Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Fusion-Based Data-Driven Strength Prediction of Bamboo Fiber-Reinforced Concrete for Sustainable Construction

G. R. Priyasri
SRM Institute of Science and Technology
Suneel Gollapalli
GMRIT Deemed to be University
P. Jagannadha Varma
LENDI Institute of Engineering & Technology
Senthil Kumar Muniasamy
senthilenvtce@gmail.com

Karpaga Vinayaga college of Engineering and Technology
Pothina Praveena
GITAM (Deemed to be University)
Swathi Annam
Kl University

Research Article

Keywords: Bamboo fiber concrete, Machine learning, Compressive strength, Sustainable materials, Predictive modelling

Posted Date: March 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9126769/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Fusion-Based Data-Driven Strength Prediction of Bamboo Fiber-Reinforced Concrete for Sustainable Construction

G. R. Priyasri[1], Suneel Gollapalli[2], P.Jagannadha Varma[3], Senthil Kumar Muniasamy[4,*], Pothina Praveena[5], Swathi Annam[6]

[1]Department of Networking and Communications, SRM Institute of Science and Technology.

[2]Department of CSE, GMRIT Deemed to be University, Rajam, Andhra Pradesh.

[3]Dept of Computer Science, LENDI Institute of Engineering & Technology.

[4]Department of Biotechnology, Karpaga Vinayaga College of Engineering and Technology Padalam, Chengalpattu 603308.

[5]Department of Computer Science and Engineering, GITAM School of Computer Science and Engineering, GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, India

[6]Department of Computer Science and Engineering, Kl University, Vaddeswaram, Guntur, Andhra Pradesh, India

Corresponding Email: senthilenvtce@gmail.com

Abstract

The study develops a machine learning-based predictive framework to estimate the compressive strength of bamboo fiber reinforced concrete (BFRC). A literature-derived dataset was compiled from key parameters influencing BFRC performance, including bamboo fiber dosage (0.25–1.00%), fine aggregate content (646–649 kg/m³), slump values (44.6–98.6 mm), and curing ages of 7, 14, and 28 days. The dataset indicates average compressive strengths of 30.18 MPa, 38.51 MPa, and 47.30 MPa at the respective curing periods, reflecting progressive hydration and densification of the cementitious matrix. Machine learning models were developed to capture nonlinear relationships between mixture parameters, fiber dosage, and strength development. The results demonstrate strong predictive capability, with model performance showing high accuracy in estimating compressive strength across different curing stages. The analysis reveals that bamboo fiber content significantly influences the mechanical behaviour of BFRC. Strength increases with fiber addition up to a moderate dosage due to effective crack bridging and stress redistribution within the matrix, while excessive fiber content reduces workability and compaction efficiency. The findings illustrate an optimal bamboo fiber dosage of approximately 0.75%, which provides improved compressive strength

while maintaining acceptable workability. The predictive framework offers a reliable decision -support tool for optimizing BFRC mix design and promoting sustainable concrete technologies.
Keywords: Bamboo fiber concrete; Machine learning; Compressive strength; Sustainable materials; Predictive modelling

1. Introduction

The construction sector is experiencing increasing pressure to incorporate sustainable materials and low-carbon building techniques as a result of the environmental implications of traditional concrete production. Cement is a major source of carbon dioxide emissions globally, and extracting raw materials for aggregates increases the ecological burden. Therefore, researchers have become interested in developing cementitious composite products containing renewable and bio-based materials, which support sustainability while minimizing structural degradation(Pickering et al. 2016)(Foraboschi 2019). Bamboo fibers appear to be one of the most viable reinforcement materials to date for such applications, primarily due to their rapid renewability, wide availability in tropical climates, relatively low density, biodegradable properties, and favourable tensile strength-to-weight ratio compared to other types of fibers(Terai and Minami 2012)(Xu et al. 2024).

Bamboo fiber reinforced concrete (BFRC) offers the possibility of providing both environmental and mechanical advantages when compared to traditionally produced cementitious composites(Lu et al. 2026)(Khadijah et al. 2024). When uniformly distributed throughout a cementitious composite matrix, individual bamboo fibers function as micro-reinforcement elements that prevent cracking, limit crack growth and improve energy absorption prior to failure(Ahmad et al. 2023)(Ali Kaveh et al. 2025). This crack-arresting ability leads to improved ductility and post-cracking properties and may contribute to greater compressive strength, depending upon how the bamboo fibers interact with the surrounding cementitious matrix(Javadian et al. 2020)(Fu and Liu 2025). Moreover, using bamboo fibers reduces dependence on synthetic fibers or steel fibers, decreases embodied energy, and promotes the reuse of biomass-based materials(You et al. 2020)(Kanagaraj et al. 2025). As a consequence, BFRC appears to be a suitable option for sustainable construction projects in areas of the world where bamboo is readily available.

However, establishing a definitive method for estimating the mechanical properties of BFRC is problematic(Widayanto et al. 2025). The relationship between the dosage of bamboo fibers added to the mixture and the compressive strength of the resulting product is

nonlinear(Kumarasamy et al. 2020)(Lin et al. 2024). Low to moderate amounts of added fibers can produce higher strengths through stress redistribution and crack bridging(Karakus-Zambak and Celik 2025)(Chen et al. 2025). However, if too many fibers are incorporated into the mixture, then fibers will begin to cluster, the mixture will become less workable, it will contain more entrapped air, the bond between fibers and the cementitious matrix will weaken, and the quality of the compaction process will degrade. All of these factors contribute to lower density matrices and subsequent reduced strength. The relationship between fiber content and compressive strength does not follow a straightforward linear or monotonic path; rather, the relationship follows a strengthening to weakening transition path.

The interaction between fiber dosage and other mix design parameters such as water - cement ratio, aggregate proportions, curing time, and workability adds another layer of complexity to the problem. Adding fibers to the mixture typically reduces the slump value of the mixture, which gives rise to a strength-workability trade-off. If the slump value of the mixture is reduced too far, it will be difficult to properly consolidate the mixture, and the amount of voids within the matrix will increase, thus reducing the beneficial properties of the fibers. The type of multi-variable, interdependent relationship between mixture parameters makes it very difficult to predict the strength of BFRC using empirically based models or traditional linear regression models(Dubey et al. 2025)(Si et al. 2025).

Machine Learning (ML) provides a more suitable method for modeling the compressive strength of BFRC(Elshaarawy et al. 2024)(Paudel et al. 2023). ML differs from traditional empirical models in that, rather than having pre-defined equations to model relationships, ML models learn the relationships from the data itself(AliKaveh and Khavaninzadeh 2023)(Li et al. 2023). Therefore, ML models have an advantage over empirical models in representing non-linear behaviour and modeling high-dimensional relationships(Pal et al. 2023)(Alghamdi 2023). Furthermore, ensemble tree-based methods and boosting algorithms are particularly suited to model the complex interactions among various mixture parameters and to find optimal fiber dosing levels for BFRC(Wei et al. 2025)(A Kaveh 2024). These types of methodologies allow for the division of the input feature space into data-driven segments and enable accurate modeling of both strength enhancements at moderate fiber levels and degradations beyond the critical limits(Muliauwan et al. 2020)(Gou et al. 2025).

The datasets of moderate size machine learning models provide a reasonable compromise between predictive accuracy, interpretability, and robustness. Once properly validated, machine learning models can minimize the risk of over-fitting and provide reliable estimates of strength for a wide range of mixture parameters(Chi et al. 2023)(Tan et al. 2026).

Therefore, integrating machine learning models with literature derived data on bamboo fiber reinforced concrete allows for the estimation of both the strength of BFRC and the understanding of the nonlinear behaviour exhibited by bio-based fiber composites(Zheng et al. 2024)(Zhao et al. 2024). A framework of this nature enables the development of evidence-based design recommendations for optimizing the composition of BFRC and contributes to the broader goal of designing low-carbon and sustainable concrete(Amin et al. 2025).

The objectives of the study are:

- To develop machine learning models for predicting the compressive strength of BFRC.
- To evaluate and compare the predictive performance of different machine learning algorithms.
- To identify the optimal bamboo fiber dosage for improved strength and workability.

2. Methodology

The methodological framework adopted to predict the performance of fiber reinforced concrete using machine learning algorithms is illustrated in Figure 1. The methodology begins with a literature-based collection of a dataset, followed by a dataset preparation, including a selection of variables and an identification of the target variable. The machine learning algorithms used in this study include GPR, SR, DKL, SADRA, and TabICL v2. The datasets were split into training and test sets to train and evaluate the different models using cross validation and a variety of performance metrics. The model development involved extensive data analysis of the dataset, including both descriptive statistics (mean, median, mode, standard deviation, variance, range, skewness, kurtosis); correlation analysis (Pearson's r); distribution analysis (histograms and density plots); and assessment of multicollinearity among all the variables in the dataset. Exploratory data analysis(EDA) was used to investigate the relationship between strength, workability, and other properties of the fiber reinforced concrete. After evaluating the performance of the predictive models and assessing their robustness, an engineering interpretation was made to assess how much each of the variables investigated influenced the strength-workability interactions of the fiber reinforced concrete and how sustainable the use of machine learning algorithms is to estimate and interpret the performance of fiber reinforced concrete.

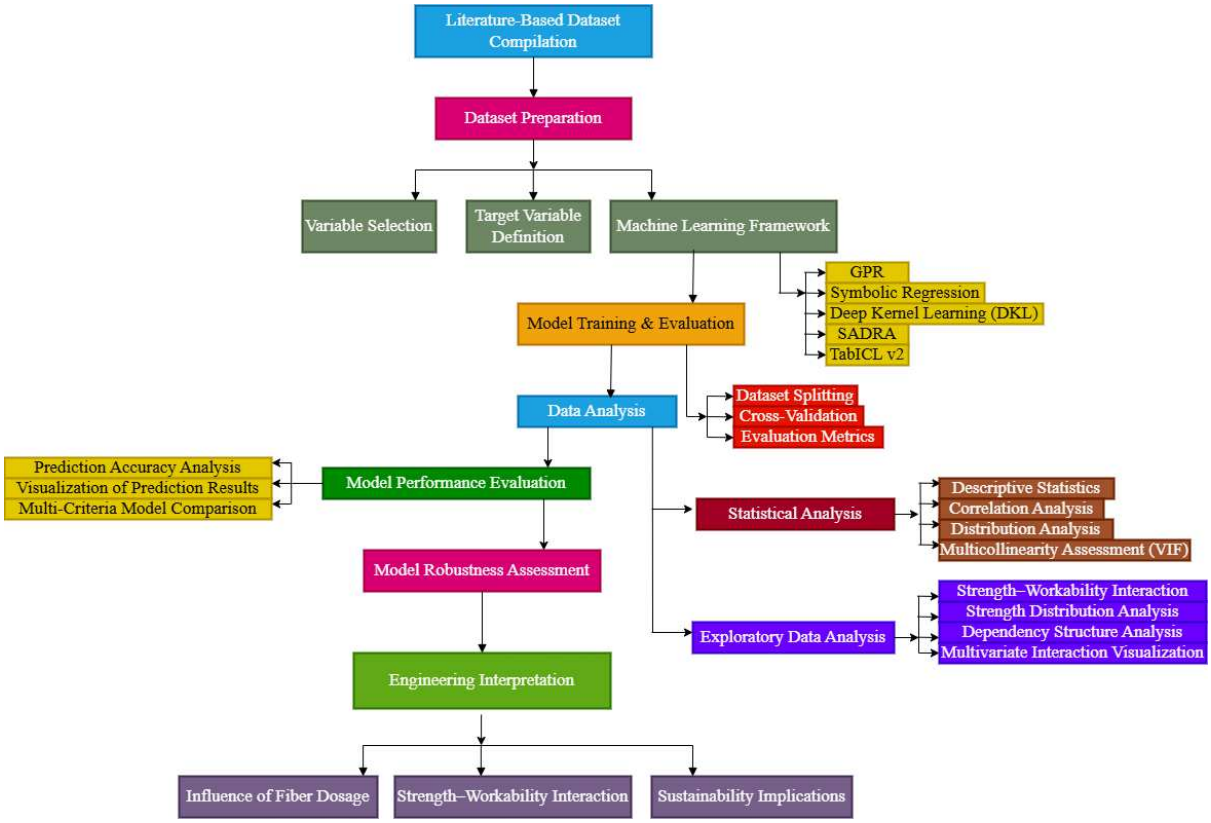


Figure 1: Methodological framework of fiber-reinforced concrete.

2.1 Dataset processing and screening

The literature data used in the study came from previous research studies of BFRC till 2025. Each of these studies reported compressive-strength data at the same three curing ages (7, 14, and 28 days). This strength data was extracted from each of those studies and organized into a consistent format to provide a structure for machine learning analysis. The data represent measured strength values that have been reported by other researchers using various mixture compositions. Input features that were selected to develop a predictive model include: cement content; water content; fine aggregate content; coarse aggregate content; fiber content (bamboo); and curing time. Each of these input variables significantly affects the hydration process, the interaction between fibers and the matrix, and the final mechanical properties of fiber-reinforced concrete. The dataset included mixes with varying amounts of fiber; thus, it is capable of representing enhanced strength at moderate fiber levels, as well as reduced performance due to high fiber dosages. Compressive strength (MPa) was used as the dependent variable for predicting the strength of BFRC. The dataset was reviewed using simple statistical measures to assess the variability and consistency of the data before the development of a predictive model. Following the review, the processed dataset was used to develop and test the proposed machine-learning models to predict the compressive strength of BFRC.

2.2 Machine Learning Framework

An ML approach was used to create an algorithm that could be used to predict the compressive strength of BFRC based on the mixture parameters of materials in the mix; these mixture parameters were taken from the same type of mixtures that had been tested in the literature. The modeling process includes six input parameters: cement content, water content, fine aggregate, coarse aggregate, fiber dosage, and curing age as inputs, with compressive strength being the predicted output. Once the training database has been created, the machine learning approach will apply multiple regression-based algorithms to determine how fiber dosages, workability, and strength progress over time. Each developed model will then be trained and validated by using data partitioning and cross validation so that each model can accurately predict strengths after it has been trained. In order to evaluate and compare the accuracy of the developed models as well as their ability to generalize beyond the data they have been trained on, each developed model will also be evaluated based on established statistical criteria. These models will provide designers with a reliable method for designing mixtures of BFRC.

2.3 Machine Learning Models

2.3.1 Gaussian Process Regression (GPR)

The GPR is a probabilistic type of machine learning which relates input features to output values via a Gaussian Process prior on functions. Unlike most deterministic approaches (which use a single functional relationship to estimate), GPR represents a distribution of possible functions that can be used to fit observed data. The covariance function (or "kernel") utilized by GPR measures how similar different data points are and thus captures some or all of the inherent structure within a given dataset. Once a new input point is specified for prediction, GPR will provide both an estimated mean and a confidence interval associated with this mean value, and therefore provides a probability-based interpretation of the estimation(Xie et al. 2025).

The model is applied to capture the relationship between mixture parameters of bamboo fiber–reinforced concrete and the resulting compressive strength of the concrete matrix. Properties of the matrix are dependent upon complex interactions between fiber content, curing time, and fresh state parameters such as slump. The use of a kernel function, GPR, does not require pre-defined mathematical equations to model these non-linear relationships. Due to the

fact that many material sciences based experimental datasets are relatively small in size, the probabilistic nature of GPR minimizes the potential for over-fitting and improves the likelihood that the model will generalize well to other unseen datasets(Woo et al. 2023). As such, GPR allows for accurate and reliable estimates of compressive strength, as well as the confidence associated with each prediction.

2.3.2 Symbolic Regression

Symbolic regression is an evolutionary modeling process to find mathematical representations of the relationship between variables in data. Unlike traditional regression techniques that rely upon a priori knowledge of how to represent the relationship, symbolic regression finds the structure of the equation and its coefficients at the same time. Symbolic regression is generally accomplished using evolutionary algorithms, specifically genetic programming, that generate new candidate equations by combining mathematical operators, variables, and constants during each generation of the evolution(Le et al. 2025). The models are evaluated based on both the model's ability to predict values from the training set and the model's complexity. The process continues until it converges upon a simple, interpretable mathematical expression that accurately represents the training data.

Symbolic regression has an advantage over purely numerical models, such as those produced with neural networks, in that symbolic regression produces explicit analytical equations that allow engineers to understand the influence of individual factors, such as bamboo fiber content, curing age and workability upon the predicted value of concrete compressive strength. Therefore, symbolic regression bridges the gap between machine learning predictions and traditional engineering modeling approaches(Yan et al. 2026). Symbolic regression will be able to identify the governing relationships in nonlinear systems, such as bamboo fiber-reinforced concrete systems, and develop compact equations that approximate these relationships. These compact equations can then be used in engineering applications, including mix design and engineering calculations.

2.3.3 Deep Kernel Learning (DKL)

Deep Kernel Learning offers the capabilities of a deep neural network to learn the representations of data, combined with the probabilistic modeling strengths of Gaussian Processes. A Deep Neural Network first maps the input variables to a higher dimensional latent feature space, where these learned features are then mapped by a Gaussian Process model to predict outputs. The Neural Network portion of the Deep Kernel Learning framework can be

thought of as the "feature extractor", which learns complex feature transformations in an automatic way. The GP model captures the uncertainty in the predictions and the smooth functional relationship between the features extracted from the input variables and the predicted output(Sun and Du 2024). In this way, the Deep Kernel Learning approach to modeling provides a flexible, yet probabilistically interpretable model, to accurately make predictions based on a set of inputs.

DKL will provide a significant improvement over traditional regression methods when predicting the compressive strength of BFRC. The use of fiber reinforcement in this type of composite material results in non-linear behavior, primarily due to crack bridging, matrix interaction, and reduced workability. Due to the non-linear behavior and complexity of the problem, it is difficult to accurately model the relationships between the input variables using traditional regression techniques. However, DKL is capable of extracting the underlying patterns or relationships among variables such as fiber percentage, slump, and curing age, through its neural network component, and subsequently mapping those relationships in a probabilistic framework via its Gaussian Process model. The combination of both components results in improved predictive accuracy and the ability to generate accurate predictions based on smaller datasets commonly found in materials research.

2.3.4 SADRA (Self-Adaptive Data-Driven Regression Analysis)

The SADRA is a framework for regression that includes optimization to enhance predictive accuracy as well as data-driven learning methods and adaptive model selection to tune model parameters based on the input and output from previous iterations of the algorithm. SADRA improves upon conventional regression techniques by using optimization algorithms to adjust model parameters during each iteration of the process to reduce prediction errors. SADRA uses iterative refinement to optimize the parameters in order to find the optimal parameter configuration to most accurately represent the underlying relationships in the dataset(Shaban et al. 2025). As such, the ability of SADRA to adaptively learn and refine its relationship with the inputs will allow the model to effectively respond to changes in the inputs and non-linear relationships inherent in many types of engineering datasets.

SADRA can effectively determine how both the fiber dosage and curing time affect the compressive strength of concrete. The data-driven and adaptive learning capability of SADRA will allow the model to develop an optimal relationship between the parameters of the mixture and the strength gain without having to rely on pre-defined and rigid functional assumptions. Since the effects of fiber addition can produce threshold dependent behaviors where strength

gains are directly related to dosage up to an optimal point and decrease after this point, SADRA's ability to adaptively and iteratively learn and update its relationship with the inputs and outputs will enable SADRA to more effectively identify and capture these types of non-linear relationships than a static regression model. SADRA will be able to provide reliable predictions of strength performance at varying levels of fiber percentage and curing stage.

2.3.5 TabICL v2 (Tabular In-Context Learning Model)

TabICL v2 was developed as an alternative to more traditional architectures of machine learning, specifically for analyzing tabular datasets (structured). Many other models, such as TabICL uses a transformer-style architecture with an attention mechanism, which enables the TabICL to derive relationships from structured data directly instead of requiring an enormous amount of feature engineering. As opposed to using traditional methods of training, such as batch training, TabICL implements in-context learning, where the model derives patterns from example data within the dataset. As a result of using attention-based layers, TabICL can focus on the most relevant features when making predictions, allowing for the analysis of complex interactions of variables found in many tabular engineering datasets.

TabICL v2 has particular advantages for use in applications of bamboo fiber concrete due to the fact that tabular datasets typically have non-linear dependencies and are comprised of a variety of variable types. Due to the ability of the model's architecture to examine the interaction among fiber content, curing age, slump, and strength development at the same time, it is capable of examining all possible interactions and identifying potential threshold effects or subtle relationships that would not be identified by more traditional regression methods. TabICL v2 has been shown to have strong generalization performance and high predictive accuracy in various materials engineering datasets, and as a result, it could be used for the creation of reliable decision support systems for the sustainable design of concrete mixes.

2.4 Model Training and Evaluation

The machine learning models developed in this study were trained based on the compiled literature database of BFRC mixes. To be able to build reliable models, the compiled literature database was reviewed before model development to ensure that it contained complete data, a consistent format of all variables, and no missing values. Next, the dataset was randomly split into two datasets: a training dataset and a test dataset. The training dataset was utilized to develop the relationship between the input parameters and the compressive strength of the mixes. The test dataset was used as an independent dataset to validate the predictions made by

the models. Separating the data into these two datasets prevents information from leaking to the other side of the model development process and provides a fair method to assess how well a model will generalize to new data.

A reliable way to assess how well the models would perform and minimize the chance of overfitting, cross-validation methods were employed during the training process of the models. Cross-validation is a method in which the training dataset is broken down into k-folds, and then each fold is iteratively selected as the validation set while the remaining k-1 folds are selected as the training sets. The accuracy of the developed models was evaluated using a variety of statistical metrics such as the Coefficient of Determination (R^2), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Weighted Mean Absolute Percentage Error (WMAPE), and the Root Mean Square Error-Observations Standard Deviation Ratio (RSR). These metrics provide a means to assess the accuracy of a model predictions, the magnitude of the errors in a model prediction, and the difference between the model predictions and actual values. Models that produced higher R^2 values and lower error metrics were considered to have greater predictive ability for estimating the compressive strength of BFRC.

- **Coefficient of Determination**

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

- **Root Mean Square Error**

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \ldots\ldots (2)$$

- **Mean Absolute Error**

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \ldots\ldots (3)$$

- **Weighted Mean Absolute Percentage Error**

$$WMAPE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{\sum_{i=1}^{n} y_i} \ldots\ldots (4)$$

- **RMSE–Observations Standard Deviation Ratio**

$$RSR = \frac{RMSE}{SD_{obs}} \ldots\ldots (5)$$

Where:

- y_i= observed compressive strength
- $\hat{y}_i$= predicted compressive strength
- $\bar{y}$= mean of observed values
- n= number of observations

- SD_{obs}= standard deviation of observed values

2.5 Statistical and Exploratory Analysis

Employing an ML model, we performed Exploratory Data Analysis (EDA) to understand the properties and associations existing in our compiled literature dataset. We conducted descriptive statistical calculations (mean, standard deviation, min, max, and Coefficient of Variation) on each of the variables to measure how the data is spread out or varied. Through this initial review, we could assess the range and consistency that existed for the mix design parameters across different curing ages. The knowledge of the statistical properties of the data allows us to confirm that the experimental conditions from the literature data are being accurately represented by the data and ultimately support reliable model development.

Graphical EDA methods were also employed to investigate the patterns of data distribution and possible variable interactions. A variety of visualization tools, including violin plots, box plots, scatter matrices, correlation matrices, and density distributions, were utilized to illustrate central tendency, variability, and potentially non-linear associations between variables. Graphical EDA provides a visual representation of strength development with curing time, the effect of varying levels of bamboo fiber dosages on mechanical performance, and the interaction between compressive strength and workability. The correlation analysis was used to quantify the degree of relationship among variables and to determine if any dependency exists in the dataset. The exploratory review helps us gain an understanding of the underlying behaviour of BFRC and aids in the selection of an ML model that can capture the complexity.

3. Exploratory Data Analysis and Statistical Characteristics

3.1 Descriptive statistical analysis

The descriptive statistical results indicated a level of control in data variability within the literature dataset. The Coefficient of Variation (COV) of Compressive Strength as measured over curing ages was found to be relatively consistent with values ranging from 5.8 % to 6.5 %. Low COV values indicate limited variability in the dataset, reflecting consistent strength development due to progressive hydration and densification of the cementitious matrix, as demonstrated by the increase in average strength from 30.18 MPa at 7 days to 47.30 MPa at 28 days. Although the standard deviations are reasonable, they reflect some measurable stability in variation that can be attributed to differences in fiber dosages as opposed to random or uncontrollable variations in the materials used.

Slump exhibited greater variability as reflected in the COV value of approximately 17.4 %. The significant influence of fiber on the properties of fresh concrete was confirmed by the large range of slump reductions (44.6–98.6 mm). The reduced workability due to increased internal friction and demand for additional surface area contributed by fibers has been documented previously in the literature. Fine aggregate content exhibited minimal variability as reflected in the COV value of 0.002, indicating that the original mix design did not vary greatly except for those adjustments made for the inclusion of fibers. The same degree of dispersion was evident in Fiber-% and Bamboo Fiber (kg) data; therefore, this indicates a deterministic relationship that may result in multicollinearity if the relationships were modeled directly without proper examination. The issue will be further investigated using VIF analysis to develop models that include all of the relevant variables while minimizing multicollinearity issues.

Table 1: Descriptive Statistics of Input and Output Variables

Variable	Mean	Std. Dev	Min	Max	COV
Fiber-%	0.623	0.280	0.25	1.00	0.449
Bamboo Fiber (kg)	2.493	1.119	1.00	4.00	0.449
Fine Aggregate (kg/m³)	647.507	1.119	646	649	0.002
CS-7 days (MPa)	30.179	1.825	26.23	34.36	0.060
CS-14 days (MPa)	38.506	2.232	33.06	44.11	0.058
CS-28 days (MPa)	47.303	3.059	38.76	56.16	0.065
Slump (mm)	70.177	12.240	44.60	98.60	0.174

3.2 Distribution of Input and Output Variables

3.2.1 Statistical Distribution and Density Analysis of Strength and Slump

Figure 2 displays violin plots showing how the statistical distribution and probability density of compressive strength at 7, 14 & 28 days as well as slump values, change as bamboo fiber content increases (0.25%, 0.50%, 0.75% & 1.00%).The violin geometry is based on kernel density estimation of the dataset, while the embedded box plot shows median, inter-quartile range (IQR) & overall dispersion. A combined visualization of central tendency, variability, and symmetry of distribution allows for simultaneous assessment.

The strength of concrete increases consistently with age of curing for all dosages of fiber, showing progressive hydration and densification of cement matrix microstructure. Non-

linear response to fiber content is observed; strength values increase significantly from 0.25% to approximately 0.75% fiber, most notably at 28 days, where both median and density concentration are highest. Improved strength values are due to effective crack bridging action, stress redistribution, and enhanced resistance to propagation of microcracks. At 1.0% fiber content, slight reductions in median strength values and increased dispersions are observed, indicating potential aggregation of fibers, reduced workability, and increased entrapped air, which contribute to impaired compactness of the matrix.

A monotonic decreasing trend of slump values is observed as the percentage of fiber content increases. Lower workability results from increased internal friction between fibers, greater demand for paste coating, and higher surface area of fiber when a high percentage of fiber content is used. The lowered slump values at 0.75% & 1.00% confirm the existence of trade-off between workability and strength of compressive properties. The figure quantitatively demonstrates the threshold dependent behaviour of reinforcement using bamboo fiber, confirms the need for non-linear predictive modeling approaches to accurately model coupled effects of fiber dosage, curing age, and fresh-state properties on compressive strength performance.

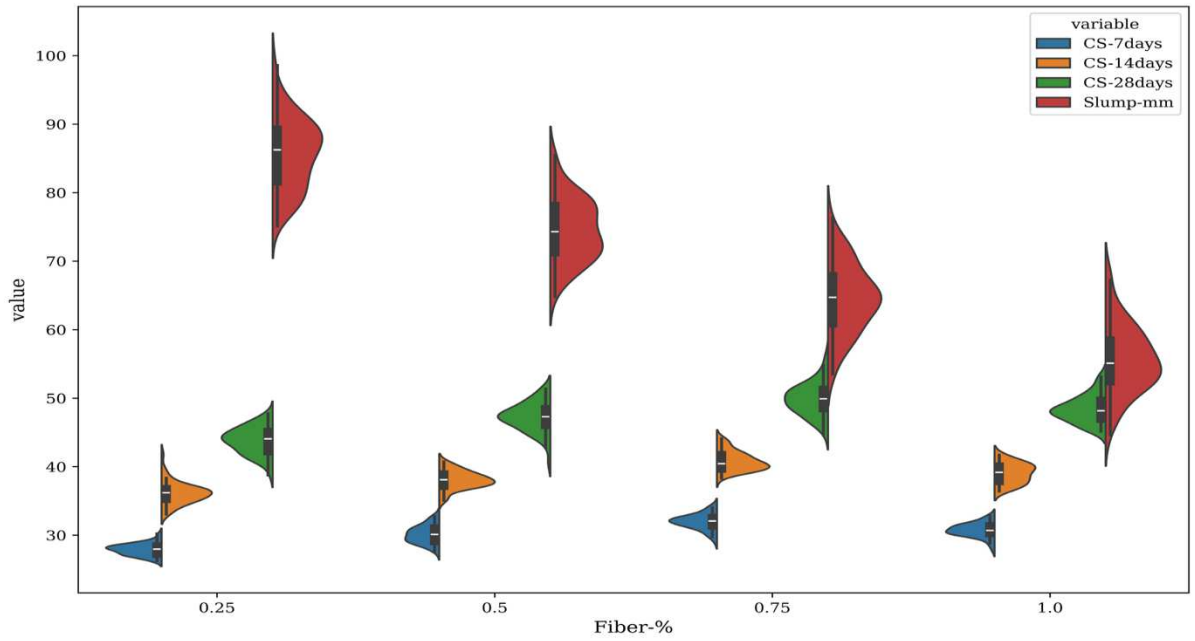


Figure 2: Violin Plot Showing the Distribution of Compressive Strength and Slump

3.2.2 Interquartile Range and Variability Assessment

Figure 3 illustrates, through box plots, the variation in compressive strength (7, 14, and 28 days) and slump as a function of percentage of bamboo fibers used in the mixes (0.25%, 0.50%, 0.75%, 1.00%). The box plots show median compressive strengths, IQRs, whiskers,

and outliers, which provide an effective means to compare the central tendency and variance of compressive strengths as well as slumps for the various percentages of bamboo fibers used. Compressive strengths increase from 0.25% to 0.75% for all three curing periods; however, the highest central tendency is found at 28 days as a result of continued hydration processes. A relatively small IQR for the 0.75% mixture indicates that the mechanical properties of this mixture are consistent and that the fibers are uniformly dispersed throughout the matrix. Conversely, at 1.00%, there is a slight reduction in the median compressive strength compared to the 0.75% mixture, along with a slight increase in the IQR, which can be attributed to decreased matrix compaction and an increase in heterogeneity due to high fiber content.

The median and upper quartile values of the slump values are monotonically decreasing functions of increasing percentage of fibers in the mixtures, which confirms a reduction in workability resulting from the increased internal friction and paste requirements. Although there is some degree of variation and individual outlier values present at the higher percentages of fibers, these variations likely occur as a result of variable interactions between fibers and the surrounding matrix and/or variations in uniformity of mixing. Overall, the use of box plots provides a quantitative basis to support the concept of threshold-dependent mechanical behaviour and the previously discussed strength-workability trade-offs associated with BFRC.

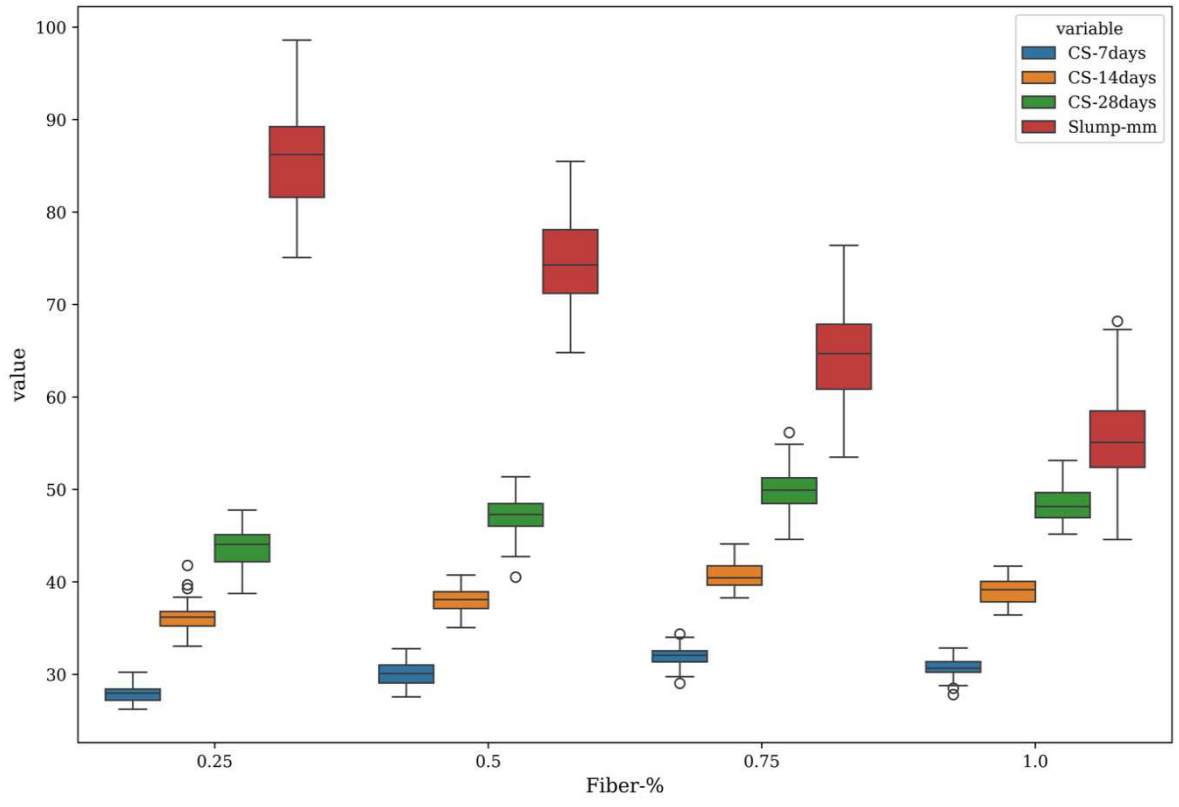


Figure 3: Box plot of central tendency, dispersion, and outliers of compressive strength and slump

3.2.3 Sample Distribution Across Fiber Percentages

The distribution of samples of bamboo fibers present at 0.25%, 0.50%, 0.75%, and 1.00% is shown in Figure 4. There appears to be a uniform distribution of samples among all four percent levels, as each level has contributed approximately one-fourth of the total number of samples measured. A uniformly distributed sample is important because it provides all fiber content levels with equal representation in the dataset; therefore, no one fiber level will dominate any results produced from the models or statistical analysis conducted in the study. Uniformly distributing samples helps to avoid class imbalance effects that may skew model learning toward more frequently occurring inputs. Since there are relatively equal amounts of samples representing each of the fiber levels, the algorithms developed will be able to identify consistent performance trends throughout all fiber content levels to accurately identify any potential non-linear threshold behaviours where performance changes occur over a narrow fiber content range.

The use of balanced datasets also produces better correlations and reduces variability associated with multiple linear regression analyses and cross-validation procedures; this is accomplished by providing similar statistical weights to each fiber content category. Therefore, the balanced nature of the dataset allows for reliable model development and enhances the ability of the overall predictive framework to generalize well to other bamboo fiber–reinforced concretes.

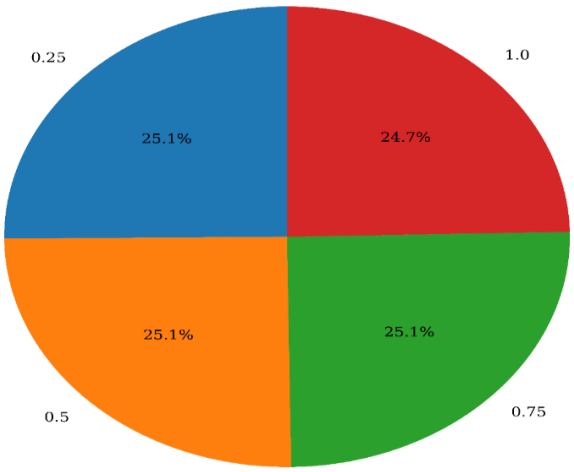


Figure 4: Distribution of samples across different bamboo fiber percentages

3.3 Strength Development with Curing Age

The combined probability density distributions of compressive strength and slump values are shown in Fig. 5 for each of the four fiber dosages (0.25%, 0.50%, 0.75%, and 1.00%) at 7, 14, and 28 days. The use of density curves provides an evaluation of modal response (strength), the degree of dispersion, and the amount of overlap between the different mixes, thus enabling the quantification of both the improvement in performance with time and the increase in variability between mixes as the fiber dosage increases. At 7 days, the mean of the strength distribution clearly increases as the fiber dosage increases from 0.25% to 0.75%, indicating improved early age strength due to improved fiber bridging efficiency and stress redistribution. However, the strength distribution of the 1.00% mix exhibits greater density spread than the 0.75% mix but no corresponding increase in mode, indicating increased heterogeneity as a result of increased fiber agglomerate formation and decreased compactive efficiency. Similar trends continue into the 14 days; however, the hydration process has resulted in an overall increase in strength at all three fiber dosages examined, and the 0.75% mix continues to exhibit the most concentrated and elevated distribution, indicating that the fiber-matrix interaction remains relatively stable over this long term exposure to water.

However, by 28 days, the threshold-dependent response of strength becomes much clearer. An increase in strength occurs with an increase in fiber dosage up to 0.75%, after which there is little or no further increase in strength, and there is actually some increase in the spread of the distribution. The addition of too high an amount of fiber will cause disruption to the

continuity of the matrix and the development of a good bond between fibers and matrix, resulting in no further increase in strength. The slump distributions show a consistently decreasing value of slump with an increase in fiber dosage, indicating a decrease in workability of the fresh concrete due to an increase in the internal frictional resistance of the fibers and an increase in the volume of paste required to achieve the same level of slump as the plain concrete. The combined distributions of slump and strength therefore illustrate a non-linear relationship between strength and workability, and indicate that approximately 0.75% fiber dosage represents an optimum balance between the fresh state properties and the mechanical performance of the fiber reinforced concrete. These results support the need for a non-linear model to be developed for the prediction of the complex relationships that govern the behaviour of BFRC.

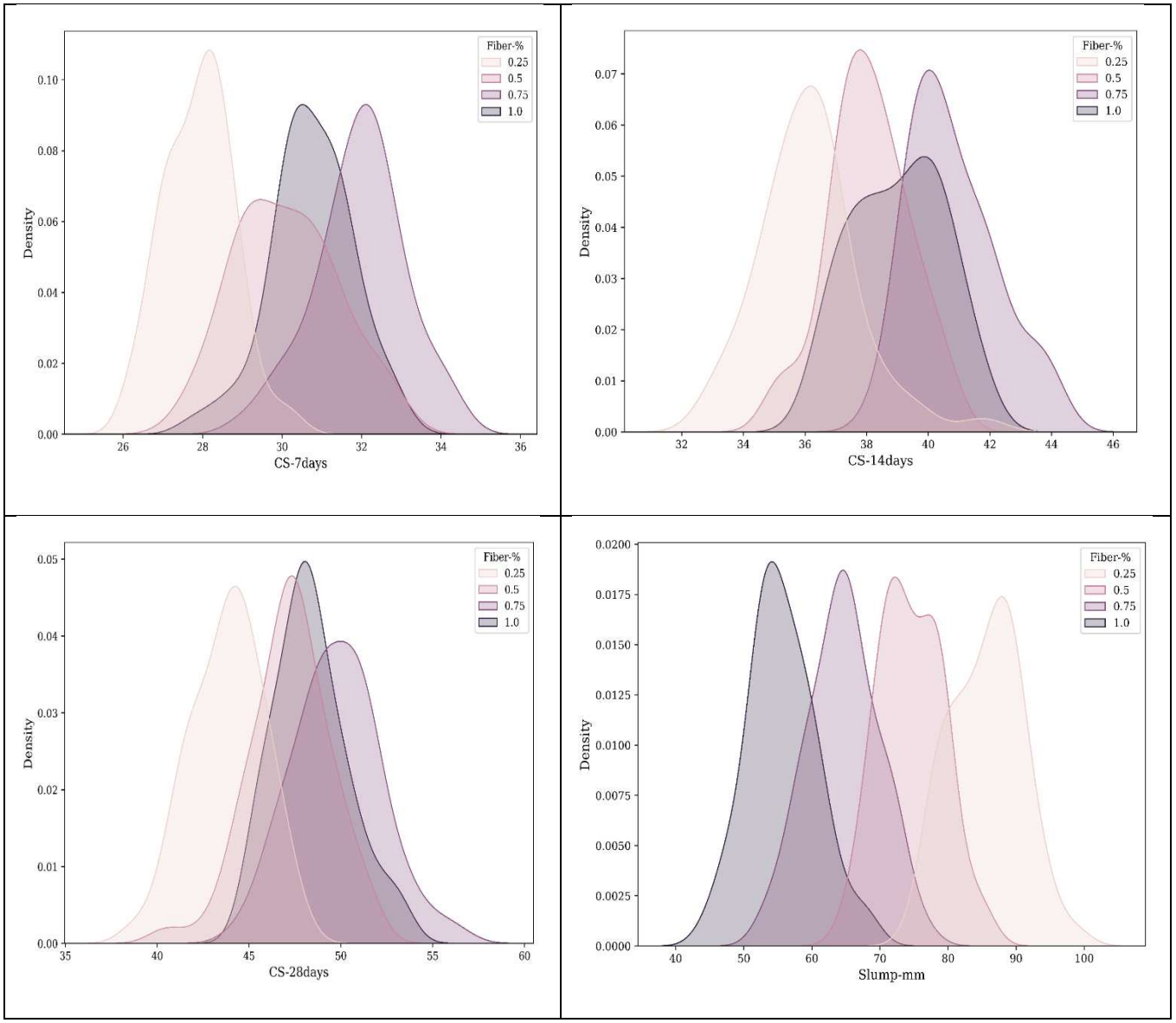


Figure 5: Probability Density Distributions of compressive strength and slump at varying Bamboo Fiber Percentages

3.4 Correlation and Dependency Analysis

3.4.1 Pearson Correlation Matrix Analysis

The Pearson Correlation Matrix in Figure 6 illustrates the linear relationship between bamboo fiber content, compressive strength at different curing times, and slump. Positive correlations were found for compressive strength (r= 0.58-0.64) as a function of fiber content, which indicated that moderate amounts of fibers improve the compressive strength. The positive correlations (r ≈ 0.57-0.60) between 7 day and 14 day and 28 day compressive strength indicate that strength develops consistently over time.

The fiber content has a strong negative correlation with slump (r = –0.91). This confirms that an increase in fiber content results in a significant decrease in the workability of the mixture. The compressive strength also has a moderate negative correlation with slump (r ≈ –0.52 to –0.59). This indicates a strength-workability trade-off associated with the mixture characteristics. The Pearson correlation matrix illustrates the interaction of the three variables and the coupling and partial opposition between the three variables. As a result of the moderate linear correlations along with the previously noted non-linear threshold effect of the interaction of the three variables, it was determined that advanced machine learning models can be applied to capture both the linear and non-linear interactions in the BFRC system.

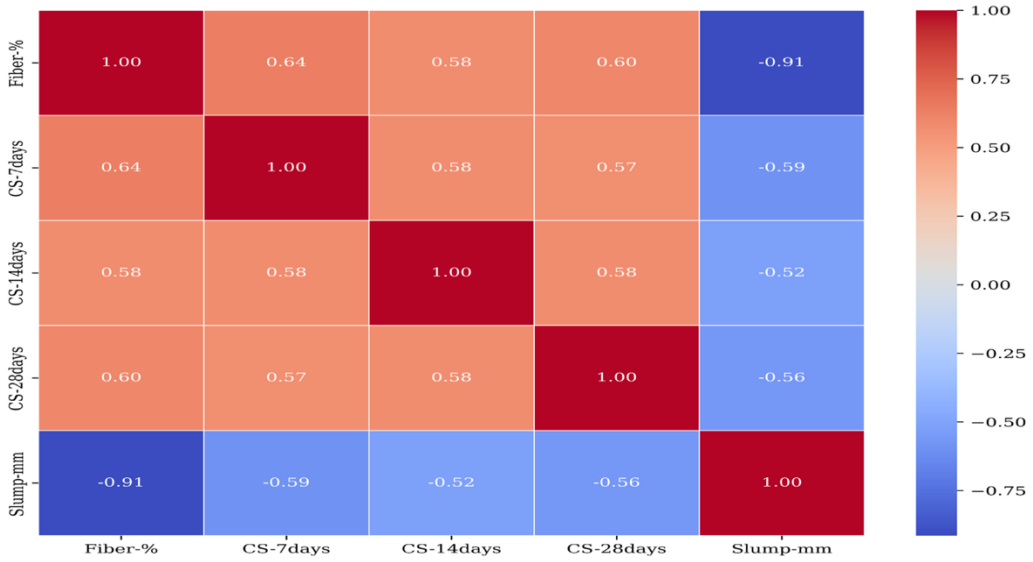


Figure 6: Pearson correlation matrix of bamboo fiber content, compressive strength, and slump

3.4.2 Multivariate Scatter Matrix Analysis

A multivariate scatter plot was used to present the relationships of each pair of variables shown in Figure 7, including compressive strength at varying curing ages, slump, and fiber content. Each diagonal element of the matrix represents the marginal distribution of each variable, and each of the off-diagonal elements of the matrix shows the bivariate scatter pattern for that pair of variables and can be analyzed to determine if there is evidence of linearity or clustering of the data points, as well as other types of interaction trends. The results from the scatter plots show that there is a moderate positive relationship between fiber content and compressive strength up to an intermediate value of fiber content. At this point, it appears that there is a significant amount of clustering of the data around approximately 0.75% fiber content.

The results clearly illustrate a very high degree of positive linear trend relationships among the inter-age strength values. These results confirm that as the material cures over time, the strength of the material consistently increases. As previously stated, the results also show a very high degree of negative association between slump and both fiber content and compressive strength. This result is likely due to a decrease in the workability of the mix as the level of fiber content increases. The matrix also suggests a slight increase in the dispersion of the data at a fiber content of 1.0%, which may be indicative of increased heterogeneity and behavior within the system. The multivariate scatter matrix clearly illustrates the relationships among the mechanical and fresh state properties of the material and clearly shows that the system has both linear and threshold dependent behavior, thus reinforcing the need for a machine learning model that can capture multi-dimensional and non-linear relationships within the BFRC dataset.

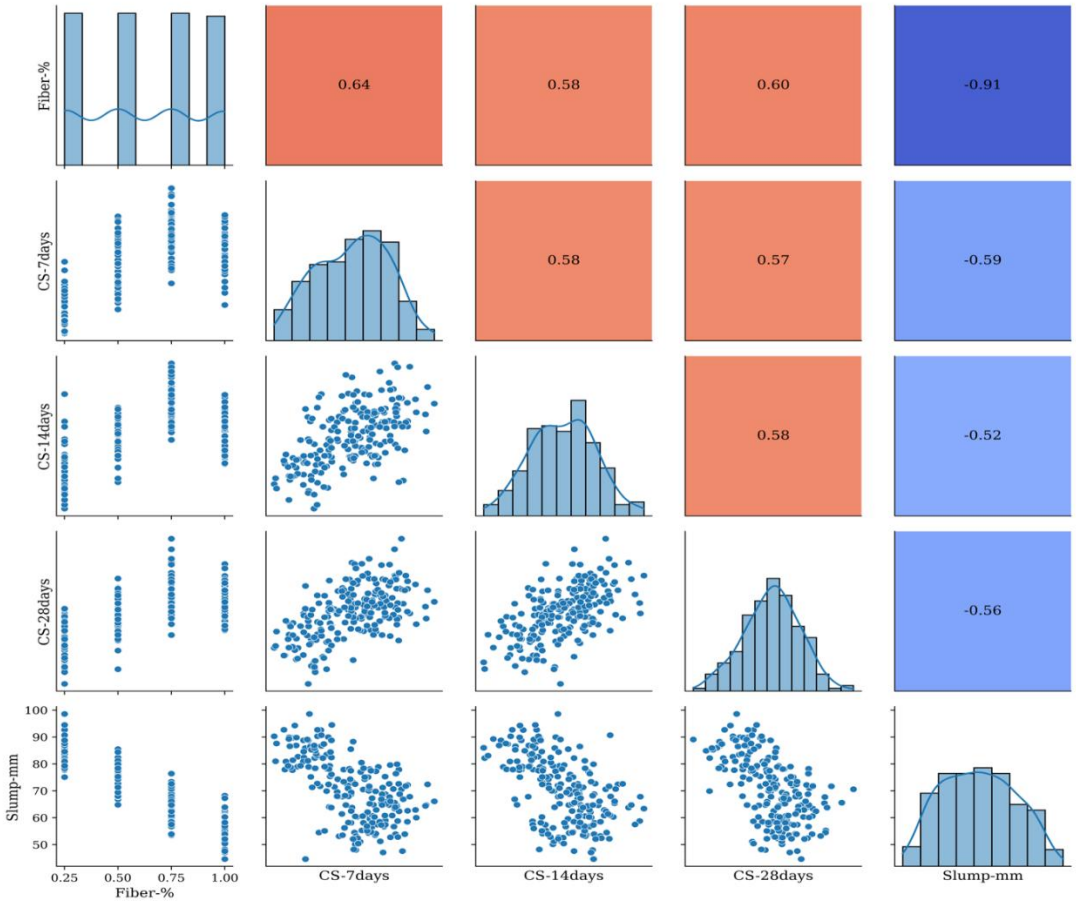


Figure 7: Scatter matrix showing pairwise relationships and marginal distributions

3.4.3 Curing-Age–Dependent Strength Progression

The progress in compressive strength versus curing age (7, 14 and 28 days) for various concentrations of bamboo fibers (0.25%, 0.5%, 0.75%, and 1.0%) is illustrated by Fig. 8. All of the tested mixtures show a progressive increase in strength with increasing curing time, which corresponds to the ongoing hydration of the binder and to the densification of the cementitious matrix. Furthermore, a consistent increase in strength can be observed with increasing fiber content up to 0.75% at every curing time, indicating that the effectiveness of crack bridging and the ability of transferring stresses within the cementitious matrix are both improved. At the same curing time, the mixture containing 0.75% fibers has always the highest strength value, confirming that this is the optimum fiber volume to obtain the best mechanical properties.

Although the mixture with 1% fibers presents higher strength values than those obtained with lower fiber contents, the difference in strength between 1% and 0.75% is not proportional to the difference in fiber content, indicating a decreasing return in terms of the mechanical performances of the mixture with increasing fiber volume. The fact that the strength

development trends of the different mixtures are very close to being parallel indicates that the hydration processes occurring in the different mixtures have similar kinetic characteristics; the differences in the strength values measured for the different mixtures reflect the effects of fiber dosages on the magnitude of the strength. These results confirm the previously identified dependence of the mechanical properties of the hardened concrete on the dose of fibers used during mixing and support the conclusion that there is an optimum range of fiber dosages to be used in order to optimize the mechanical performances of the concrete.

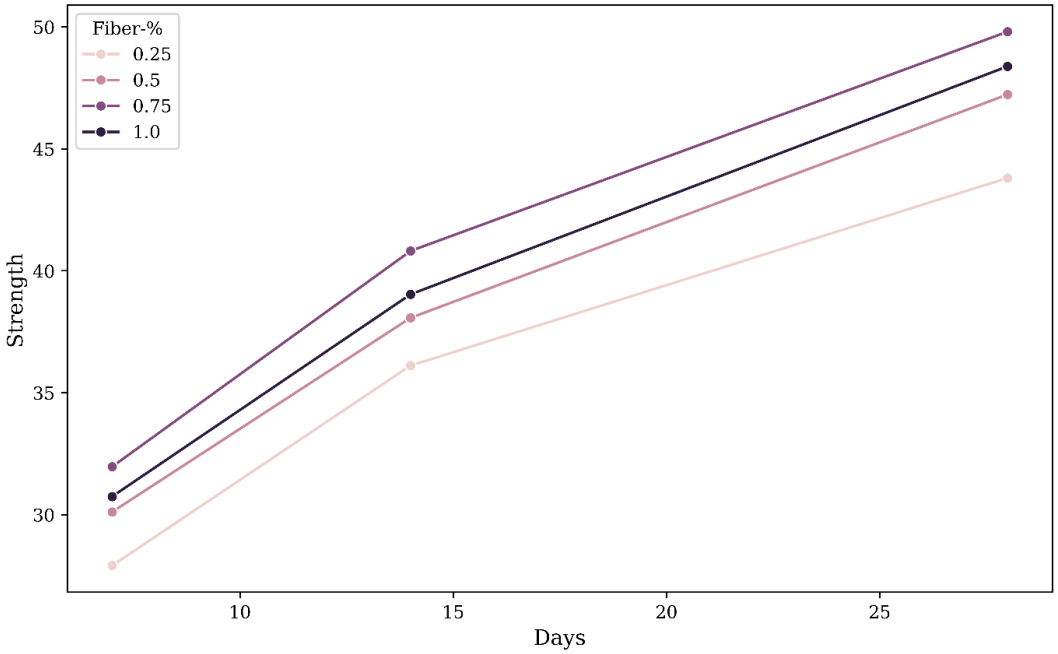


Figure 8: Curing age dependent compressive strength development for different bamboo fiber percentages

3.4.4 Multicollinearity Assessment Using VIF

The VIF analysis found a substantial degree of multicollinearity with respect to selected predictor variables. The infinite VIF values observed for both "Fiber-%" and "Bamboo Fiber (kg)" clearly illustrate a direct, perfect relationship due to the fact that the weight of fiber (in kg) is simply equal to the percent of fiber in the control mix design, thus indicating perfect linear dependence. The very high VIF value obtained for fine aggregate (>1000) is indicative of the deterministic inverse adjustment to compensate for the increase in fiber content while maintaining the overall volumetric ratio of constituents. This type of dependency could lead to instability in the estimation of parameters using regression-based methods and artificially inflate the variance of the coefficients.

The compressive strength variables show relatively low VIF values (<2), indicating acceptable levels of independence and low redundancy. However, there appears to be a moderate degree of multicollinearity (VIF = 6.1) associated with slump; this level of multicollinearity falls within the tolerance limits for many machine learning algorithms, but would likely need to be carefully considered when using linear regression. Therefore, to help achieve numerical stability and to avoid biased learning, redundant variables were removed from the data before model development. The feature selection was performed based on statistical diagnostic results rather than based on an arbitrary decision to exclude certain features, thereby enabling enhanced model reliability and interpretability.

Table 2: Variance Inflation Factor (VIF) Analysis

Feature	VIF
Fiber(%)	∞
Bamboo Fiber (kg)	∞
Fine Aggregate	1049.07
Slump	6.11
CS-7d	1.99
CS-14d	1.83
CS-28d	1.87

Feature selection was conducted using a statistically driven approach to eliminate deterministic and redundant predictors. Variables exhibiting perfect linear dependence were excluded to prevent numerical instability and inflated variance in regression coefficients. Since primary mixture constituents such as cement, water, and coarse aggregate remained constant, the predictive modeling primarily captures the nonlinear interaction between fiber dosage, early-age strength progression, and workability behavior. This dimensional reduction ensures that the developed machine learning models focus on physically meaningful variability rather than artificial statistical dependencies.

3.5 Multidimensional Dependency Structure

3.5.1 Parallel Coordinate Based Multivariate Interaction Analysis

Figure 9 shows a parallel coordinates plot for the multivariable relationship between compressive strength at 7, 14, and 28 days and slump for different percentages of bamboo

fibers added to the concrete (0.25%, 0.5%, 0.75%, and 1.0%). The polyline for each point allows for the simultaneous analysis of how variables interact with one another, as well as to identify patterns of clustering in high dimensional space. The same increasing trend from day 7 to 28 was shown in each of the fiber percentages, and therefore confirms that the increase in strength over time is due to hydration of the concrete.

The 0.75% fiber percentage has the highest number of polylines located in the upper part of the graph, which indicates that these had the best mechanical performance. The 1.0% fiber mixture showed much larger scatter in the slump dimension compared to the other two fiber percentages. This scatter is due to the increased variability or heterogeneity in the fiber distribution and, therefore, the decreased ability of the fibers to be mixed into a homogeneous paste at a fiber dosage of 1.0%. The separation of the polylines between the compressive strength axis and the slump axis clearly illustrates the inverse relationship between mechanical performance and workability. The plot further illustrates how the addition of fibers affects both slump and compressive strength in a coupled and multidimensional manner, further supporting the fact that the behaviour of the BFRC is nonlinear and driven by the interactions of the different components.

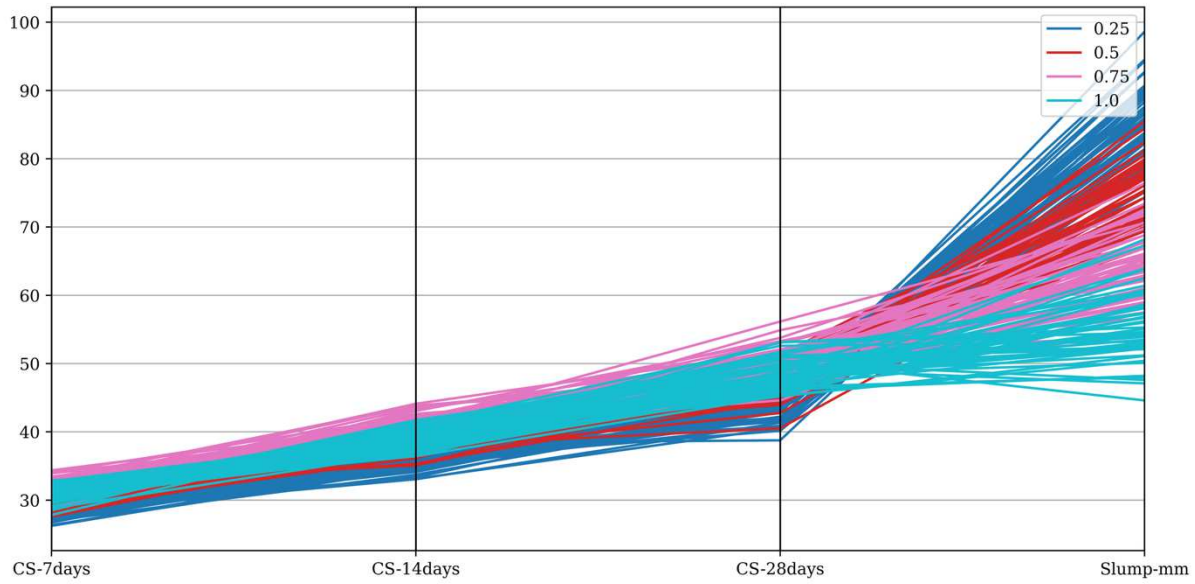


Figure 9: Parallel Coordinate Plot of Multivariate interactions among compressive strength and slump

4. Results and Model Performance Evaluation

4.1 Predictive Performance of ML Models

Comparative evaluation of the machine learning models shows consistent, high levels of prediction for all algorithms with Test R^2 values greater than 0.93. TabICL v2 achieved the

greatest generalization accuracy (Test $R^2 = 0.962$) and had the least error metrics (Test RMSE = 1.035 MPa; Test MAE = 0.823 MPa). This indicates an enhanced ability to capture non-linear relationships in the dataset by TabICL v2. The deep learning approach showed a significant degree of predictive performance, and therefore, it is likely that the latent features extracted via the kernel-based method are effectively modeling the interactions between fibers and strength. The small RSR values (<0.25) across all models are indicative of the reliable predictive nature of each model, whereas the small WMAPE values (<2.2%) are indicative of very little relative deviation. As such, although the Gaussian process regression was able to achieve high training accuracy, as evidenced by the R^2 values, it may be exhibiting some level of mild overfitting relative to the foundational models. These results demonstrate that advanced tabular learning architectures are capable of providing an enhanced representation of the threshold dependent fiber effects and the multi-dimensional interaction patterns in BFRC.

Table 3: Comparative Test Performance of Machine Learning Models

Model	**R^2**	**RMSE**	**MAE**	**WMAPE**	**RSR**
GPR	0.947	1.257	0.987	0.0204	0.230
Symbolic Regression	0.935	1.379	1.046	0.0216	0.254
Deep Kernel Learning	0.959	1.079	0.857	0.0177	0.203
SADRA (Optimized GPR)	0.942	1.312	1.023	0.0211	0.240
TabICL v2	0.962	1.035	0.823	0.0170	0.195

4.2 Visual Assessment of Prediction Accuracy

4.2.1 Bivariate Response Analysis

Figure 10 illustrates the bivariate distribution of 28 day compressive strength that was recorded for each of the four different fiber dosages (0.25% - 1%) as reported by researchers using this literature database. Each point represents one test result from the numerous tests conducted; they represent the two aspects of each test result: the magnitude of the compressive strength and the degree of spread or dispersion within each fiber percentage group. There is a consistent positive trend in strength from the lowest fiber dosage (0.25%) through to the second lowest fiber dosage (0.75%), and then a large number of test results are concentrated at the highest level of strength at the 0.75% fiber dosage level. This trend is indicative of an improvement in the mechanical properties of the cementitious material at a moderate level of fiber addition. The trend also indicates that at lower levels of fiber addition, the material is less

able to redistribute stresses effectively and develop cracks in the cement paste. However, at higher levels of fiber addition (> 0.75%), the material has an improved ability to distribute stresses and resist cracking, resulting in a significant improvement in the mechanical properties of the cementitious material. The systematic trend in strength across the four different fiber percentages, indicated by the cluster pattern, provides evidence of a systematic relationship between the fiber dosage and the mechanical properties of the cementitious material.

At 1% fiber content, while the strength values were similar to those obtained at 0.75%, there was no indication of an increase in strength that would indicate a linear increase in strength as a function of increasing fiber content above 0.75%. In fact, the slightly greater vertical dispersion of the data at the 1% fiber content, when compared to the 0.75% fiber content, indicates an increase in the variability of the strength values. While it is likely that a portion of this increase in variability is due to decreased workability and localized heterogeneities developed in the cementitious matrix at the higher fiber content, it is possible that some of the variation in strength value could be related to the effects of differences in testing procedures. Nonetheless, the pattern of the data does provide evidence of diminishing returns at high fiber contents. Furthermore, the data support an optimum fiber content of approximately 0.75% for achieving maximum 28 day compressive strength.

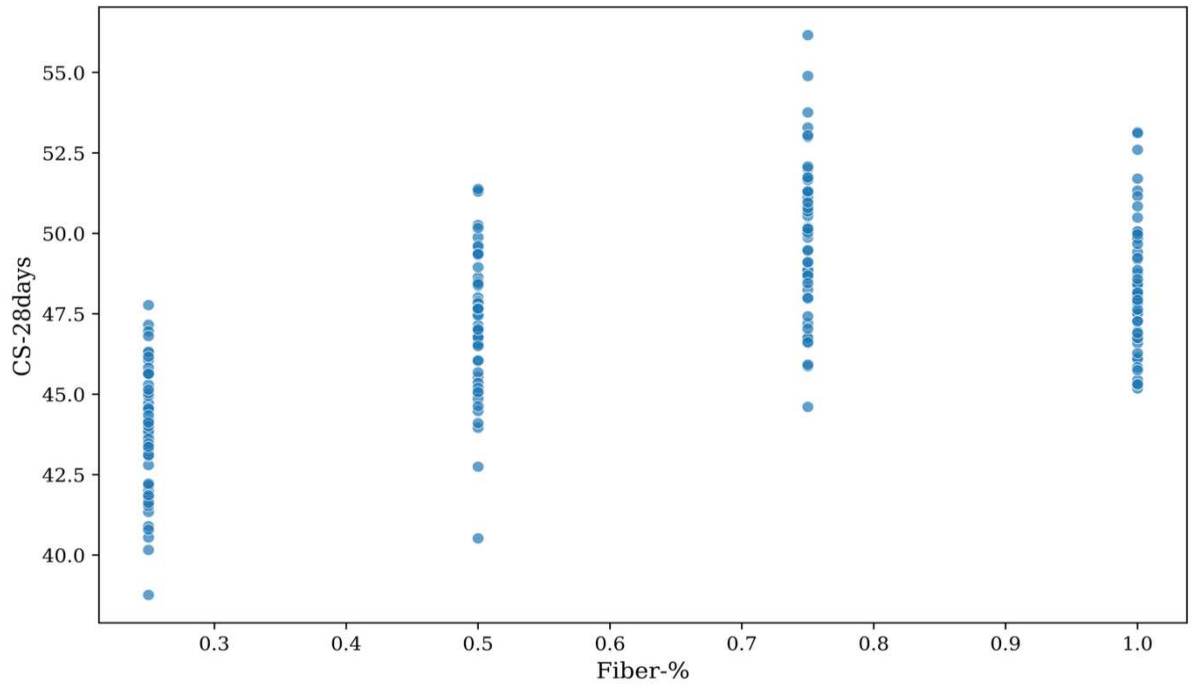


Figure 10: Bivariate distribution of 28-Day compressive strength across different bamboo fiber percentages

4.2.2 Hexagonal Density Distribution Analysis

A hexagonal binning density plot for 28 day compressive strength and slump as shown in Figure 11, displays how all of these properties are distributed together. Instead of plotting individual data points like an ordinary scatter plot, a hexbin density plot groups the individual data points by the number they fall into for their respective hexagon cells, and the colour intensity indicates the density of the data in that particular hexagon cell. Therefore, dark hexagon cells have high densities of data, which will make it easier to see if there are any clusters or relationships between fresh and hard properties.

There is also an inverse relationship between slump and compressive strength as can be seen from the figure. Approximately 80-95 mm slump values are generally associated with low strength, and conversely, high strength (48-52 MPa) is generally associated with low slump values. The greatest amount of data density falls into the middle of both the compressive strength and slump values. In this way, the density distribution shows that there is a practical compromise between the workability of the concrete and its mechanical performance. The density distribution also demonstrates the strength-workability trade off and supports the non-linear dependency structure identified in the earlier exploratory analysis.

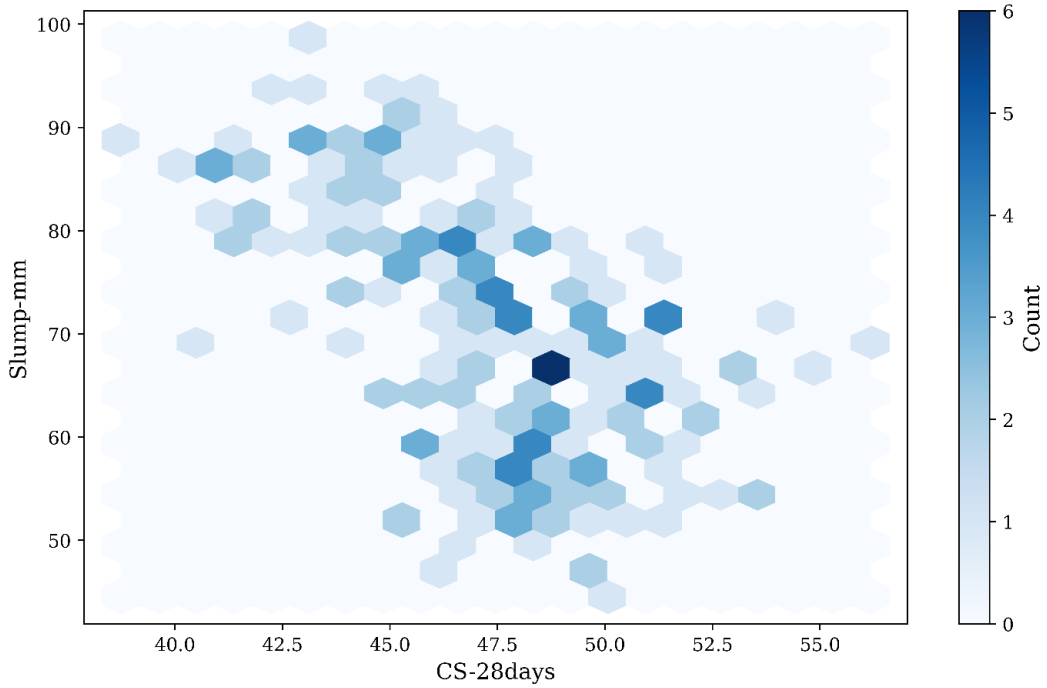


Figure 11: Hexagonal Density Plot Showing the Relationship Between 28-Day Compressive Strength and Slump

4.2.3 Three-Dimensional Strength Progression Visualization

Figure 12 illustrates how 7-day, 14-day and 28 day compressive strength relate to one another using a 3D scatter plot to display the relationship between the compressive strength at

different times and the amount of bamboo fibers used by varying shades of colour in a gradient. Each point represents the data for an individual test, thus allowing for the comparison of the early stage of development of compressive strength with the progress of compressive strength as it develops towards 28 days. An upward trend can be seen when moving from the lowest to the highest curing age, which shows that hydration of the matrix and densification of the matrix will occur progressively throughout all fiber content levels.

Colour variation also shows that the middle range of fiber content levels (around 0.75%), is consistently related to higher compressive strength values during all curing stages. Lower fiber content (0.25%) primarily occupied the lower strength zone, while high fiber content (1.00%) had a large number of data points with no increase in strength as compared to other data points. The progression from 7 days to 28 days, in terms of compressive strength, is almost perfectly linear; however, the vertical separation of the colour zones shows that there is a threshold dependent effect of fiber content on mechanical properties. Thus, this multi-dimensional representation supports the non-linear behaviour and optimal fiber range previously analysed.

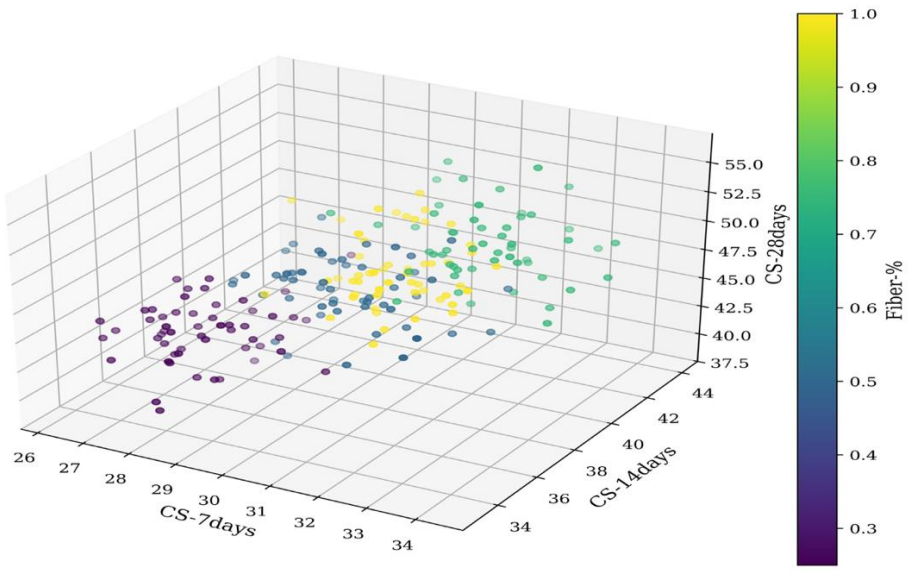


Figure 12. Three-Dimensional Scatter Plot among 7-Day, 14-Day and 28-Day Compressive Strength with Fiber Content

4.3 Multi-Criteria Model Comparison

4.3.1Radar Chart–Based Comparative Analysis

Figure 13 is a radar chart comparing the performance of bamboo fiber mixtures (0.25%, 0.50%, 0.75%, and 1.00%), using the same four response parameters as previously mentioned (7-day compressive strength, 14-day compressive strength, 28-day compressive strength, and

slump). A radial axis for each parameter represents the performance of each parameter, and an enclosed polygon for each fiber dose indicates how it compares with respect to each parameter. Therefore, the radar chart allows for a simultaneous evaluation of both strength development and workability characteristics.

Each enclosed polygon within the radar chart indicates the relative performance for each fiber dose. While the 0.75% fiber mixture displays the largest, or most expansive, enclosed polygon in the strength dimension, this indicates that it has demonstrated the greatest overall compressive performance at every age of cure. Although the 0.25% fiber mixture exhibits the greatest slump value, it also exhibits the lowest values for the three strength parameters, which indicates that it has greater workability but less mechanical performance than the other three fiber dosages. While the 1.00% fiber mixture exhibits the least amount of slump of the four fiber dosages, there was no proportionate increase in the compressive strength of the concrete beyond the 0.75% fiber dosage. Increasing fiber content beyond 0.75% leads to diminishing improvements in compressive strength. The results indicate that the optimal balance of mechanical strength and workability can be achieved by maintaining a fiber content of approximately 0.75%. These results support the findings of previous studies, which indicated that fiber content has a threshold dependent effect on the mechanical properties and workability of high strength concrete.

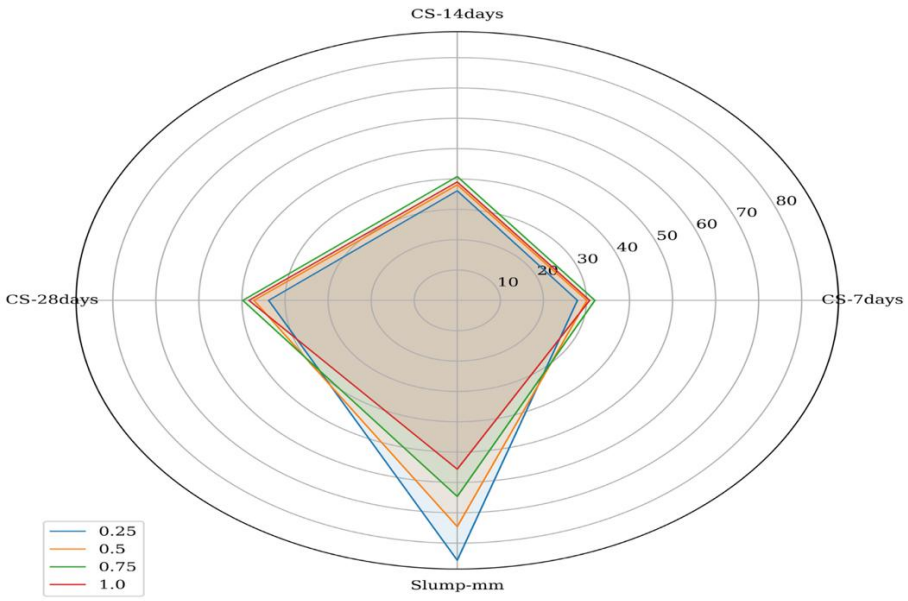


Figure 13: Radar chart comparing compressive strength at different curing ages and slump

4.4 Model Robustness Analysis

The reliability of the developed machine learning model was examined through the evaluation of its prediction stability, the performance metrics, and the validation reliability of the model predictions on both datasets. It is critical to evaluate a predictive model's reliability as it is essential to ensure that the developed model will extract the underlying relationship between the input variables and the output variable. The relatively close agreements in the training and testing performance metrics indicate that there is little evidence of model overfitting or unstable learning behaviour, which supports the reliability of the predictive model. The relatively small range of error metrics among the models further supports the idea that the predictive framework has captured the non-linear interactions between the bamboo fiber dosage, curing time, and the workability of the cement paste.

Cross-validation in this research also added to the reliability of the predictive models by evaluating the performance of the models with respect to different subsets of the dataset. Cross-validation is an iterative process, where each subset of the dataset is used to train and test the models, and then each subset is used to evaluate the performance of the models. This process assures that the models will be stable if they are trained and tested on multiple subsets of the dataset, and that they can generalize the results to other subsets of the dataset. High coefficients of determination and low prediction errors were demonstrated during testing of the models. Therefore, the results demonstrate that the models have made accurate predictions using information outside of the training dataset. The models evaluated, TabICL v2 and Deep Kernel Learning, appear to be the most stable, and therefore, these two models appear to be capable of capturing the complex and threshold dependent behaviour of the bamboo fiber reinforced concrete system. The reliability analysis demonstrates that the developed machine learning models are able to make dependable and repeatable predictions that are beneficial to support performance based mix design.

5. Discussion and Engineering Interpretation

5.1 Influence of Bamboo Fiber Content on Compressive Strength

The findings show that the amount of bamboo fiber affects how well it controls the compressive strength of the concrete by causing microstructural reinforcement(Wang et al. 2026). Compressive strength improvements found at moderate levels of fiber content are mainly due to the ability of the fibers to bridge cracks. Once micro-cracking occurs in the hydrated cement paste, the fibers provide a means of transferring tensile loads across the crack planes and inhibit cracking from spreading by delaying crack coalescence and providing

greater overall integrity of the matrix under compressive loading, ultimately increasing load transfer and thus improving compressive strength(Althoey et al. 2022). Also, the fibers promote local stress redistributions within the cement paste and reduce the magnitude of localized stress concentrations, which contribute to better composite properties for this material.

The effectiveness of the bamboo fibers as reinforcement depends significantly on the dosage level. As fiber contents approach the upper end, there is an increased likelihood of fiber agglomerations and poor dispersion; these conditions may cause disruption to the continuous matrix and the formation of localized voids. High amounts of fiber also produce excessive internal friction during mixing and impede compaction efficiency and weaken the fiber-matrix bond. These factors negate the positive crack bridging capability and limit additional strength improvement above the optimum dosage range. Therefore, the strength response of bamboo fiber-reinforced concrete is threshold-dependent, where mechanical performance will improve up to an intermediate fiber content and stabilize or decline slightly with increasing fiber content. Thus, it emphasizes the need for control of fiber inclusion to achieve an efficient reinforcement while maintaining the proper density of the matrix.

5.2 Strength-Workability Interaction

The relationship between mechanical performance and workability is a major attribute of composite materials made from reinforced concrete containing bamboo fibers. The addition of fiber affects the rheology of fresh state concrete. Bamboo fibers cause increased internal frictional resistance and add new surfaces for the cement paste to coat. Increased fiber dosages will result in the collective effect of increasing the internal frictional resistance and increasing the amount of surface area available for coating; therefore, the mobility of the fresh mix decreases, and there are decreased slump values associated with increased fiber dosages. Decreased workability may make it difficult to properly consolidate the fresh concrete, and the possibility of trapped air in the hardened matrix will increase.

Mechanical properties indicate that moderate (and generally acceptable) decreases in slump will likely have little or no adverse impact on compressive strength because the fiber's role as a structural component improves cracking resistance and transmits stresses. When excessive decreases in slump occur, however, the potential exists for reduced compaction and poor fiber distribution; which results in local porosity and poor quality interfacial bond zones. Such microstructure failures will negate many of the reinforcing functions of the fibers and restrict the ability to improve compressive strength. Thus, it is critical to maintain a reasonable

balance between workability and the ability to reinforce effectively to obtain the maximum compressive strength. The observed interaction indicates that the dosage of fiber should be determined about both mechanical reinforcement and the rheological characteristics of the fresh state material.

5.3 Reliability and Practical Applicability of ML Predictions

ML models were able to predict compressive strength of the bamboo fiber reinforced concrete with high accuracy. Thus, it was possible to determine that the input parameters that were used in the models are sufficient to describe all relevant factors (governing mechanisms) affecting the mechanical properties of this type of material. Good values of coefficients of determination together with small values of prediction error show that models correctly found non-linear relations between fiber content, time of curing, and workability of the mixture. In particular, models that included a more complex, non-linear representation of the analysed problem showed the best results. This suggests that using a data-driven approach, it is possible to model the dependence of the material's mechanical properties on its micro-structure as long as there exists some kind of threshold in this relationship.

The engineering aspect of the developed models' ability to predict the material's compressive strength based on the mixture parameters allows for an advantage in the process of designing new mixtures and evaluating their properties. Using machine learning methods, it is possible to quickly evaluate how different combinations of mixture parameters would affect the predicted value of the compressive strength. A good prediction accuracy in both cases (training and testing sets), indicates that the developed models have a stable generalization behaviour. Therefore, they can be used to predict the compressive strength of similar types of materials under similar environmental and curing conditions. The use of machine learning methods with literature data provides the engineers with practical decision-making tools to optimize the mixture compositions of fiber reinforced concretes and reduce the time spent.

6. Sustainability and Practical Implications

6.1 Resource Efficiency and Material Sustainability

A new way to create sustainable composite building products that are made from bamboo and have less embodied energy than traditional composite products is being developed. Bamboo is a fast-growing plant that requires very little processing energy and is found in abundance in many parts of the world where people live. By using bamboo fibers as part of the

reinforcing materials in concrete, we can partially replace synthetic fibers or steel-based reinforcing materials used in the past. This will lower the amount of embodied energy related to the production of these materials. It will also help to eliminate some of the greenhouse gas emissions associated with transporting large amounts of raw materials from distant locations.

An alternative to traditional materials, using fiber-reinforced composites may provide additional indirect benefits to sustainability due to reduced cement consumption. When bamboo fibers are added to concrete, they improve its ability to resist cracking (post-cracking), and this can result in improved mechanical performance at equal cement content levels. Since better crack control results in fewer cracks forming in concrete, it also extends the structure's lifespan and reduces the need for repairs. The composite products made of bio-based materials can decrease the overall environmental impact of a structure throughout its lifetime.

6.2 Implications for Sustainable Concrete Design

The potential for the use of machine learning combined with literature data is an exciting area for the advancement of the design process for sustainable concretes. ML models allow for the rapid estimation of compressive strength of concretes based on the input parameters of the concrete mixture. This enables the engineer to rapidly compare many different possible combinations of mixture proportions or fiber contents in order to identify the most efficient and cost-effective combinations of these two factors that provide the desired mechanical properties while minimizing excess raw materials. In this regard, ML-assisted modeling has significant potential to reduce the time required to develop sustainable concrete formulations that are environmentally optimal.

The use of machine learning as part of a data-driven approach to model development facilitates the transition to low carbon construction by providing the ability to optimize mix designs on the basis of performance criteria. The predictive frameworks that utilize machine learning can be used to guide the substitution of renewable fibers and other sustainable materials into concrete systems while still ensuring that the engineering reliability of the system is maintained. The combination of materials with advanced predictive modeling techniques may represent a viable means of facilitating the widespread implementation of low emission, resource efficient concrete technology in the development of new and renovated infrastructures.

7. Limitations and Future Scope

The predictive model presented here has been constructed using a collection of data from prior research concerning reinforced concrete fiber made from bamboo fibers. The dataset used in this study was compiled from published literature, which may introduce variability due to differences in experimental conditions, curing environments, and material properties. The amount of variation within the dataset may affect the reliability of the predicted relationships between inputs and outputs from the developed machine learning models, and therefore, the output of the developed models should be viewed with a great deal of caution when predicting outside the range of the parameter values included in the dataset used for training.

Further research will need to provide larger datasets related to all of the above-mentioned properties, including tensile strength, flexural strength, crack resistance, shrinkage, etc., and develop multi-target models to predict both mechanical and durability performance to fully evaluate the potential of fiber-reinforced composites made from bamboo. Developing predictive models and incorporating them into lifecycle assessments and optimization methods, it will further enhance the development of sustainable and high-performance composite materials made from bamboo fibers.

8. Conclusion

The study developed a ML-based predictive framework to estimate the compressive strength of BFRC using a dataset compiled from literature sources. The analysis confirmed that bamboo fiber incorporation significantly influences mechanical performance through fiber-matrix interaction and crack-bridging mechanisms. Strength improvement was observed with increasing fiber dosage up to an intermediate range, beyond which further addition provided no proportional benefit due to reduced workability and potential fiber clustering. The results indicate the presence of an optimal fiber dosage range that balances reinforcement efficiency with matrix compactness.

Comparative evaluation of ML models demonstrated strong predictive capability across all algorithms, with advanced nonlinear models achieving higher accuracy in capturing the complex interactions between mixture parameters and strength development. Among the evaluated approaches, the transformer-based tabular learning model exhibited the highest prediction accuracy and lowest error metrics, indicating superior capability in representing nonlinear and threshold-dependent behaviour within the dataset. Kernel-based models also showed reliable performance, confirming their suitability for modeling material behaviour in moderate-sized datasets.

From a mechanical standpoint, the predictive models are a valuable resource that may be used to estimate the compressive strength of bamboo fiber concrete through a variety of mixture parameter inputs. This data-driven estimation model will allow researchers to quickly test various mixture designs to determine which ones meet design specifications and reduce the need for time-consuming experimental testing in the design optimization phase. Demonstrating the potential of the predictive models as a tool for directing performance-based design of fiber reinforced concrete mixtures, the non-linear relationship demonstrated between the fiber content, workability, and strength development of the mixes demonstrates the ability of these models to capture complex interactions. Furthermore, the integration of bio-based bamboo fibers into machine learning assisted predictions is contributing to the advancement of sustainable concrete technologies. Utilizing bio-based reinforcement materials supports sustainable resource efficient building practices, and utilizing data-driven modeling facilitates informed decisions within the design of materials. Therefore, the integrated experimental and computational approach described in this research serves as a basis for the creation of sustainable, performance optimized, cementitious composite materials.

Author Contributions: G. R. Priyasri: Conceptualization, Data curation, Methodology, Writing – original draft preparation. Suneel Gollapalli: Software development, Machine learning model implementation, Data analysis, and Visualization. P. Jagannadha Varma: Formal analysis, Validation, Investigation, and interpretation of results. Senthil Kumar Muniasamy: Supervision, Project administration, Methodology design, Writing – review and editing. Pothina Praveena: Data processing, Statistical analysis, Literature review, Validation. Swathi Annam: Visualization, Graphical representation, Editing, and proofreading. All authors have read and approved the final version of the manuscript.

Funding: No funding was provided for the study.

Data availability statement: Data will be made available on request by the corresponding author.

References

Ahmad, J., Zhou, Z., & Deifalla, A. F. (2023). Structural properties of concrete reinforced with bamboo fibers: a review. *Journal of Materials Research and Technology*, *24*, 844–865. https://doi.org/https://doi.org/10.1016/j.jmrt.2023.03.038

Alghamdi, S. J. (2023). Prediction of Concrete's Compressive Strength via Artificial Neural Network Trained on Synthetic Data. *Engineering, Technology & Applied Science Research*, *13*(6), 12404–12408. https://doi.org/10.48084/etasr.6560

Althoey, F., Awoyera, P. O., Inyama, K., Khan, M. A., Mursaleen, M., Hadidi, H. M., & Najm, H. M. (2022). Strength and microscale properties of bamboo fiber-reinforced concrete modified with natural rubber latex. *Frontiers in Materials*, *Volume 9-2022*. https://doi.org/10.3389/fmats.2022.1064885

Amin, F., Ali, I., Husnain, A., Javed, M. F., Alabduljabbar, H., & Junaid, A. (2025). Sustainable strengthening of concrete deep beams with openings using ECC and Bamboo: An equation and data-driven approach through abaqus modeling and GEP. *Results in Engineering*, *26*, 104813. https://doi.org/https://doi.org/10.1016/j.rineng.2025.104813

Chen, W., Tian, Z., Yuan, Q., Zeng, Y., Zou, Z., Ma, H., & Ye, Z. (2025). Experimental investigation on shear cracking resistance behavior of bamboo fiber reinforced concrete composite beams with steel wire mesh and basalt fiber reinforced polymer bar. *Structures*, *79*, 109595. https://doi.org/https://doi.org/10.1016/j.istruc.2025.109595

Chi, L., Wang, M., Liu, K., Lu, S., Kan, L., Xia, X., & Huang, C. (2023). Machine learning prediction of compressive strength of concrete with resistivity modification. *Materials Today Communications*, *36*, 106470. https://doi.org/https://doi.org/10.1016/j.mtcomm.2023.106470

Dubey, S., Gupta, D., & Mallik, M. (2025). Extreme learning machine to foretell the deflection in plain cement, steel-reinforced, and bamboo-reinforced concrete beams. *Emerging Materials Research*, *14*(1), 12–27. https://doi.org/https://doi.org/10.1680/jemmr.24.00046

Elshaarawy, M. K., Alsaadawi, M. M., & Hamed, A. K. (2024). Machine learning and interactive GUI for concrete compressive strength prediction. *Scientific Reports*, *14*(1), 16694. https://doi.org/10.1038/s41598-024-66957-3

Foraboschi, P. (2019). Masonry does not limit itself to only one structural material: Interlocked masonry versus cohesive masonry. *Journal of Building Engineering*, *26*, 100831. https://doi.org/https://doi.org/10.1016/j.jobe.2019.100831

Fu, X., & Liu, P. (2025). Mechanical properties of raw bamboo fiber- reinforced phosphogypsum-based composite cementitious materials and their strengthening mechanism, 1–19. https://doi.org/10.1371/journal.pone.0323011

Gou, J., Zaman, A., & Farooq, F. (2025). Machine learning-based prediction of compressive

strength in sustainable self-compacting concrete. *Engineering Applications of Artificial Intelligence*, *161*, 112190. https://doi.org/https://doi.org/10.1016/j.engappai.2025.112190

Javadian, A., Smith, I. F. C., & Hebel, D. E. (2020). Application of Sustainable Bamboo-Based Composite Reinforcement in Structural-Concrete Beams : Design and Evaluation.

Kanagaraj, B., Sohliya, A., Jayakumar, G., Anand, N., & Lubloy, E. (2025). Performance evaluation of bamboo reinforced concrete slab: A comparative study between fibre reinforced mix and shape optimised composite. *Cleaner Waste Systems*, *12*, 100409. https://doi.org/https://doi.org/10.1016/j.clwas.2025.100409

Karakus-Zambak, O., & Celik, O. C. (2025). Behavior of bamboo reinforced concrete (BRC) beams under monotonic and dynamic loads. *Construction and Building Materials*, *458*, 139683. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2024.139683

Kaveh, A. (2024). Efficient Training of Two ANNs Using Four Meta-Heuristic Algorithms for Predicting the FRP Strength (pp. 359–388). https://doi.org/10.1007/978-3-031-66051-1_19

Kaveh, Ali, & Khavaninzadeh, N. (2023). Efficient training of two ANNs using four meta-heuristic algorithms for predicting the FRP strength. *Structures*, *52*, 256–272. https://doi.org/https://doi.org/10.1016/j.istruc.2023.03.178

Kaveh, Ali, Shahini, M. H., Zhang, X., Xu, W., Firouzabadi, P. Z., Zhang, B., & Meng, W. J. (2025). From a notch to a crack: Monitoring crack initiation from a notch through cyclic bending of microscale cantilevers. *Materials Science and Engineering: A*, *931*, 148187. https://doi.org/https://doi.org/10.1016/j.msea.2025.148187

Khadijah, S., Osmi, C., Husen, H., Othman, M., Nor, N. M., Sojipto, S., et al. (2024). Performance of Bamboo Fibre on Modifying Mechanical Properties of Concrete Abstract :, 1–9. https://doi.org/10.2174/0118741495288340240312050220

Kumarasamy, K., Shyamala, G., & Gebreyowhanse, H. (2020). Strength Properties of Bamboo Fiber Reinforced Concrete. *IOP Conference Series: Materials Science and Engineering*, *981*, 32063. https://doi.org/10.1088/1757-899X/981/3/032063

Le, B. A., Nguyen, H. Q., Tran, B. V., Vu, T. S., & Bui, T. L. (2025). Symbolic regression model for predicting compression strength of steel fiber-reinforced concrete. *Transportation Research Procedia*, *85*, 94–101. https://doi.org/https://doi.org/10.1016/j.trpro.2025.03.138

Li, D., Tang, Z., Kang, Q., Zhang, X., & Li, Y. (2023). Machine Learning-Based Method for Predicting Compressive Strength of Concrete. *Processes*.

https://doi.org/10.3390/pr11020390

Lin, Y., Ding, M., Wang, L., & Wei, Y. (2024). Experimental study on mechanical properties of raw bamboo fibre-reinforced concrete. *Case Studies in Construction Materials*, *21*, e04003. https://doi.org/https://doi.org/10.1016/j.cscm.2024.e04003

Lu, J., Wei, Y., Yan, Z., Du, H., Chen, J., & Li, G. (2026). Experimental and numerical study on the flexural behavior of bamboo-concrete composite beams reinforced with bottom-anchored rebars. *Construction and Building Materials*, *506*, 144921. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2025.144921

Muliauwan, H., Prayogo, D., Gaby, G., & Harsono, K. (2020). Prediction of Concrete Compressive Strength Using Artificial Intelligence Methods. *Journal of Physics: Conference Series*, *1625*, 12018. https://doi.org/10.1088/1742-6596/1625/1/012018

Pal, A., Ahmed, K. S., Hossain, F. M. Z., & Alam, M. S. (2023). Machine learning models for predicting compressive strength of fiber-reinforced concrete containing waste rubber and recycled aggregate. *Journal of Cleaner Production*, *423*, 138673. https://doi.org/https://doi.org/10.1016/j.jclepro.2023.138673

Paudel, S., Pudasaini, A., Shrestha, R. K., & Kharel, E. (2023). Compressive strength of concrete material using machine learning techniques. *Cleaner Engineering and Technology*, *15*, 100661. https://doi.org/https://doi.org/10.1016/j.clet.2023.100661

Pickering, K. L., Efendy, M. G. A., & Le, T. M. (2016). A review of recent developments in natural fibre composites and their mechanical performance. *Composites Part A: Applied Science and Manufacturing*, *83*, 98–112. https://doi.org/https://doi.org/10.1016/j.compositesa.2015.08.038

Shaban, W. M., Shen, S.-L., Ashour, A. G., Zhou, A., & Elbaz, K. (2025). Physics-based self-adaptive algorithm for estimating the long-term performance of concrete shrinkage. *Engineering Applications of Artificial Intelligence*, *142*, 109945. https://doi.org/https://doi.org/10.1016/j.engappai.2024.109945

Si, Q., Zhang, W., Pan, Z., Zheng, J., Yu, C., & Zhang, G. (2025). Study on the mechanical and durability properties of 3D-printed bamboo fiber-reinforced concrete. *Construction and Building Materials*, *478*, 141464. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2025.141464

Sun, B., & Du, S. (2024). Kernel extreme learning machine and finite element method fusion fire damage prediction of concrete structures. *Structures*, *68*, 107172. https://doi.org/https://doi.org/10.1016/j.istruc.2024.107172

Tan, X., Xing, J., Wang, Y., Qiu, H., Mahjoubi, S., & Guo, P. (2026). Explainable machine

learning for predicting compressive strength of rubberized concrete: SHAP interpretation, lifecycle assessment, and design recommendations. *Journal of Cleaner Production*, *538*, 147338. https://doi.org/https://doi.org/10.1016/j.jclepro.2025.147338

Terai, M., & Minami, K. (2012). Basic Study on Mechanical Properties of Bamboo Fiber Reinforced Concrete. *IABSE Symposium Report*.

Wang, Y., Wei, Y., Wei, B., Li, Q., Zhu, B., & Ding, M. (2026). Mechanical and microstructural properties of bamboo fiber-reinforced bamboo aggregate ceramsite concrete. *Construction and Building Materials*, *514*, 145556. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2026.145556

Wei, B., Wei, Y., Yi, J., Chen, J., Lin, Y., & Ding, Y. (2025). Machine learning-driven prediction of axial capacity of concrete-filled steel tubes columns with built-in bamboo or timber cores. *Engineering Structures*, *341*, 120817. https://doi.org/https://doi.org/10.1016/j.engstruct.2025.120817

Widayanto, E., Wiswamitra, K. A., & Subchan, B. R. (2025). Flexural behaviour of bamboo-reinforced concrete beams soaked in salt solution and molasse treatments and wrapped in fibreglass tape: An experimental and numerical study. *Advances in Bamboo Science*, *11*, 100147. https://doi.org/https://doi.org/10.1016/j.bamboo.2025.100147

Woo, B.-H., Ryou, J.-S., Kim, J. Y., Lee, B., Gi Kim, H., & Kim, J.-S. (2023). Freeze-thaw durability estimation for concrete through the Gaussian process regression with kernel convolution. *Construction and Building Materials*, *400*, 132825. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2023.132825

Xie, L., He, J., Wang, C., Zhu, S., Zhang, H., & Yang, X. (2025). Enhanced bayesian gaussian process regression for compressive strength prediction of multi-binder concrete. *Journal of Building Engineering*, *111*, 113308. https://doi.org/https://doi.org/10.1016/j.jobe.2025.113308

Xu, B., Tian, R., Wang, Y., Zhang, Z., & Zhang, Z. (2024). Preparation and Properties of Natural Bamboo Fiber-Reinforced Recycled Aggregate Concrete. *Materials*. https://doi.org/10.3390/ma17122972

Yan, S., Liu, Y., Lin, J., & Liang, Y. (2026). Deep learning physical symbolic regression models for chloride ion transport in cracked concrete. *Journal of Building Engineering*, *117*, 114763. https://doi.org/https://doi.org/10.1016/j.jobe.2025.114763

You, A., Be, M., & In, I. (2020). The Use of Bamboo Fiber in Reinforced Concrete Beam to, *020003*(September 2017). https://doi.org/10.1063/1.5003486

Zhao, H., Tang, J., Li, Z., Zhou, T., & Xiong, T. (2024). Experimental study on the static

properties of bamboo fiber reinforced ultra-high performance concrete (UHPC). *Construction and Building Materials*, *453*, 138974. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2024.138974

Zheng, X., Ban, Y., Chen, Z., Fu, T., Fei, M., Liu, W., & Qiu, R. (2024). Modification of bamboo fiber for reinforcing cement-based composites and durability improvement. *Construction and Building Materials*, *451*, 138802. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2024.138802