

Journal of Materials and Engineering Structures

Research Paper

Soft Computing Approaches for Tensile Strength Prediction of Sustainable Concrete Mixtures

Rawa Aree Jamal^a, Taghreed Khaleefa Mohammed Ali^{b,*}, Halo Rebwar Nor^a, Sivand Sarkawt Saeed^a, Ahmed Salih Mohammed^{c, d}

^a Civil Engineering Department, College of Engineering, the American University of Iraq Sulaimani, Kurdistan, ACI-Kurdistan Chapter (AUIS-Student Chapter).

^b Department of Architecture Engineering, Faculty of Engineering, Koya University, Danielle Mitterrand Boulevard, Koya KOY45, Kurdistan Region – F.R. Iraq.

^c Civil Engineering Department, College of Engineering, University of Sulaimani, Kurdistan, Iraq.

^d American University of Iraq, Sulaimani, Iraq.

ARTICLE INFO

Article history :

Received : 6 August 2025

Revised : 20 September 2025

Accepted : 10 November 2025

Keywords:

Eco-friendly concrete

Tensile Strength

Soft Computing Models

Supplemental Cementitious Materials

ABSTRACT

Conventional concrete consumes large amounts of cement and aggregates, contributing to environmental degradation. This study investigates eco-friendly concrete mixtures incorporating industrial by-products such as fly ash, ground granulated blast furnace slag (GGBFS), and granite powder to enhance tensile strength and promote sustainable construction. A dataset of 252 experimental samples was analyzed using five statistical models—Linear Regression (LR), Pure Quadratic (PQ), Logarithmic (LOG), Interaction (INT), and Full Quadratic (FQ)—to predict tensile strength based on eight input parameters, including curing time, pull-off layer thickness, and mix ratios. Among the models, the FQ model exhibited the highest accuracy with the lowest residual error, followed by the INT model. Sensitivity analysis identified curing time, pull-off layer thickness, and the cement-to-sand ratio as key parameters affecting tensile strength. Results confirm that optimizing the water-to-cement and cement-to-sand ratios can significantly improve the performance of eco-friendly concrete. Overall, this study demonstrates that incorporating industrial by-products into concrete mixtures effectively balances mechanical performance with environmental sustainability, supporting the transition toward greener construction practices.

1 Introduction

Due to its durability, accessibility, and adaptability in construction materials globally, conventional concrete significantly contributes to environmental deterioration and the depletion of natural resources, as it consists primarily of 80% aggregates and approximately 12% cement by weight [1]. This output level raises significant sustainability concerns, including the

* Corresponding author. Tel.: +964 7714729728.
E-mail address: taghreed.khaleefa@koyauniversity.org

overexploitation of natural aggregates, substantial energy demands in cement manufacturing, and increased greenhouse gas emissions [2]. The construction industry is undergoing a transformative shift toward sustainability, driven by the urgent need to reduce its environmental impact. Traditional concrete production accounts for 8% of global CO₂ emissions, primarily due to the energy-intensive manufacturing of Portland cement, which requires 1.5 tons of raw materials per ton of cement and emits 0.9 tons of CO₂ [3]. In response, researchers and industry leaders are developing eco-friendly alternatives that utilize industrial byproducts, advanced technologies, and principles of the circular economy. One of the most promising developments is green concrete, which incorporates supplementary cementitious materials (SCMs) such as fly ash (FA), ground granulated blast-furnace slag (GGBFS), and waste granite powder (GP) to replace a portion of the cement. These materials not only reduce CO₂ emissions by 20–40% but also enhance the durability and workability of concrete. For instance, fly ash improves long-term strength and reduces permeability, while GGBFS offers superior resistance to sulfate attacks, making it ideal for marine environments [4]. Waste granite powder, a byproduct of stone processing, enhances surface adhesion and reduces porosity in flooring applications, demonstrating how industrial waste can be repurposed effectively [4]. Recycled materials are also playing a crucial role in sustainable concrete production. Recycled coarse aggregates (RCA) from demolished structures can replace up to 60% of natural aggregates, maintaining compressive strengths of 35–40 MPa while reducing landfill waste [4]. Pretreatments, such as pozzolan slurries or sodium silicate coatings, improve RCA's density and water absorption, addressing one of its primary limitations. Similarly, recycled plastic concrete incorporates waste plastics as lightweight aggregates, offering a dual solution to plastic pollution and reducing the high carbon footprint of construction [5]. A lower W/C ratio leads to denser microstructures, reducing capillary voids and enhancing both compressive and tensile strength due to improved interfacial bonding between the cement paste and aggregate [6]. Conversely, higher W/C ratios, while improving workability, increase pore volume and reduce tensile capacity, especially in early-age concrete [7]. Curing time is another crucial factor, as it directly affects the hydration of cementitious materials and the formation of calcium silicate hydrate (C-S-H), which contributes to strength gain. Prolonged curing, typically lasting up to 28 days, enables deeper hydration, resulting in enhanced tensile strength and overall durability of eco-concretes [8]. Inadequate curing, on the other hand, can result in poor tensile behavior due to incomplete hydration and increased microcracking [9]. An optimal C/S ratio enhances cohesion and contributes positively to tensile strength, while excessively low or high ratios may compromise workability or introduce micro-defects in the matrix [10]. Balancing C/S alongside W/C is critical in mix design to ensure that both strength and sustainability goals are met in eco-friendly concrete formulations [10]. These materials, typically by-products of industrial processes, not only improve environmental performance but also introduce complex interactions that affect the behavior of concrete, especially its tensile strength, which is crucial for long-term structural integrity and crack mitigation [5]. Historically, tensile strength has been the primary performance criterion in concrete design; however, tensile strength is equally crucial for resisting cracking and failure under flexural or oblique loads. Estimating tensile strength is more complex for ecologically friendly concrete systems due to the nonlinear interactions among variables, including curing time, water-to-cement ratio, and mineral additives, which arise from material replacements and chemical admixtures [11]. The traditional experimental or empirical techniques for determining tensile strength are labor-intensive, need costly laboratory equipment, and are typically not scalable across various mixing circumstances. Predictive modeling, particularly using soft computing and statistical regression methods, has emerged as a crucial tool for accurately assessing mechanical performance based on mix design factors [12]. Commonly employed strategies for this purpose include linear regression (LR) and multiple linear regression (MLR) models, which are straightforward to build and can discern first-order correlations between mechanical responses and input parameters [13]. The linear models, however, are inadequate for accurately representing the actual behavior of modern eco-concretes, which are significantly influenced by extremely nonlinear phenomena. Consequently, nonlinear modeling approaches, such as Pure Quadratic (PQ), Logarithmic (LOG), and Fully Quadratic Regression (FQ) models, are being widely employed in materials research [14]. The PQ model is highly effective in describing curvature effects in data, as it mitigates interaction and incorporates squared terms [15]. Conversely, the Interaction Regression (INT) model incorporates cross-terms to demonstrate the synergistic effects of factors, such as fly ash and slag ratios, on tensile strength, a relationship thoroughly examined in concrete science [16]. Five regression prediction models, including Linear Regression (LR), Polynomial Quadratic (PQ), Logistic Regression (LOG), Interaction (INT), and Fully Quadratic (FQ), are employed to describe and predict the tensile strength of eco-friendly concrete. Although several studies have employed soft computing techniques to predict the mechanical properties of eco-friendly concrete, the majority focus on compressive strength, with limited attention given to tensile strength, a key parameter for structural durability and crack resistance. Furthermore, few works have systematically compared different regression-based soft computing models, such as linear, logarithmic, quadratic, and interaction models, to evaluate their predictive performance for tensile strength using a unified dataset. Additionally, the influence of individual mix design parameters on tensile strength has not been

extensively analyzed through sensitivity analysis. This study addresses these limitations by developing and comparing five regression-based soft computing models on a diverse experimental dataset and identifying the most critical variables that influence tensile performance.

2 Research Objectives

This study aims to examine the impact of fly ash on the tensile strength of eco-friendly concrete. Referring to previous literature, this study utilizes a dataset of 252 experimental samples (Table 1) with the precise objectives being as follows:

- To conduct a comprehensive sensitivity analysis on the tensile strength of eco-friendly concrete mixes to find the main factors that affect it.
- Investigate the impact of the fly ash to binder ratio on the tensile strength of eco-friendly concrete.
- Develop multiple soft computing models to predict the tensile strength of 252 various eco-friendly concrete mixtures.
- This research also provides insight into the utility of environmentally friendly concrete mixes, thereby reducing the impact of construction operations on the environment.

Table 1. Statistical Information for Input and Output Variables

Variable	Units	Subscript	Min.	Max	Mode	Mean	Standard deviation
Curing Time	Days	CT	7	28	7	17.5	10.47
Cement to Sand Ratio	-	C/S	0.23	0.33	0.23	0.26	0.031
Water to Cement Ratio	-	W/C	0.5	0.71	0.71	0.65	0.072
GP to Binder Ratio	-	GP/B	0	0.3	0	0.071	0.089
FA to Binder Ratio	-	FA/B	0	0.3	0	0.077	0.084
GGFBS to Binder Ratio	-	GGFBS/Binder	0	0.3	0	0.070	0.084
Plasticizer to Cement Ratio	-	P/C	2.34	3.35	3.35	3.04	0.34
Thickness of the pull-off layer	mm	TPOL	1	14	8	7.28	3.57
Tensile Strength	MPa	f_h	1.17	2.28	1.47	1.63	0.25

3 Methodology

Figure 1 illustrates the activities undertaken in this study, highlighting the processes implemented. Research on eco-friendly concrete is essential, and 252 of 368 different samples and combinations were collected from a previous study [17]. The database has been divided into training and testing datasets to develop precise and reliable models. A total of 168 different combinations were used as training data (70%), while the remaining 84 were utilized as testing data (30%). The gathered data includes the following attributes: curing time, pull-off layer thickness, cement-to-sand ratio, water-to-cement ratio, granite powder-to-binder ratio, fly ash-to-binder ratio, ground granulated blast furnace slag-to-binder ratio, and plasticizer-to-cement ratio. The primary objective of this study is to predict the value of f_h and to develop realistic and reliable models. The primary predictive models employed were Linear Regression (LR), Pure Quadratic Regression (PQ), Logarithmic Regression (LOG), and Fully Quadratic Regression (FQ).

1. **Model Development:** Four regression models (LR, PQ, INT, FQ) were constructed using the dataset.
2. **Performance Evaluation:** Evaluation metrics (R^2 , RMSE, MAE, OBJ, SI) were then applied to all models to identify the most accurate and reliable predictive model.
3. **Model Selection:** Based on the performance metrics, the best-performing model was identified.
4. **Sensitivity Analysis:** Sensitivity analysis was subsequently performed on this selected model to determine the relative influence of input parameters on the predicted tensile strength.

Figures 2 and 3 present the scatter of the dependent and independent values used in the study.

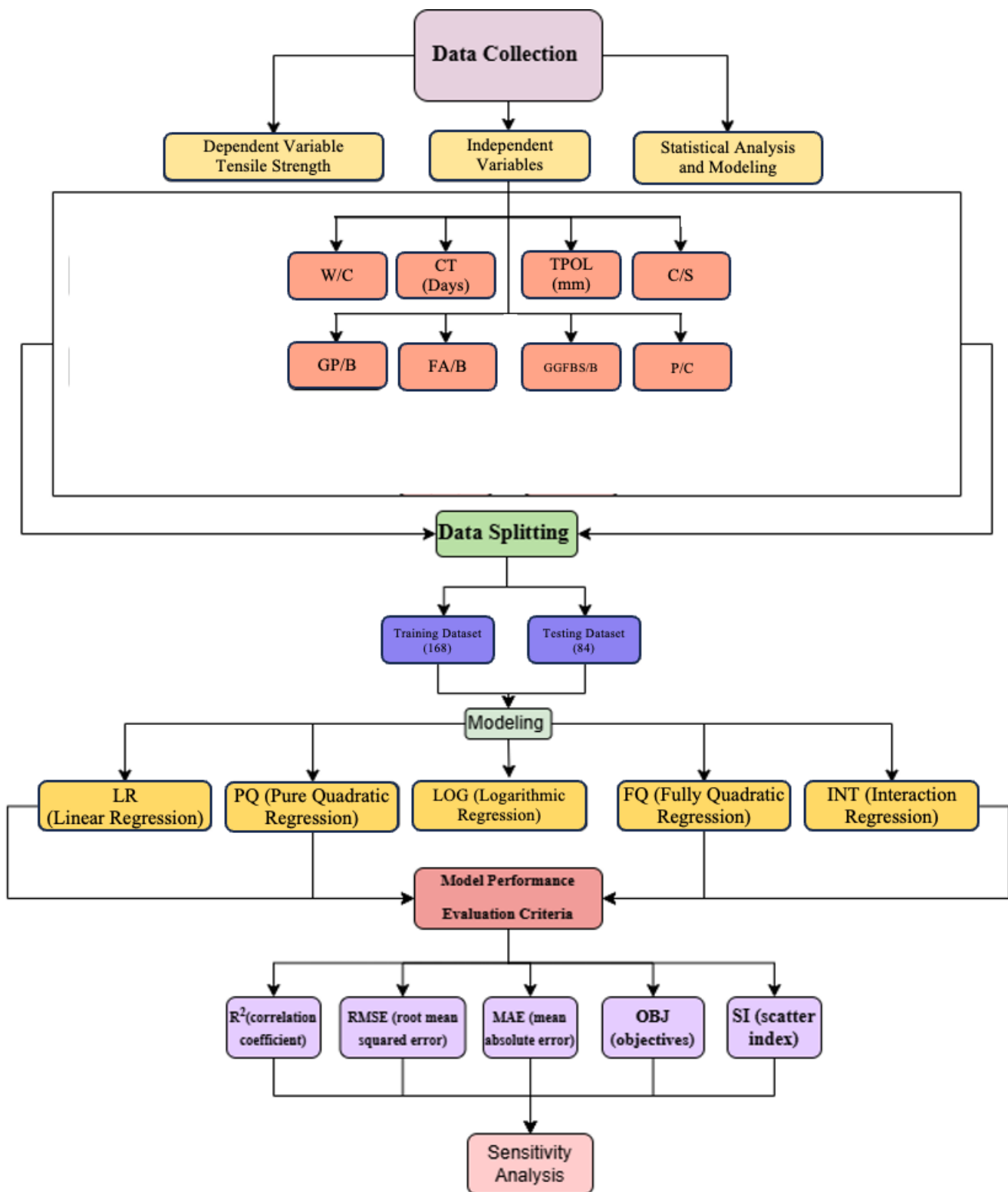


Fig. 1- Flowchart Diagram for the Conducted Study

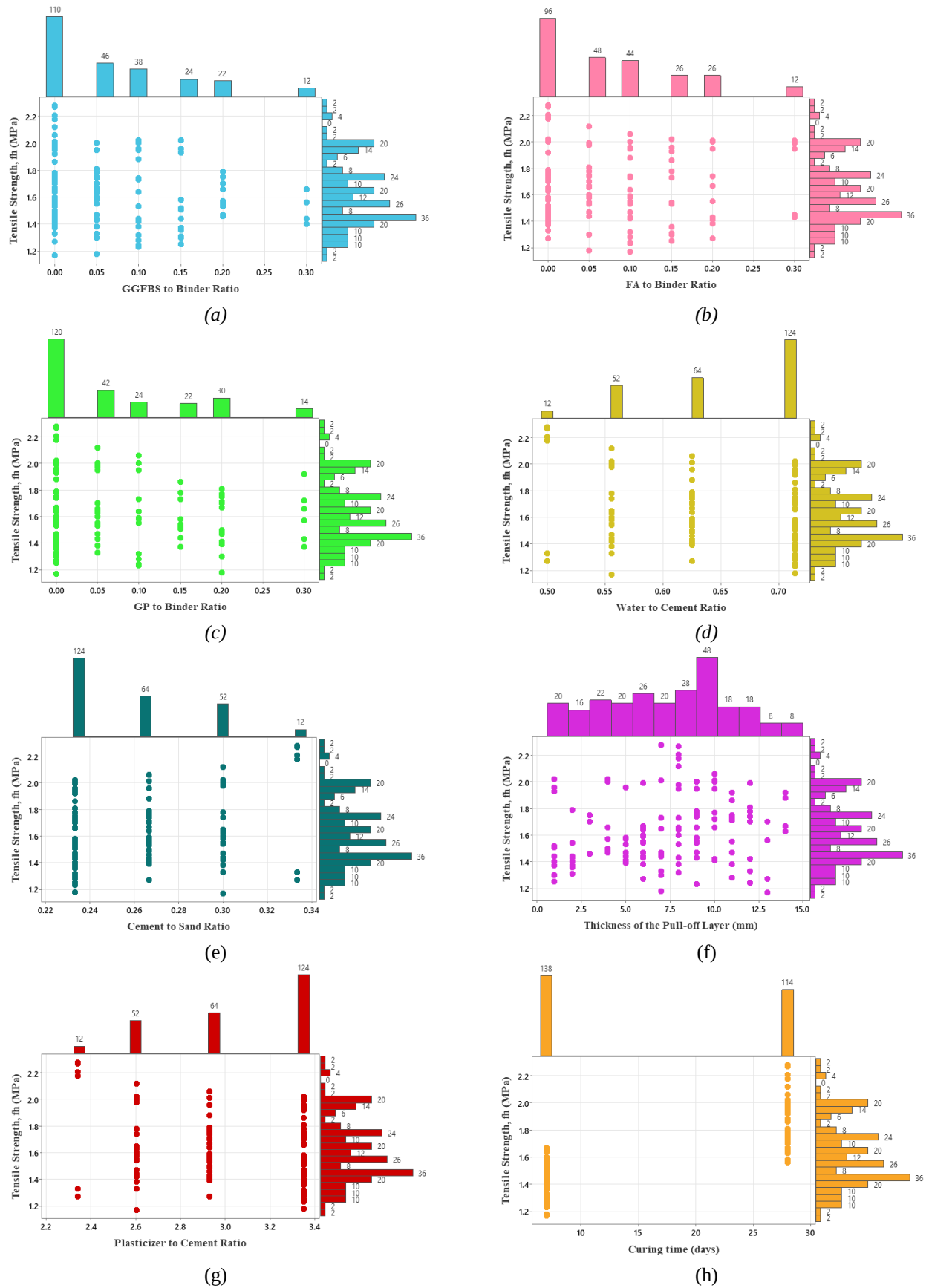


Fig. 2- Marginal plot between tensile strength of eco-friendly concrete and (a) GGFBS/B ratio; (b) FA/B ratio; (c) GP/B ratio; (d) W/C ratio; (e) C/S ratio; (f) TPOL (mm); (g) P/C ratio; (h) Curing time (days).

	CT	fh	TPOL	C/S	W/C	GP/B	FA/B	GGFBS/B	P/C
CT	1								
fh	0.82	1							1
TPOL	0.34	0.17	1						0.5
C/S	-0.05	0.15	0.21	1					0
W/C	0.06	-0.14	-0.20	-1.00	1				-0.5
GP/B	0.05	-0.06	0.12	-0.40	0.40	1			-1
FA/B	0.04	0.07	0.09	-0.35	0.34	-0.30	1		
GGFBS/B	-0.03	-0.17	-0.45	-0.33	0.33	-0.32	-0.29	1	
P/C	0.06	-0.14	-0.20	-1.00	1	0.40	0.34	0.33	1

Fig. 3- Correlation Matrix for the Correlation Between all the Variables

3.1 Linear Regression Model (LR)

Using a linear equation, a statistical method known as linear regression models the relationship between a dependent variable and one or more independent variables. Its simplicity makes it common in engineering for prediction [18]. LR models, classified by the number of independent and dependent variables, exist in four types. Simple linear regression occurs when there is just one independent variable; multiple linear regression occurs when there is more than one independent variable [13]. In terms of dependent variables, univariate linear regression is used when there is only one dependent variable, and multivariate linear regression is used when there are more than one [19]. In this study, multiple linear regression is used to predict tensile strength using various independent variables. This model may exhibit lower accuracy compared to nonlinear or advanced machine learning methods, as it assumes a linear relationship. [20] The relationship between the independent variables and tensile strength is expressed as:

$$f_h(\text{MPa}) = \beta_0 + \beta_1(CT) + \beta_2(TPOL) + \beta_3\left(\frac{C}{S}\right) + \beta_4\left(\frac{W}{C}\right) + \beta_5\left(\frac{GP}{B}\right) + \beta_6\left(\frac{FA}{B}\right) + \beta_7\left(\frac{GGFBS}{B}\right) + \beta_8\left(\frac{P}{C}\right) \quad (1)$$

The independent parameters are curing time (CT), thickness of pull-off layer (TPOL), cement to sand ratio (C/S), water to cement ratio (W/C), granite powder to binder ratio (GP/B), fly ash to binder ratio (FA/B), slag to binder ratio (GGFBS/B), plasticizer to cement ratio (P/C). The dependent variable is tensile strength (f_h). Also, β_0 until β_8 are the parameters.

3.2 Pure Quadratic Regression Model (PQ)

The third model used in this research, the pure quadratic model, is a specialized form of polynomial regression that includes the squared terms of the independent variables only [21]. While quite similar to a full quadratic model, the PQ model retains quadratic effects and excludes interaction terms, making it suitable for capturing nonlinear relationships without excessive complexity [22]. In equation 2, the PQ model predicts the tensile strength of the concrete. PQ is expected to yield more accurate predictions than linear regression since the proper relationship between the variables is nonlinear [21]. Since it employs higher-order terms, it may encounter calculation issues if not fitted correctly [23]. The PQ model is mathematically expressed as:

$$f_h(\text{MPa}) = \beta_0 + \beta_1(CT) + \beta_2(TPOL) + \beta_3\left(\frac{C}{S}\right) + \beta_4\left(\frac{W}{C}\right) + \beta_5\left(\frac{GP}{B}\right) + \beta_6\left(\frac{FA}{B}\right) + \beta_7\left(\frac{GGFBS}{B}\right) + \beta_8\left(\frac{P}{C}\right) + \beta_9(CT)^2 + \beta_{10}(TPOL)^2 + \beta_{11}\left(\frac{C}{S}\right)^2 + \beta_{12}\left(\frac{W}{C}\right)^2 + \beta_{13}\left(\frac{GP}{B}\right)^2 + \beta_{14}\left(\frac{FA}{B}\right)^2 + \beta_{15}\left(\frac{GGFBS}{B}\right)^2 + \beta_{16}\left(\frac{P}{C}\right)^2 \quad (2)$$

Whereas β_0 until β_{16} are the parameters.

3.3 Logarithmic Regression Model (LOG)

The logarithmic regression model (LOG) is a nonlinear approach to represent the relationship between the dependent variable and the natural logarithm of the independent variables [24]. This model works best when the rate of change of the response variable decreases as the predictor variable increases. In many engineering and material science applications, the

(LOG) model is commonly used [25]. The logarithmic model is employed in this study to illustrate the diminishing returns effect that often occurs as concrete strength develops [26]. When the connection between variables follows a logarithmic pattern, this approach should yield better results than linear models [27]. However, this model requires all of its input variables to be strictly positive, and it cannot function effectively with zero or negative variables [28]. The LOG model is mathematically represented as:

$$f_h(\text{MPa}) = \beta_0 + \beta_1 \ln(CT) + \beta_2 \ln(TPOL) + \beta_3 \ln\left(\frac{C}{S}\right) + \beta_4 \ln\left(\frac{W}{C}\right) + \beta_5 \ln\left(\frac{GP}{B}\right) + \beta_6 \ln\left(\frac{FA}{B}\right) + \beta_7 \ln\left(\frac{GGFBS}{B}\right) + \beta_8 \ln\left(\frac{P}{C}\right) \quad (3)$$

Whereas β_0 until β_8 are the parameters.

3.4 Interaction Regression Model (INT)

The interaction regression model incorporates product terms among independent variables into a linear regression model, allowing for the analysis of cooperative effects among predictors [29]. The (INT) method is especially useful in concrete technology, where the material's components are interdependent. The model may show situations in which the influence of one variable depends on the amount of another [30]. Equation 4 presents the interaction regression model, which can be used to predict tensile strength and provides a more detailed picture than the basic linear regression model. However, it requires careful interpretation, as interaction effects may complicate the estimation of individual parameters [31]. The (INT) model is particularly beneficial when prior information indicates that specific parts of a material will interact in a particular manner [32]. The INT model is mathematically represented as:

$$\begin{aligned} f_h(\text{MPa}) = & \beta_0 + \beta_1(CT) + \beta_2(TPOL) + \beta_3\left(\frac{C}{S}\right) + \beta_4\left(\frac{W}{C}\right) + \beta_5\left(\frac{GP}{B}\right) + \beta_6\left(\frac{FA}{B}\right) + \beta_7\left(\frac{GGFBS}{B}\right) + \beta_8\left(\frac{P}{C}\right) + \\ & \beta_9(CT)(TPOL) + \beta_{10}(CT)\left(\frac{C}{S}\right) + \beta_{11}(CT)\left(\frac{W}{C}\right) + \beta_{12}(CT)\left(\frac{GP}{B}\right) + \beta_{13}(CT)\left(\frac{FA}{B}\right) + \beta_{14}(CT)\left(\frac{GGFBS}{B}\right) + \beta_{15}(CT)\left(\frac{P}{C}\right) + \\ & \beta_{16}(TPOL)\left(\frac{C}{S}\right) + \beta_{17}(TPOL)\left(\frac{W}{C}\right) + \beta_{18}(TPOL)\left(\frac{GP}{B}\right) + \beta_{19}(TPOL)\left(\frac{FA}{B}\right) + \beta_{20}(TPOL)\left(\frac{GGFBS}{B}\right) + \beta_{21}(TPOL)\left(\frac{P}{C}\right) + \\ & \beta_{22}\left(\frac{C}{S}\right)\left(\frac{W}{C}\right) + \beta_{23}\left(\frac{C}{S}\right)\left(\frac{GP}{B}\right) + \beta_{24}\left(\frac{C}{S}\right)\left(\frac{FA}{B}\right) + \beta_{25}\left(\frac{C}{S}\right)\left(\frac{GGFBS}{B}\right) + \beta_{26}\left(\frac{C}{S}\right)\left(\frac{P}{C}\right) + \beta_{27}\left(\frac{W}{C}\right)\left(\frac{GP}{B}\right) + \beta_{28}\left(\frac{W}{C}\right)\left(\frac{FA}{B}\right) + \\ & \beta_{29}\left(\frac{W}{C}\right)\left(\frac{GGFBS}{B}\right) + \beta_{30}\left(\frac{W}{C}\right)\left(\frac{P}{C}\right) + \beta_{31}\left(\frac{GP}{B}\right)\left(\frac{FA}{B}\right) + \beta_{32}\left(\frac{GP}{B}\right)\left(\frac{GGFBS}{B}\right) + \beta_{33}\left(\frac{GP}{B}\right)\left(\frac{P}{C}\right) + \beta_{34}\left(\frac{FA}{B}\right)\left(\frac{GGFBS}{B}\right) + \beta_{35}\left(\frac{FA}{B}\right)\left(\frac{P}{C}\right) + \\ & \beta_{36}\left(\frac{GGFBS}{B}\right)\left(\frac{P}{C}\right) \end{aligned} \quad (4)$$

Whereas β_0 until β_{36} are the parameters.

3.5 Fully Quadratic Regression Model (FQ)

The fully quadratic regression (FQ) model is the fourth model used in this study. It enhances the pure quadratic model by adding squared terms and cross-product interaction terms between independent variables [33]. This approach enables the model to identify not only single nonlinear effects but also synergistic interactions across variables [34]. The FQ model is beneficial in materials science, where several interrelated elements affect mechanical characteristics [35]. The completely quadratic model for forecasting the tensile strength of concrete is given in Equation 5. This model is expected to provide more accurate predictions than simpler linear and pure quadratic models, especially when there are strong interactions among the material components [36]. To avoid overfitting, the model's complexity must be carefully validated [37]. The FQ model is mathematically represented as:

$$\begin{aligned} f_h(\text{MPa}) = & \beta_0 + \beta_1(CT) + \beta_2(TPOL) + \beta_3\left(\frac{C}{S}\right) + \beta_4\left(\frac{W}{C}\right) + \beta_5\left(\frac{GP}{B}\right) + \beta_6\left(\frac{FA}{B}\right) + \beta_7\left(\frac{GGFBS}{B}\right) + \beta_8\left(\frac{P}{C}\right) + \\ & \beta_9(CT)(TPOL) + \beta_{10}(CT)\left(\frac{C}{S}\right) + \beta_{11}(CT)\left(\frac{W}{C}\right) + \beta_{12}(CT)\left(\frac{GP}{B}\right) + \beta_{13}(CT)\left(\frac{FA}{B}\right) + \beta_{14}(CT)\left(\frac{GGFBS}{B}\right) + \beta_{15}(CT)\left(\frac{P}{C}\right) + \\ & \beta_{16}(TPOL)\left(\frac{C}{S}\right) + \beta_{17}(TPOL)\left(\frac{W}{C}\right) + \beta_{18}(TPOL)\left(\frac{GP}{B}\right) + \beta_{19}(TPOL)\left(\frac{FA}{B}\right) + \beta_{20}(TPOL)\left(\frac{GGFBS}{B}\right) + \beta_{21}(TPOL)\left(\frac{P}{C}\right) + \\ & \beta_{22}\left(\frac{C}{S}\right)\left(\frac{W}{C}\right) + \beta_{23}\left(\frac{C}{S}\right)\left(\frac{GP}{B}\right) + \beta_{24}\left(\frac{C}{S}\right)\left(\frac{FA}{B}\right) + \beta_{25}\left(\frac{C}{S}\right)\left(\frac{GGFBS}{B}\right) + \beta_{26}\left(\frac{C}{S}\right)\left(\frac{P}{C}\right) + \beta_{27}\left(\frac{W}{C}\right)\left(\frac{GP}{B}\right) + \beta_{28}\left(\frac{W}{C}\right)\left(\frac{FA}{B}\right) + \\ & \beta_{29}\left(\frac{W}{C}\right)\left(\frac{GGFBS}{B}\right) + \beta_{30}\left(\frac{W}{C}\right)\left(\frac{P}{C}\right) + \beta_{31}\left(\frac{GP}{B}\right)\left(\frac{FA}{B}\right) + \beta_{32}\left(\frac{GP}{B}\right)\left(\frac{GGFBS}{B}\right) + \beta_{33}\left(\frac{GP}{B}\right)\left(\frac{P}{C}\right) + \beta_{34}\left(\frac{FA}{B}\right)\left(\frac{GGFBS}{B}\right) + \beta_{35}\left(\frac{FA}{B}\right)\left(\frac{P}{C}\right) + \\ & \beta_{36}\left(\frac{GGFBS}{B}\right)\left(\frac{P}{C}\right) + \beta_{37}(CT)^2 + \beta_{38}(TPOL)^2 + \beta_{39}\left(\frac{C}{S}\right)^2 + \beta_{40}\left(\frac{W}{C}\right)^2 + \beta_{41}\left(\frac{GP}{B}\right)^2 + \beta_{42}\left(\frac{FA}{B}\right)^2 + \beta_{43}\left(\frac{GGFBS}{B}\right)^2 + \beta_{44}\left(\frac{P}{C}\right)^2 \end{aligned} \quad (5)$$

Whereas β_0 until β_{44} are the parameters.

3.6 Performance evaluation criteria

The Numerical values used were the correlation coefficient (R^2), root mean squared error (RMSE), mean absolute error (MAE), and objective function (OBJ), along with the scatter index (SI) to observe data dependence within the models (Table 2). These values were used to examine and evaluate the models' efficiency. Table 2 shows the ranking of the five different models. The statistical factors used in this study formed the basis for this ranking. The coefficient of determination (R^2) evaluates a model's ability to explain observed variability. In contrast, the root mean squared error (RMSE) measures the average amount of prediction errors and model accuracy. That error can be the average absolute difference between measured and predicted values here. MAE, however, is defined as the absolute error, examining consistency between observations matched within the same phenomenon. SI measures the difference between the recorded and modeled values. OBJ presents an objective function, a linear function of two decision variables, to be either maximized or minimized under the given constraints. Statistical values for each criterion are shown in Equations 6–10. [38–40].

$$R^2 = 1 - \frac{\left\{ \sum_{i=1}^n (y_i - \bar{y})^2 \right\}}{\left\{ \sum_{i=1}^n (y_i - \bar{y})^2 \right\}} \quad (6)$$

$$MAE = \frac{\sum_{i=1}^n (|y_i - \bar{y}|)}{n} \quad (7)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{N}} \quad (8)$$

$$OBJ = \left(\frac{n_{tr}}{n_{all}} \times \frac{RMSE_{tr} + MAE_{tr}}{R_{tr}^2 + 1} \right) + \left(\frac{n_{tst}}{n_{all}} \times \frac{RMSE_{tst} + MAE_{tst}}{R_{tst}^2 + 1} \right) \quad (9)$$

$$SI = \frac{RMSE}{y_i} \quad (10)$$

$\bar{y}$ expresses the predicted value of CS,

y_i Denotes the tested value of CS,

$\bar{x}$ Refers to an average of the predicted values,

$\bar{y}$ Refers to an average of the tested data,

n_{all} stands for the total number of datasets (including training and testing),

n_{tr} is a value that denotes the number of training datasets,

n_{tst} is a value denoting the number of testing datasets,

N is the dataset's total number of data points.

R^2 is considered best when it's one, while its value ranges from zero to one. In MAE, RMSE, and OBJ, a blank indicates maximum efficacy, with a range of 0 to 1, meaning that lower values yield better model efficacy. Models outperforming expectations have SI values of less than 0.1. It is argued that if the positive correlation coefficient is greater than 0.8, a significant relationship exists between the actual value and the predicted value [37]. The SI value is assigned to various categorized performance levels: less than 0.03 is excellent, between 0.03 and 0.06 is good, between 0.06 and 0.09 is fair, and above 0.09 is poor [35, 36].

Table 2. A List of the Criteria Used to Evaluate the Performance of the Models

MODEL	Figure	Equation	Training			Testing			Rank
			R^2	RMSE (MPa)	MAE (MPa)	R^2	RMSE (MPa)	MAE (MPa)	
FQ	4	6	0.829	0.105	0.089	0.850	0.094	0.077	2
LR	5	2	0.788	0.117	0.098	0.794	0.111	0.090	4
PQ	6	3	0.794	0.115	0.097	0.810	0.095	0.089	3
LOG	7	4	0.769	0.122	0.102	0.758	0.120	0.096	5
INT	8	5	0.848	0.099	0.082	0.851	0.094	0.076	1

4 Results and discussion

4.1 Linear Regression Model

A linear regression model was used to compute the tensile strength of eco-friendly concrete with varying mix ratios (Figure 4). While the linear regression (LR) method is simple to use, it is generally not expected to produce highly accurate findings, as it only considers constant terms and linear parameter correlations, which are insufficient to represent the complexity of sustainable concrete mixtures.

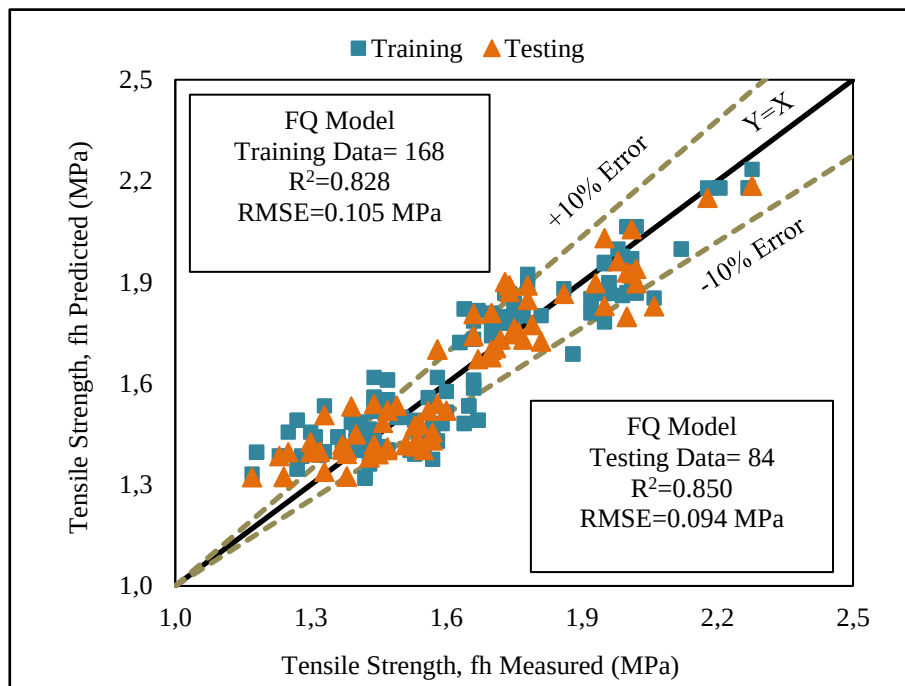


Fig. 4- Comparison between predicted and measured tensile strength using the FQ model for training and testing data

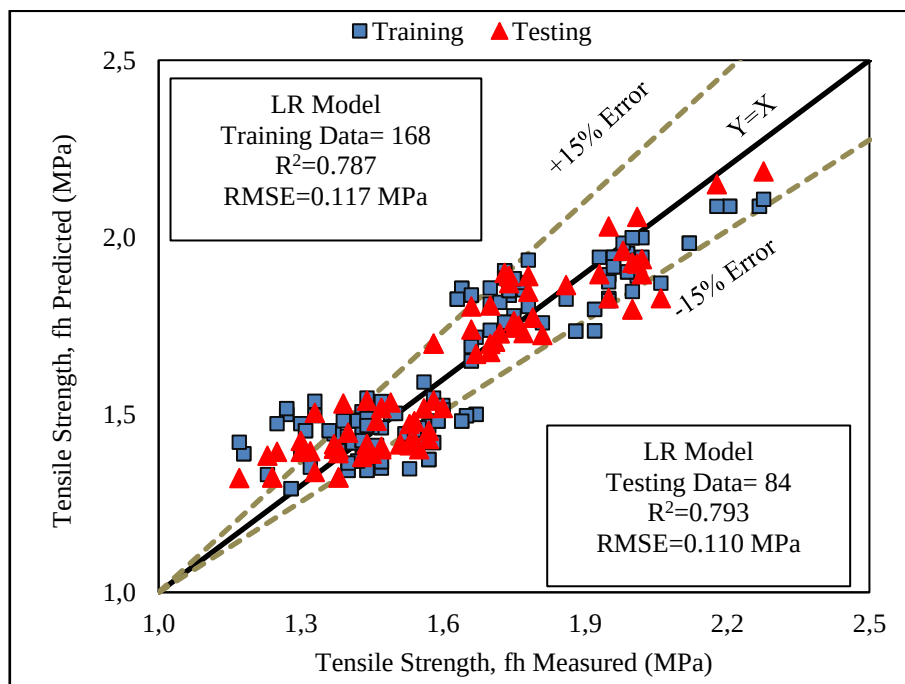


Fig. 5- comparison between predicted and measured tensile strength using the LR model for training and testing data

Equation 11 was trained on experimental datasets and then validated on testing data. The model had a R^2 value of 0.788 and an RMSE value of 0.117 MPa on training data. The performance on the testing data showed enhancement ($R^2 = 0.794$, RMSE = 0.111 MPa).

$$f_h(\text{MPa}) = -0.36 + 0.02(CT) - 0.02(TPOL) + 2.46\left(\frac{C}{S}\right) + 2.33\left(\frac{W}{C}\right) - 1.60\left(\frac{GP}{B}\right) - 1.41\left(\frac{FA}{B}\right) - 2.15\left(\frac{GGFBS}{B}\right) - 0.001\left(\frac{P}{C}\right) \quad (11)$$

Figure 5 shows that the error distribution varied from 15% to 15%. 85% of the predictions were within a 0.85–1.15 ratio of the measured tensile strength values. This shows that the linear model is a satisfactory initial approximation. Still, since it consistently underestimates higher-strength mixes (as demonstrated by the negative error skew), it implies that nonlinear influences require consideration for better accuracy.

4.2 Pure Quadratic Regression Model

A pure quadratic (PQ) regression model was used to model the tensile strength of eco-friendly concrete. Squared terms were used in this model to describe materials' nonlinear behavior better, whereas interaction terms were eliminated. Equation 12 shows the PQ model, which was trained and tested using experimental datasets. The performance metrics demonstrated a slight improvement over linear regression. However, the PQ model was not the most accurate model for our dataset in this study. The training results had an R^2 of 0.801 and an RMSE of 0.113 MPa, while the testing results had an R^2 of 0.824 and an RMSE of 0.091 MPa. The error distribution, as seen in Fig. 6, varied from –15% to 15%. 85% of the predictions were within a 0.85–1.15 ratio of the experimental values. This equation was applied to the testing dataset to improve its accuracy (Fig. 6).

$$f_h(\text{MPa}) = 0.04 - 0.006(CT) - 0.001(TPOL) + 0.01\left(\frac{C}{S}\right) - 2.61\left(\frac{W}{C}\right) - 0.12\left(\frac{GP}{B}\right) - 0.06\left(\frac{FA}{B}\right) + 1.57\left(\frac{GGFBS}{B}\right) + 0.067\left(\frac{P}{C}\right) + 0.25(CT)^2 - 0.001(TOPL)^2 - 0.001\left(\frac{C}{S}\right)^2 - 0.028\left(\frac{W}{C}\right)^2 - 1.79\left(\frac{GP}{B}\right)^2 - 1.64\left(\frac{FA}{B}\right)^2 - 2.71\left(\frac{GGFBS}{B}\right)^2 - 0.06\left(\frac{P}{C}\right)^2 \quad (12)$$

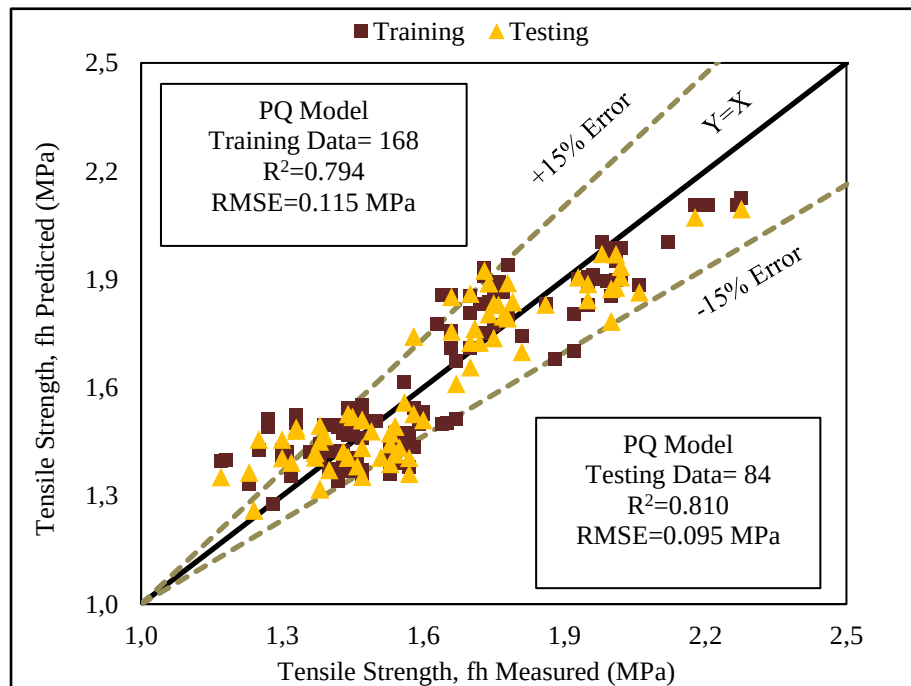


Fig. 6- Comparison between predicted and measured tensile strength using the PQ model for training and testing data

4.3 Logarithmic Regression Model

The third model employed in this study was logarithmic regression to model the tensile strength of eco-friendly concrete. The LOG model is nonlinear, and the material behavior often follows nonlinear trends. Equation 13 presents the LOG model, developed from the training dataset and subsequently tested on the test dataset. However, in this study, the LOG model had the lowest accuracy among the models used. This model on the training dataset yielded an R^2 of 0.769 and an RMSE of 0.122 MPa, while the test data produced comparable results ($R^2 = 0.7758$, RMSE = 0.120 MPa). As shown in Figure 7, 85% of results fall within $\pm 15\%$ of experimental values (Figure 7).

$$f_h(\text{MPa}) = 0.33 - 0.09 \ln(CT) - 0.12 \ln(TPOL) - 0.75 \ln\left(\frac{C}{S}\right) 0.002 \ln\left(\frac{W}{C}\right) + 0.003 \ln\left(\frac{GP}{B}\right) - 0.005 \ln\left(\frac{FA}{B}\right) + 0.013 \ln\left(\frac{GGFBS}{B}\right) + 0.43 \ln\left(\frac{P}{C}\right) \quad (13)$$

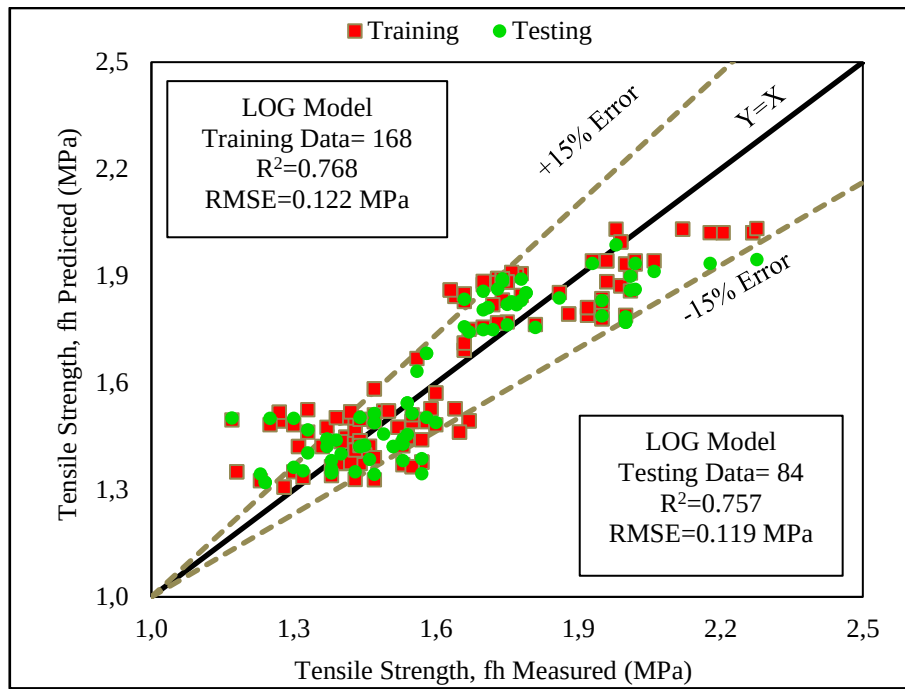


Fig. 7- Comparison between predicted and measured tensile strength using the LOG model for training and testing data

4.4 Interaction Regression Model

The interaction regression (INT) model was developed to predict the tensile strength of eco-friendly concrete. The INT model incorporates cross-product terms to show the critical interactions between the independent parameters. The INT model demonstrated the best performance among the models used in this study. The training dataset results yielded an R^2 of 0.848 and RMSE of 0.099 MPa, and while the testing results achieved an R^2 of 0.851 and RMSE of 0.094 MPa, outperforming both fully quadratic and pure quadratic models, which are the second and third-most accurate models in this study, respectively. As shown in Fig. 8, the prediction errors fell within a narrow band of $\pm 5\%$, with 95% of data points concentrated in the 0.95–1.05 ratio range, confirming its exceptional accuracy (Fig. 8).

$$f_h(\text{MPa}) = 7.11 - 0.76(CT) + 0.04(TPOL) - 12.25\left(\frac{C}{S}\right) - 1.26\left(\frac{W}{C}\right) + 0.3\left(\frac{GP}{B}\right) - 0.02\left(\frac{FA}{B}\right) + 0.03\left(\frac{GGFBS}{B}\right) + 1.45(CT)\left(\frac{C}{S}\right) - 0.04(CT)\left(\frac{GP}{B}\right) - 0.04(CT)\left(\frac{GGFBS}{B}\right) + 0.13(CT)\left(\frac{P}{C}\right) - 0.18(TPOL)\left(\frac{C}{S}\right) + 0.06(TPOL)\left(\frac{W}{C}\right) + 0.07(TPOL)\left(\frac{GP}{B}\right) + 0.02(TPOL)\left(\frac{FA}{B}\right) - 0.008(TPOL)\left(\frac{GGFBS}{B}\right) - 0.02(TPOL)\left(\frac{P}{C}\right) + 0.004\left(\frac{C}{S}\right)\left(\frac{P}{C}\right) + 0.001\left(\frac{W}{C}\right)\left(\frac{GP}{B}\right) + 0.001\left(\frac{W}{C}\right)\left(\frac{FA}{B}\right) + 0.003\left(\frac{W}{C}\right)\left(\frac{GGFBS}{B}\right) - 0.87\left(\frac{W}{C}\right)\left(\frac{P}{C}\right) - 0.075\left(\frac{FA}{B}\right)\left(\frac{GGFBS}{B}\right) - 0.014\left(\frac{FA}{B}\right)\left(\frac{P}{C}\right) + 0.056\left(\frac{GGFBS}{B}\right)\left(\frac{P}{C}\right) \quad (14)$$

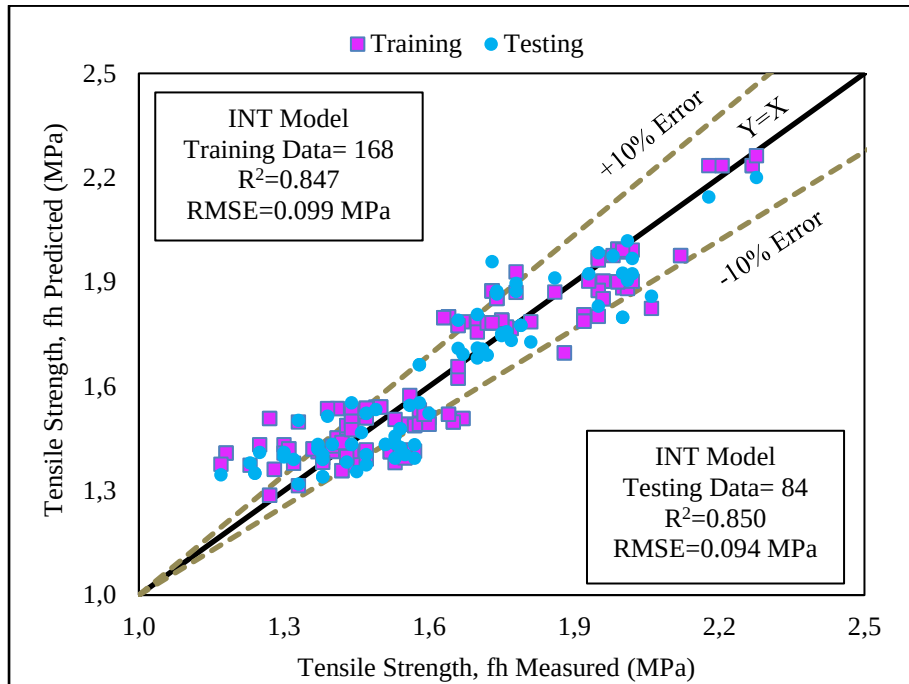


Fig. 8- Comparison between predicted and measured tensile strength using the INT model for training and testing data

4.5 Fully Quadratic Regression Model

The fifth model used in this study to predict the tensile strength of eco-friendly concrete is the fully quadratic regression model (FQ). It provides a comprehensive second-order approximation that captures both nonlinear effects and the synergistic interactions among the independent parameters. FQ is an advanced model that incorporates all quadratic terms and cross-product interaction terms, achieving training metrics of $R^2 = 0.829$ and RMSE = 0.105 MPa, with testing performance of $R^2 = 0.850$ and RMSE = 0.094 MPa.

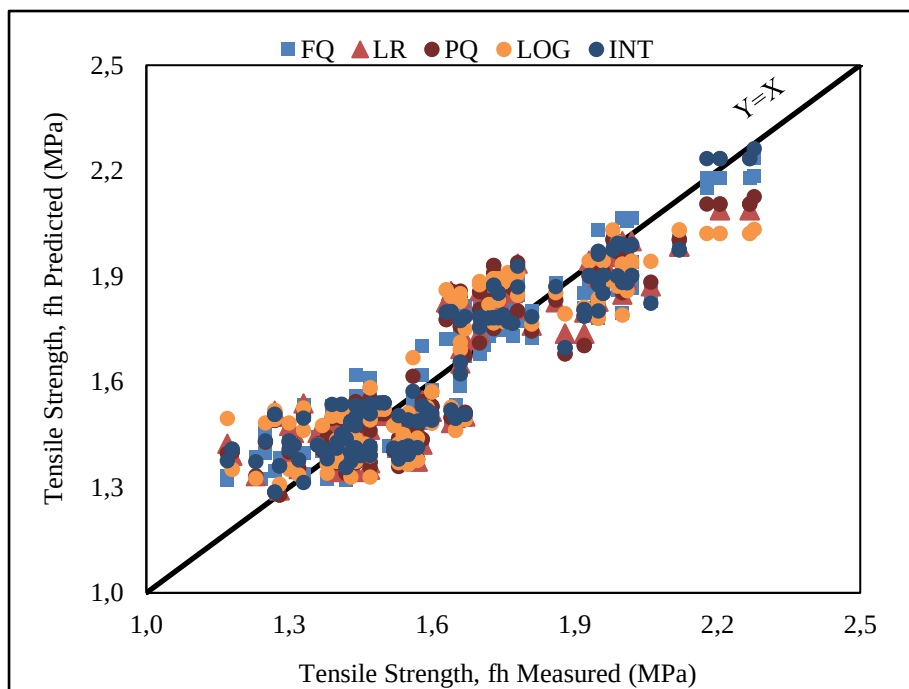


Fig. 9- Comparison between predicted and measured tensile strength using the five models for training and testing data

Fig. 4 shows the error distribution, with 90% of predictions falling within $\pm 10\%$ error margins (Fig. 9).

$$\begin{aligned}
 f_h(\text{MPa}) = & 0.02(CT) - 0.02(TPOL) + 1.38\left(\frac{C}{S}\right) + 0.96\left(\frac{W}{C}\right) - 2.16\left(\frac{GP}{B}\right) - 1.92\left(\frac{FA}{B}\right) - 2.42\left(\frac{GGFBS}{B}\right) + 0.23\left(\frac{P}{C}\right) + \\
 & 0.10(CT)\left(\frac{C}{S}\right) + 0.043(CT)\left(\frac{W}{C}\right) - 0.11(CT)\left(\frac{GP}{B}\right) - 0.077(CT)\left(\frac{FA}{B}\right) - 0.12(CT)\left(\frac{GGFBS}{B}\right) + 0.02(CT)\left(\frac{P}{C}\right) - \\
 & 0.18(TPOL)\left(\frac{C}{S}\right) + 0.008(TPOL)\left(\frac{W}{C}\right) + 0.13(TPOL)\left(\frac{GP}{B}\right) + 0.077(TPOL)\left(\frac{FA}{B}\right) + 0.055(TPOL)\left(\frac{GGFBS}{B}\right) + \\
 & 0.006(TPOL)\left(\frac{P}{C}\right) - 0.015\left(\frac{C}{S}\right)\left(\frac{W}{C}\right) - 0.012\left(\frac{C}{S}\right)\left(\frac{GP}{B}\right) - 0.008\left(\frac{C}{S}\right)\left(\frac{FA}{B}\right) + 0.004\left(\frac{C}{S}\right)\left(\frac{GGFBS}{B}\right) - 0.07\left(\frac{C}{S}\right)\left(\frac{P}{C}\right) - \\
 & 0.04\left(\frac{W}{C}\right)\left(\frac{GP}{B}\right) - 0.04\left(\frac{W}{C}\right)\left(\frac{FA}{B}\right) + 0.01\left(\frac{W}{C}\right)\left(\frac{GGFBS}{B}\right) - 0.5\left(\frac{W}{C}\right)\left(\frac{P}{C}\right) - 0.002\left(\frac{GP}{B}\right)\left(\frac{FA}{B}\right) - 0.002\left(\frac{GP}{B}\right)\left(\frac{GGFBS}{B}\right) - \\
 & 0.17\left(\frac{GP}{B}\right)\left(\frac{P}{C}\right) - 0.01\left(\frac{FA}{B}\right)\left(\frac{GGFBS}{B}\right) - 0.18\left(\frac{FA}{B}\right)\left(\frac{P}{C}\right) + 0.04\left(\frac{GGFBS}{B}\right)\left(\frac{P}{C}\right) - 0.003(CT)^2 + 0.004(TPOL)^2 - 0.11\left(\frac{W}{C}\right)^2 - \\
 & 0.01\left(\frac{GP}{B}\right)^2 - 0.01\left(\frac{FA}{B}\right)^2 + 0.017\left(\frac{GGFBS}{B}\right)^2 + 0.09\left(\frac{P}{C}\right)^2
 \end{aligned} \quad (15)$$

5 Model Comparison

This study investigated the impact of various admixtures on the surface tensile strength of eco-friendly concrete using five predictive models. To find a reliable model, several mathematical parameters were used to assess each model's performance. The Interaction Regression (INT) model yields the most accurate and reliable results, as indicated by the subsequent results: R2: 0.848, RMSE: 0.099 MPa, MAE: 0.082 MPa, SI: 0.061, and OBJ: 0.096 MPa. The Fully Quadratic Regression (FQ) yields the second most accurate results: R2: 0.829, RMSE: 0.105 MPa, MAE: 0.089 MPa, SI: 0.064, and OBJ: 0.102 MPa. The INT and FQ models yield satisfactory results, with 90% of the data points falling within $\pm 10\%$ of the target. Meanwhile, the Linear Regression, Pure Quadratic, and Logarithmic models exhibited lower accuracy (Fig. 9). The outcomes of these models fall within $\pm 15\%$ error boundaries, indicating lower accuracy and performance compared with the INT and FQ models. As shown in Table 2, the R2 and RMSE values are lower for the test datasets than for the training datasets, indicating that model performance deteriorates with new data. The lower MAE on the test datasets suggests effective generalization and indicates no overfitting. The INT model achieves the minimum OBJ value of 0.096, indicating the highest predictive precision. According to the SI value (Fig. 11), the INT model yields the smallest value, indicating the best performance. For the residual errors, the INT model's residual error falls within the range -0.236 to +0.237 MPa. The FQ model varies from -0.231 to +0.222 MPa. The LR, PQ, and LOG models have higher ranges, varying from -0.300 to +0.367 MPa, -0.244 to +0.240 MPa, and -0.331 to +0.326 MPa, respectively (Figs. 10, 11, and 12).

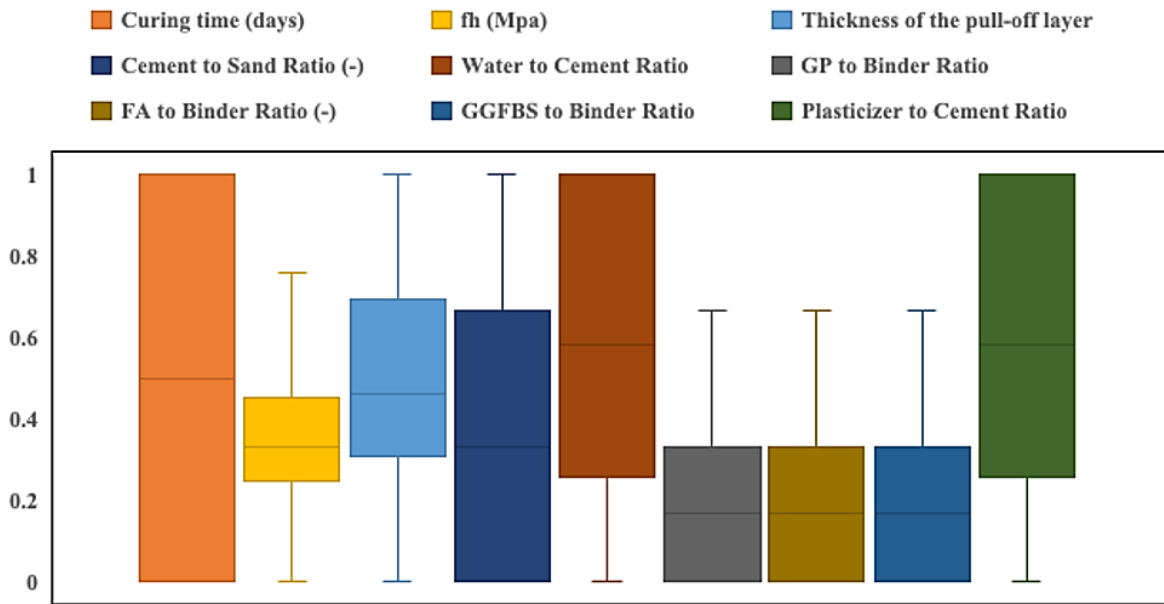
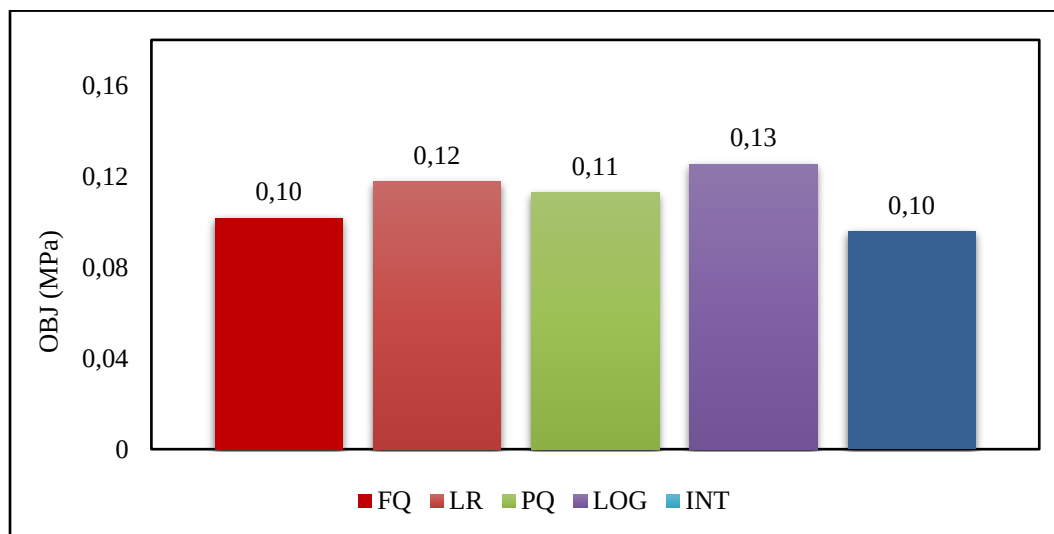
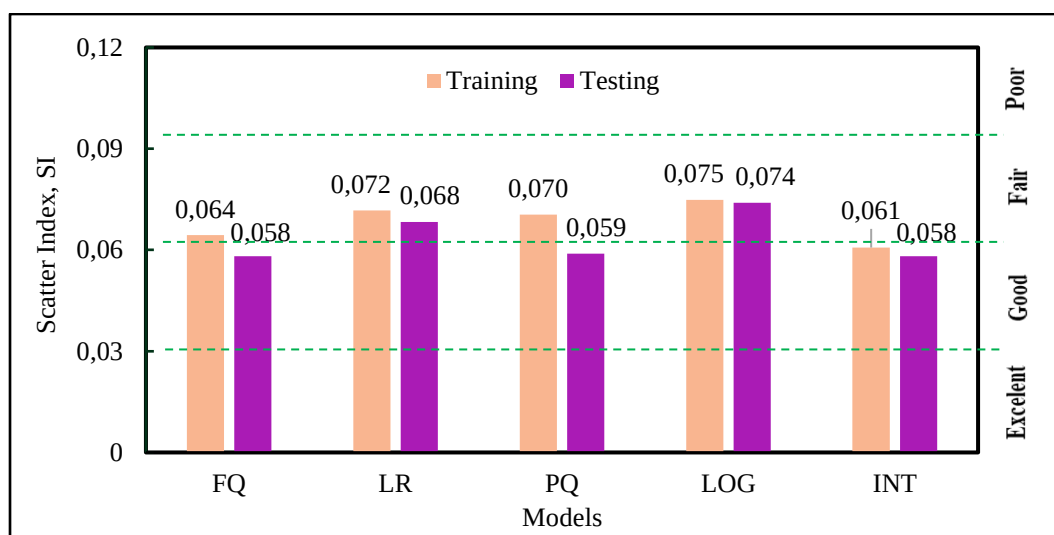


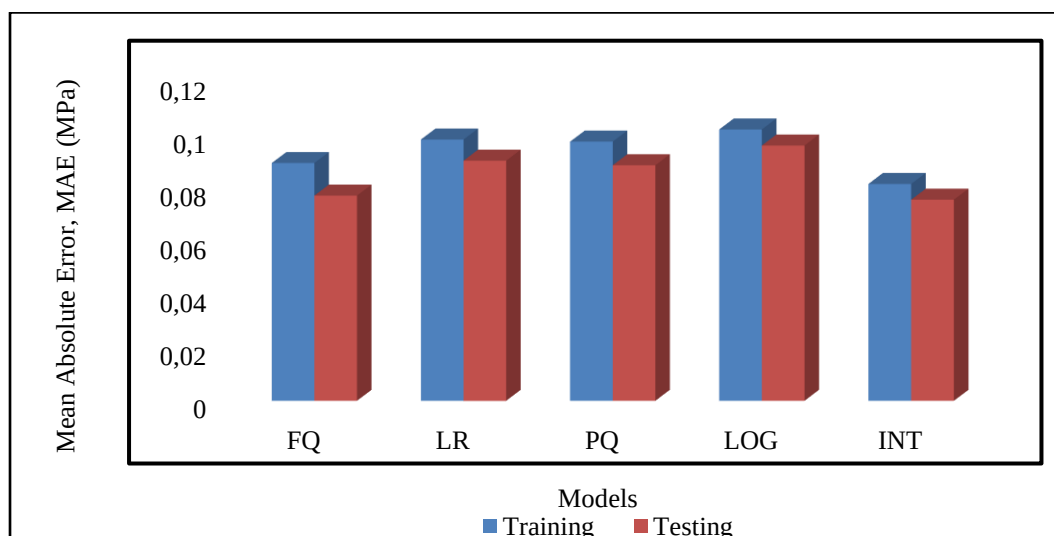
Fig. 10- Box plot of features influence on tensile strength (f_h) of eco-friendly concrete



(a)



(b)



(c)

Fig. 11- Comparison of (a) OBJ, (b) SI, and (c) MAE performance metrics across the five employed models using the training and testing datasets.

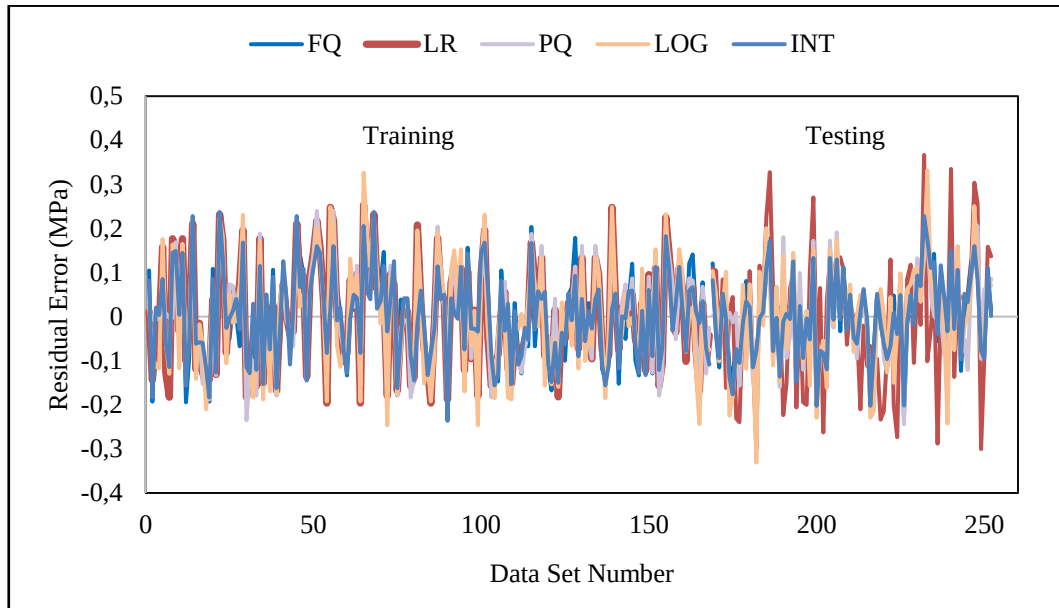


Fig. 12- Comparison of residual errors in eco-friendly concrete models (LR, FQ, PQ, LOG and INT).

6 Investigation of Sensitivity

The Interaction Regression (INT) model, which achieved the highest accuracy among all regression models used in this study, was employed to conduct a sensitivity analysis. A Pareto chart (Fig. 13), which illustrates the relative importance of features, was used to evaluate the impact of each input parameter on tensile strength. The curing time (CT) was the most important feature, followed by the thickness of the pull-off layer (TPOL), the cement-to-sand ratio (C/S), and the water-to-cement ratio (W/C). These parameters highly influenced the tensile strength of environmentally friendly concrete. Other parameters were either moderately or weakly sensitive and comprised the fly ash-to-binder ratio (FA/B), the granite powder-to-binder ratio (GP/B), the pull-off layer thickness (TPOL), and the plasticizer-to-cement ratio (P/C). This indicates that the tensile strength is predominantly controlled by the binder composition and hydration development, suggesting that these factors primarily influence the tensile strength.

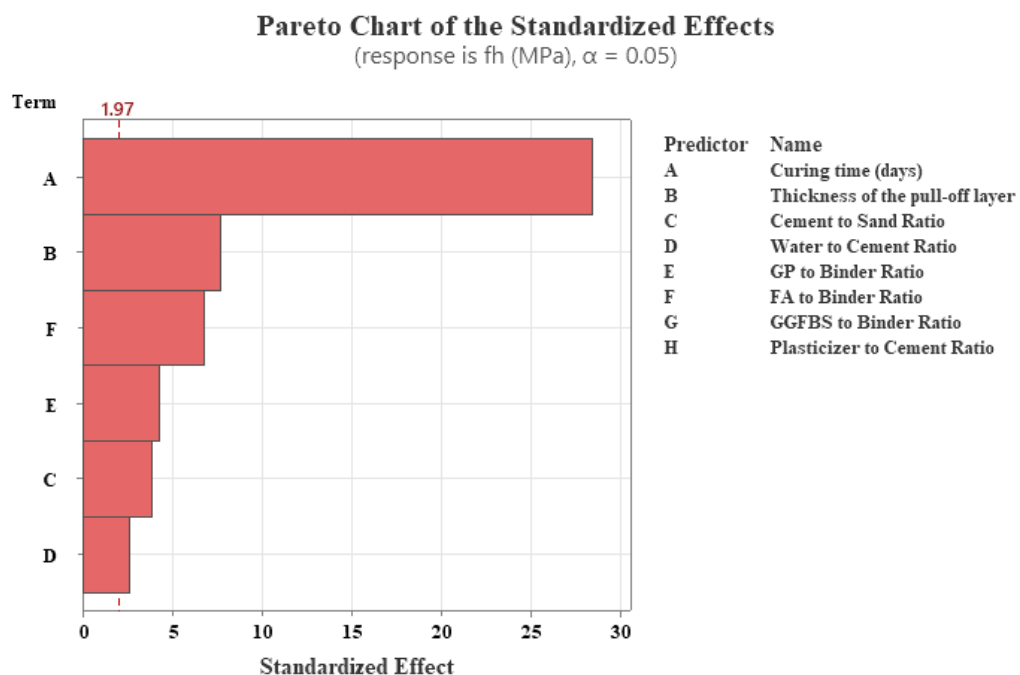


Fig.13- Pareto chart of the standardized effects of each independent variable on tensile strength (sensitivity chart).

7 Limitations of the study

- The experimental dataset employed in this study was generated under controlled laboratory conditions. While these settings allowed for precise measurement and systematic evaluation of diverse green concrete mix ratios, they cannot fully capture the heterogeneity and variability encountered in real-world construction environments. Factors such as fluctuating curing environments, variability in raw material sourcing, and site-specific practices introduce complexities that extend beyond the limitations of laboratory replication.
- Although the models incorporated critical predictors such as curing duration, binder composition, and supplementary cementitious materials, they inevitably simplify the reality of concrete behavior. Key external influences including ambient temperature, relative humidity, environmental exposure, and intrinsic material irregularities were not included in the predictive framework. Such exclusions may restrict the models' accuracy in reflecting tensile strength performance across different climatic regions and construction practices.
- The predictive models are based on assumed mathematical relationships (linear, quadratic, logarithmic, or interaction-based) between the chosen input parameters and tensile strength. While the models demonstrated strong statistical performance, they may not fully capture the complex, nonlinear, and multi-physics interactions that govern concrete behavior in the field, particularly under long-term service conditions. Consequently, the generalizability of the models may be limited when extrapolated beyond the scope of the experimental parameters.
- Additional limitations include the absence of certain influential material parameters, such as aggregate morphology, mineralogical composition, or microstructural evolution under exposure conditions (e.g. carbonation, freeze–thaw cycles, chloride ingress). Excluding these aspects may introduce biases or constrain the predictive range when the models are applied to concretes with different aggregate sources or durability requirements.
- While the overfitting risk was mitigated through rigorous dataset partitioning (training and testing phases), the complex structure of Fully Quadratic and Interaction models inherently increases susceptibility to overfitting. This remains a potential constraint, particularly when models are applied to datasets outside the calibration range.

8 Conclusions

This study systematically developed and evaluated five regression models—Linear Regression (LR), Pure Quadratic (PQ), Logarithmic (LOG), Interaction (INT), and Fully Quadratic (FQ)—to predict the tensile strength of durable concrete. The analysis was based on an experimental dataset comprising 252 distinct mixtures, incorporating key input variables such as curing time, cement-to-sand ratio, water-to-cement ratio, and supplementary cementitious materials (fly ash, slag, and granite powder). The dataset was divided into training and testing subsets to ensure reliable evaluation of model performance.

- Sensitivity analysis identified curing duration, fly ash-to-binder ratio, and granite powder-to-binder ratio as the most significant variables governing tensile strength. These findings highlight that optimizing the dosage of fly ash and granite powder can be crucial for enhancing the performance of sustainable concrete mixtures.
- Among the five investigated models, the Interaction Regression (INT) model demonstrated superior predictive capability. It achieved the highest coefficient of determination ($R^2 = 0.851$) and the lowest root-mean-square error (RMSE = 0.094 MPa) on the test dataset. Furthermore, approximately 90% of predictions fell within a $\pm 10\%$ error margin, confirming the model's robustness and reliability.
- The INT model showed narrow residual error bounds, ranging from -0.236 MPa to $+0.237$ MPa. This indicates effective generalization across datasets, minimal risk of overfitting, and a strong alignment between observed and predicted tensile strength values.
- When evaluated using additional statistical indicators, the INT model consistently ranked highest. It achieved the lowest objective function (OBJ = 0.10) and the smallest scatter index (SI = 0.061 for the training dataset and

0.058 for the testing dataset), further reinforcing its status as the most accurate and reliable model among those considered.

- Overall, this study demonstrates that regression-based modeling can effectively capture the complex relationships between mix design parameters and tensile strength of concrete. The findings emphasize the critical role of curing duration and supplementary binders (particularly fly ash and granite powder) in optimizing concrete performance. The superior accuracy of the INT model suggests it can serve as a practical and reliable tool for predicting tensile strength in sustainable concrete design and for guiding future experimental and field applications.

List of Abbreviations

CT: Curing Time

C/S: Cement to Sand Ratio

FA/B: Fly Ash to Binder Ratio

f_t : Tensile Strength

GGFBS/B: Slag to Binder Ratio

GP/B: Granite Powder to Binder Ratio

P/C: Plasticizer to Cement Ratio

TPOL: Thickness of Pull-Off Layer

W/C: Water to Cement Ratio

Declaration of AI and AI-Assisted Technologies in the Writing Process

The authors utilized **OpenAI ChatGPT-4o** to enhance readability during the preparation of this work. They thoroughly reviewed and edited all content generated by these tools and take full responsibility for the final version of the published article.

Availability of data and materials

The data supporting the conclusions of this article are included with the article. All the data was extracted from this paper (<https://doi.org/10.1016/j.advengsoft.2024.103710>).

REFERENCES

- [1]- Y.V. Fan, P. Jiang, J.J. Klemeš, P.Y. Liew, C.T. Lee, Integrated regional waste management to minimise the environmental footprints in circular economy transition. *Resour. Conserv. Recycl.* 168 (2021) 105292. <https://doi.org/10.1016/j.resconrec.2020.105292>
- [2]- H. Asakura, K. Endo, M. Yamada, Y. Inoue, Y. Ono, Y. Ono, Influence of oxygen flow rate on reaction rate of organic matter in leachate from aerated waste layer containing mainly incineration ash. *Waste Manage.* 30(11) (2010) 2185–2193. <https://doi.org/10.1016/j.wasman.2010.06.001>
- [3]- P. Monteiro, S. Miller, A. Horvath, Towards sustainable concrete. *Nature Mater.* 16 (2017) 698–699. <https://doi.org/10.1038/nmat4930>
- [4]- P. Partonia, S. Heidarian, G.A. Dezvareh, A.K. Darban, Investigating Eco-Friendly Concrete in the Construction Industry in Order to Improve the Circular Economy - A Literature Review. *Chem. Eng. Process Tech.* 9(2) (2024) 1093.
- [5]- I. Almeshal, B.A. Tayeh, R. Alyousef, H. Alabduljabbar, A. Mustafa Mohamed, Eco-friendly concrete containing recycled plastic as partial replacement for sand. *J. Mater. Res. Technol.* 9(3) (2020) 4631–4643. <https://doi.org/10.1016/j.jmrt.2020.02.090>
- [6]- H.W. Reinhardt, Factors affecting the tensile properties of concrete. Eds Jaap Weerheijm, In Woodhead Publishing Series in Civil and Structural Engineering, Understanding the Tensile Properties of Concrete, Woodhead Publishing,

- (2013) 19-51. <https://doi.org/10.1533/9780857097538.1.19>
- [7]- X. Li, X. Gu, X. Xia, E. Madenci, X. Chen, Q. Zhang, Effect of water-cement ratio and size on tensile damage in hardened cement paste: Insight from peridynamic simulations. *Constr. Build. Mater.* 356 (2022) 129256. <https://doi.org/10.1016/j.conbuildmat.2022.129256>
- [8]- Y. Pamu, S. Alugubelli, A comparative study of environmental impacts due to conventional and sustainable concrete. *Mater. Today: Proc.* 91(1) (2023) 112-120. <https://doi.org/10.1016/j.matpr.2023.03.817>
- [9]- F. Niu, Y. Liu, F. Xue, H. Sun, T. Liu, H. He, X. Kong, Y. Chen, H. Liao, Ultra-high performance concrete: A review of its material properties and usage in shield tunnel segment. *Case Stud. Constr. Mater.* 22 (2025) e04194. <https://doi.org/10.1016/j.cscm.2024.e04194>
- [10]- I.-C. Yeh, Modeling slump flow of concrete using second-order regressions and artificial neural networks. *Cem. Concr. Compos.* 29(6) (2007) 474–480. <https://doi.org/10.1016/j.cemconcomp.2007.02.001>
- [11]- T. Roy, P. Das, R. Jagirdar, M. Shehabat, M.S. Abdullah, A. Kashem, R. Rahman, Prediction of mechanical properties of eco-friendly concrete using machine learning algorithms and partial dependence plot analysis. *Smart Constr. Sustainable Cities*, 3(2) (2025) <https://doi.org/10.1007/s44268-025-00048-8>.
- [12]- P.G. Asteris, V.G. Mokos, Concrete compressive strength using artificial neural networks. *Neural Comput. Appl.* 32 (2020) 11807–11826. <https://doi.org/10.1007/s00521-019-04663-2>
- [13]- N.R. Draper, H. Smith, *Applied regression analysis* (3rd ed.). (1998) Wiley. <https://doi.org/10.1002/9781118625590>
- [14]- R. Jin, Q. Chen, A.B. Soboyejo, Non-linear and mixed regression models in predicting sustainable concrete strength. *Constr. Build. Mater.* 170 (2018) 142–152. <https://doi.org/10.1016/j.conbuildmat.2018.03.063>
- [15]- A. Dean, D. Voss, D. Draguljić, *Design and analysis of experiments* (2nd ed.). (2017) Springer. <https://doi.org/10.1007/978-3-319-52250-0>
- [16]- R. Karim, M.H. Islam, S.D. Datta, A. Kashem, Synergistic effects of supplementary cementitious materials and compressive strength prediction of concrete using machine learning algorithms with SHAP and PDP analyses. *Case Stud. Constr. Mater.* 20 (2023) e02828. <https://doi.org/10.1016/j.cscm.2023.e02828>
- [17]- M. Moj, S. Czarnecki, Comparative analysis of selected machine learning techniques for predicting the pull-off strength of the surface layer of eco-friendly concrete. *Adv. Eng. Softw.* 195 (2024) 103710. <https://doi.org/10.1016/j.advengsoft.2024.103710>
- [18]- G. James, D. Witten, T. Hastie, R. Tibshirani, *An introduction to statistical learning*. Springer Texts Stat. (2013). <https://doi.org/10.1007/978-1-4614-7138-7>
- [19]- N.H. Timm, *Applied multivariate analysis*. (2002) Springer. <https://doi.org/10.1007/b98969>
- [20]- T. Hastie, R. Tibshirani, J. Friedman, *The elements of statistical learning* (2nd ed.). (2009) Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- [21]- G.A.F. Seber, C.J. Wild, *Nonlinear regression*. (2003) Wiley. <https://doi.org/10.1002/0471725315>
- [22]- A. Dean, D. Voss, D. Draguljić, *Design and analysis of experiments* (2nd ed.). (2017) Springer. <https://doi.org/10.1007/978-3-319-52250-0>
- [23]- D.A. Belsley, E. Kuh, R.E. Welsch, *Regression diagnostics*. (2005) Wiley. <https://doi.org/10.1002/0471725153>
- [24]- S. Weisberg, *Applied linear regression* (4th ed.). (2014) Wiley. <https://doi.org/10.1002/9781118614733>
- [25]- G.A.F. Seber, A.J. Lee, *Linear regression analysis* (2nd ed.). (2012) Wiley. <https://doi.org/10.1002/9780470824566>
- [26]- I.-C. Yeh, Modeling slump flow of concrete using second-order regressions and artificial neural networks. *Cem. Concr. Compos.* 29(6) (2007) 474-480. <https://doi.org/10.1016/j.cemconcomp.2007.02.001>
- [27]- J.W. Osborne, Improving your data transformations: Applying the Box-Cox transformation. *Pract. Assess. Res. Eval.* 15(12) (2010) 1-9. <https://doi.org/10.7275/qbpc-gk17>
- [28]- J.W. Osborne, Notes on the use of data transformations. *Pract. Assess. Res. Eval.* 8(6) (2002). <https://doi.org/10.7275/4vng-5608>
- [29]- T. Brambor, W.R. Clark, M. Golder, Understanding interaction models: Improving empirical analyses. *Polit. Anal.* 14(1) (2006) 63-82. <https://doi.org/10.1093/pan/mpi014>
- [30]- Y. Zheng, P. Zhang, J. Wang, Interaction effects of mineral admixtures on high-performance concrete properties. *Cem. Concr. Res.* 107 (2018) 1-12. <https://doi.org/10.1016/j.cemconres.2018.02.013>
- [31]- J. Zhou, X. Chen, S. Chen, Interaction effects of steel fibers and expansive additives on tensile behavior of ultra-high performance concrete. *Constr. Build. Mater.* 223 (2019) 120-132. <https://doi.org/10.1016/j.conbuildmat.2019.06.181>
- [32]- B.S. Thomas, J. Yang, A. Bahrudeen, Modeling the interactive effects of sugarcane bagasse ash and silica fume in

- cement composites. *J. Build. Eng.* 44 (2021) 103297. <https://doi.org/10.1016/j.jobbe.2021.103297>
- [33]- L. Zhang, et al. Optimization of geopolymer concrete mixtures using fully quadratic response surface models. *Cem. Concr. Compos.* 114 (2020) 103741. <https://doi.org/10.1016/j.cemconcomp.2020.103741>
- [34]- G.E.P. Box, N.R. Draper, *Response surfaces, mixtures, and ridge analyses* (2nd ed.). (2007) Wiley. <https://doi.org/10.1002/0470072768z>
- [35]- O. Giannini, P. Bassani, M. Altosole, Fully quadratic response surfaces for multi-scale material optimization: A comparative study on steel alloys. *J. Mater. Res. Technol.* 24 (2023) 5678-5692. <https://doi.org/10.1016/j.jmrt.2023.04.123>
- [36]- P. Zhang, Q. Li, J. Wang, Fully quadratic response surface modeling of fiber-reinforced concrete tensile strength considering matrix-fiber interactions. *Constr. Build. Mater.* 342 (2022) 127982. <https://doi.org/10.1016/j.conbuildmat.2022.127982>
- [37]- Y. Huo, H. Sun, Z. Chen, Y. Yang, Mechanical properties and its reliability prediction of engineered /strain-hardening cementitious composites (ECC/SHCC) with different moisture contents at negative temperatures. *Cem. Concr. Compos.* 134 (2022) 104812. <https://doi.org/10.1016/j.cemconcomp.2022.104812>
- [38]- P.G. Asteris, M. Apostolopoulou, D.J. Armaghani, et al., On the metaheuristic models for the prediction of cement-metakaolin mortars compressive strength. *Math. Comput. Appl.* 1(1) (2020) 063. <https://doi.org/10.12989/mca.2020.1.1.063>
- [39]- A. Kooshkaki, H. Eskandari-Naddaf, Effect of porosity on predicting compressive and flexural strength of cement mortar containing micro and nano-silica by multi-objective ANN modeling. *Constr. Build. Mater.* 212 (2019) 176-191. <https://doi.org/10.1016/j.conbuildmat.2019.03.243>
- [40]- M.H. Akeed, S. Qaidi, H.U. Ahmed, W. Emad, R.H. Faraj, A.S. Mohammed, B.A. Tayeh, A.R. Azevedo, Ultra-high-performance fiber-reinforced concrete. Part III: Fresh and hardened properties. *Case Stud. Constr. Mater.* 17 (2022) e01265. <https://doi.org/10.1016/j.cscm.2022.e01265>