

Artificial Intelligence-Based Prediction of Compressive Strength of Metakaolin-Saw Dust Geopolymer Concrete for Sustainable Construction Applications

Hyginus Obinna Ozioko^{a,*} Emmanuel Ebube Eze^a

^a Department of Civil Engineering, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria.

Keywords:

Workability
ANN
OPC
Sawdust
Slump

* Corresponding author:

Hyginus Obinna Ozioko 
E-mail:
ho.ozioko@mouau.edu.ng

Received: 02.09.2025.

Revised: 22.10.2025.

Accepted: 29.10.2025.



ABSTRACT

This study examines the fresh and hardened characteristics of metakaolin-sawdust geopolymer concrete (MSGC) and develops models to predict its compressive strength. MSGC mixes were prepared with sawdust replacing fine aggregates at levels from 0% to 40%. Evaluations covered workability, setting time, bulk density, water absorption, and compressive strength, alongside artificial intelligence-based prediction. Increasing sawdust levels led to marked reductions in slump (172 mm at 0% to 0 mm at 30–40%) and substantial delays in initial setting time (53 minutes at 0% to 242 minutes at 40%). Bulk density fell from 2350 kg/m³ to 1400 kg/m³, while water absorption rose sharply from 3.5% to 25% as sawdust content increased. MSGC compressive strength decreased from 36.1 MPa (0%) to 3.8 MPa (40%) at 28 days. The control mix outperformed ordinary Portland cement concrete (OPC), and MSGC with up to 10% sawdust remained competitive (28.5 MPa). Predictive models developed using Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), and Gene Expression Programming (GEP) showed the ANN model provided best accuracy, with $R^2 = 0.9423$. Overall, findings confirm MSGC's potential as a sustainable alternative for construction.

© 2025 Journal of Management and Engineering Sciences

1. INTRODUCTION

Geopolymer concrete (GPC) has gained prominence as an environmentally friendly and sustainable alternative to conventional cement-based concrete due to its reduced carbon footprint. Instead of using ordinary Portland

cement, GPC utilizes industrial by-products such as fly ash, ground granulated blast furnace slag (GGBFS), rice husk ash, and metakaolin as binders, significantly lowering the environmental impact of concrete production [1,2]. The compressive strength of geopolymer concrete is a crucial parameter for its use in structural

applications, and its prediction is essential for optimizing mix designs and improving the sustainability of construction materials.

Among various alternative materials for geopolymer concrete, metakaolin and sawdust are gaining attention. Metakaolin, a kaolin-based material, is known for its high reactivity, which enhances the mechanical properties of geopolymer concrete [1]. Sawdust, a widely available waste material, has also been used in geopolymer concrete, often as a replacement for natural aggregates. Sawdust-based geopolymer concrete exhibits improved thermal insulation and sound absorption properties, making it suitable for lightweight construction applications [3]. However, the inclusion of sawdust can reduce the compressive strength of the concrete, especially at higher replacement levels [4].

The combination of metakaolin and sawdust in geopolymer concrete can offer a dual benefit of improved strength and sustainability. Metakaolin acts as a strong binder, while sawdust contributes to the reduction of environmental impact by replacing natural aggregates and enhancing the material's insulation properties [3]. Therefore, the prediction of compressive strength in such geopolymer mixes is vital to ensure optimal performance and structural reliability.

Recent studies have employed machine learning models to predict the compressive strength of geopolymer concrete. For instance, Support Vector Regression (SVR) and Grey Wolf Optimization (GWO) models have been used to predict the strength of GGBFS-based geopolymer concrete [5]. Additionally, Artificial Neural Networks (ANNs) have shown potential in predicting the compressive strength of fly ash-based geopolymer concrete [6]. These predictive models can incorporate various factors, such as the ratio of alkaline liquids, curing conditions, and binder contents, to forecast the concrete's strength with high accuracy.

The present study aims to develop predictive models for the compressive strength of metakaolin-saw dust geopolymer concrete by leveraging machine learning techniques. By analyzing the impact of various mix design parameters, including the content of metakaolin and sawdust, and curing conditions, the study

seeks to establish a reliable model for predicting the compressive strength of these eco-efficient concretes, thereby contributing to sustainable construction practices [7].

This work aligns with the growing need to optimize the use of industrial by-products in construction materials while ensuring that the mechanical properties of the resulting concrete meet the required standards for structural integrity [1,2]. Through predictive analysis, this study aims to improve the formulation of geopolymer concrete mixes, thus advancing the goal of sustainable construction with reduced environmental impact.

2. MATERIALS AND METHOD

Fig. 1 presents the study's methodology, from material preparation and experimental testing to machine learning model development and evaluation using R^2 , RMSE, MSE, and MAE.

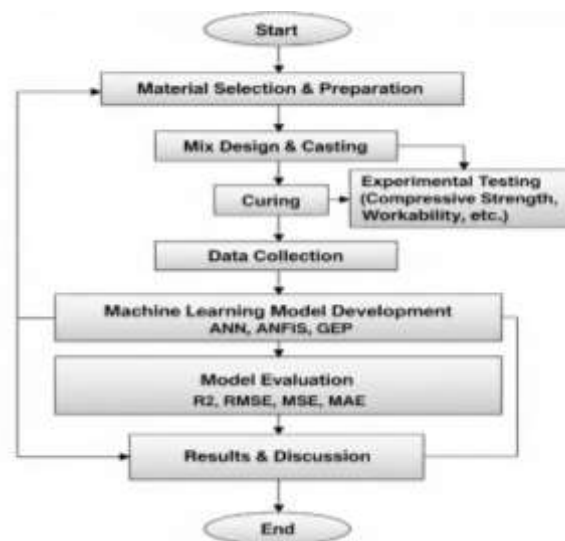


Fig. 1. Methodology Flow Chart.

2.1 Metakaolin and Sawdust Preparation

Metakaolin was procured from trusted suppliers in Ahiaeke Market, Abia State, ensuring it was free from impurities such as quartz, mica, and other minerals that could compromise its pozzolanic reactivity. To enhance surface area and improve reactivity during geopolymerization, the metakaolin was ground into a fine powder, targeting a fineness of 325 mesh (45 microns) in line with ASTM C618 requirements for pozzolanic materials. Sawdust, on the other hand, was collected from Ahiaeke

Building Material Market, carefully selected to be free of oils, resins, and paints that could interfere with the mix. To ensure consistency, the sawdust was oven-dried at 100–110°C for 24 hours until a constant weight was reached, thereby eliminating excess moisture that could affect the water-to-solid ratio. The dried sawdust was sieved through a 1.18 mm standard sieve to achieve uniform particle size distribution, conforming to ASTM D3164-09, and prepared for blending with the binder.

2.2 Alkaline Activators: Sodium Hydroxide and Sodium Silicate Solutions

The alkaline activators, sodium hydroxide (NaOH) and sodium silicate (Na_2SiO_3), were prepared at the Chemistry Laboratory of Michael Okpara University of Agriculture, Umudike, to serve as the primary agents for initiating geopolymerization. Sodium hydroxide flakes were dissolved in distilled water to obtain a concentrated solution of 8–16 M, prepared cautiously due to its caustic nature and in compliance with [8]. Sodium silicate, also known as water glass, was prepared with an $\text{SiO}_2/\text{Na}_2\text{O}$ ratio between 1.5 and 3.0 to enhance bond formation within the polymeric matrix. The standard mixing ratio of sodium silicate to sodium hydroxide was maintained at approximately 2.5:1, ensuring optimum balance between strength and workability. The prepared solutions were combined immediately before mixing with the dry ingredients to maximize reactivity. All preparation procedures conformed to [9] to guarantee uniformity, safety, and consistency in the geopolymerization process.

2.3 Aggregates and Mixing Water

Both fine and coarse aggregates were selected to meet the necessary standards for strength, durability, and grading. Fine aggregates were sourced from Imo River and confirmed to be well-graded, free from clay, silt, and organic matter, in accordance with [10]. Coarse aggregates, consisting of crushed granite obtained from Ishiagu Quarry in Ebonyi State, were graded within 5–20 mm and tested for hardness, texture, and moisture content in line with [11]. These aggregates provided the essential bulk and stability needed for the geopolymer concrete. Additionally, clean borehole water was obtained from the College of Engineering and Engineering

Technology, MOUAU. The water was clear, odorless, and colorless, meeting the quality requirements for mixing concrete. Together, the fine aggregates, coarse aggregates, and water formed the structural backbone of the mix, complementing the binder and activators to produce durable and workable geopolymer concrete.

2.4 Tests on Metakaolin–Sawdust Geopolymer Concrete

A series of tests were carried out to evaluate the fresh and hardened properties of the metakaolin–sawdust geopolymer concrete in accordance with relevant standards. The compressive strength test, conducted using standard cube specimens in line with [12], assessed the concrete's ability to withstand axial loads after curing for 7, 14, and 28 days. This test provided crucial data on the strength development and structural reliability of the mix. To complement this, the workability of fresh concrete was examined using the slump test as prescribed by [13]. The slump values offered insight into the ease of placement, compaction, and handling of the fresh geopolymer concrete.

Durability and quality indicators were further investigated through density and void content tests in accordance with [14]. These tests determined the bulk density, water absorption, and volume of voids, which are essential parameters for predicting long-term performance and resistance to environmental attack. Additionally, the setting time of the concrete was measured using a Vicat apparatus following [15]. This provided information on both initial and final setting times, which are critical for understanding the hardening behavior and workability duration of the mix. Collectively, the tests offered a comprehensive evaluation of the mechanical, fresh state, and durability properties of the metakaolin–sawdust geopolymer concrete, ensuring that the material meets performance expectations for structural and sustainable construction applications.

2.5 Methodology for Predicting Compressive Strength of Geopolymer Concrete using AI Models

This study employs advanced Artificial Intelligence (AI) techniques, specifically Artificial

Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), and Gene Expression Programming (GEP), to predict the compressive strength of geopolymer concrete. The methodology encompasses data collection and preprocessing, model construction and training, performance assessment, and visualization of results. The dataset used in this study consists of 44 experimental observations that were manually compiled. This compilation focused on identifying key parameters that influence the compressive strength of geopolymer concrete. The selected input parameters (X) include Metakaolin content (MK, kg/m³), Sawdust Replacement (SD, %), Slump (S, mm), and Curing Days (CD, days). The output parameter (Y), which is the primary focus of the study, is Compressive Strength (CS, MPa).

Data Cleaning

Prior to analysis, the dataset was inspected for missing values or outliers. For this specific dataset, all values were assumed to be complete and accurate, thus no explicit cleaning operations such as imputation or outlier removal were performed.

Data Normalization

To ensure that all input and output parameters contribute equally to the model training and to improve the convergence and performance of the AI algorithms, the data were normalized to a range of [0, 1]. The Min-Max scaling method was applied using Eq. 1.

For an individual data point x_i in a feature column X:

$$x_{norm,i} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Where x_{min} and x_{max} are the minimum and maximum values of the respective feature in the entire dataset. This transformation was applied to both the input features (X) and the output target (Y). The normalized input and output matrices are denoted as x_{norm} and y_{norm} respectively.

Data Splitting

The normalized dataset was divided into three subsets for model development and evaluation.

The training set, comprising 70% of the data, was used for model learning. The validation set, with 15% of the data, served to tune model hyperparameters and prevent overfitting during training, particularly for ANNs. The testing set, also 15% of the data, was reserved for an unbiased evaluation of the final model's performance on new data. Random partitioning ensured representative subsets. Indices for these sets are $trainInd$, $valInd$, and $testInd$.

$$(x_{norm,i}, y_{norm,i}), \quad i = 1, 2, \dots, N \quad (2)$$

Where $x_{norm,i}$ and $y_{norm,i}$ are column vectors representing the i -th normalized input and output samples, respectively.

The splitting process can be formally expressed as:

$$\text{Data Points} = \{(x_{norm,i}, i, y_{norm,i}) \mid i = 1, \dots, N\} \quad (3)$$

$$\text{Training Set} = \{(x_{norm,i}, i, y_{norm,i}) \mid i \in trainInd\} \quad (4)$$

$$\text{Validation Set} = \{(x_{norm,i}, i, y_{norm,i}) \mid i \in valInd\} \quad (5)$$

$$\text{Testing Set} = \{(x_{norm,i}, i, y_{norm,i}) \mid i \in testInd\} \quad (6)$$

Artificial Neural Networks (ANN)

A feed-forward backpropagation ANN architecture was employed for compressive strength prediction. This type of network consists of an input layer, one or more hidden layers, and an output layer.

- **Architecture:** A single hidden layer with 10 neurons was selected after experimentation. The network structure can be represented as 4-10-1, corresponding to 4 input features, 10 hidden neurons, and 1 output (compressive strength) and 1 output (compressive strength).
- **Activation Functions:** The *tansig* (hyperbolic tangent sigmoid) transfer function was used for the hidden layer, providing non-linearity crucial for learning complex relationships. A *purelin* (linear) transfer function was used for the output layer, suitable for regression tasks.
- **Hidden Layer Output:** For a neuron j in the hidden layer with inputs x_i from the input layer:

$$h_j = \tanh \sum_{i=1}^{n_{inputs}} w_{ij}^{(1)} x_i + b_j^1 \quad (7)$$

Where $n_{inputs} = 4$ is the number of input features, $w_{ij}^{(1)}$ are the weights connecting input i to hidden neuron j , and $b_j^{(1)}$ is the bias for hidden neuron j .

- Output Layer Output: For the output neuron:

$$y_{pred,norm} = \sum_{i=1}^{n_{hidden}} w_j^{(2)} h_j + b^{(2)} \quad (8)$$

Where $n_{hidden} = 10$ the number of hidden neurons is, $w_j^{(2)}$ are the weights connecting hidden neuron j to the output, and $b^{(2)}$ is the bias for the output neuron.

- Training Algorithm: The Bayesian Regularization backpropagation algorithm (*trainbr*) was utilized. This algorithm is robust for small to medium-sized datasets, often leading to better generalization by preventing overfitting through regularization. It updates weights and biases according to Levenberg-Marquardt optimization and minimizes a combination of squared errors and weights. The performance function minimized by *trainbr* is:

$$F(w) = \beta ED(w) + \alpha Ew(w) \quad (9)$$

Where:

$$ED(w) = \sum_{k=1}^{N_{train}} (y_{normk} - y_{predicted,normk})^2 \quad (10)$$

$$Ew(w) = \sum_i w_i^2 \quad (11)$$

- Training Parameters: The model training process was configured with the following key parameters: a maximum limit of 1000 epochs was set for the training iterations. The performance objective for the model was defined by a Mean Squared Error (MSE) goal of 1×10^{-5} .

Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS functions as a hybrid intelligent system by combining the adaptive learning capabilities of Artificial Neural Networks (ANNs) with the rule-based interpretability of Fuzzy Inference Systems (FIS). For this study, a first-order Sugeno-type FIS was implemented. An initial FIS structure was generated from the training data using the grid partitioning method, specifically via the *genfis1* function in MATLAB. The ANFIS model was trained for 100 epochs.

- Membership Functions (MFs): For each input variable, three membership functions were assigned. Gaussian bell-shaped membership

functions (*gbellmf*) were selected as the input MF type due to their smooth and differentiable properties, which are beneficial for gradient-based learning algorithms. The generalized bell-shaped membership function for a given input x and parameters a , b , and c is defined as:

$$\mu(x; a, b, c) = \frac{1}{1 + \left| \frac{x-c}{a} \right|^{2b}} \quad (12)$$

Here, a controls the width, b controls the slope, and c defines the center of the membership function.

- Output MF Type: Linear membership functions (linear) were used, meaning the output of each rule is a linear combination of the inputs. For a rule k :

$$fk = pkMK + qkSD + rkS + skCD + tk \quad (13)$$

Where pk, qk, rk, sk , and tk are consequent parameters optimized during training.

- Training Algorithm: The ANFIS function employs a hybrid learning algorithm. This algorithm combines the least-squares method to optimize the consequent parameters and the gradient descent method to optimize the premise parameters (Membership Function parameters).

○ Rule Firing Strength Calculation

For a given set of normalized inputs $X_{norm} = [MK_{norm}, SD_{norm}, S_{norm}, CD_{norm}]$, the firing strength of each rule k , denoted as w_k , is computed as the product of the membership degrees of the inputs to their respective membership functions:

$$w_k = \mu_{MK,k}(MK_{norm}) \times \mu_{SD,k}(SD_{norm}) \times \mu_{S,k}(S_{norm}) \times \mu_{CD,k}(CD_{norm}) \quad (14)$$

○ Normalized Firing Strength

The normalized firing strength ϖ_k , for each rule k is calculated by dividing its firing strength by the sum of the firing strengths of all rules:

$$\varpi_k = \frac{w_k}{\sum_{j=1}^{N_{rules}} w_j} \quad (15)$$

Where N_{rules} is the total number of fuzzy rules.

- Overall Output: The final ANFIS predicted output, $y_{predicted,norm}$, is the weighted average of the individual rule outputs:

$$y_{predicted,norm} = \sum_{k=1}^{N_{rules}} \varpi_k \cdot fk \quad (16)$$

Where fk is the output of rule k .

Gene Expression Programming (GEP)

Gene Expression Programming (GEP) is an evolutionary algorithm designed to discover explicit mathematical expressions or computer programs. In contrast to opaque models generated by ANNs and ANFIS, GEP aims to produce clear mathematical equations that define the relationship between input and output variables.

For model initialization, the GEP module (*gplearn.genetic.SymbolicRegressor*) in Python was utilized to implement the GEP algorithm. The population size for each generation was set to 5000 individuals (programs), and the evolutionary process spanned 20 generations. A predefined function set, including elementary mathematical operations such as addition (add), subtraction (sub), multiplication (*mul*), division (div), square root (*sqr*t), natural logarithm2 (log), absolute value (abs), and negation (*neg*), was provided. The GEP algorithm constructs expressions by combining these functions with the input variables. Mean Squared Error (MSE) served as the fitness function, guiding the evolutionary process by measuring model performance. A parsimony coefficient of 0.01 was applied to penalize overly complex models, thereby promoting the evolution of simpler, more interpretable expressions. GEP's evolutionary process incorporates genetic operators like mutation, recombination, and transposition to refine programs across generations, continuously improving their fitness by minimizing MSE. The objective is to identify a mathematical expression $f(\text{MK}, \text{SD}, \text{S}, \text{CD})$ such that:

$$CS_{\text{predicted}} = f(\text{MK}, \text{SD}, \text{S}, \text{CD}) \quad (17)$$

The final evolved equation represents the outcome of this optimization process. The model was trained using the normalized training data to uncover the inherent mathematical relationship.

Evaluation Metrics

Model performance was quantitatively assessed using widely accepted metrics as given by Ozioko and Eze [16] and represented by Eqs. 18-21.

Mean Squared Error (MSE):

$$MSE = \frac{1}{n} \sum_{i=0}^n (y_i - \hat{y}_i)^2 \quad (18)$$

Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{MSE} \quad (19)$$

Mean Absolute Error (MAE):

$$MAE = \left(\frac{1}{n} \right) \sum_{i=0}^n |y_i - \hat{y}_i| \quad (20)$$

R-squared (R^2):

$$R^2 = 1 - \frac{\frac{1}{n} \sum_{i=0}^n (y_i - \hat{y}_i)^2}{\frac{1}{n} \sum_{i=0}^n (y_i - \bar{y})^2} \quad (21)$$

Where: y_i represents the i -th actual (observed) value, $\hat{y}_i$ represents the i -th predicted value, $\bar{y}$ represents the mean of the actual (observed) values and N represents the total number of data points.

3. RESULTS AND DISCUSSION

3.1 Material Characterization

The physical properties and gradation profiles of the materials used are summarized in Table 1. Metakaolin and sawdust exhibited low densities (740 kg/m^3 and 215.3 kg/m^3 , respectively) and high porosity, particularly sawdust, which showed a water absorption of 31.2%. Granite and sand displayed typical aggregate characteristics with higher densities and lower moisture contents. The particle size distribution curves presented in Fig. 2 illustrate the grading behavior of the materials. Granite showed a gap-graded profile with a Coefficient of Uniformity (Cu) of 1.84 and Coefficient of Curvature (Cc) of 0.68, indicating poor gradation. Sand, in contrast, was well-graded with Cu and Cc values of 20.35 and 0.72, respectively. Metakaolin and sawdust, though non-aggregates, exhibited moderate gradation (Cu = 12.63 and 10.00), suggesting a wide range of fine particles.

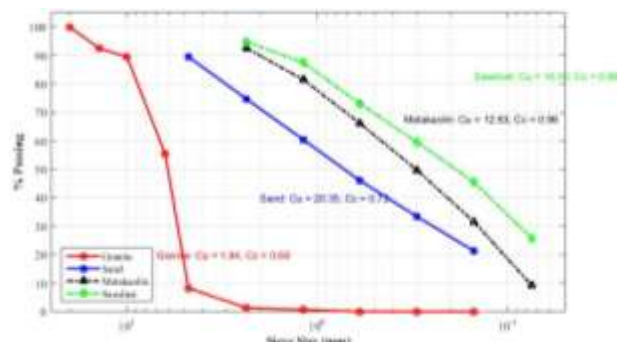
The weighing of metakaolin and sawdust prior to mixing is shown in Fig. 3, highlighting the use of an electronic balance for precise measurement. This ensures consistency in mix design and material batching. Properties of the alkaline activators used in the geopolymer system are detailed in Table 2. A 10 M NaOH solution and commercial-grade sodium silicate ($\text{SiO}_2/\text{Na}_2\text{O}$ ratio = 3.22) were combined in a 40:60 weight ratio to facilitate effective geopolymerization.

Table 1. Material Properties and Particle Size Distribution.

Property	Metakaolin	Sawdust	Granite	Sand
Density (kg/m ³)	740.0	215.3	2665.0	1652.4
Bulk Density (kg/m ³)	980.1	139.0	1512.3	1548.7
Specific Gravity	2.52	0.62	2.68	2.59
Moisture Content (%)	0.5	8.1	0.3	1.5
Water Absorption (%)	–	31.2	0.8	1.1
Particle Size Distribution (% Passing)				
Sieve Size				
20.0 mm	–	–	100.0	–
14.00mm	–	–	92.5	–
10.0 mm	–	–	89.5	–
6.3mm	–	–	55.4	–
4.75 mm	–	–	8.2	89.5
2.36 mm	92.6	94.8	1.3	74.7
1.18 mm	81.4	87.5	0.6	60.3
600 µm	66.2	73.1	0.0	46.1
300 µm	49.7	59.6	0.0	33.4
150 µm	31.5	45.4	0.0	21.2
0.075 mm	9.3	25.8	–	–
Pan				

Table 2. Material Properties of Alkaline Activators.

Property	Sodium Hydroxide (NaOH)	Sodium Silicate (Na ₂ SiO ₃)
Physical Form	White pellets (solid)	Viscous liquid (gel-like)
Concentration Used	10 M (molar)	Commercial-grade, 3.22 SiO ₂ /Na ₂ O ratio
Density (kg/m ³)	~1310 (10M solution at 25 °C)	1510
Specific Gravity	1.31	1.51
Viscosity @ 25°C	–	400 – 600 mPa·s
pH (at 25°C)	13.5 – 14.0	~11.3
Water Content (%)	~70 (in solution form)	~55
Mixing Ratio (NaOH:Na ₂ SiO ₃ by wt.)	40:60	40:60

**Fig. 2.** Gradation Characteristics of the Materials.**Fig. 3.** Metakaolin and Sawdust being measured.

3.2 Workability of Fresh Geopolymer Concrete

The results presented in Table 3 consistently show that an increase in sawdust replacement percentage leads to a notable decrease in the workability of fresh geopolymer concrete mixes. The control mix (0% sawdust) exhibited a high slump of 172 mm, indicating excellent flowability. As sawdust content increased, the slump values progressively reduced, with mixes containing 30% and 40% sawdust exhibiting zero slump. This phenomenon is largely attributed to the high-water absorption capacity of sawdust (31.2%, as indicated in Table 2), which significantly reduces the free water available for lubrication within the mix. Furthermore, the irregular, fibrous nature and relatively large surface area of sawdust particles, even after sieving to 1.18 mm as per the methodology, likely increase the internal friction and inter-particle resistance, thereby impeding the flow and compaction of the fresh concrete.

Table 3. Slump Test Results for Metakaolin-Sawdust Geopolymer Concrete.

Mix ID	Metakaolin (%)	Sawdust (%)	Slump (mm)	Description of Workability
Control (0% SD)	100	0	172	High workability
GPC-SD4	96	4	164	Moderate-high workability
GPC-SD7	93	7	155	Moderate workability
GPC-SD10	90	10	137	Moderate workability
GPC-SD13	87	13	105	Low-moderate workability
GPC-SD15	85	15	78	Low workability
GPC-SD20	80	20	56	Very low workability
GPC-SD22	78	22	33	Extremely low workability
GPC-SD25	75	25	14	Nearly no slump
GPC-SD30	70	30	0	Zero slump
GPC-SD40	60	40	0	Zero slump

These findings align with previous research on the influence of sawdust on concrete workability. For instance, Duan et al. [4] observed that sawdust addition inversely influences the workability of fly ash geopolymer paste. Similarly, Oyedepo et al. [17] reported a decrease in workability (slump values of 40 mm, 9 mm, and 5 mm for 0%, 25%, and 50% sawdust as partial replacement for fine sand in OPC concrete, respectively), which is consistent with the trend observed in the present study. Although their study used sawdust as a sand replacement in OPC concrete, the underlying principle of sawdust's high absorbency and irregular shape impacting fresh properties holds true for geopolymer systems. Onyechere et al. [18], in their review on sawdust ash in concrete, also generally noted a reduction in concrete workability as sawdust ash content increases, reinforcing the idea that sawdust-derived materials tend to absorb water and stiffen the mix.

While the exact slump values vary between studies due to differences in mix proportions, binder types (geopolymer vs. OPC), sawdust characteristics, and experimental procedures, the general trend of reduced workability with increased sawdust content is consistently reported. The current study's observation of zero slump at higher sawdust percentages (30% and 40%) underscores the significant challenge sawdust poses to workability, potentially requiring the use of superplasticizers to maintain practical consistency for real-world applications, as hinted by Ikumapayi et al. [19] who used a superplasticizer to enhance properties of sawdust ash concrete. This emphasizes that while sawdust offers sustainability benefits, its application requires careful mix design adjustments to ensure adequate workability for placement and compaction. Fig. 4 shows the slump being measured with a meter rule.

**Fig. 4.** Measuring the slump with meter rule.

3.3 Setting Time of Geopolymer Concrete

The setting time results, as presented in Table 4, clearly indicate that the inclusion of sawdust significantly prolonged both the initial and final setting times of the metakaolin-sawdust geopolymer concrete. The control mix (0% sawdust) exhibited an initial setting time of 53 minutes and a final setting time of 143 minutes. As the sawdust content increased, these times consistently extended. For example, the mix with 10% sawdust had an initial setting time of 82 minutes and a final setting time of 193 minutes, while the 40% sawdust mix showed substantially longer setting times of 242 minutes and 425 minutes, respectively.

Table 4. Initial and Final Setting Times of Geopolymer Concrete Mixes.

Mix ID	Metakaolin (%)	Sawdust (%)	Initial Setting Time (min.)	Final Setting Time (min.)
Control (0% SD)	100	0	53	143
GPC-SD4	96	4	58	147
GPC-SD7	93	7	63	174
GPC-SD10	90	10	82	193
GPC-SD13	87	13	96	221
GPC-SD15	85	15	114	247
GPC-SD20	80	20	140	286
GPC-SD22	78	22	153	298
GPC-SD25	75	25	171	330
GPC-SD30	70	30	213	373
GPC-SD40	60	40	242	425

This observed retardation in setting is primarily attributed to the chemical composition and physical properties of sawdust. Sawdust, being a lignocellulosic material, contains organic compounds such as sugars, cellulose, hemicellulose, and lignin. These organic constituents can interfere with the geopolymerization process by retarding the dissolution of aluminosilicate precursors from metakaolin or by adsorbing the alkaline activators (sodium hydroxide and sodium silicate), thereby reducing their effective concentration in the mix. Additionally, the high water absorption capacity of sawdust (as detailed in Table 1) means it competes for the available water in the alkaline solution, which can further slowdown the chemical reactions critical for geopolymer formation and hardening.

While Duan et al. [4] also investigated the fresh properties of sawdust-reinforced geopolymer, their summary states that 'The sawdust content is inversely proportional to the setting time.' This specific observation from Duan et al. [4] does not align with the current study's findings, which consistently demonstrate a direct relationship between increasing sawdust content and a prolongation of both initial and final setting times (i.e., higher sawdust percentages lead to longer setting times). Fig. 5 captures the geopolymer

concrete in its unhardened, malleable phase, ready for or just subjected to molding into specimens.

**Fig. 5.** Geopolymer Concrete during Mixing and Molding.

3.4 Bulk Density and Water Absorption Capacity

In the current study (Table 5), the bulk density of geopolymer concrete progressively declined from 2350 kg/m³ (0% SD) to 1400 kg/m³ (40% SD) with increasing sawdust content. Conversely, water absorption rose from 3.5% to 25%, indicating increasing porosity and reduced compactness at higher sawdust proportions. These findings align with Mehdi et al. [3], who observed that complete replacement of natural aggregate with sawdust in FA-GBFS-based geopolymer led to a significant reduction in weight and improved thermal and acoustic properties. Though strength reduced, the concrete became more porous and lighter supporting the trends in density and water absorption observed in this study.

Similarly, Duan et al. [4] reported a linear inverse relationship between sawdust content and density, confirming that sawdust addition increases porosity and reduces compactness. They also noted improved shrinkage resistance and microstructure refinement at moderate additions. Oyedepo et al. [20] and Osei and Jackson [21] also documented a steady reduction in density with rising sawdust content in OPC-based mixes. Osei and Jackson [21] quantified this with a 17.93% drop in density at 100%

replacement, mirroring the lighter concrete observed in this study.

Table 5. Bulk Density and Water Absorption of Hardened Geopolymer Concrete at 28 Days.

Mix ID	Metakaolin (%)	Sawdust (%)	Bulk Density (kg/m ³)	Water Absorption (%)
Control (0% SD)	100	0	2350.0	3.5
GPC-SD4	96	4	2280.0	4.2
GPC-SD7	93	7	2210.0	5.1
GPC-SD10	90	10	2130.0	6.5
GPC-SD13	87	13	2050.0	8.0
GPC-SD15	85	15	1970.0	9.5
GPC-SD20	80	20	1800.0	12.5
GPC-SD22	78	22	1720.0	14.0
GPC-SD25	75	25	1650.0	16.0
GPC-SD30	70	30	1550.0	19.5
GPC-SD40	60	40	1400.0	25.0

Overall, the current results corroborate existing literature that sawdust inclusion reduces bulk density and increases water absorption, especially beyond 15–20% replacement, making it suitable primarily for non-structural or lightweight applications where thermal and acoustic insulation may be desirable. Fig. 6 depicts the influence of Sawdust Replacement Percentage on the Properties of Geopolymer Concrete: (a) Slump, (b) Initial Setting Time, (c) Final Setting Time, (d) Bulk Density, and (e) Water Absorption.

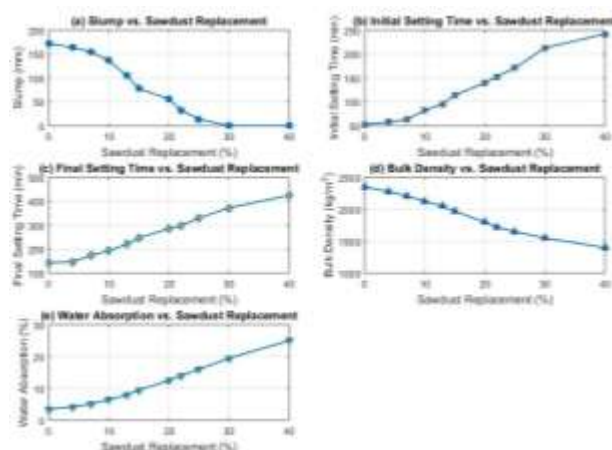


Fig. 6. Influence of Sawdust Replacement Percentage on the Properties of Geopolymer.

3.5 Compressive Strength Results of Geopolymer Concrete

The current study (Table 6) and Fig. 7 recorded a gradual reduction in compressive strength with increasing sawdust content. At 28 days, strength dropped from 36.1 MPa (0% SD) to 21.9 MPa (15% SD) and further to 3.8 MPa at 40% sawdust replacement. In comparison, Mehdi et al. [3] found that substituting natural aggregate with 100% sawdust in FA-GBFS-based geopolymer concrete led to an approximate 35% reduction in compressive strength, but improved sound absorption and thermal performance, emphasizing sawdust's multifunctional impact.

Table 6. Compressive Strength of Metakaolin-Sawdust Geopolymer Concrete.

Mix ID	Metakaolin (%)	Sawdust (%)	7 Days	14 Days	21 Days	28 Days
Control (0% SD)	100	0	25.2	31.5	34.8	36.1
GPC-SD4	96	4	23.8	29.7	32.5	33.9
GPC-SD7	93	7	21.5	27.0	29.8	31.2
GPC-SD10	90	10	19.3	24.5	27.1	28.5
GPC-SD13	87	13	17.0	21.8	24.2	25.5
GPC-SD15	85	15	14.5	18.5	20.8	21.9
GPC-SD20	80	20	11.2	14.0	16.0	17.2
GPC-SD22	78	22	9.5	11.8	13.5	14.5
GPC-SD25	75	25	7.8	9.5	10.8	11.5
GPC-SD30	70	30	5.0	6.2	7.1	7.5
GPC-SD40	60	40	2.5	3.1	3.5	3.8
OPC (for comparison)	N/A	N/A	20.0	26.0	28.0	30.0

Oyedepo et al. [20] reported compressive strengths of 14.15 MPa, 12.96 MPa, and

11.93 MPa at 25%, 75%, and 100% sawdust replacement of sand in OPC concrete, all below

the 17 MPa minimum for lightweight structural concrete highlighting strength limitations at higher replacements. Duan et al. [4] showed that sawdust had little effect on compressive strength before 14 days, but improved later-age strength and microstructure, especially at $\leq 20\%$ addition, which is consistent with the strength gain pattern observed in this study after 14 and 28 days. Meanwhile, Ezeagu & Agbo-Anike [22] and Osei & Jackson [21] reported that increasing sawdust content generally reduced workability, density, and strength, but recognized its economic and environmental potential as a lightweight material. Overall, the findings of the current study confirm that while sawdust reduces strength beyond 15%, moderate inclusion ($\leq 10\text{--}15\%$) remains viable for lightweight or non-structural geopolymer applications.

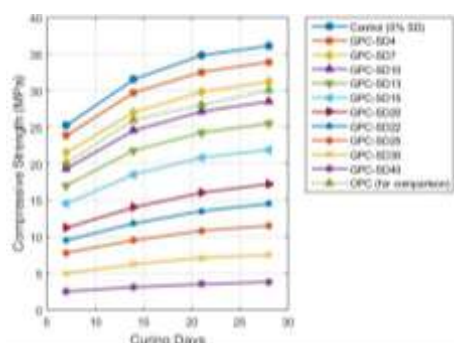


Table 7. Experimental Data (where S.R = Sawdust Replacement and C.S = Compressive Strength).

Mix ID	W/B	Water (kg/m ³)	Metakaolin (kg/m ³)	Sawdust (kg/m ³)	Sand (kg/m ³)	Granite (kg/m ³)	Alkaline Activators (kg/m ³)	S. R (%)	Slump (mm)	Curing Days	C.S (MPa)
Control	0.4	92.5	231.25	0	619.65	832.8	416.4	0	172	7	25.2
Control	0.4	92.5	231.25	0	619.65	832.8	416.4	0	172	14	31.5
Control	0.4	92.5	231.25	0	619.65	832.8	416.4	0	172	21	34.8
Control	0.4	92.5	231.25	0	619.65	832.8	416.4	0	172	28	36.1
GPC-SD4	0.4	92.5	222	9.25	619.65	832.8	416.4	4	164	7	23.8
GPC-SD4	0.4	92.5	222	9.25	619.65	832.8	416.4	4	164	14	29.7
GPC-SD4	0.4	92.5	222	9.25	619.65	832.8	416.4	4	164	21	32.5
GPC-SD4	0.4	92.5	222	9.25	619.65	832.8	416.4	4	164	28	33.9
GPC-SD7	0.4	92.5	215.06	16.19	619.65	832.8	416.4	7	155	7	21.5
GPC-SD7	0.4	92.5	215.06	16.19	619.65	832.8	416.4	7	155	14	27
GPC-SD7	0.4	92.5	215.06	16.19	619.65	832.8	416.4	7	155	21	29.8
GPC-SD7	0.4	92.5	215.06	16.19	619.65	832.8	416.4	7	155	28	31.2
GPC-SD10	0.4	92.5	208.12	23.12	619.65	832.8	416.4	10	137	7	19.3
GPC-SD10	0.4	92.5	208.12	23.12	619.65	832.8	416.4	10	137	14	24.5
GPC-SD10	0.4	92.5	208.12	23.12	619.65	832.8	416.4	10	137	21	27.1
GPC-SD10	0.4	92.5	208.12	23.12	619.65	832.8	416.4	10	137	28	28.5
GPC-SD13	0.4	92.5	201.19	30.06	619.65	832.8	416.4	13	105	7	17
GPC-SD13	0.4	92.5	201.19	30.06	619.65	832.8	416.4	13	105	14	21.8
GPC-SD13	0.4	92.5	201.19	30.06	619.65	832.8	416.4	13	105	21	24.2
GPC-SD13	0.4	92.5	201.19	30.06	619.65	832.8	416.4	13	105	28	25.5
GPC-SD15	0.4	92.5	196.56	34.69	619.65	832.8	416.4	15	78	7	14.5
GPC-SD15	0.4	92.5	196.56	34.69	619.65	832.8	416.4	15	78	14	18.5
GPC-SD15	0.4	92.5	196.56	34.69	619.65	832.8	416.4	15	78	21	20.8
GPC-SD15	0.4	92.5	196.56	34.69	619.65	832.8	416.4	15	78	28	21.9
GPC-SD20	0.4	92.5	185	46.25	619.65	832.8	416.4	20	56	7	11.2
GPC-SD20	0.4	92.5	185	46.25	619.65	832.8	416.4	20	56	14	14
GPC-SD20	0.4	92.5	185	46.25	619.65	832.8	416.4	20	56	21	16
GPC-SD20	0.4	92.5	185	46.25	619.65	832.8	416.4	20	56	28	17.2
GPC-SD22	0.4	92.5	180.38	50.88	619.65	832.8	416.4	22	33	7	9.5
GPC-SD22	0.4	92.5	180.38	50.88	619.65	832.8	416.4	22	33	14	11.8
GPC-SD22	0.4	92.5	180.38	50.88	619.65	832.8	416.4	22	33	21	13.5
GPC-SD22	0.4	92.5	180.38	50.88	619.65	832.8	416.4	22	33	28	14.5
GPC-SD25	0.4	92.5	173.44	57.81	619.65	832.8	416.4	25	14	7	7.8
GPC-SD25	0.4	92.5	173.44	57.81	619.65	832.8	416.4	25	14	14	9.5
GPC-SD25	0.4	92.5	173.44	57.81	619.65	832.8	416.4	25	14	21	10.8
GPC-SD25	0.4	92.5	173.44	57.81	619.65	832.8	416.4	25	14	28	11.5
GPC-SD30	0.4	92.5	161.88	69.38	619.65	832.8	416.4	30	0	7	5
GPC-SD30	0.4	92.5	161.88	69.38	619.65	832.8	416.4	30	0	14	6.2
GPC-SD30	0.4	92.5	161.88	69.38	619.65	832.8	416.4	30	0	21	7.1
GPC-SD30	0.4	92.5	161.88	69.38	619.65	832.8	416.4	30	0	28	7.5
GPC-SD40	0.4	92.5	138.75	92.5	619.65	832.8	416.4	40	0	7	2.5
GPC-SD40	0.4	92.5	138.75	92.5	619.65	832.8	416.4	40	0	14	3.1
GPC-SD40	0.4	92.5	138.75	92.5	619.65	832.8	416.4	40	0	21	3.5
GPC-SD40	0.4	92.5	138.75	92.5	619.65	832.8	416.4	40	0	28	3.8

3.7 MATLAB Neural Network Training Summary

Fig. 10 shows a 3-layer neural network (4-10-1 architecture) trained with Bayesian Regularization to minimize Mean Squared Error. After 157 iterations, the model achieved a low MSE of $2.37\text{e-}05$ with no validation failures. A gradient of 1.60 and Mu of 0.005 suggest stable and effective learning. The setup reflects efficient training and strong generalization.



Fig. 10. Neural Network Training Progress and Configuration in MATLAB.

ANN General Performance

The Artificial Neural Network (ANN) model employed in this study exhibited excellent predictive performance in estimating the compressive strength of geopolymer concrete incorporating sawdust and metakaolin. During training, the model achieved a coefficient of determination (R^2) of 0.9982, with a Mean Squared Error (MSE) of 0.2632, Root Mean Squared Error (RMSE) of 0.5130, and Mean Absolute Error (MAE) of 0.3462. These values indicate a highly accurate fit to the training data, demonstrating the model's ability to capture the complex nonlinear relationships among the input variables (metakaolin percentage, sawdust percentage, and curing days). In the testing phase, the ANN maintained strong generalization capabilities, achieving an R^2 of 0.9423, MSE of 2.1898, RMSE of 1.4798, and MAE of 1.2297. Although the error metrics were slightly higher compared to the training phase, the model still delivered reliable predictions on unseen data. The relatively small difference between training and testing R^2 values suggests that overfitting

was effectively minimized. Furthermore, the network architecture, consisting of three input neurons, ten hidden neurons, and one output neuron, proved to be suitably optimized for the dataset. The use of the Levenberg-Marquardt algorithm during training contributed to the model's rapid convergence and high accuracy. Overall, the ANN model outperformed the ANFIS and GEP model (presented in my colleagues work) in terms of generalization and error metrics, confirming its suitability as a robust and reliable predictive tool for compressive strength modeling in geopolymer concrete systems.

Model Prediction Accuracy across Data Sets

Fig. 11 presents four scatter plots illustrating the performance of a predictive model across different data sets: training, validation, test, and all combined. Each plot shows predicted values versus actual target values with a regression line and accompanying equation. The high correlation coefficients (close to 0.999) and slope values near 1 indicate strong linear alignment and minimal prediction error, confirming that the model performs consistently and accurately, not just on the data it was trained on, but also on unseen data. This reinforces its generalization capability and reliability.

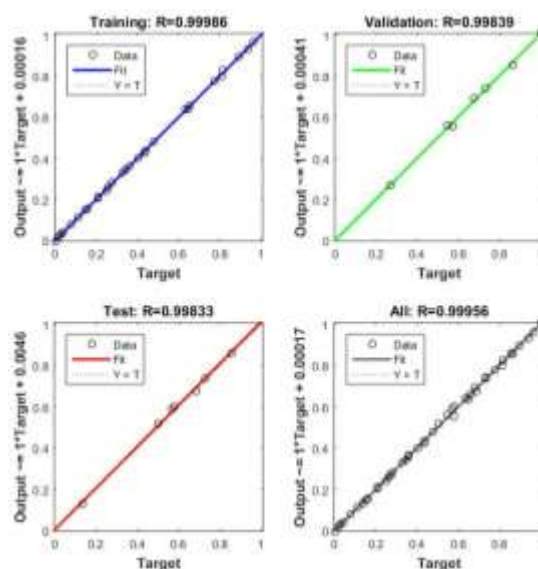


Fig. 11. Regression Analysis of Model Performance across Data Sets.

Model Error Distribution across Datasets

Fig. 12 illustrates the distribution of prediction errors and differences between targets and

outputs, across training, validation, and test sets using a 20-bin histogram. The dominant clustering of errors near zero, indicated by the height of bars around the center and the orange marker at zero error, signals strong model accuracy overall. The presence of slightly wider error margins in the validation and test sets, compared to training, hints at potential generalization gaps that may warrant attention. Color coding enhances interpretability: blue for training, green for validation, and red for test data. Overall, the error pattern supports the model's predictive reliability, while also guiding future improvements to minimize deviation and reinforce robustness across datasets.

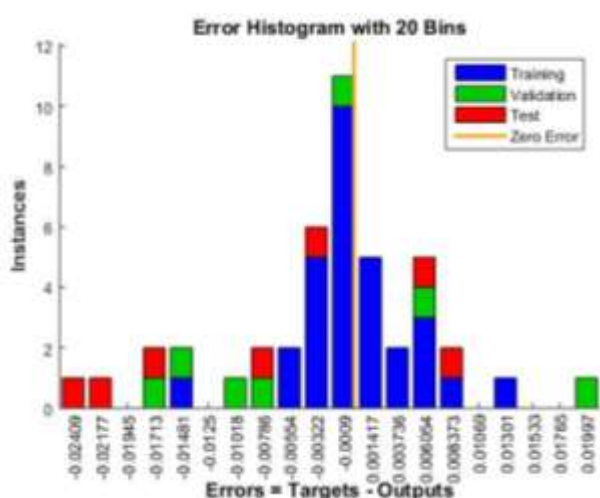


Fig. 12. Error Distribution with 20 Bins.

ANN Training Dynamics and Convergence Behavior

Fig. 13 shows the ANN training process over 157 epochs. The gradient reduced to 2.3584×10^{-5} , indicating convergence. The mu value increased only at the final epoch, showing stable learning. The number of parameters stabilized at 15.937, while the sum of squared weights (ssX) settled at 5.1855, confirming weight stability. No validation checks occurred, indicating consistent generalization and no overfitting.

ANN Model Error Performance Evaluation

Fig. 14 illustrates the Mean Squared Error (MSE) during training, validation, and testing over 157 epochs. The best validation performance of 0.00016768 was achieved at epoch 156, showing excellent model generalization. The validation and test curves closely follow each other,

indicating minimal overfitting. The early and sharp drop in MSE confirms fast learning, while the final convergence near the minimum value highlights training stability and good predictive accuracy.

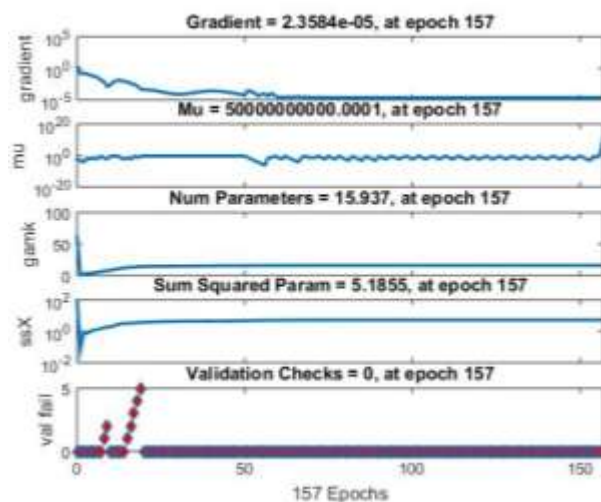


Fig. 13. ANN Training Metrics over 157 Epochs.

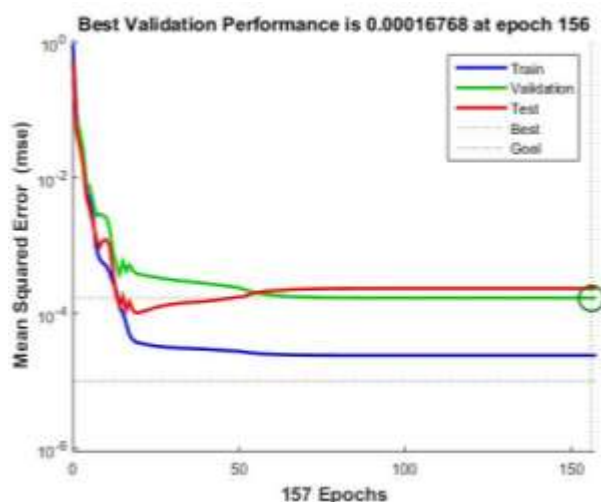


Fig. 14. ANN Training, Validation, and Test MSE Performance Curve.

ANN-Based Prediction of Compressive Strength Response to Material Variation

Fig. 15 presents a 3D surface plot showing the ANN-predicted compressive strength of geopolymer concrete at 28 days as a function of sawdust replacement (%) and metakaolin content (kg/m^3), with slump held constant at 83.1 mm. The compressive strength increases with higher metakaolin content and decreases as sawdust replacement rises. This trend reflects the strengthening role of metakaolin and the diluting effect of sawdust, aligning with the

experimental observations. The smooth gradient confirms the ANN model’s strong predictive capacity across the input variables.

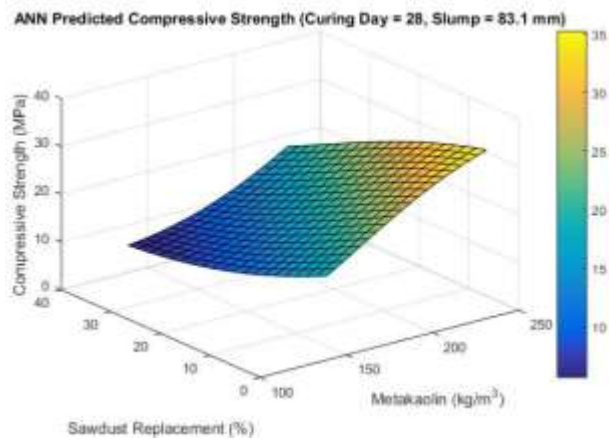


Fig. 15. ANN Surface Plot of Predicted Compressive Strength at 28 Days (Slump = 83.1 mm).

3.8 Adaptive Neuro Fuzzy Inference System (ANFIS)

Fig. 16 details the ANFIS Model architecture and parameters, displaying a Sugeno-type system with its three inputs, single output, and specific fuzzy operators. Complementing this, the Figure also shows the Fuzzy Inference System Rule Viewer, which visually traces the ANFIS's inference for given inputs (all 0.5), showing rule activation and the derivation of the crisp output (0.506). Together, these figures comprehensively illustrate both the design and the detailed operational mechanics of the ANFIS model.

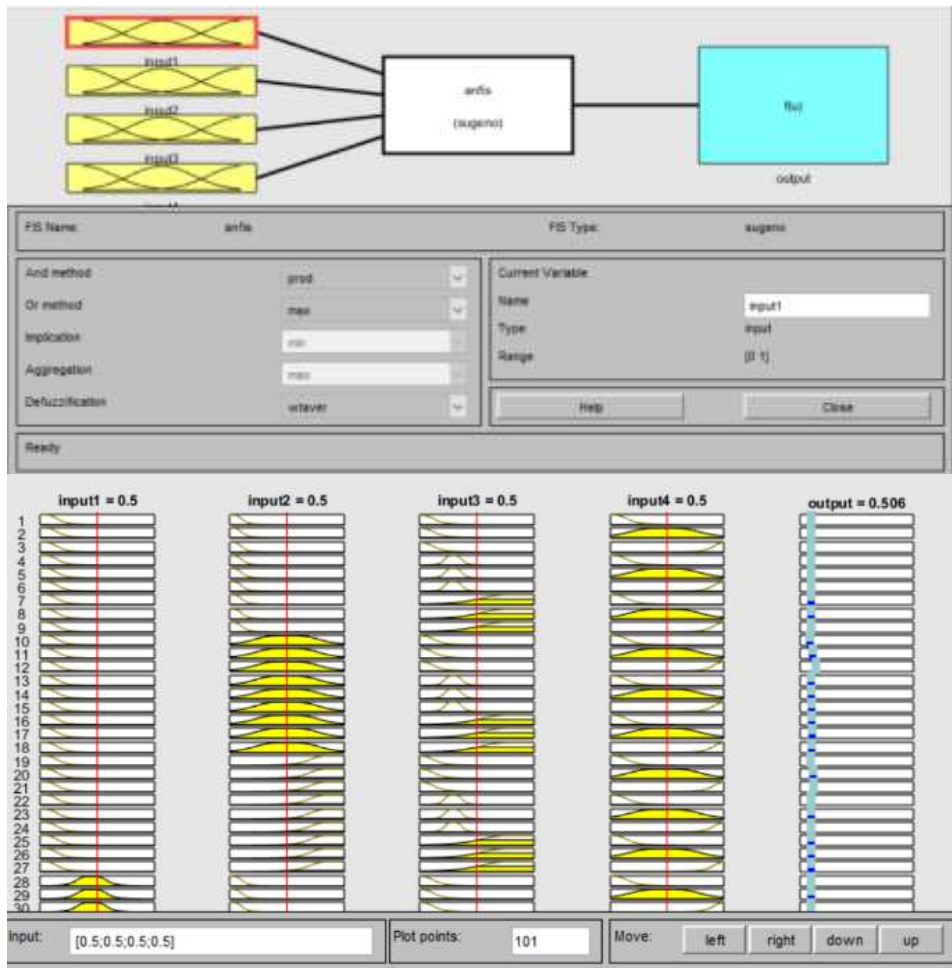


Fig. 16. ANFIS Model Architecture and Inference Process.

ANFIS General Performance

The Adaptive Neuro Fuzzy Inference System (ANFIS) was implemented using 30 training data

pairs and 7 checking (testing) data pairs to model the relationship between the selected input variables and the output. The model architecture consisted of 193 nodes, 405 linear parameters, 36

nonlinear parameters, totaling 441 modifiable parameters, and a complex rule base made up of 81 fuzzy rules. A warning was issued indicating that the number of training data points was lower than the number of parameters, suggesting a potential risk of overfitting. This concern was reflected in the training performance expressed in Fig., where the model achieved a perfect goodness of fit with an R^2 value of 1.0000 and zero errors across the MSE, RMSE, and MAE metrics. This result indicates that the model fit the training data exceptionally well, possibly memorizing it rather than learning generalized patterns.

However, the performance on the checking (test) set provided a more realistic measure of generalization capability. The testing results showed an R^2 of 0.7584, MSE of 12.1267, RMSE of 3.4823, and MAE of 2.0608 (Fig. 17). These values suggest that although the model still performs fairly well on unseen data, the accuracy drops noticeably compared to the training set, confirming the initial concern about overfitting due to the high number of parameters relative to available data. Additionally, the training process displayed progressive learning, with the step size increasing steadily over epochs from 0.0110 after epoch 5 to 0.0214 after epoch 36, indicating stable convergence behavior.

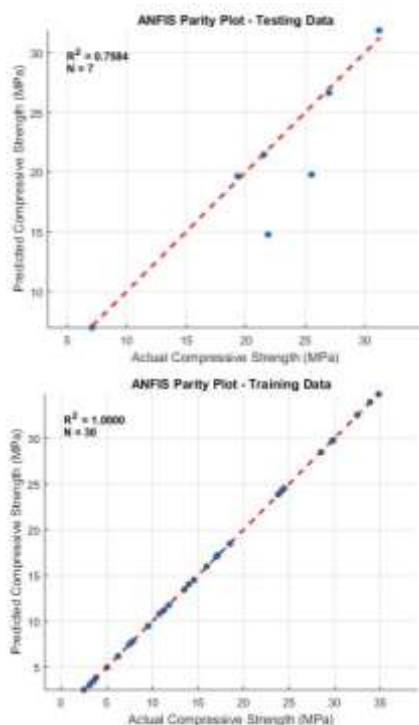


Fig. 17. ANFIS Parity Plots.

Nevertheless, the generated 3D surface plots represented by Fig. 20 provide meaningful visualization of the interactions between input variables and their influence on the predicted output. In overall, the ANFIS model demonstrated excellent performance on training data and acceptable generalization on the testing set, but improvements could be made by simplifying the model structure or increasing the dataset size to enhance robustness and reduce overfitting.

ANFIS Model Error Analysis and Membership Functions

Fig. 18 displays histograms of ANFIS prediction errors (Actual - Predicted) for both training (left) and testing (right) datasets, with a superimposed normal distribution curve. Both plots show errors largely centered around zero, indicating accurate model predictions for both seen and unseen data. The distributions highlight the model's performance and generalization ability. Fig. 19 shows provides the input membership functions, each variable is characterized by three fuzzy sets (e.g., low, medium, high), and indicating the degree of belongingness for values within a 0-1 normalized range.

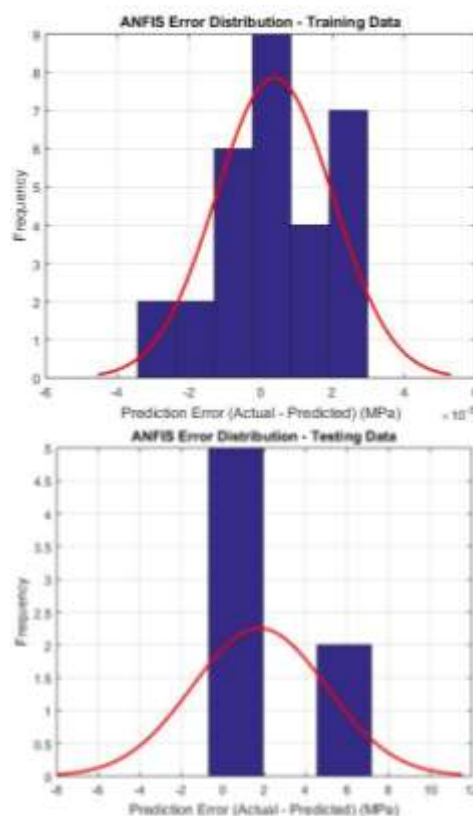


Fig. 18. ANFIS Prediction Error Distribution.

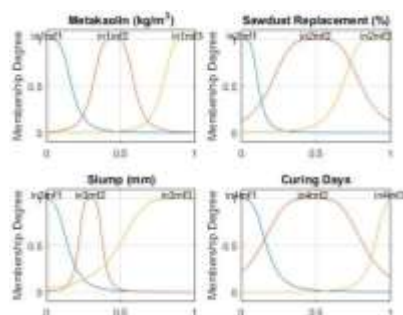


Fig. 19. Input Membership Functions.

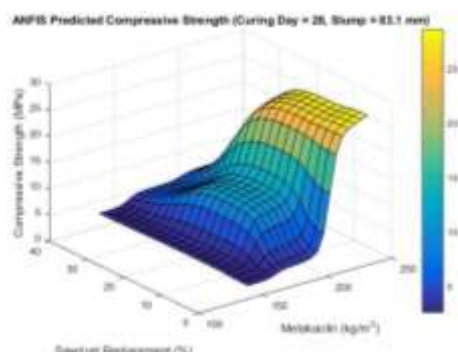


Fig. 20. ANFIS 3D Surface Plots.

3.9 Performance Evaluation of GEP Model Training

The Gene Expression Programming (GEP) model was trained over 20 generations, showing rapid

convergence to an optimal solution. Initially, both the population average fitness and the best individual fitness improved significantly from generation 0 to 1, with best fitness reducing from 0.00449 to 0.00397. By generation 6, a consistent minimum fitness value of 0.0136981 was achieved and maintained across subsequent generations, indicating convergence and model stability. The chromosome length of the best individual also stabilized at 1, showing that the model had minimized redundancy and complexity while preserving predictive accuracy. Table 8 shows the evolution of fitness and chromosome length during GEP training.

GEP Model Performance Discussion

The Gene Expression Programming (GEP) model exhibited moderate-to-good predictive performance across the training and testing phases. On the training set, the model achieved an R^2 of 0.8354, indicating that approximately 83.5% of the variance in compressive strength is captured by the evolved expression. Correspondingly, the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) values were 3.93 MPa and 3.20 MPa, respectively, suggesting reasonably low deviation between predicted and actual values.

Table 8. Evolution of Fitness and Chromosome Length during GEP Training.

Generation	Avg. Chromosome Length	Avg. Fitness	Best Chromosome Length	Best Fitness	OOB Fitness	Time Left
0	11.88	164249	9	0.00449082	N/A	1.72m
1	4.53	2.40773	10	0.00396707	N/A	1.23m
2	1.73	0.0574911	5	0.00823218	N/A	1.45m
3	1.04	0.0828585	6	0.013697	N/A	1.02m
4	1.03	0.234424	7	0.008207	N/A	1.00m
5	1.03	0.232288	1	0.0136981	N/A	1.22m
6	1.05	0.0310682	1	0.0136981	N/A	50.40s
7	1.04	0.0272813	1	0.0136981	N/A	1.13m
8	1.04	0.211432	1	0.0136981	N/A	45.94s
9	1.03	0.0303183	1	0.0136981	N/A	37.32s
10	1.04	0.0261529	1	0.0136981	N/A	37.33s
11	1.03	0.0251979	1	0.0136981	N/A	37.47s
12	1.03	0.0329538	1	0.0136981	N/A	28.43s
13	1.04	2.99697	1	0.0136981	N/A	22.87s
14	1.03	0.0277006	1	0.0136981	N/A	25.06s
15	1.05	0.576413	1	0.0136981	N/A	15.30s
16	1.04	0.122667	1	0.0136981	N/A	12.23s
17	1.04	0.0306872	1	0.0136981	N/A	10.40s
18	1.03	0.0288972	1	0.0136981	N/A	3.81s
19	1.04	0.108793	1	0.0136981	N/A	0.00s

However, on the testing set, the performance slightly declined with an R^2 of 0.6532, showing that the model explains around 65.3% of the variation in unseen data. The RMSE and MAE increased to 5.69 MPa and 4.31 MPa, respectively, which indicates a drop in generalization accuracy. This reduction in predictive accuracy on the test set is common in symbolic regression models and may be attributed to overfitting or the structural simplicity of the evolved expression (X_2 , in this case), which might not fully capture the nonlinear interactions among all variables. The parity plots shown in Fig. 21 support these findings: training predictions align well with actual values, while testing predictions show greater scatter. Overall, the model performs acceptably but may benefit from further tuning or feature expansion.

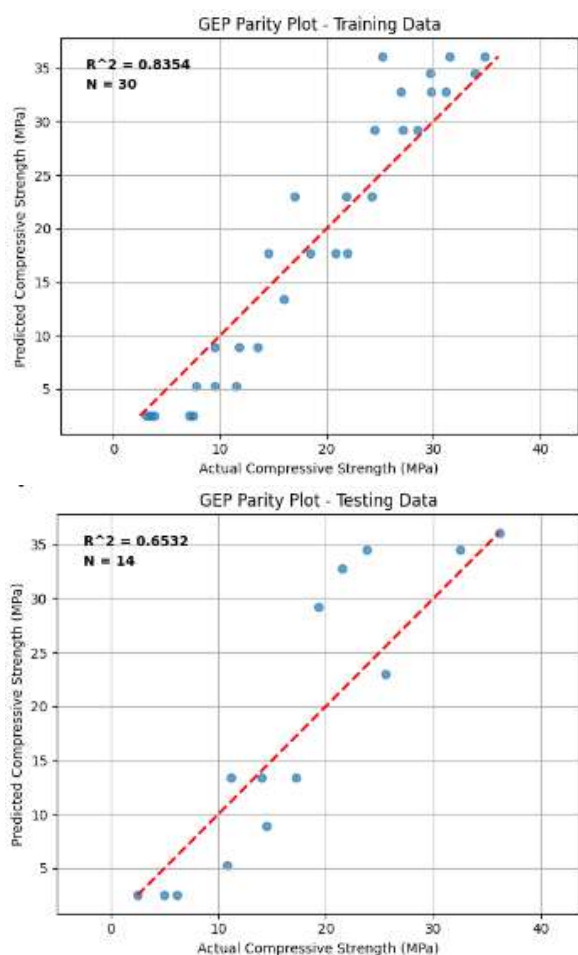


Fig. 21. GEP Parity Plots.

GEP Model Error Analysis

Fig. 22 displays histograms of the GEP model's prediction errors (Actual - Predicted) for both training (left) and testing (right) datasets. The

errors are broadly distributed, spanning a significant range for both sets. The training errors show peaks, and the testing errors exhibit noticeable peaks away from zero (e.g., around -10 MPa and between 2.5-5 MPa). This indicates that the GEP model's predictions are not consistently centered around zero and show a wider, more scattered error distribution, particularly for unseen data.

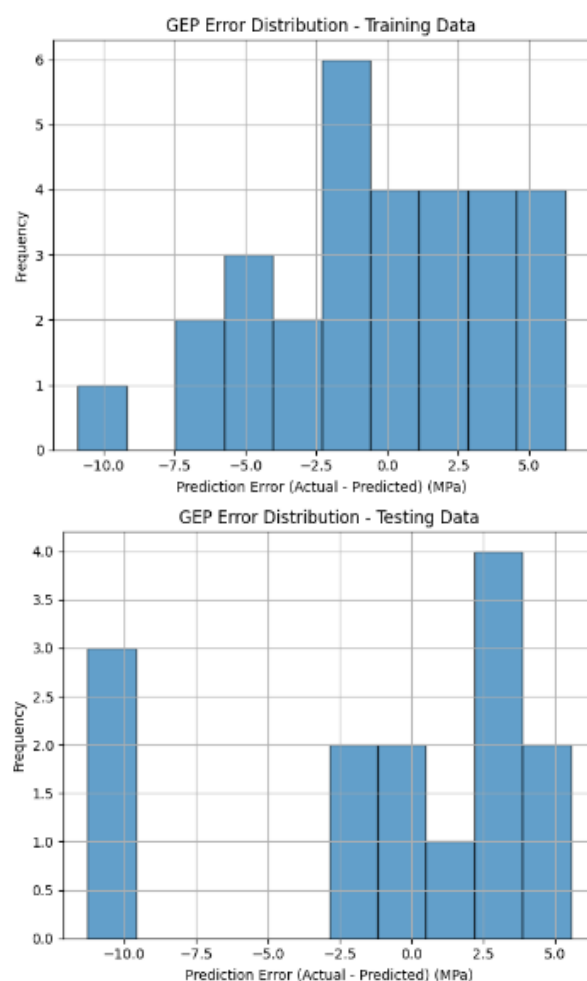


Fig. 22. GEP Prediction Error Distribution.

GEP Model Response Surface Analysis

Fig. 23 (GEP Predicted Compressive Strength, Curing Day = 28, Slump = 17.5 mm) visualizes the GEP model's predicted compressive strength. It shows how strength subtly varies with Metakaolin content and Sawdust Replacement percentage, holding Curing Day and Slump constant. The largely flat surface indicates a consistent, albeit low-ranging (5.4-6.4 MPa), predicted strength within the shown input ranges under these fixed conditions.

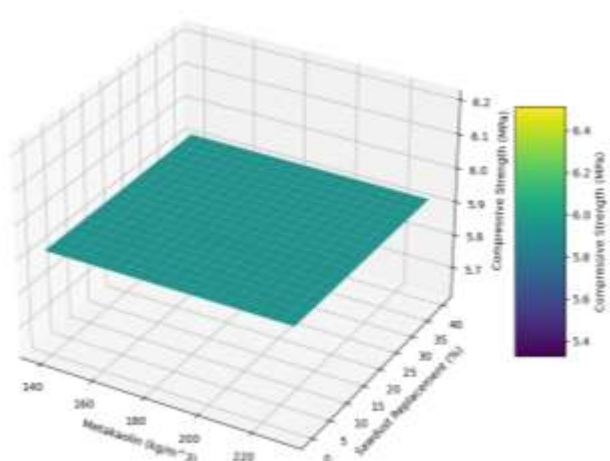


Fig. 23. GEP 3D Surface plots (GEP Predicted Compressive Strength, Curing Day = 28, Slump = 17.5 mm).

3.10 Model Performance Comparison

This study generally compared Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), and Gene Expression Programming (GEP) for predicting geopolymer concrete compressive strength, revealing distinct characteristics and performance trade-offs. The ANN, a robust "black-box" model, demonstrated exceptional predictive capability, achieving an R^2 of 0.9982 on the training data and maintaining strong generalization with an R^2 of 0.9423 and an RMSE of 1.48 MPa on the unseen testing data, effectively minimizing overfitting. In contrast, the ANFIS, a complex hybrid system, perfectly fit the training data with an R^2 of 1.0000 (RMSE 0.00 MPa), but this indicated significant overfitting, as its performance notably declined to an R^2 of 0.7584 and an RMSE of 3.48 MPa on the testing set. Finally, the GEP model provided highly interpretable explicit mathematical expressions, yielding an R^2 of 0.8354 on training. However, it exhibited the lowest generalization capability among the three, with an R^2 of 0.6532 and an RMSE of 5.69 MPa on the testing data, a likely consequence of its evolved simplicity trading off some predictive accuracy (see Table 9). Overall, while ANFIS and GEP offered varying degrees of interpretability, the ANN model proved to be the most accurate and reliable predictive tool for this specific application due to its superior generalization performance. The results of all the models are provided in Table 9.

Table 9. Model metrics comparison.

Metric	Dataset	ANN	ANFIS	GEP
R^2	Training	0.9982	1.0000	0.8354
	Testing	0.9423	0.7584	0.6532
MSE	Training	0.2632	0.0000	15.4449 (from RMSE)
	Testing	2.1898	12.1267	32.3761 (from RMSE)
RMSE	Training	0.5130	0.0000	3.93
	Testing	1.4798	3.4823	5.69
MAE	Training	0.3462	0.0000	3.20
	Testing	1.2297	2.0608	4.31

3.11 Practical Application of the study

The metakaolin-sawdust geopolymer concrete (MSGC) developed in this study presents diverse and valuable applications in real-life engineering practices, contingent upon the sawdust content and specific project requirements. While higher sawdust percentages yield genuinely lightweight concrete suitable for non-structural uses, lower sawdust inclusions result in materials capable of fulfilling structural demands. Specifically, MSGC mixes with up to approximately 20% sawdust can achieve or exceed the minimum 17 N/mm² compressive strength requirement for lightweight concrete, with the control mix (0% sawdust) notably reaching 36.1 MPa at 28 days, surpassing conventional Ordinary Portland Cement (OPC) concrete. This makes these mixes suitable for structural elements such as beams, columns, and slabs, offering a sustainable alternative that leverages waste materials. Conversely, MSGC with higher sawdust content (e.g., 22% to 40%) provides ultra-lightweight concrete (densities as low as 1400 kg/m³), which, despite lower compressive strengths (falling below 17 N/mm²), excels in non-load-bearing applications. These include insulating blocks and panels, partition walls, and void-filling materials, where reduced weight, enhanced thermal insulation, and improved acoustic properties are highly desirable. It is important to acknowledge that the study identified challenges such as significant reductions in workability (slump decreasing from 172 mm to zero at higher sawdust contents) and prolonged setting times, which would necessitate careful mix design adjustments, potentially including the use of superplasticizers, to ensure practical placement

and compaction in construction. Fundamentally, this research contributes to sustainable construction by demonstrating the viability of utilizing metakaolin and sawdust as eco-friendly alternatives, thereby reducing reliance on virgin resources and lowering the industry's carbon footprint. Furthermore, the development of a highly accurate Artificial Neural Network (ANN) model ($R^2=0.9423$) provides engineers with a robust tool to predict compressive strength, enabling optimized mix designs and more efficient resource allocation in practical applications.

4. CONCLUSION

This study comprehensively investigated the influence of metakaolin and sawdust on the fresh and hardened properties of geopolymer concrete, alongside developing predictive models for its compressive strength. The inclusion of sawdust profoundly impacted fresh concrete characteristics, with workability (slump) consistently decreasing from a high of 172 mm for the control mix (0% sawdust) to zero slump at 30% and 40% sawdust replacement, primarily due to sawdust's high water absorption capacity (31.2%). Setting times were significantly prolonged by sawdust; initial setting time increased from 53 minutes (0% sawdust) to 242 minutes (40% sawdust), and final setting time from 143 minutes to 425 minutes, attributed to organic compounds interfering with geopolymerization. In the hardened state, an inverse relationship was observed between sawdust content and bulk density, which declined from 2350 kg/m³ (0% sawdust) to 1400 kg/m³ (40% sawdust), while water absorption progressively rose from 3.5% to 25%, indicating increased porosity and reduced compactness. Regarding compressive strength, a general reduction was noted with higher sawdust content, decreasing from 36.1 MPa (0% sawdust) to 3.8 MPa (40% sawdust) at 28 days. Notably, the control mix (0% sawdust) achieved a superior 28-day strength (36.1 MPa) compared to traditional Ordinary Portland Cement (OPC) concrete (30.0 MPa), and mixes with up to 10% sawdust (28.5 MPa) remained competitive.

Regarding predictive modeling, Artificial Neural Networks (ANN) demonstrated superior performance in forecasting compressive strength, achieving a coefficient of determination

(R^2) of 0.9423 on the testing dataset, with a Mean Squared Error (MSE) of 2.1898 and a Root Mean Squared Error (RMSE) of 1.4798. While Adaptive Neuro-Fuzzy Inference Systems (ANFIS) showed perfect fit on the training data ($R^2 = 1.0000$), its testing performance ($R^2 = 0.7584$) was less robust than ANN. Gene Expression Programming (GEP) performed the least effectively with a testing R^2 of 0.6532. These quantitative results confirm that metakaolin-sawdust geopolymer concrete, particularly at lower sawdust inclusion rates, offers a viable and high-strength sustainable construction material, and ANN models are highly effective for its strength prediction.

Acknowledgement

The team is thankful to the management of Michael Okpara University of Agriculture Umudike for providing enabling environment to conduct this research work. Finally, we return all glory to God Almighty for the good health, wisdom and intelligent he provided to the team throughout the period of this research work.

REFERENCES

- [1] H. U. Ahmed, A. S. Mohammed, A. A. Mohammed, and R. H. Faraj, "Systematic multiscale models to predict the compressive strength of fly ash-based geopolymer concrete at various mixture proportions and curing regimes," *PLOS ONE*, vol. 16, no. 6, p. e0253006, 2021, doi: 10.1371/journal.pone.0253006.
- [2] H. Alabduljabbar, G. F. Huseien, A. R. M. Sam, R. Alyouef, H. A. Algaifi, and A. Alaskar, "Engineering Properties of Waste Sawdust-Based Lightweight Alkali-Activated Concrete: Experimental Assessment and Numerical Prediction," *Materials*, vol. 13, no. 23, p. 5490, 2020, doi: 10.3390/ma13235490.
- [3] N. Mehdi, H. Ghazanfarah, F. Iman, G. F. Huseien, and C. Bedon, "Investigating the fresh and mechanical properties of wood sawdust-modified lightweight geopolymer concrete," *Advances in Structural Engineering*, vol. 26, no. 7, pp. 1287–1306, 2023, doi: 10.1177/13694332231161103.
- [4] P. Duan, C. Yan, W. Zhou, and W. Luo, "Fresh properties, mechanical strength and microstructure of fly ash geopolymer paste reinforced with sawdust," *Construction and Building Materials*, vol. 111, pp. 600–610, 2016, doi: 10.1016/j.conbuildmat.2016.02.091.

- [5] H. U. Ahmed, R. R. Mostafa, A. Mohammed, P. Sihag, and A. Qadir, "Support vector regression (SVR) and grey wolf optimization (GWO) to predict the compressive strength of GGBFS-based geopolymer concrete," *Neural Computing and Applications*, vol. 35, no. 3, pp. 2909–2926, Sep. 2022, doi: 10.1007/s00521-022-07724-1.
- [6] D. A. Emarah, "Compressive strength analysis of fly ash-based geopolymer concrete using machine learning approaches," *Results in Materials*, vol. 16, p. 100347, 2022, doi: 10.1016/j.rinma.2022.100347.
- [7] O. Gift, N. Temple, and Dr. S. Samuel, "Optimization of Flexural Strength of Sawdust Ash Blended Geopolymer Concrete," *Saudi Journal of Engineering and Technology*, vol. 9, no. 09, pp. 419–432, 2024, doi: 10.36348/sjet.2024.v09i09.001.
- [8] ASTM D1784, Standard Classification System and Basis for Specification for Rigid Poly(Vinyl Chloride) (PVC) Compounds and Chlorinated Poly(Vinyl Chloride) (CPVC) Compounds, ASTM International, West Conshohocken, 2013.
- [9] BS EN 196-1, British Standards Institution. Methods of testing cement - Part 1: Determination of strength, British Standards Institution London, 2016.
- [10] ASTM C33, Standard specification for concrete aggregates. West Conshohocken, ASTM International, ASTM International, West Conshohocken, 2018.
- [11] ASTM C136, Standard test method for sieve analysis of fine and coarse aggregates, ASTM International, West Conshohocken 2014.
- [12] BS EN 12390-3, Testing hardened concrete - Part 3: Compressive strength of test specimens, British Standards Institution London, 2019.
- [13] ASTM C143/C143M, Standard test method for slump of hydraulic-cement concrete, ASTM International, West Conshohocken 2021.
- [14] ASTM C642, Standard test method for density, absorption, and voids in hardened concrete, ASTM International, West Conshohocken 2021.
- [15] BS EN 196-3, Methods of testing cement - Part 3: Determination of setting times and soundness, British Standards Institution London, 2016.
- [16] H. O. Ozioko and E. E. Eze, "Predictive modeling of Cbr and compressibility in lime stabilized lateritic soil using machine learning and Pchip data augmentation," *Discover Civil Engineering*, vol. 2, no. 1, 2025, doi: 10.1007/s44290-025-00304-x.
- [17] O. Oyedepo, S. Oluwajana, and S. Akande, "Investigation of Properties of Concrete Using Sawdust as Partial Replacement for Sand. Civil and environmental research," vol. 6, no. 2, pp. 35–42, 2014.
- [18] I. C. Onyechere, C. U. Anya, L. Anyaogu, and L. Ezeamaku, "Review on the Suitability of Bamboo as a Building Material in Nigeria," *International Journal of Research and Innovation in Applied Science*, vol. 8, no. 7, pp. 268–273, 2023, doi: 10.51584/ijrias.2023.8731.
- [19] C. M. Ikumapayi, J. A. Ajayi, and A. Fasuba, "Enhancement of Concrete Properties Using Sawdust Ash and Superplasticizer," *Advances in Science and Technology*, vol. 154, pp. 67–72, 2024, doi: 10.4028/p-b4lomt.
- [20] O. Oyedepo, S. Oluwajana, S. Akande, "Investigation of Properties of Concrete Using Sawdust as Partial Replacement for Sand," *Civil and environmental research*, vol. 6, pp. 35–42, 2014.
- [21] D.Y. Osei, and E.N. Jackson, "Compressive Strength of Concrete Using Sawdust as Aggregate," *International Journal of Scientific & Engineering Research*, vol.7, no. 4, pp. 1349–1353, 2016.
- [22] C. A. Ezeagu, and O. J. Agbo-Anike, Experimental Investigation Using Design Expert Software For Lightweight Concrete Production with Sawdust as Partial Replacement for fine Aggregates, *Nigerian Journal of Engineering*, vol. 27, no. 2, pp. 9–16, 2020.