated in Fig. 2. Across all six subplots, a consistent and monotonic reduction in compressive strength is observed as the W/B ratio increases from 0.30 to 0.40, irrespective of SCM type and M-sand content. This behavior reflects the fundamental role of water content in governing capillary porosity and paste densification, wherein higher W/B ratios promote the formation of continuous pore networks that weaken the load-bearing cementitious matrix.

For FA-based systems (Fig. 2a–b), the strength hierarchy among M-sand replacement levels remains stable across curing ages, with the 60% M-sand mixes consistently exhibiting the highest compressive strength. At W/B = 0.30, the enhanced performance is primarily attributed to improved particle packing and stronger interfacial bonding between the aggregate and paste phases. The observed increase in strength from 7 to 28 days further reflects the delayed pozzolanic reaction of fly ash, which gradually consumes calcium hydroxide and contributes to secondary C–S–H gel formation, thereby densifying the microstructure.

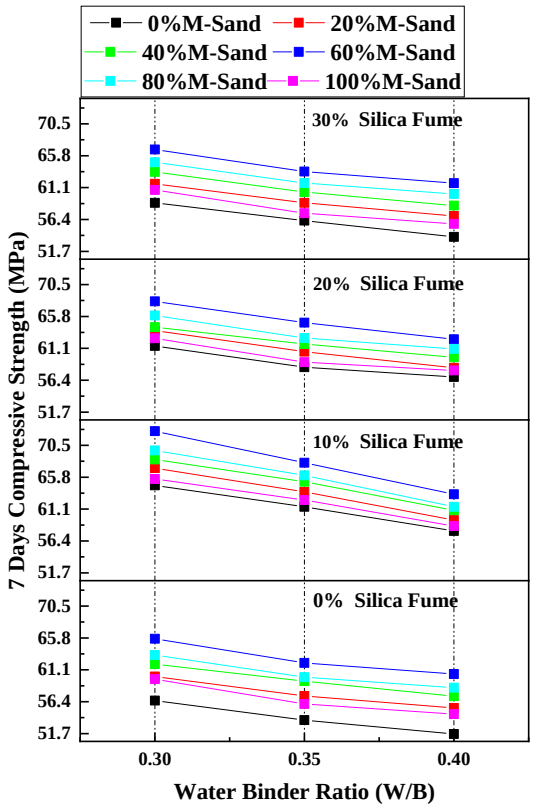
In SF-based mixes (Fig. 2c–d), the compressive strength at early age is comparatively higher than that of FA systems, particularly at lower W/B ratios. This enhancement is associated with the ultrafine particle size and high specific surface area of silica fume, which accelerates nucleation of hydration products and promotes rapid pore refinement. The beneficial role of 60% M-sand replacement is again evident, indicating a synergistic interaction between optimized fine aggregate grading and matrix densification.


0%M-Sand
20%M-Sand
40%M-Sand
60%M-Sand
80%M-Sand
100%M-Sand
30% Fly ash
20% Fly ash
10% Fly ash
0% Fly ash
7 Days Compressive Strength (MPa)
Water Binder Ratio (W/B)

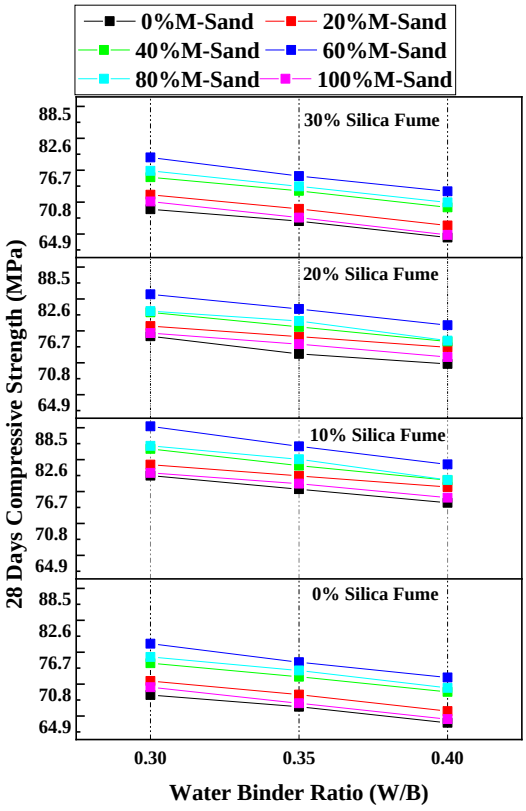
(a)


0%M-Sand
20%M-Sand
40%M-Sand
60%M-Sand
80%M-Sand
100%M-Sand
30% Flyash
20% Flyash
10% Flyash
0% Flyash
28 Days Compressive Strength (MPa)
Water Binder Ratio (W/B)

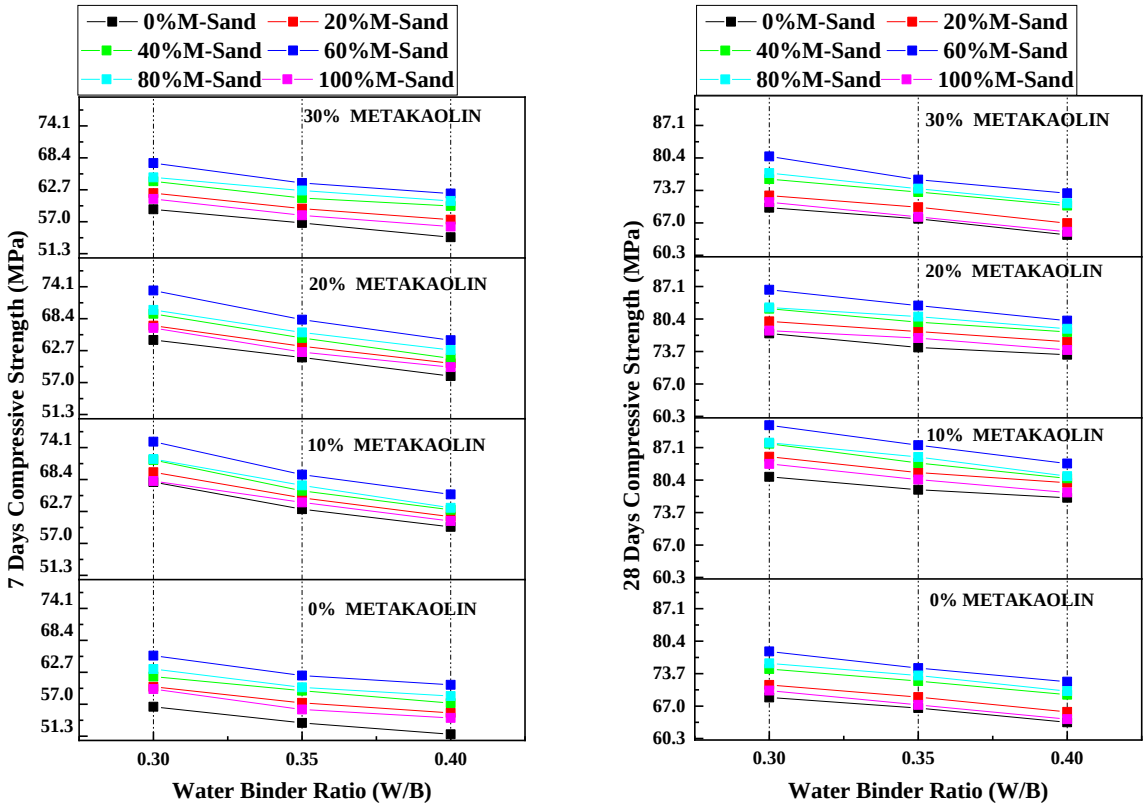
(b)


0%M-Sand
20%M-Sand
40%M-Sand
60%M-Sand
80%M-Sand
100%M-Sand
30% Silica Fume
20% Silica Fume
10% Silica Fume
0% Silica Fume
7 Days Compressive Strength (MPa)
Water Binder Ratio (W/B)

(c)


0%M-Sand
20%M-Sand
40%M-Sand
60%M-Sand
80%M-Sand
100%M-Sand
30% Silica Fume
20% Silica Fume
10% Silica Fume
0% Silica Fume
28 Days Compressive Strength (MPa)
Water Binder Ratio (W/B)

(d)




(e) (f)

Fig. 2. Compressive strength vs W/B Ratio: (a) FA: 7-day CS vs W/B, (b) FA: 28-day CS vs W/B, (c) SF: 7-day CS vs W/B, (d) SF: 28-day CS vs W/B, (e) MK: 7-day CS vs W/B, and (f) MK: 28-day CS vs W/B.

The MK-based MSHPC mixes (Fig. 2e–f) exhibit the highest compressive strength across all W/B ratios and curing ages, underscoring the superior reactivity of metakaolin as a highly pozzolanic and alumina-rich SCM. The combination of 10% metakaolin substitution and 60% M-sand replacement at W/B = 0.30 yields the most pronounced strength enhancement. The compressive strength increased from 68.79 MPa (i.e., M0MA0A) for the control mix to 91.72 MPa for the optimal mix (i.e., M60M10A), corresponding to a 33.33% improvement for 28 days. This improvement is attributed to accelerated formation of C–S–H and C–A–S–H gels, along with enhanced matrix compactness and improved interfacial characteristics, leading to efficient load transfer within the composite.

The trends in Fig. 2 are consistent with prior studies on sustainable high-performance concretes incorporating manufactured sand and supplementary cementitious materials. The reduction in compressive strength with increasing W/B ratio is attributed to increased capillary porosity and weakened paste–aggregate bonding [39, 40]. The optimal performance at approximately 60% M-sand replacement agrees with findings by Vardhan et al. [41]. The superior performance of silica fume and metakaolin-based systems is consistent with their well-known pozzolanic and micro filler effects [42–44]with similar threshold behavior beyond 10% SCM substitution [45–48].

### *3.1.2 Effect of SCM Replacement Level*

Fig.3 presents the variation in 7-day and 28-day compressive strength of MSHPC as a function of SCM replacement level at a constant W/B ratio of 0.30. Across all SCM types and M-sand replacement levels, compressive strength increases as the SCM content rises from 0% to 10%, followed by a progressive decline at higher substitution levels (20–30%), indicating a distinct optimum dosage. For FA–based mixes, the initial improvement is attributed to the micro filler effect and delayed pozzolanic reaction, which enhances particle packing and promotes secondary C–S–H gel formation at later curing ages. The reduction beyond 10% replacement reflects a dilution effect,

where reduced clinker content limits early hydration and binder continuity [49, 50]. In SF–based systems, a more pronounced early-age strength gain is observed due to the ultrafine particle size and high specific surface area of silica fume, which accelerates hydration nucleation and refines pore structure. The decline at higher replacement levels is associated with increased water demand and reduced workability, which may lead to incomplete compaction and microstructural discontinuities [42–44].

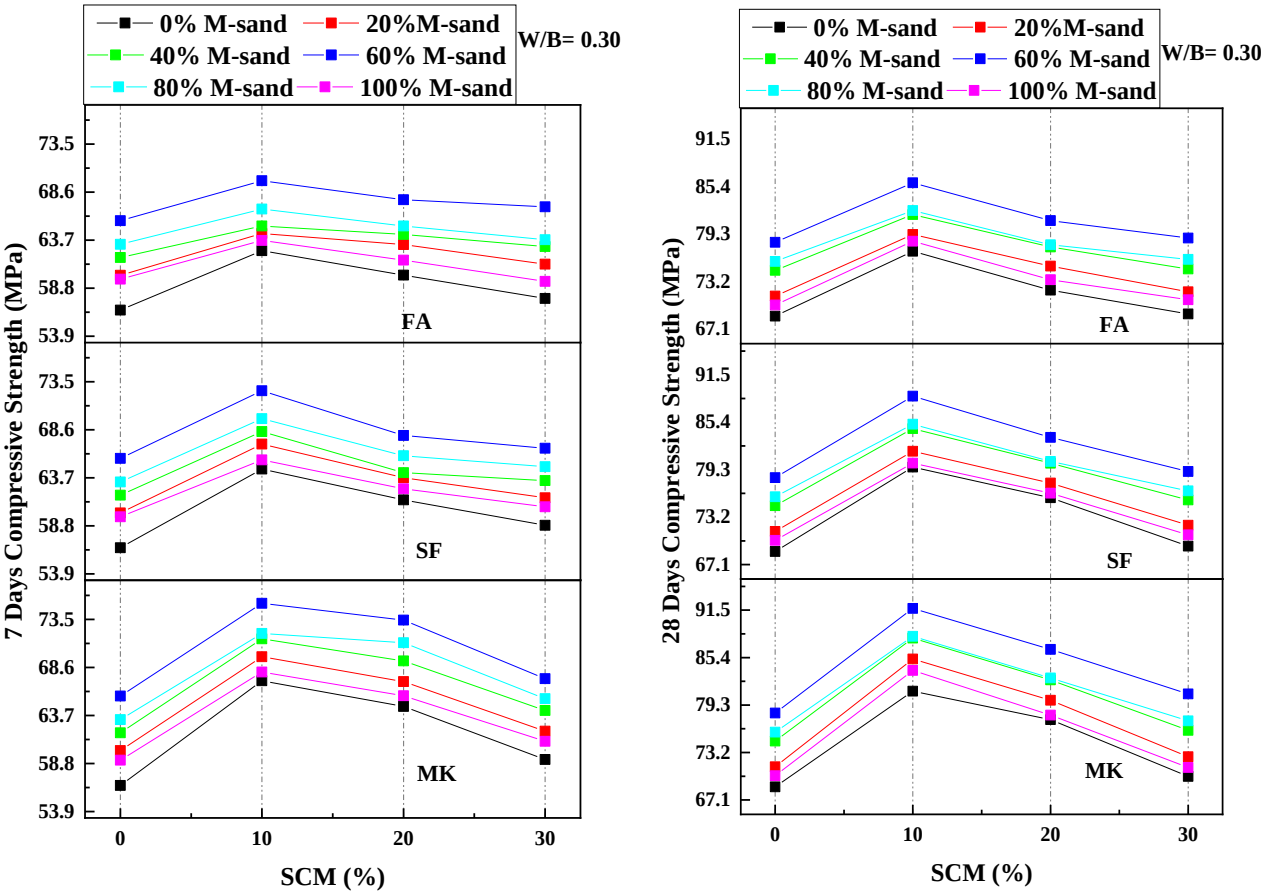



Fig. 3. Compressive strength Vs. % of mineral admixture (a) 7-days (b) 28-days

The MK-based mixes consistently exhibit the highest compressive strength at both curing ages, confirming the superior pozzolanic reactivity of metakaolin. The optimal performance at approximately 10% substitution is attributed to the formation of additional C–S–H and C–A–S–H phases and improved ITZ quality. At higher replacement levels, insufficient $Ca(OH)_2$ availability limits further pozzolanic reaction, resulting in reduced matrix densification reported[44–47]

### *3.1.3 Effect of Manufactured Sand Replacement Level*

Fig. 4 illustrates the influence of M-sand replacement level (0–100%) on the 7-day and 28-day compressive strength of FA, SF, and MK-based MSHPC mixes at a constant W/B ratio of 0.30. For all SCM systems, compressive strength increases progressively as the M-sand content rises from 0% to approximately 60%, beyond which a compressive strength gradual decline is observed.

The improvement in strength up to 60% M-sand replacement can be attributed to enhanced packing efficiency and improved aggregate–paste interaction, resulting in a denser cementitious matrix and more effective load transfer. Similar trends have been reported by[41], who demonstrated that partial substitution of natural sand with M-sand enhances compressive strength and microstructural homogeneity. The reduction in strength beyond the optimal replacement level is primarily associated with increased surface area and water demand of M-sand-rich mixes, which adversely affect workability and compaction. This leads to higher void content and reduced matrix

continuity, ultimately lowering strength. This behavior is consistent with observations reported in studies on fine aggregate grading optimization in high-performance concrete [41,42].

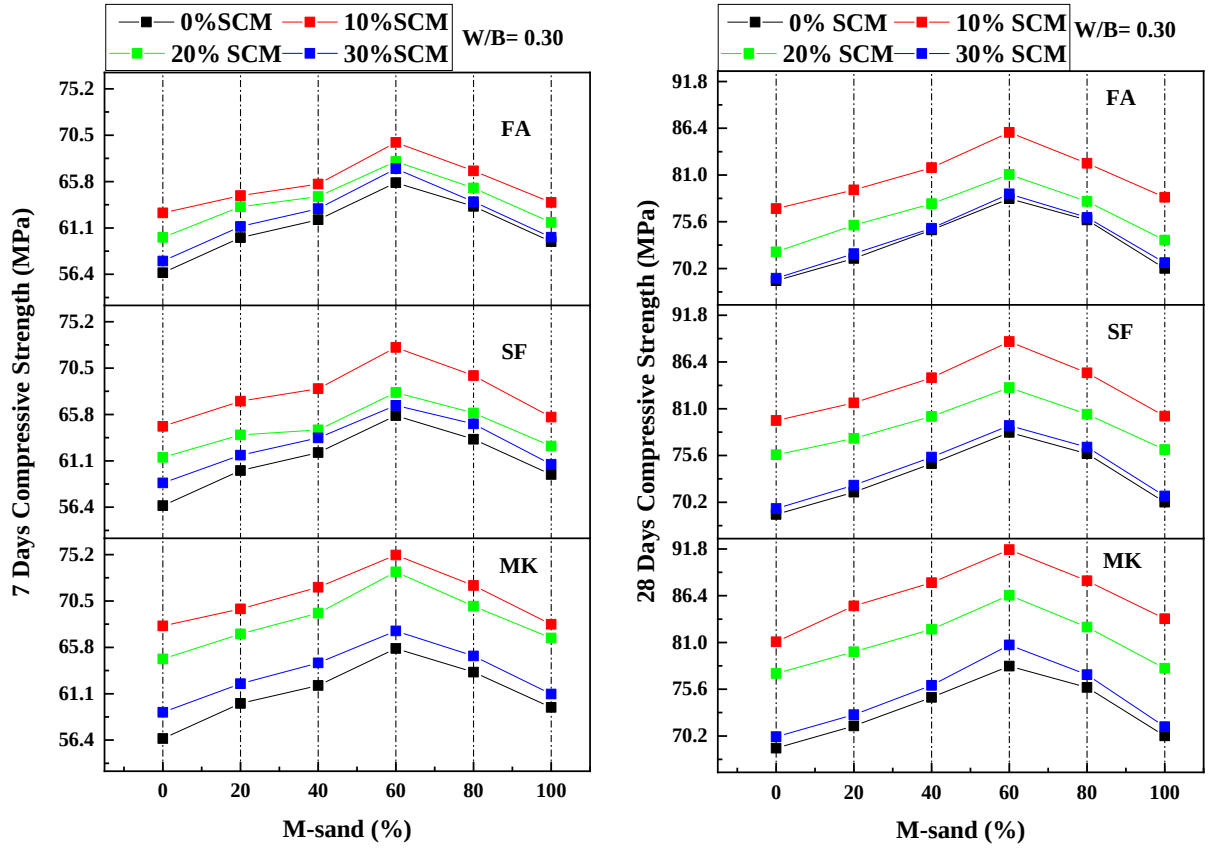



Fig. 4. Compressive strength Vs. % M-Sand (a) 7-days (b) 28-days

*3.1.4 Microstructure–Strength Mechanism of Manufactured Sand High-Performance Concrete*

The superior mechanical performance observed at approximately 60% manufactured sand replacement can be attributed to an optimal balance between particle packing density and interfacial transition zone (ITZ) characteristics. The use of manufactured sand improves aggregate–paste interaction and reduces inter-particle voids, resulting in a denser and more cohesive cementitious matrix with enhanced load transfer capability. At moderate replacement levels, improved packing efficiency reduces capillary porosity and strengthens the ITZ. In contrast, excessive M-sand content increases surface area and water demand, which adversely affects workability and leads to localized void formation, thereby reducing strength. Furthermore, the incorporation of metakaolin significantly enhances the microstructural properties through accelerated pozzolanic reactions. The consumption of calcium hydroxide and the formation of additional C–S–H and C–A–S–H gel phases contribute to pore refinement, increased matrix stiffness, and improved bonding between paste and aggregates. Overall, the observed strength enhancement is governed by the combined effects of optimized packing, improved interfacial characteristics, and binder-chemistry-driven microstructural refinement.

## 3.2 Split Tensile and Flexural Strength

Fig.5 and 6 illustrate the variation in 28-day split tensile strength (STS) and flexural strength (FS) of MSHPC as a role of manufactured sand (M-sand) replacement level at a constant W/B ratio of 0.30 for FA, SF and MK-based systems. Although split tensile strength and flexural strength tests were conducted at both 7 and 28 days, only the 28-day results are presented here for brevity and comparative discussion. For all SCM types, both STS and FS increase progressively as the M-sand

content rises from 0% to approximately 60%, beyond which a gradual decline is observed, indicating an optimal replacement level.

The strength enhancement up to 60% M-sand is attributed to the angular particle morphology and rough surface texture of manufactured sand, which improve mechanical interlocking and ITZ quality, thereby enhancing crack-bridging capacity and load transfer efficiency within the cementitious matrix. Similar trends have been reported in studies on optimized fine aggregate grading in high-performance concretes [11,41,51]

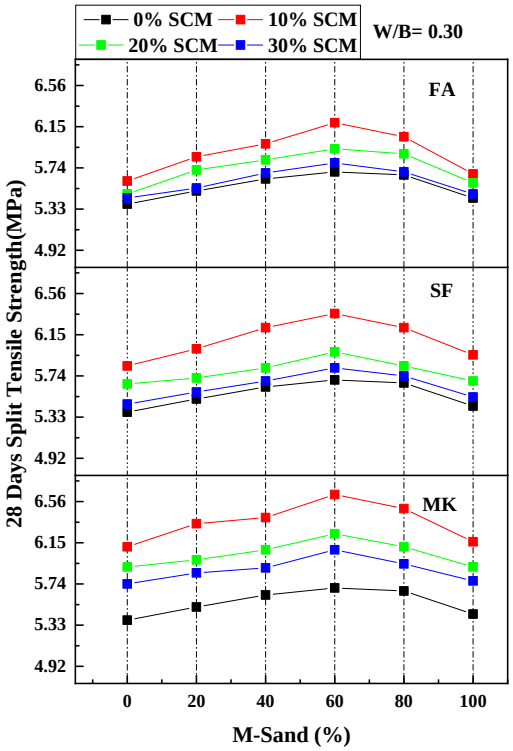



Fig. 5. Split tensile strength vs M-sand% (28-day, W/B=0.30)

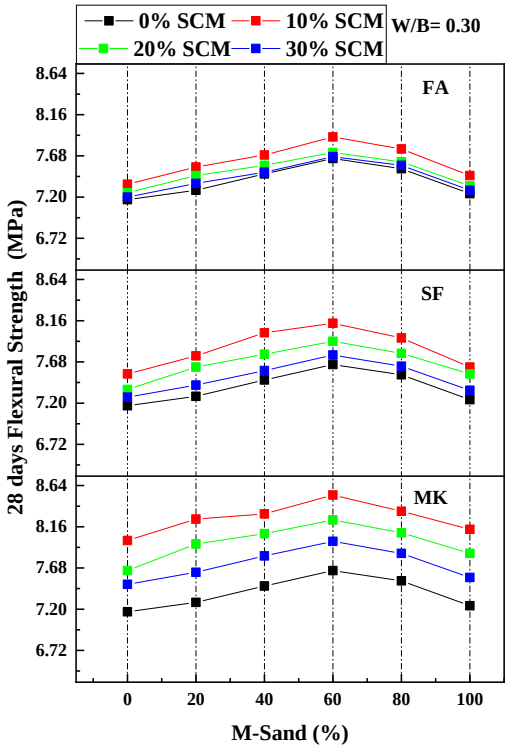



Fig. 6. Flexural strength vs M-sand% (28-day, W/B=0.30)

Among the SCM systems, the metakaolin-based mix, i.e., M60M10A (a mix containing 60% manufactured sand, 10% metakaolin, and a water–binder ratio of 0.30), consistently exhibits the highest tensile and flexural performance, achieving peak values of approximately 6.6 MPa (STS) and 8.5 MPa (FS) its improved split tensile strength (23.23%) and flexural strength (18.96%), compared to control mix i.e.M0MA0A.This enhanced performance is attributed to the high pozzolanic reactivity and micro filler effect of metakaolin, which contribute to pore refinement and improved bonding between paste and aggregates through the formation of additional C–S–H and C–A–S–H phases, as reported in [44, 46].

The reduction in strength beyond 60% M-sand replacement is associated with the increased specific surface area and water demand of M-sand-rich mixes, which can adversely affect workability and compaction, leading to higher void content and reduced binder continuity[39, 40].

### 3.3 Empirical Relationships Between Mechanical Properties

The relationships between cube compressive strength and the corresponding split tensile and flexural strengths of FA–MSHPC based MSHPC were evaluated for specimens cured for 28 days. The square-root form recommended by BIS code IS:456–2000 for conventional concrete,

$$f_t = 0.7\sqrt{f_{ck}}\ \ MPa \tag{1}$$

Equation 1 was adopted as a reference to examine the applicability of similar functional behavior in MSHPC systems. Regression analysis was performed to quantify the strength interdependence and to establish reliable predictive relationships.

### 3.3.1 FA–MSHPC

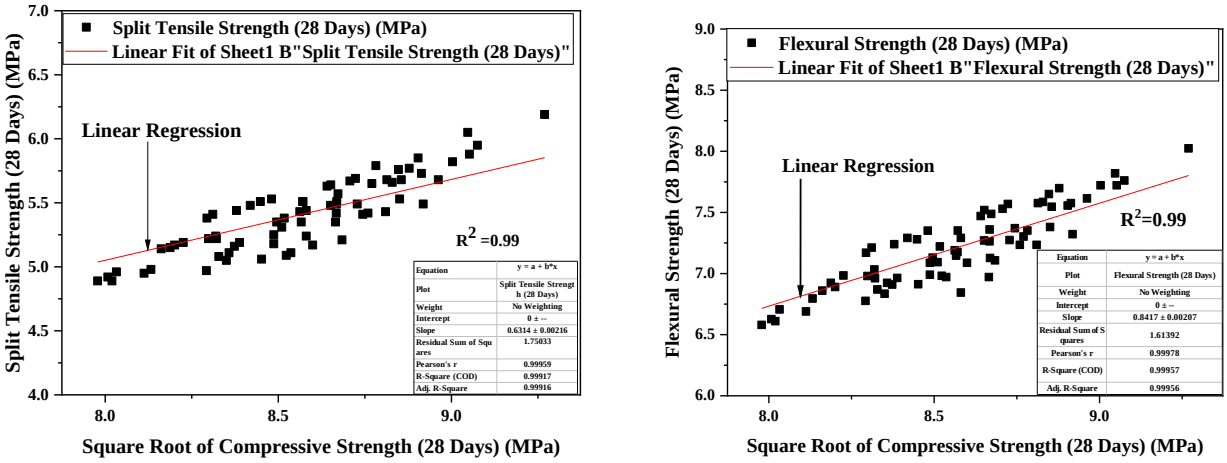
The empirical relationships for FA-MSHPC at 28 days are shown in Fig.7 (a–b). A strong dependency between compressive strength and split tensile strength was observed, and the best-fit relationship is:

$$f_{ct} = 0.631\sqrt{f_{ck}} \quad MPa \tag{2}$$

with a high correlation coefficient ($R^2$ = 0.99) indicating excellent agreement between predicted and experimental results. The slightly lower coefficient compared to conventional OPC concrete reflects the influence of M-sand and fly ash content to continuity and pore structure on tensile behavior. Similarly, the relationship between compressive strength and flexural strength is expressed as:

$$f_r = 0.841\sqrt{f_{ck}} \quad MPa \tag{3}$$

The strong correlation ($R^2$ = 0.99) indicates that flexural resistance in FA-MSHPC is primarily governed by matrix stiffness, aggregate interlock, and crack-bridging mechanisms.




(a) (b)

Fig. 7. Empirical relationships for FA-MSHPC at 28 days: (a) cube compressive strength versus split tensile strength, (b) cube compressive strength versus flexural strength

### 3.3.2 SF–MSHPC

The empirical relationships for SF-MSHPC at 28 days are presented in Fig.8(a–b). The correlation between compressive strength and split tensile strength is given by:

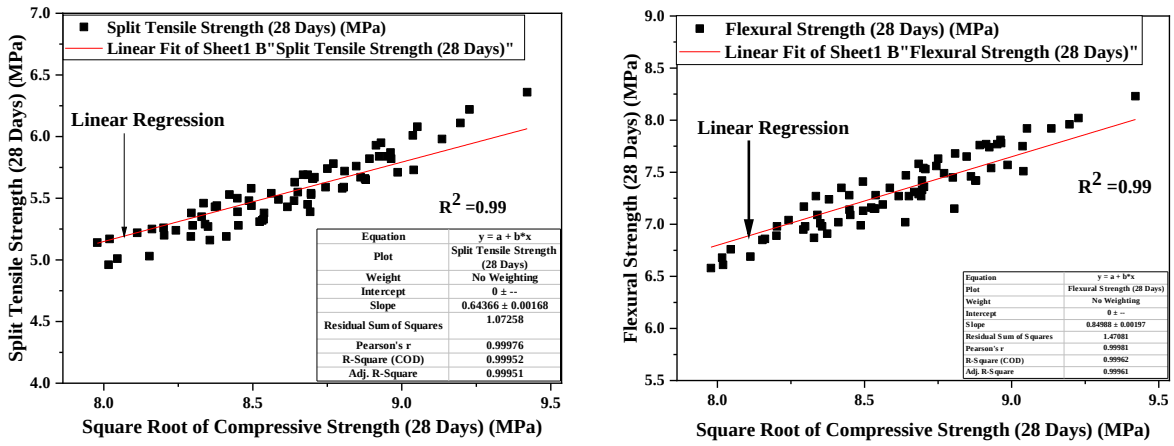
$$f_{ct} = 0.643\sqrt{f_{ck}} \quad MPa \tag{4}$$

with a high coefficient of determination ($R^2 \approx 0.99$) indicating strong predictive reliability. The improved proportionality constant compared to FA-MSHPC suggests enhanced tensile stress transfer due to refined pore structure and improved interfacial bonding.

The flexural strength relationship is expressed as:

$$f_r = 0.849\sqrt{f_{ck}} \quad MPa \tag{5}$$

The high correlation confirms the role of silica-rich gel densification and improved matrix stiffness in enhancing bending resistance.




(a) (b)

Fig. 8. Empirical relationships for SF-MSHPC at 28 days: (a) cube compressive strength versus split tensile strength, (b) cube compressive strength versus flexural strength

*3.3.3 MK–MSHPC*

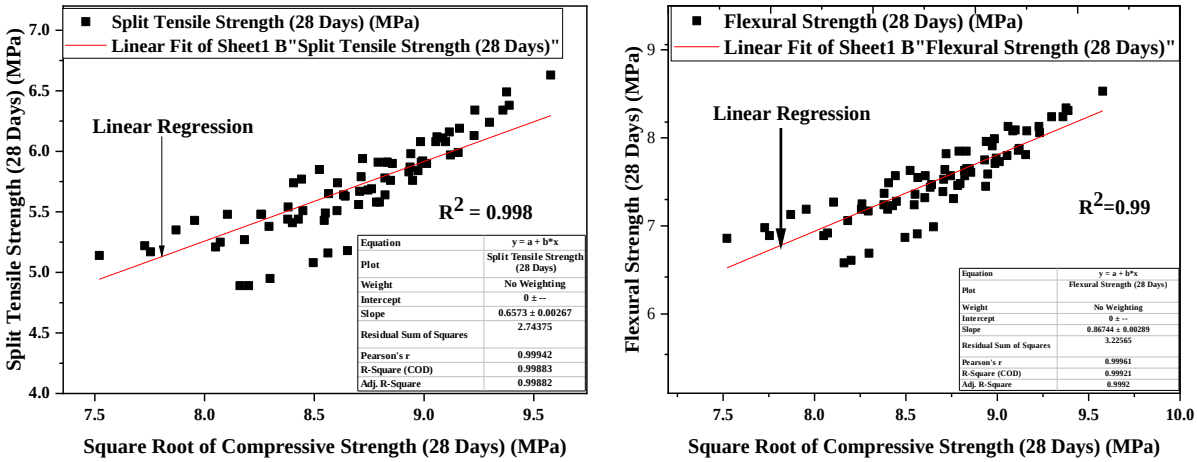
The empirical relationships for MK-MSHPC at 28 days are shown in Fig. 9(a–b). The relationship between compressive strength and split tensile strength is:

$$f_{ct} = 0.657\sqrt{f_{ck}} \quad MPa \tag{6}$$

with an excellent correlation ($R^2$ = 0.99), indicating strong agreement between experimental and predicted values. The higher coefficient reflects improved tensile resistance due to dense aluminosilicate gel formation in metakaolin-based systems. The flexural strength relationship is expressed as:

$$f_r = 0.867\sqrt{f_{ck}}\ MPa \tag{7}$$

The strong correlation ($R^2 \approx 0.99$) confirms that flexural performance in MK-MSHPC is significantly influenced by matrix densification and enhanced crack-bridging behavior.




(a) (b)

Fig. 9. Empirical relationships for MK-MSHPC at 28 days: (a) cube compressive strength versus split tensile strength, (b) cube compressive strength versus flexural strength

## 3.4 Structural Implications of Optimized Manufactured Sand High-Performance Concrete

The improved mechanical performance of the optimized MSHPC mix has important implications for structural applications. The enhanced tensile and flexural strengths indicate improved resistance to crack initiation, which is beneficial for structural serviceability. The refined interfacial

characteristics and denser cementitious matrix contribute to more effective stress transfer between aggregate and paste, resulting in improved stiffness and load distribution under compressive and bending stresses. The strong correlation observed between compressive, split tensile, and flexural strengths suggests improved overall structural integrity and reduced susceptibility to brittle failure. Furthermore, the reduction in capillary porosity associated with lower water–binder ratios and optimized mineral admixture content indicates the potential for improved durability-related properties, as reported in the literature [52, 53]. However, it should be noted that durability-related aspects such as permeability, long-term performance, and microcrack propagation were not directly evaluated in the present study and require dedicated experimental investigation. Therefore, the above implications are inferred based on mechanical performance and supported by existing literature[54]. These characteristics suggest that the optimized MSHPC may be suitable for high-performance structural concrete applications where enhanced strength and mechanical reliability are required.

### 3.5 Correlation Analysis and Multicollinearity Assessment

The Pearson correlation heatmap (Fig. 10) illustrates the linear relationships among input and output variables. A strong positive correlation is observed between curing age and compressive strength ($r \approx 0.80$), indicating the significant influence of hydration and pozzolanic reactions on strength development. In contrast, the water-to-binder (W/B) ratio exhibits a moderate negative correlation with strength properties ($r \approx -0.30$ to $-0.40$), confirming that increased water content leads to higher porosity and reduced matrix densification. Additionally, strong inter-correlation is observed among compressive, split tensile, and flexural strengths ($r > 0.90$), indicating consistent mechanical behavior across different strength measures. The Variance Inflation Factor (VIF) analysis (Table 5) indicates the presence of moderate to high multicollinearity among certain predictors.

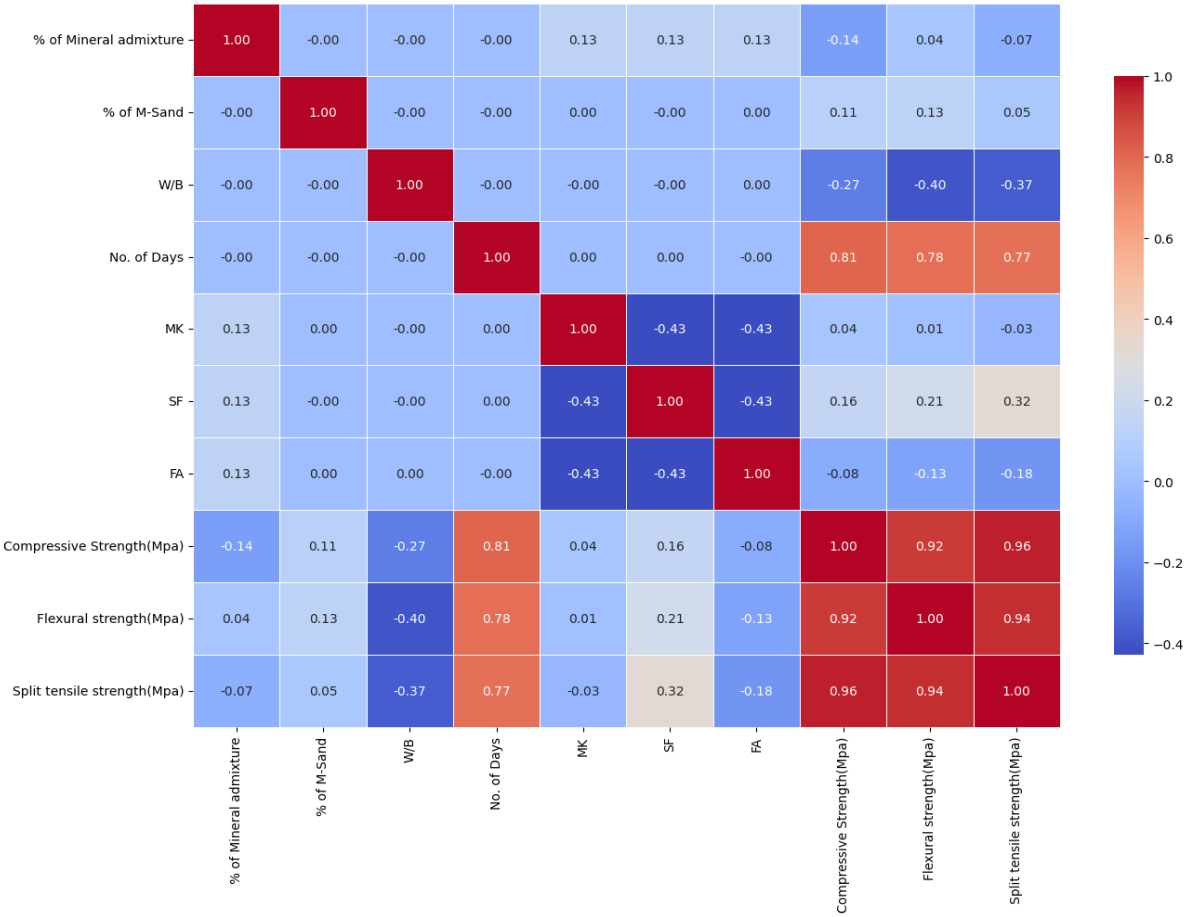



Fig. 10. Pearson correlation heatmap of input and output features

In particular, the W/B ratio shows high multicollinearity (VIF > 10), while the percentage of mineral admixture exhibits moderate multicollinearity (VIF ≈ 7). The one-hot encoded variables (MK, SF, FA) also present relatively elevated VIF values (≈5–6), which is expected due to the inherent dependency among dummy variables. However, the machine learning models employed in this study, particularly tree-based ensemble methods such as Random Forest and XGBoost, are inherently less sensitive to multicollinearity, as they rely on feature splitting rather than coefficient estimation. Therefore, all input variables were retained in the model to preserve their physical interpretability and contribution to concrete behavior.

Table 5. Variance inflation factor (VIF) values of input variables

| Feature | VIF (Compressive Strength) | VIF (Split Tensile & Flexural Strength) |
|---|---|---|
| % of Mineral admixture | 7.0000 | 7.0000 |
| % of M-Sand | 3.0667 | 3.0642 |
| W/B | 12.5716 | 10.5630 |
| No. of Days | 3.6591 | NA |
| MK | 5.6607 | 5.6497 |
| SF | 5.6607 | 5.6497 |
| FA | 5.6607 | 5.6497 |

## 3.6 Machine Learning

### *3.6.1 Machine Learning Prediction and Comparative Analysis*

Parity plots (Fig.11 (a–f)) compare experimentally measured and predicted compressive strength values obtained using ANN, NLR, RF, AdaB, GB, and XGBoost models. Among all models, XGBoost demonstrated the highest predictive accuracy, with data points tightly clustered around the 1:1 reference line. This is further supported by the highest testing $R^2$ value (0.9904) and the lowest MAE and RMSE.

The high training $R^2$ value observed for the XGBoost model suggests a strong fitting capability, which may indicate slight overfitting due to model complexity. However, the consistently high-test performance ($R^2 \approx 0.99$) indicates satisfactory generalization capability. The use of cross-validation further supports the reliability and robustness of the developed model. Future work may involve additional hyperparameter optimization, such as controlling tree depth and regularization parameters, to further reduce overfitting while maintaining predictive accuracy. The superior performance of XGBoost is attributed to its regularized gradient-boosting framework, which effectively captures complex nonlinear interactions among mix design parameters. GB and RF models also demonstrated strong predictive capability with slightly higher dispersion, whereas ANN and AdaBoost exhibited comparatively greater scatter, particularly at higher strength levels, indicating relatively lower generalization performance for the present dataset.

The obtained results are consistent with recent studies reporting the superior performance of ensemble-based models for predicting concrete strength properties. For instance,[55, 56] reported that GB-based algorithms outperform ANN and conventional regression models in predicting compressive strength of high-performance and sustainable concretes. Huang et al[57] further demonstrated that XGBoost effectively captures the variability associated with manufactured sand concrete. Compared to these studies, the present work achieved marginally higher prediction accuracy ($R^2 > 0.99$), indicating improved robustness and generalization capability of the proposed ML framework.

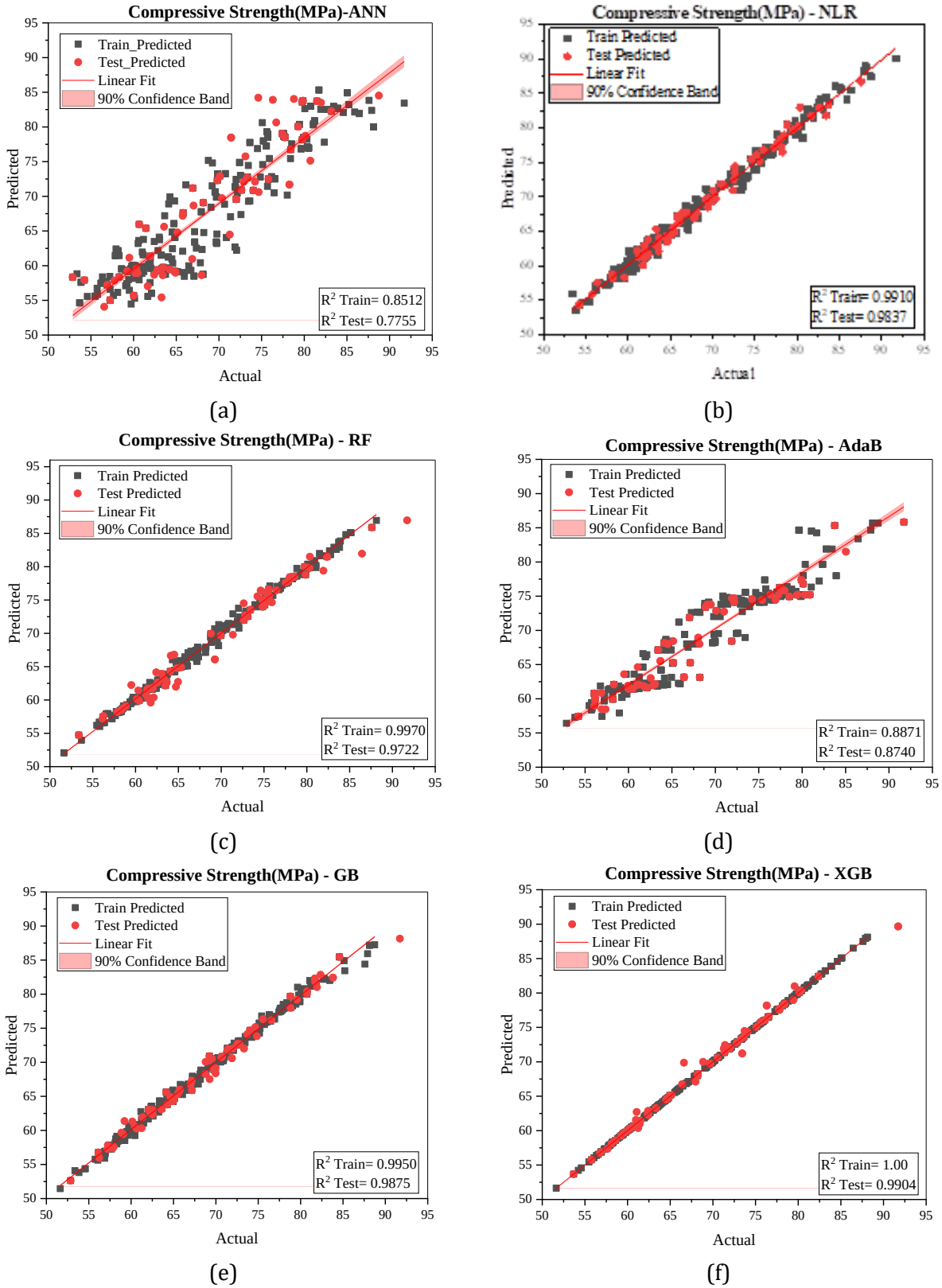



Fig. 11. Parity plots comparing experimentally and predicted compressive strength of MSHPC using different models: (a) ANN, (b) NLR, (c) RF, (d) AdaB, (e) GB, and (f) XGB

### *3.6.2 Split Tensile Strength*

Parity plots (Fig. 12(a–f)) illustrate the predictive performance of different machine learning models for split tensile strength (STS). Among the models, XGBoost demonstrates the closest agreement between experimental and predicted values, followed by Gradient Boosting (GB) and

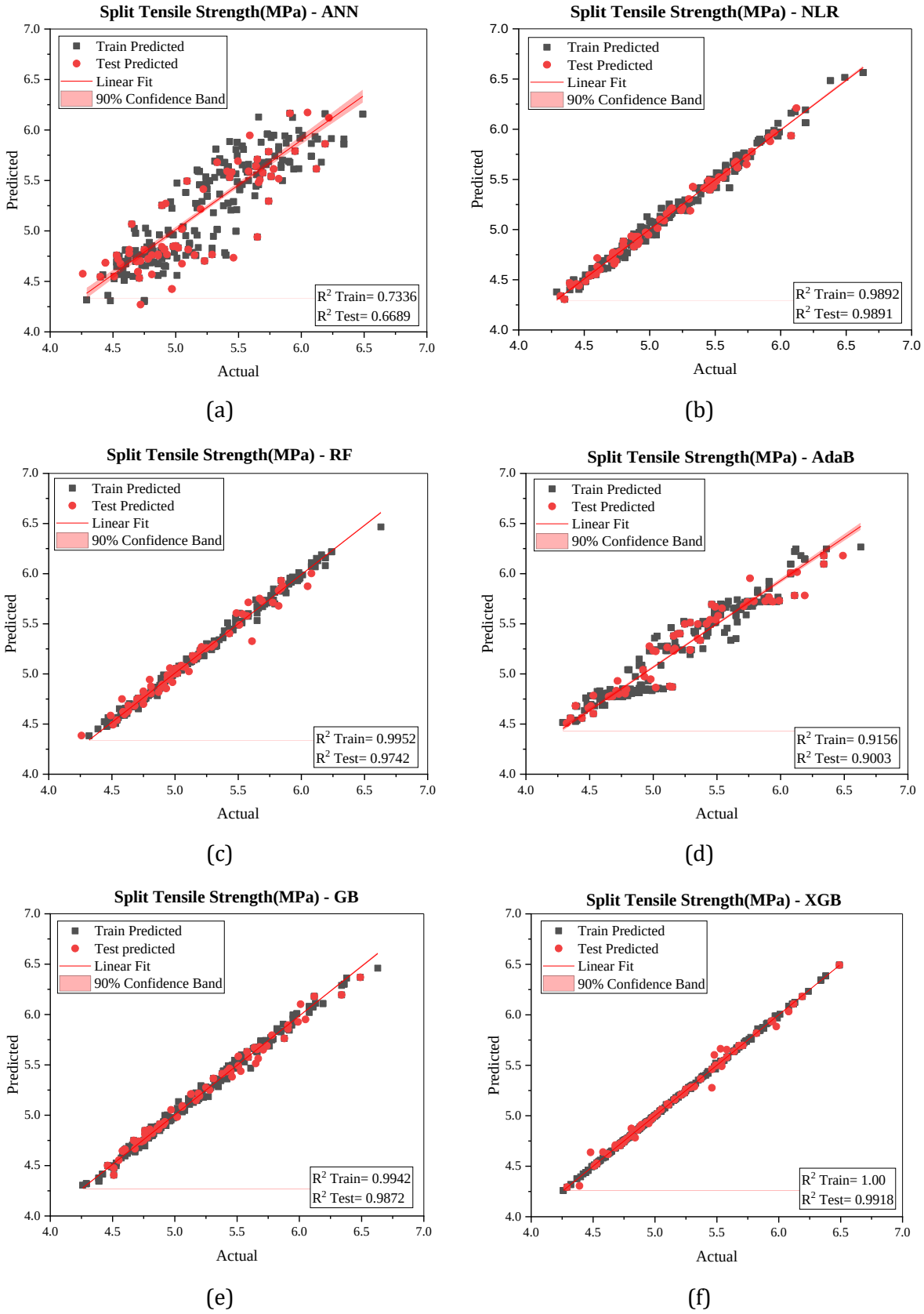



Fig. 12. Parity plots comparing experimentally and predicted split tensile strength of MSHPC using different models: (a) ANN, (b) NLR, (c) RF, (d) AdaB, (e) GB, and (f) XGB

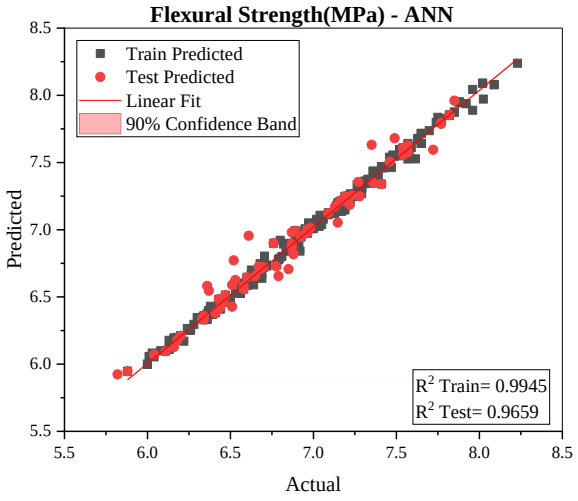
Random Forest (RF) models. In contrast, ANN and AdaBoost (AdaB) exhibit comparatively higher dispersion, particularly at higher tensile strength levels. The high training $R^2$ value observed for the XGBoost model suggests a strong fitting capability, which may indicate a degree of overfitting

due to model complexity. However, the consistently high-test performance indicates that the model maintains good generalization ability. The use of cross-validation further supports the reliability of the model predictions, although careful tuning of model complexity is necessary to ensure optimal performance.

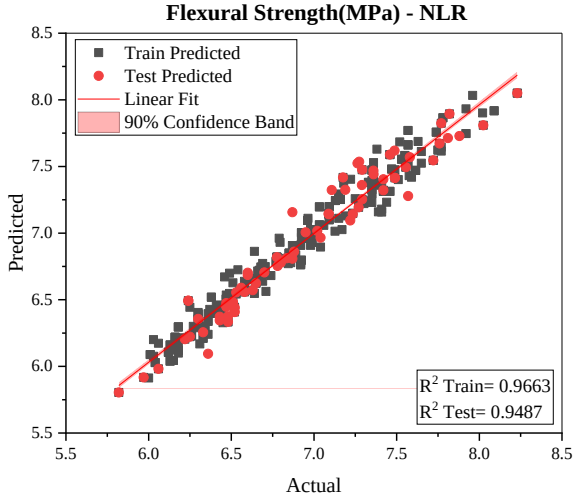
Future work may involve further hyperparameter tuning, such as controlling tree depth and regularization parameters, to minimize overfitting while maintaining predictive accuracy. The observed prediction trends are consistent with previous machine learning studies. Zhang et al. [49] reported improved predictive capability for tensile strength in high-performance concrete incorporating silica fume and metakaolin, while Huang et al. [51] highlighted the effectiveness of manufactured sand-based systems in data-driven modeling. Furthermore, Paudel et al. [50] demonstrated that ensemble-based models such as Gradient Boosting outperform ANN-based approaches in predicting split tensile strength. The close agreement of the present results with these findings further validates the robustness of the proposed modeling framework.

### 3.7 Flexural Strength

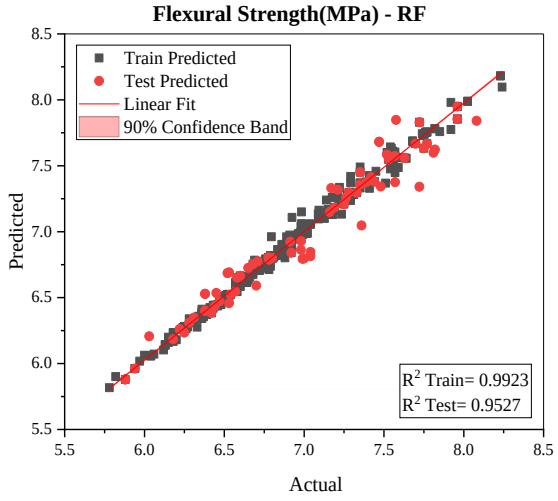
FS exhibited trends similar to compressive and split tensile strengths. Replacement of natural sand with manufactured sand improved flexural performance up to an optimum level, primarily due to enhanced stress transfer and stronger matrix–aggregate bonding. Beyond the optimum replacement level, a marginal reduction in flexural strength was observed, likely due to increased fines content and initiation of microcracks. SF and MK significantly enhanced FS by refining pore structure and improving ITZ characteristics. Lower W/B ratios consistently produced higher FS by reducing porosity and limiting microcrack propagation under bending stresses.



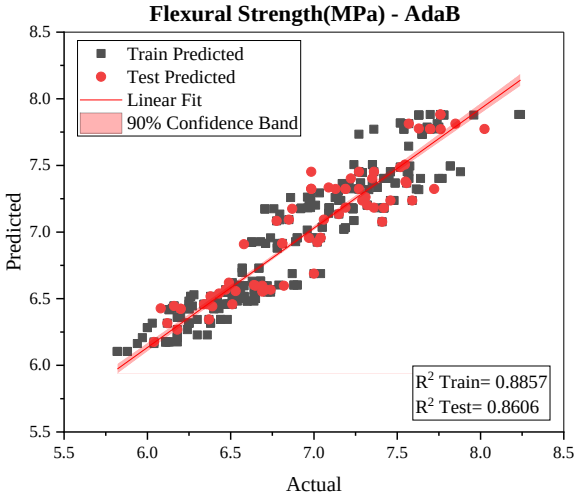

(a)



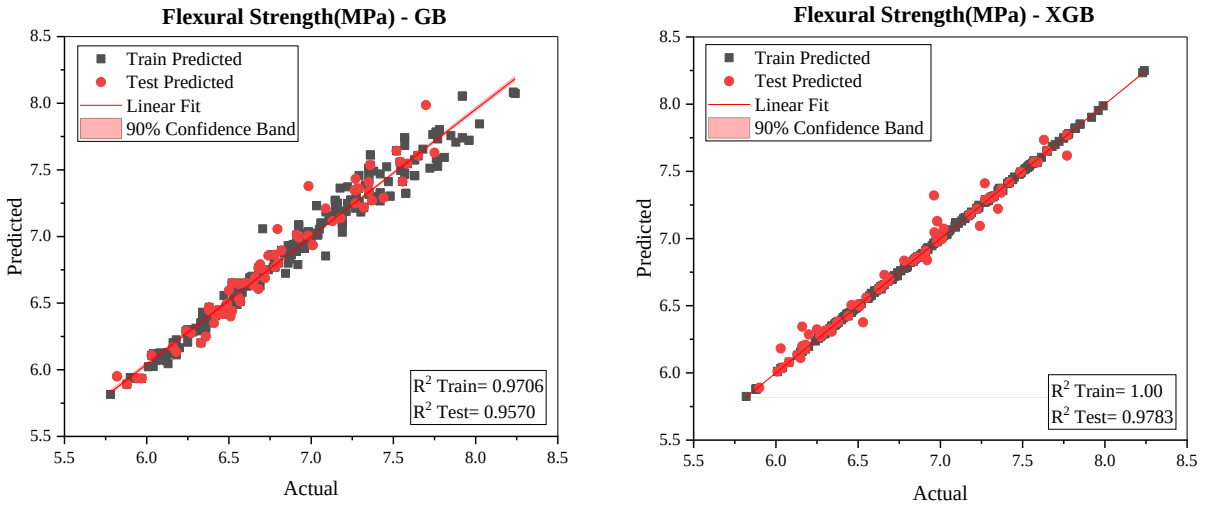

(b)




(c)




(d)



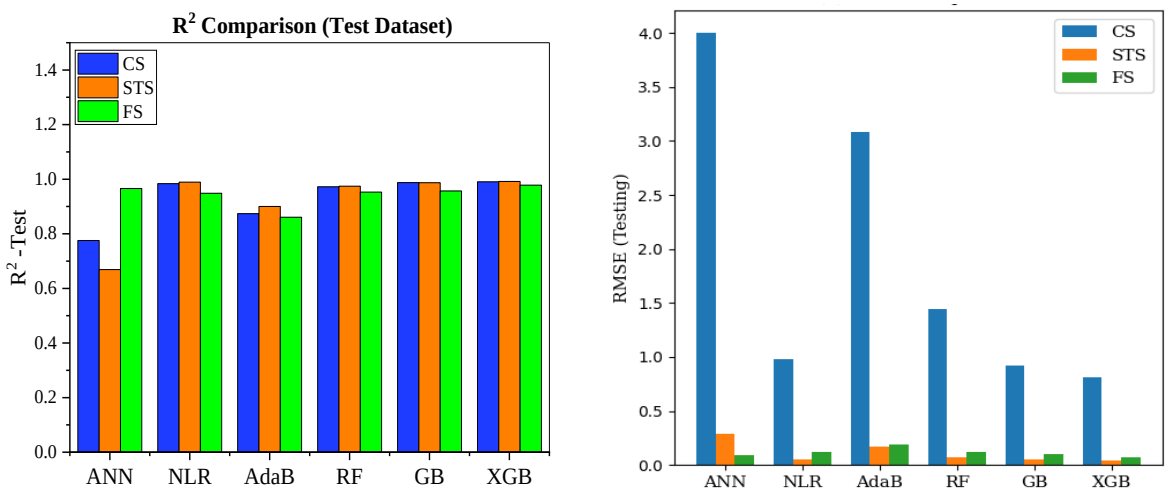

(e) (f)

Fig. 13. Parity plots comparing experimentally measured and machine learning–predicted flexural strength of MSHPC using different models: (a) ANN, (b) NLR, (c) RF, (d) AdaB, (e) GB, and (f) XGB

Parity plots (Fig. 13(a–f)) reveal that the XGBoost model provides the most accurate flexural strength predictions, followed by GB and RF models, whereas ANN and AdaB exhibited comparatively lower predictive capability. The high training $R^2$ value observed for the XGBoost model suggests a strong fitting capability, which may indicate slight overfitting due to model complexity. However, the consistently high-test performance indicates satisfactory generalization capability. The use of cross-validation further supports the reliability and robustness of the developed model. The superior performance of XGBoost is attributed to its regularized gradient-boosting framework, which effectively captures complex nonlinear relationships among the mix design parameters.

### 3.8 Comparative Performance of Machine Learning Models

Table 6 and Fig.14 present a comprehensive comparison of ML model performance for predicting the compressive, split tensile and flexural strength of MSHPC. The testing $R^2$ shown in Fig.14(a) indicate that ensemble-based models consistently outperform single and regression-based approaches across all mechanical properties. Among all models, XGBoost exhibits the highest testing $R^2$ values, exceeding 99% for compressive and split tensile strength and approaching 98% for flexural strength.




(a) (b)

Fig. 14. Comparison of machine learning models for predicting mechanical properties of MSHPC: (a) testing $R^2$ and (b) testing RMSE

The RMSE comparison in Fig.14(b), together with the MAE and MSE values reported in Table 6, further confirms the superiority of XGBoost, which yields the lowest prediction errors across all mechanical properties. GB and RF models also demonstrate high accuracy, whereas ANN and AdaB show comparatively lower generalization capability. The close agreement between graphical trends and statistical indicators validates the robustness and reliability of the proposed ML framework. To statistically validate the superior performance of the XGBoost model, a paired t-test was conducted on the prediction errors obtained from XGBoost and the second-test performing model (GB). At a 90% confidence level, the results indicate a statistically significant improvement ($p < 0.05$), confirming that the observed performance gain is not attributable to random variation.

Table 6. Statistical evaluation of machine learning models for prediction of mechanical properties of MSHPC

| Mechanical property | Model | $R^2$ (Training) | $R^2$ (Testing) | MAE | MSE | RMSE |
|---|---|---|---|---|---|---|
| Compressive Strength | ANN | 0.8512 | 0.7755 | 3.2585 | 16.0381 | 4.0048 |
| | NLR | 0.9910 | 0.9837 | 0.7754 | 0.9687 | 0.9843 |
| | AdaB | 0.8871 | 0.8740 | 2.6784 | 9.5199 | 3.0854 |
| | RF | 0.997 | 0.9722 | 1.0055 | 2.0944 | 1.4472 |
| | GB | 0.9950 | 0.9875 | 0.7182 | 0.8540 | 0.9241 |
| | XGB | 1.00 | 0.9904 | 0.4048 | 0.6545 | 0.8090 |
| Split Tensile Strength | ANN | 0.7336 | 0.6689 | 0.2280 | 0.0861 | 0.2934 |
| | NLR | 0.9892 | 0.9891 | 0.0383 | 0.0023 | 0.0489 |
| | AdaB | 0.9156 | 0.9003 | 0.1463 | 0.0294 | 0.1714 |
| | RF | 0.9952 | 0.9742 | 0.0464 | 0.0047 | 0.0687 |
| | GB | 0.9942 | 0.9872 | 0.0430 | 0.0031 | 0.0560 |
| | XGB | 1.00 | 0.9918 | 0.0225 | 0.0019 | 0.0441 |
| Flexural Strength | ANN | 0.9945 | 0.9659 | 0.0623 | 0.0084 | 0.0917 |
| | NLR | 0.9663 | 0.9487 | 0.0949 | 0.0149 | 0.1224 |
| | AdaB | 0.8857 | 0.8606 | 0.1544 | 0.0354 | 0.1883 |
| | RF | 0.9923 | 0.9527 | 0.0886 | 0.0151 | 0.1232 |
| | GB | 0.9706 | 0.9570 | 0.0744 | 0.0101 | 0.1007 |
| | XGB | 1.00 | 0.9783 | 0.0392 | 0.0054 | 0.0739 |

Overall, the close agreement between the graphical trends in Fig.14 and the numerical performance metrics in Table 6 validates the robustness and generalization capability of the developed machine learning framework. The combined analysis clearly establishes that ensemble learning models, particularly XGBoost, provide the most reliable and accurate predictions for the mechanical properties of MSHPC, thereby offering an effective data-driven tool for optimizing high-performance concrete mix designs.

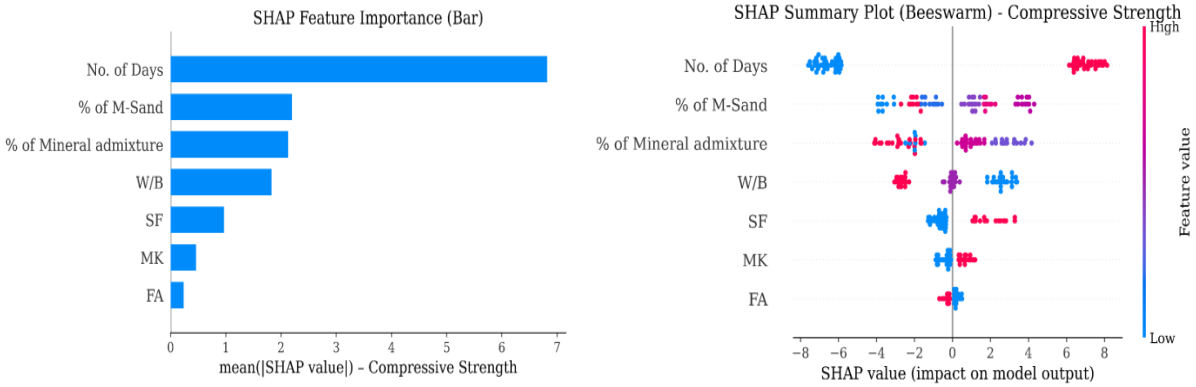
### 3.9 Explainable AI: Global Feature Attribution

Figures 15–17 present the SHAP-based interpretation of the XGBoost models developed for predicting compressive, split tensile, and flexural strengths of MSHPC. The mean absolute SHAP value plots (Figs. 15a, 16a, and 17a) quantify the relative importance of each input variable, while the corresponding summary (beeswarm/violin) plots (Figs. 15b, 16b, and 17b) illustrate both the magnitude and direction of feature influence across the dataset. The SHAP force plots (Figs. 15c, 16c, and 17c) provide local interpretability by decomposing individual predictions into contributing feature effects. For compressive strength (Fig. 15), curing age is identified as the most dominant parameter, exhibiting the highest mean SHAP value. The summary plot clearly indicates that higher curing durations contribute positively to strength development, whereas lower curing ages yield negative contributions. This behavior is consistent with progressive hydration and continued pozzolanic reactions leading to densification of the cementitious matrix.

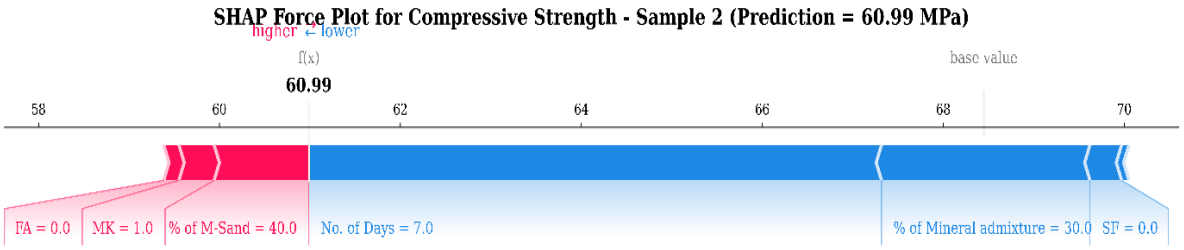
For split tensile and flexural strength (Figs. 16 and 17), the relative importance of curing age is reduced due to limited variability within the dataset, and the predictions are primarily governed by mixture design parameters. Among these, the water-to-binder (W/B) ratio emerges as the most

influential variable, consistently showing negative SHAP contributions at higher values. This reflects the increase in capillary porosity and reduction in matrix compactness associated with higher water content. Conversely, lower W/B ratios enhance particle packing and improve mechanical performance.

The percentage of manufactured sand (% M-sand) exhibits a nonlinear influence on strength, as observed from the spread of SHAP values in the summary plots. Intermediate replacement levels (approximately 40–60%) tend to produce positive contributions, whereas very high or very low replacement levels result in reduced performance. This trend aligns with the experimental observations reported in Section 3.1. The effect of mineral admixture is represented through one-hot encoded variables (FA, SF, MK), enabling independent evaluation of each material without imposing ordinal bias. Among these, silica fume (SF) generally demonstrates the most significant positive contribution, particularly in tensile and flexural strength models, due to its high pozzolanic reactivity and micro-filling ability. Metakaolin (MK) also contributes positively but to a comparatively moderate extent, while fly ash (FA) exhibits relatively lower or slightly negative contributions under certain conditions, especially at early ages.

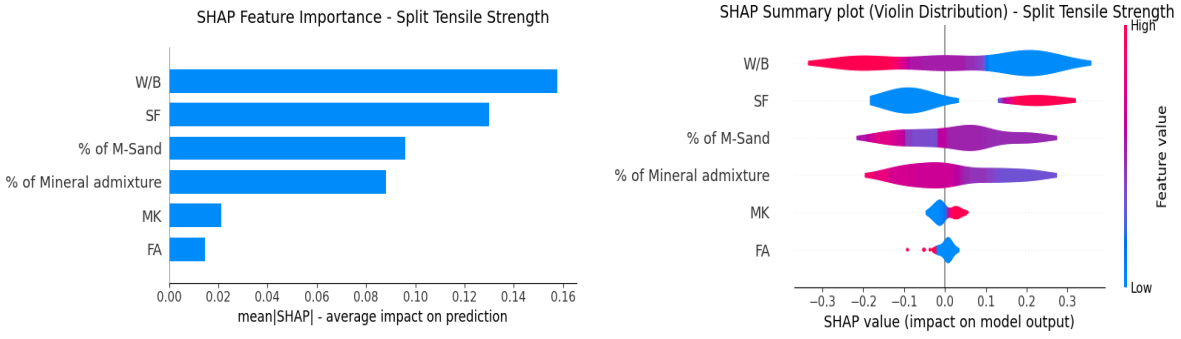



(a) (b)




(c)

Fig. 15. SHAP interpretation for compressive strength prediction:(a) Global feature importance (mean |SHAP|), (b) SHAP summary (violin distribution), (c) SHAP force plot illustrating local feature contribution




(a) (b)

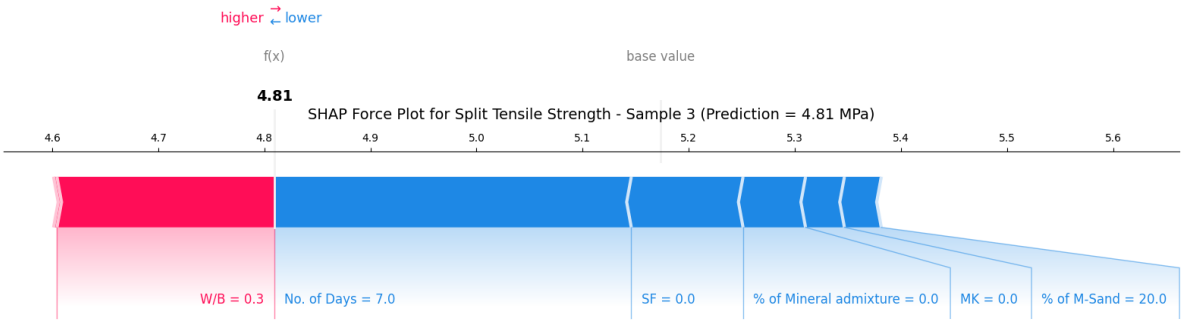
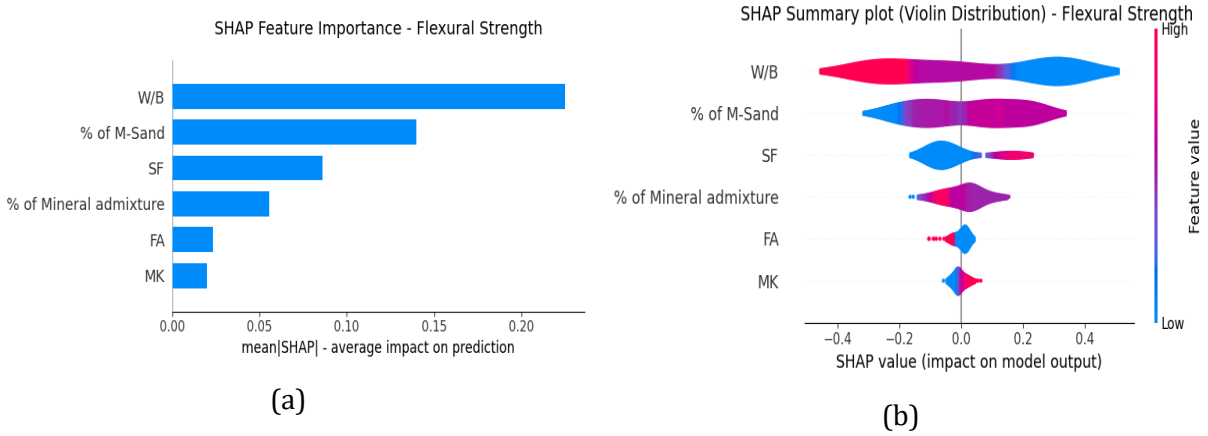
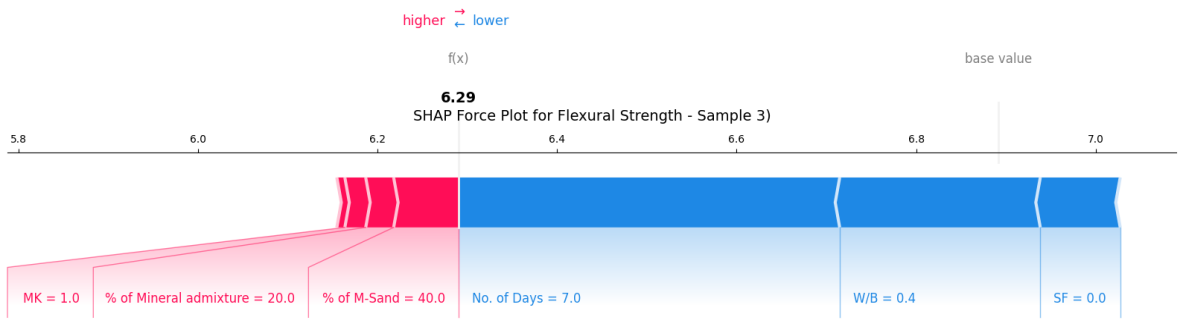



Fig. 16 SHAP interpretation for split tensile strength prediction:(a) Global feature importance (mean |SHAP|),(b) SHAP summary (violin distribution),(c) SHAP force plot illustrating local feature contribution




(a) (b)




(c)

Fig. 17 SHAP interpretation for flexural strength prediction:(a) Global feature importance (mean |SHAP|),(b) SHAP summary (violin distribution),(c) SHAP force plot illustrating local feature contribution

The SHAP force plots further validate these findings at the individual prediction level. For instance, higher W/B ratios consistently drive predictions downward, whereas optimal M-sand replacement and reactive mineral admixtures (particularly SF and MK) contribute positively toward strength enhancement. Overall, the SHAP analysis reveals that compressive strength is strongly governed by curing kinetics, whereas tensile and flexural properties are predominantly controlled by matrix quality parameters such as W/B ratio, mineral admixture type, and M-sand replacement. These findings demonstrate the capability of explainable AI to provide physically meaningful insights into material behavior beyond conventional statistical analysis.

## 4. Conclusions

- This study demonstrates that the combined use of manufactured sand and mineral admixtures significantly enhances the mechanical performance of high-performance concrete. The optimal mix proportion was identified at 60% manufactured sand replacement with 10% mineral admixture, which consistently yielded superior strength characteristics.
- Among the mineral admixtures investigated, metakaolin-based MSHPC exhibited the highest experimental strength performance, followed by silica fume and fly ash systems. This improvement is attributed to the high pozzolanic reactivity of metakaolin, leading to enhanced formation of C–S–H and C–A–S–H phases, improved pore refinement, and stronger interfacial transition zone (ITZ) bonding.
- The optimized mix (M60M10A) achieved a maximum compressive strength improvement of 33.33% compared to the control mix, with peak performance observed at 28 days of curing, highlighting the critical role of curing kinetics in strength development.
- The water-to-binder (W/B) ratio was identified as a key governing parameter, with strength consistently decreasing as W/B increased from 0.30 to 0.40. Lower W/B ratios resulted in improved matrix densification and reduced capillary porosity, leading to enhanced mechanical performance confirming the importance of maintaining a low W/B ratio for high-performance concrete.
- The combination of metakaolin at 10% and 60% of M-Sand i.e.M60M10A, replacement significantly improved split tensile strength (23.23%) and flexural strength (18.96%), compared to control mix, confirming its effectiveness in improving crack resistance and load transfer through microstructural refinement.
- The effect of manufactured sand replacement was found to be nonlinear, with approximately 60% replacement providing optimal performance. This is attributed to improved particle packing, reduced void content, and enhanced aggregate–paste interaction, whereas higher replacement levels led to reduced workability and increased porosity.
- SHAP-based explainable machine learning analysis revealed that curing age exhibited a strong positive contribution toward strength development, whereas higher W/B ratios consistently produced negative contributions due to increased porosity and reduced matrix densification. Manufactured sand replacement exhibited a nonlinear threshold-dependent behavior, with approximately 60% replacement providing the most beneficial contribution to mechanical performance.
- The XGBoost model achieved the highest predictive accuracy ($R^2 > 0.99$) among all models, demonstrating its effectiveness in capturing complex nonlinear relationships in MSHPC systems. The integration of SHAP further enabled transparent and physically interpretable insights into model predictions.
- This study presents an integrated experimental and explainable machine learning framework for optimizing sustainable high-performance concrete while reducing dependence on natural river sand and Portland cement. The incorporation of 60% manufactured sand and partial cement replacement using mineral admixtures such as fly ash, silica fume, and metakaolin contributes toward sustainable construction by reducing ecological degradation associated with river sand mining and lowering the environmental burden associated with cement production.
- Future research should focus on long-term durability performance, microstructural characterization, and structural-scale validation of optimized MSHPC mixes under varying environmental and loading conditions.

## References

[1] Ozmen HB, Tanrıverdi B. Concrete paradox: Economic importance, environmental impacts, and the sustainability of concrete material. Research and design. 2025;2(2):66-84. https://doi.org/10.17515/rede2025-008en0801rs

[2] Kandilli C. Exploring the potential of natural zeolites to enhance solar energy systems: A concise review. Research and design. 2024;1(1):13-21. https://doi.org/10.17515/rede2024-001en1101rv

[3] Sathvik S, Oyebisi S, Kumar R, et al. Analyzing the influence of manufactured sand and fly ash on concrete strength through experimental and machine learning methods. Scientific Reports. 2025;15(1):4978. https://doi.org/10.1038/s41598-025-88923-3
[4] Anh LH, Mihai FC, Belousova A, et al. Life Cycle Assessment of River Sand and Aggregates Alternatives in Concrete. Materials. 2023;16(5):2064. https://doi.org/10.3390/ma16052064
[5] Patel V, Shah N. A Survey of High-Performance Concrete Developments in Civil Engineering Field. Open Journal of Civil Engineering. 2013;03(02):69-79. https://doi.org/10.4236/ojce.2013.32007
[6] Juenger MCG, Siddique R. Recent advances in understanding the role of supplementary cementitious materials in concrete. Cement and Concrete Research. Elsevier Ltd. 2015; 78:71-80. https://doi.org/10.1016/j.cemconres.2015.03.018
[7] Kumar S, Park E, Tran DD, et al. A deep learning framework to map riverbed sand mining budgets in large tropical deltas. GI Science & Remote Sensing. 2023;61(1):2285178. https://doi.org/10.31223/X55M39
[8] Mathur A, Mathur M. An Experimental Study on the Effect of Replacement of Natural Sand with Manufacture Sand. International Journal of Engineering Research & Technology (IJERT).2018;6(11):1-9.
[9] Suresh S, Revathi J. Effect of M-Sand on Setting Time of High-Performance Concrete. Asian Journal of Research in Social Sciences and Humanities. 2016;6(10):1648. https://doi.org/10.5958/2249-7315.2016.01118.7
[10] Khan SU, Nuruddin MF, Ayub T, Shafiq N. Effects of different mineral admixtures on the properties of fresh concrete. The Scientific World Journal. 2014; 986567:1-11. https://doi.org/10.1155/2014/986567
[11] Kavya A, Venkateshwara Rao A. Experimental investigation on mechanical properties of concrete with M-sand. In: Materials Today: Proceedings.2020;13(A):663-667. https://doi.org/10.1016/j.matpr.2020.05.774
[12] Umamaheswaran V, Sudha C, Ravichandran PT, Kannan Rajkumar PR. Use of M Sand in High Strength and High-Performance Concrete. Indian Journal of Science and Technology. 2015;8(28). https://doi.org/10.17485/ijst/2015/v8i28/84018
[13] Rajendra T N, Surendra. B.V, An Experimental Investigation on Strength Properties of Concrete with Partial Replacement of Cement by Fly Ash and Metakaolin & with M. Sand as Fine Aggregate, International Research Journal of Engineering and Technology, 2016;3(7):2150-2153.
[14] Patel D, Modhera C, Gaurav J, Patel VN. Experimental investigation on mechanical properties of ternary blended concrete using manufactured sand. Innovative Infrastructure Solutions. 2023;8(5). https://doi.org/10.1007/s41062-023-01087-9
[15] Mane KM, Kulkarni DK, Prakash KB. Prediction of Flexural Strength of Concrete Produced by Using Pozzolanic Materials and Partly Replacing NFA by MS. Journal of Soft Computing in Civil Engineering. 2019;3(2):65-75. https://doi.org/10.1108/JEDT-12-2019-0346
[16] Liu G, Guo Y, Li Q, Wang L. The Threshold Value of Effective Replacement Ratio of Fly Ash Mortar Based on Amount of Calcium Hydroxide. In: IOP Conference Series: Earth and Environmental Science. Institute of Physics Publishing; 2020; 455:012125. https://doi.org/10.1088/1755-1315/455/1/012125
[17] Venu Malagavelli, Srinivas Angadi and J S R Prasad. Influence Of Metakaolin in Concrete as Partial Replacement of Cement. International Journal of Civil Engineering and Technology.2018;9(7):105-111.
[18] Mahalakshmi SHV, Khed VC. Experimental study on M-sand in self-compacting concrete with and without silica fume. In: Materials Today: Proceedings.2020:27:1061-1065. https://doi.org/10.1016/j.matpr.2020.01.432
[19] Alla S, Jayaram M, Asadi SS. An experimental investigation for replacements of river sand and cement with Robo-sand, fly-ash and silica fume in concrete to evaluate the influence in durability properties. Materials Today: Proceedings.2020;43(2):954-961. https://doi.org/10.1016/j.matpr.2020.07.340
[20] Abdul-Jabbar H, Khan M, Awan HH, Eldin SM, Alyousef R, Mohamed AM. Predicting ultra-high-performance concrete compressive strength using gene expression programming method. Case Studies in Construction Materials. 2023;18: e02074. https://doi.org/10.1016/j.cscm.2023.e02074
[21] Chen L, Wang Z, Khan AA, et al. Development of predictive models for sustainable concrete via genetic programming-based algorithms. Journal of Materials Research and Technology. 2023; 24:6391-6410. https://doi.org/10.1016/j.jmrt.2023.04.180
[22] Islam N, Kashem A, Das P, Ali MdN, Paul S. Prediction of high-performance concrete compressive strength using deep learning techniques. Asian Journal of Civil Engineering. 2024;25(1):327-341. https://doi.org/10.1007/s42107-023-00778-z
[23] Albostami AS, Al-Hamd RKS, Alzabeebee S, Minto A, Keawsawasvong S. Application of soft computing in predicting the compressive strength of self-compacted concrete containing recyclable aggregate. Asian Journal of Civil Engineering. 2024;25(1):183-196. https://doi.org/10.1007/s42107-023-00767-2
[24] Shahani NM, Ullah B, Shah KS, et al. Predicting Angle of Internal Friction and Cohesion of Rocks Based on Machine Learning Algorithms. Mathematics. 2022;10(20). https://doi.org/10.3390/math10203875

[25] Hulipalled P, Algur V, Lokesha V, Sharanappa DS. A statistical and neural network approach to investigate the tribological behavior of ZA-27 alloy. Multiscale and Multidisciplinary Modelling, Experiments and Design. 2024;7(6):5855-5870. https://doi.org/10.1007/s41939-024-00544-x
[26] Caglar N, Pala M, Elmas M, Mercan Eryilmaz D. A new approach to determine the base shear of steel frame structures. Journal of Constructional Steel Research. 2009;65(1):188-195. https://doi.org/10.1016/j.jcsr.2008.07.012
[27] Solhmirzaei R, Salehi H, Kodur V, Naser MZ. Machine learning framework for predicting failure mode and shear capacity of ultra-high performance concrete beams. Engineering Structures. 2020; 224:111221. https://doi.org/10.1016/j.engstruct.2020.111221
[28] Behnood A, Golafshani EM. Machine learning study of the mechanical properties of concretes containing waste foundry sand. Construction and Building Materials. 2020; 243:118152. https://doi.org/10.1016/j.conbuildmat.2020.118152
[29] Shang M, Li H, Ahmad A, et al. Predicting the Mechanical Properties of RCA-Based Concrete Using Supervised Machine Learning Algorithms. Materials. 2022;15(2):647. https://doi.org/10.3390/ma15020647
[30] Hong-Guang N, Wang JZ. Prediction of Compressive Strength of Concrete by Neural Networks. Cement and Concrete Research.2000;30:1245-1250. https://doi.org/10.1016/S0008-8846(00)00345-8
[31] Saridemir M. Predicting the compressive strength of mortars containing metakaolin by artificial neural networks and fuzzy logic. Advances in Engineering Software. 2009;40(9):920-927. https://doi.org/10.1016/j.advengsoft.2008.12.008
[32] Sobhani J, Najimi M, Pourkhorshidi AR, Parhizkar T. Prediction of the compressive strength of no-slump concrete: A comparative study of regression, neural network and ANFIS models. Construction and Building Materials.2010;24(5):709-718. https://doi.org/10.1016/j.conbuildmat.2009.10.037
[33] Hodhod OA, Ahmed HI. Modeling the corrosion initiation time of slag concrete using the artificial neural network. HBRC Journal. 2014;10(3):231-234. https://doi.org/10.1016/j.hbrcj.2013.12.002
[34] Hulipalled P, Algur V, Lokesha V, Saumya S, Satyanarayan. Interpretable ensemble machine learning framework to predict wear rate of modified ZA-27 alloy. Tribology International. 2023; 188:108783. https://doi.org/10.1016/j.triboint.2023.108783
[35] Hulipalled P, Algur V, Lokesha V. An Approach of Data Science for the Prediction of Wear Behaviour of Hypereutectoid Steel. Journal of Bio- and Tribo-Corrosion.2022;8(3):69. https://doi.org/10.1007/s40735-022-00668-y
[36] Das B, Hulipalled P, Algur V, Jain M, Karmakar A, Barat K. Enhancing Prediction Accuracy of Machine Learning Models for Materials Informatics Problems in Alloy Design: A Case Study on Dual-Phase Steel.Journal of Materials Engineering and Performance,2026;35(11):11234-11249. https://doi.org/10.1007/s11665-025-12401-0
[37] Arun BR, Srishaila JM, Md Khalid S, Algur V, Kavyashree K, Tanu HM. Prediction of machine learning application in the development of novel sustainable self-compacting geopolymer concrete. Research on Engineering Structures and Materials. 2025;11(4):1469-1490.
[38] Algur V, Hulipalled P, Kumar KS, Srishaila JM, Sharvani V, Ahamed HMZ. Explainable machine learning for predicting the mechanical strength of eco-friendly geopolymer concrete. ENGINEERING Structure and Civil Engineering. Published online 2026. https://doi.org/10.1007/s11709-026-1274-z
[39] Neville A M. Properties of concrete.5th edition, Pearson Education Limited, 2011.
[40] Mehta PM. Concrete Microstructure, Properties, and Materials.3rd edition, McGraw-Hill Companies 2006.
[41] Vardhan K, Siddique R, Goyal S. Strength, permeation and micro-structural characteristics of concrete incorporating waste marble. Construction and Building Materials. 2019; 203:45-55. https://doi.org/10.1016/j.conbuildmat.2019.01.079
[42] ACI COMMITTEE REPORT. Guide for the Use of Silica Fume in Concrete 234R-2.
[43] Bhanja S, Sengupta B. Influence of silica fume on the tensile strength of concrete. Cement and Concrete Research. 2005;35(4):743-747. https://doi.org/10.1016/j.cemconres.2004.05.024
[44] Poon CS, Lam L, Kou SC, Wong YL, Wong R. Rate of pozzolanic reaction of metakaolin in high-performance cement pastes. Cement and Concrete Research. 2001;31(9):1301-1306. https://doi.org/10.1016/S0008-8846(01)00581-6
[45] Adanagouda, Somasekharaiah HM. Strength and Durability Studies on Hybrid Fiber Reinforced High-Performance Concrete for Silica Fume Based Mineral Admixture. In: IOP Conference Series: Earth and Environmental Science. IOP Publishing Ltd; 2021; 822:012041. https://doi.org/10.1088/1755-1315/822/1/012041
[46] Adanagouda, Somasekharaiah HM, Shobha MS, Mallikarjuna HM. Combined effect of metakaolin and hybrid fibers on the strength properties of high-performance concrete. Materials Today: Proceedings.2022;49:1527-1536. https://doi.org/10.1016/j.matpr.2021.07.310

[47] Adanagouda, Somasekharaiah HM, Shobha MS, Mallikarjuna HM. Mechanical properties and acid attack test of hybrid fibre reinforced high performance concrete for fly ash based mineral admixture. Materials Today: Proceedings.2022; 51:742-749. https://doi.org/10.1016/j.matpr.2021.06.220
[48] Mahesh, Shobha MS, Hulipalled P, Algur V. ML driven optimization of mechanical properties in hybrid fiber-reinforced tertiary blended high-performance concrete. Research on Engineering Structures and Materials. 2025;11(5):2013-2039.
[49] Siddique R. Performance characteristics of high-volume Class F fly ash concrete. cement concrete research. 2004; 34:487-493. https://doi.org/10.1016/j.cemconres.2003.09.002
[50] Thomas M. Optimizing the use of fly ash in concrete. Skokie (IL): Portland Cement Association; 2007.
[51] Patil MSS. Experimental Investigation on Strength and Durability Properties of Concrete by Replacing Natural Sand by Manufacture Sand and Fly Ash. International Journal for Research in Applied Science & Engineering Technology. 2018;6(6):1080-1086. https://doi.org/10.22214/ijraset.2018.6159
[52] Siddique R, Klaus J. Influence of metakaolin on the properties of mortar and concrete: A review. Applied Clay Science.2009;43(3-4):392-400. https://doi.org/10.1016/j.clay.2008.11.007
[53] Pillay D L, Olalusi OB, Kiliswa MW, Awoyera PO, Kolawole JT, Babafemi AJ. Engineering performance of metakaolin based concrete. Cleaner Engineering and Technology.2022;6:100383. https://doi.org/10.1016/j.clet.2021.100383
[54] Zhang M H, Malhotra VM. Characteristics of a thermally activated alumino-silicate pozzolanic material and its use in concrete. Cement and Concrete Research.1995;25(8):1713-1725. https://doi.org/10.1016/0008-8846(95)00167-0
[55] Zhang Y, Ren W, Chen Y, Mi Y, Lei J, Sun L. Predicting the compressive strength of high-performance concrete using an interpretable machine learning model. Scientific Reports. 2024;14(1):28346. https://doi.org/10.1038/s41598-024-79502-z
[56] Paudel S, Pudasaini A, Shrestha RK, Kharel E. Compressive strength of concrete material using machine learning techniques. Cleaner Engineering and Technology. 2023; 15:100661. https://doi.org/10.1016/j.clet.2023.100661
[57] Huang P, Mei X, Sheng H, Li K, Di S, Cui Z. Prediction of Manufactured-Sand Concrete Compressive Strength Using Hybrid ML Models and Dream Optimization Algorithm. Mathematics. 2025;13(23):3792. https://doi.org/10.3390/math13233792