Advances in Technology Innovation, vol. 11, no. 2, 2026, pp. 124-139

Assessment of Hot-Water-Alkali Treated Bagasse Fiber in Metakaolin-Based Geopolymer Using Machine Learning

Fransiskus Xaverius Maradona Manteiro[1], Tavio[1,*], Hosta Ardhyananta[2]

[1]Department of Civil Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

[2]Department of Materials and Metallurgical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Received 03 October 2025; received in revised form 08 December 2025; accepted 19 December 2025

DOI: https://doi.org/10.46604/aiti.2026.15755

Abstract

This study aims to develop metakaolin-based geopolymers reinforced with sugarcane bagasse fiber (BF) and to evaluate the effect of BF treatment on the composite's performance. The BF is pretreated with hot water and sodium hydroxide before being incorporated into the geopolymer matrix. Metakaolin-based geopolymer specimens containing 0%, 3%, 4%, and 5% BF by weight are prepared, and their mechanical properties and water absorption are analyzed. Scanning electron microscopy and Fourier transform infrared spectroscopy analyses reveal that the combined hot-water-alkali treatment significantly modifies the fiber surface. The treatment removes impurities and increases surface roughness, thereby enhancing fiber–matrix bonding. As a result, this treatment improves compressive and splitting tensile strength (STS) while reducing water absorption compared to untreated BF. Furthermore, machine learning algorithms, including random forest, AdaBoost, and XGBoost, are applied to predict STS. Among the three models, XGBoost demonstrates the highest predictive accuracy (R^² = 0.95, MAE = 0.28), indicating reliable predictions of mechanical strength.

Keywords: geopolymer, metakaolin, bagasse fiber, hot-water-alkali treatment, machine learning.

1. Introduction

Efforts have been made to introduce innovative cementitious materials to improve the performance of structural concrete. These advances leverage advancements in material technology and are primarily driven by the need to enhance mechanical properties, reduce self-weight, increase durability, and minimize the environmental impact of conventional cement-based materials. Consequently, researchers explored various approaches, including material modification, alternative constituents, and reinforcement techniques to achieve higher performance and more sustainable concrete systems [1-2].

In recent years, further innovation has focused on reducing the environmental footprint of concrete materials by replacing ordinary Portland cement with alternative binders. Geopolymer is a concrete mixture based on a non-cementitious binder containing silica (Si) and aluminum (Al). This binder requires less energy to produce than Portland cement. For example, 1 ton of geopolymer binder generates 0.184 tons of carbon dioxide (CO_2), compared to 1 ton from portland cement. Metakaolin (MK), a pozzolanic material obtained by calcining kaolin at high temperatures to particles smaller than 5 μm [1], is commonly used as a precursor. Sodium silicate (Na_2SiO_3) and sodium hydroxide (NaOH) are generally used as alkaline activators because they are less expensive and more widely available on the market compared to potassium silicate [4].

Agriculture generates substantial amounts of waste, with rice residues alone totaling 497.2 million tons globally in 2020 [5]. In recent years, increasing attention has been directed toward valorizing agricultural by-products as alternative raw

---

[*] Corresponding author. E-mail address: tavio@its.ac.id

materials in construction composites. Such materials include cotton, maize stalks, peach shells, miscanthus, wood chips, bagasse, apricot shells, hemp, and rice husks [6-8]. Among these, sugarcane bagasse fiber (BF) has emerged as a promising green reinforcement material. Previous studies have found that incorporating sugarcane BF into geopolymers reduced density while increasing sonic velocity [9], which benefits the acoustic and thermal insulation [10]. Moreover, elevated curing temperatures have been shown to accelerate geopolymerization, producing denser gels [11] and higher compressive strength, particularly in fly ash-based geopolymers [12]. However, high-temperature curing is energy-intensive and achieves only a modest 9% reduction in CO_2 emissions compared with cement [13]. Therefore, producing geopolymers with adequate strength under ambient curing conditions remains a major challenge.

Despite its potential advantages, BF presents several limitations when incorporated into geopolymers. Due to its hydrophilic nature, BF exhibits high water absorption, which increases porosity and reduces mechanical strength [14]. To address this issue, researchers have investigated alkali treatment with sodium hydroxide (NaOH) [15] and hot water washing at 80 ℃ for multiple cycles to remove impurities and improve fiber–matrix compatibility [16]. Nonetheless, little research has explored the combined effect of hot-water-alkali (HWA) treatment, particularly at varying fiber dosages in MK-based geopolymers.

This study investigates the influence of HWA-treated BF on the performance of MK-based geopolymer composites (BFG). BF is incorporated into BFG at dosages of 0%, 3%, 4%, and 5% under two conditions: untreated (UT3-UT5) and HWA-treated (T3-T5). The mixture containing 0% BF, referred to as plain BFG (NS), serves as the control. Compressive strength, splitting tensile strength (STS), density, and water absorption are assessed as the key mechanical properties. Fourier transform infrared spectroscopy (FTIR) and scanning electron microscopy (SEM) are used to characterize BF, while digital microscopy is employed to assess pore structure and fiber–matrix bonding. Additionally, machine learning (ML) algorithms, namely random forest (RF), AdaBoost (ADA), and XGBoost (XGB), are employed to predict STS, and feature contributions are interpreted using Shapley additive explanations (SHAP).

2. Material and Methodology

This section describes the raw materials, specimen preparation, and testing methods employed in this study. Metakaolin-based geopolymers are synthesized using commercial metakaolin and reinforced with sugarcane BF obtained from Kediri, East Java. The alkali activators (AA) consist of sodium silicate and sodium hydroxide. A series of physical, mechanical, and microstructural tests is conducted to evaluate the properties of the BF and BFG. These tests include water absorption, density, mechanical strength, microstructural analysis, and machine learning implementation.

2.1. Raw material and preparation.

The AA and distilled water used in this study are described as follows: sodium silicate ($Na_2SiO_3.9H_2O$), with a molar mass of 284.2 g/mol and a molar ratio (Si/Na) of 0.61; sodium hydroxide (NaOH), with a molar mass of 40 g/mol and purity of 99%, in flake form. Metakaolin used in this study is a commercial product. The chemical and physical compositions of this MK are shown in Tables 1 and 2.

Table 1 Chemical properties of the metakaolin

SiO_2 (%)	Al_2O_3 (%)	Fe_2O_3 (%)	CaO (%)	MgO (%)	Na_2O (%)	K_2O (%)	TiO_2 (%)	Loss on ignition (%)
55.00	40.00	1.41	0.30	0.21	0.00	0.82	1.50	1.00

Table 2 Physical properties of the metakaolin

Density ($g.cm^{-3}$)	Particle Diameter			Specific surface area (BET), m^2/g
	d_{10} (μm)	d_{50} (μm)	d_{90} (μm)	
2.42	2.67	8.14	8.14	20.00

The bagasse is extracted using a milling machine and subsequently washed with water. A portion of the bagasse is further treated by washing with hot water at 80°C for five cycles, while the remaining bagasse is left untreated. The weight ratio of bagasse to hot water is 1:20. Subsequently, 1 g of bagasse is soaked in 40 mL of a 1.5 wt% NaOH solution at room temperature for 120 minutes. After soaking, the bagasse is rinsed with distilled water repeatedly until the washing water reaches a neutral pH, as confirmed by litmus paper, and then dried in an oven at 50°C. The dried bagasse is processed into fibers using a crushing machine to produce the treated BF (T). The same procedure, without alkali treatment, is also applied to produce untreated BF (UT). To produce the geopolymer paste, a mix design adapted from previous studies [17] is used with slight modifications. The weight ratios of BF to MK are set at 0%, 3%, 4%, and 5%, while the weight ratio of MK to the AA is maintained at 1:1, with the alkali solution having a concentration of 8 M. The detailed mix design for the BFG specimens is presented in Table 3.

Table 3 Mix design of the BFG specimens

Specimens	MK:AA	Alkali Cons.	Na_2SiO_3/ NaOH	Si/Na of Sodium Silicate	BF:MK	Curing Temp.	Humidity
NS/Control	1:1	8M	1.5:1	0.61	0%	34°C	63%
UT3 ; T3	1:1	8M	1.5:1	0.61	3%	34°C	63%
UT4 ; T4	1:1	8M	1.5:1	0.61	4%	34°C	63%
UT5 : T5	1:1	8M	1.5:1	0.61	5%	34°C	63%

The liquid AA is prepared by blending the NaOH and Na_2SiO_3 solutions. The NaOH solution is first obtained by dissolving the NaOH flakes in distilled water and allowing it to cool to room temperature. The MK powder is then added to the AA solution and mixed by an electric mixer at a speed of 200 rpm for 8 min. BF is subsequently incorporated into the geopolymer paste and mixed for 3 min at a speed of 50 rpm. The resulting BFG mixture is cast into molds and manually vibrated. Curing is conducted at ambient temperature, approximately 34°C, and 63% humidity (as observed in Surabaya). The molds are fully sealed with duct tape to maintain an airtight environment throughout the curing period. The BFG specimens are produced as illustrated in Fig. 1, and the resulting specimens are shown in Fig. 2.

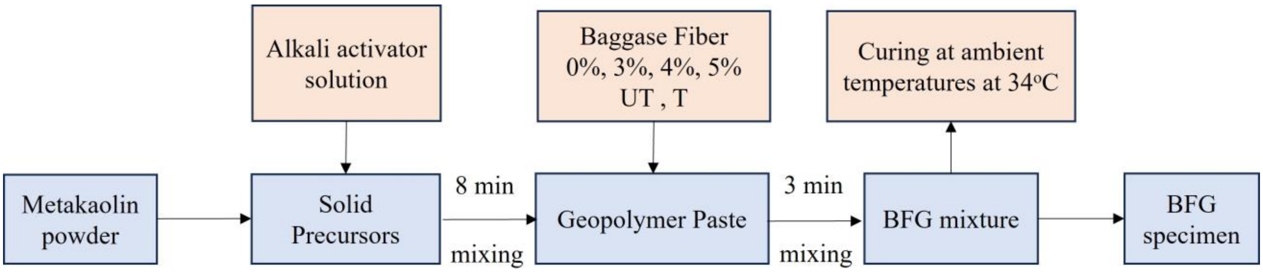


Fig. 1 Production of BFG specimens

(a) MK (b) BF (c) BFG

Fig. 2 Raw material and specimens

2.2. Methodology

(1) Water absorption of BF and the effectiveness of the treatment

The water absorption of both the HWA-treated and untreated BF is evaluated according to the American Society for Testing and Materials (ASTM) D570-22. Since cellulose degrades at temperatures above 110℃ and BF exhibits high water absorption, the fibers are dried in an oven at 50℃ ± 3℃ for 24 hours and then immersed in water for 120 minutes ± 4 minutes, following the standard test procedures outlined in the ASTM (Sections 7.1.1 and 8.2). A portion of the BF has a diameter of 1.5 to 2 mm and a length of 30 mm. The effectiveness of the treatment is assessed by the color change in the wastewater at the end of each cycle.

(2) Fourier transform infrared spectroscopy (FTIR) and scanning electron microscope (SEM)

FTIR is performed to analyze the chemical content of BF, while SEM is used to examine its morphology. Both tests are conducted in the Laboratory in the Department of Material and Metallurgical Engineering, Institut Teknologi Sepuluh Nopember.

(3) Measurement of water absorption and density of BFG

The water absorption and density of the BFG are evaluated following the procedures outlined in ASTM C642-21. For each geopolymer mixture, three cylindrical specimens with a diameter of 100 mm and a height of 50 mm are prepared and tested. Water absorption measurements are conducted at 7, 14, and 21 days.

(4) Compressive strength and STS tests

The compressive strength of the BFG is assessed using cubic specimens with a side length of 50 mm, following ASTM C579 Test B load rate I. For each composition, the compressive strength is reported as the mean value of the three specimens. Tests are conducted at 28 days using a universal testing machine (HT-2101) provided by the Bhirawa Laboratory in Surabaya, with a loading rate of 5 mm/min. The STS is determined according to ASTM C496 using three cylindrical specimens, each with a diameter of 50 mm and a height of 100 mm, with a load rate of 0.7 MPa/min.

(5) Digital microscope inspection

Visual inspection is conducted using a digital microscope with a magnification of 500X. The inspection focused on the surface of the specimens.

(6) Machine learning implementation

The models are executed on a Windows 11 system powered by an i7 processor with 10 cores running at 2.40 GHz, 16 GB of DDR5 RAM, and a GPU with 4 GB of memory. Model construction is carried out using Python in the Jupyter Lab environment. The ML algorithms used in this research are RF, ADA, and XGB. RF is an interpretable ensemble learning algorithm that improves prediction accuracy by averaging the outputs of multiple decision trees [18]. ADA is a robust ML algorithm that combines multiple weak learners to build a strong regressor, enhancing underperforming models and reducing bias [20]. The XGB is implemented using the XGBoost package in Python. In the XGB, decision trees are built progressively through iterations. During each iteration, the algorithm evaluates the bias of the current model and trains a new decision tree to minimize it [21].

Previous studies applied XGB to predict the compressive and flexural strength of coal gangue-based geopolymers and the compressive strength of calcium-based geopolymers [22]. Both XGB and ADA models in this study employ a linear approach, with the booster and the loss function set to ‘gblinear’ and ‘linear’ forms, respectively. The performance of each model is evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). Feature importance is assessed using SHAP, which fairly distributes each feature's contribution to the predictions [24]. Positive SHAP values indicate that a feature increases the prediction, while negative values indicate a decrease. The absolute SHAP values reflect the strength of the effect, and the mean absolute SHAP values represent overall feature importance. The R^2, RMSE, and MAE are computed following the approach of Zhang et al. [19], while the SHAP values are derived according to Zhao et al. [25]. All of the ML models are implemented using the open-source Python library scikit-learn. The workflow of the modeling investigation is shown in Fig. 3.

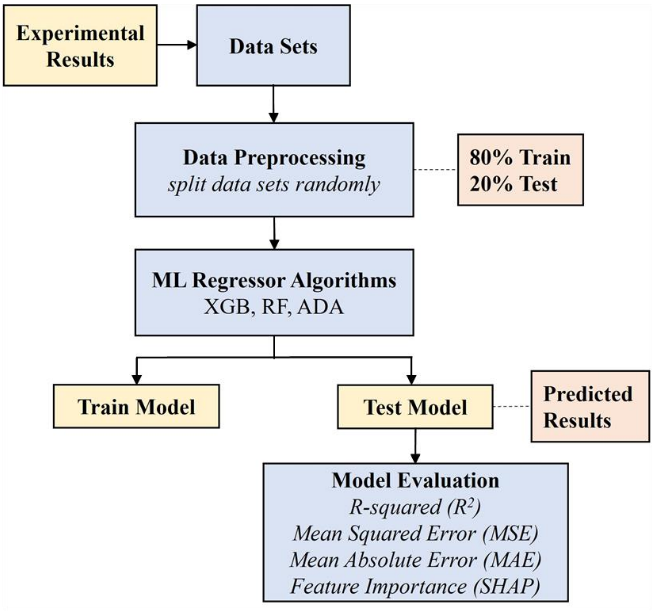


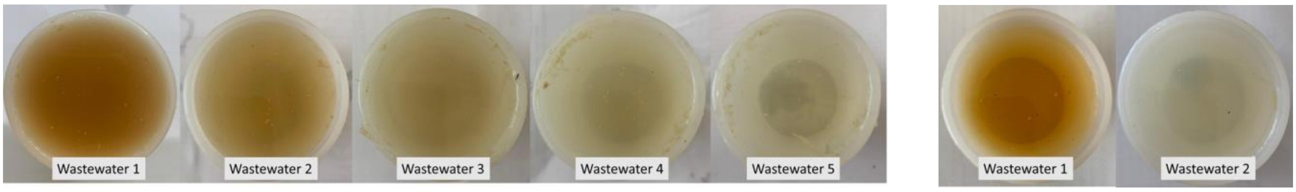
Fig. 3 Flowchart of the ML algorithms

3. Results and Discussion

This section presents the experimental results and discusses the effects of HWA treatments on BF and BFG. The analysis includes fiber treatment effectiveness, chemical and morphological changes, and their influence on physical and mechanical properties. In addition, statistical and data-driven modeling approaches are used to evaluate the relationships between material parameters and splitting tensile performance, supported by microstructural observations.

3.1. HWA treatment and water absorption of BF

As shown in Fig. 4, both the hot water treatment (a) and the alkali treatment (b) clearly show a change in the color of the wastewater from each cycle. The darker water indicates the dissolution of the organic impurities originating from the bagasse. The darker the water, the higher the concentration of the dissolved impurities. The wastewater from cycle 1 is very dark, which aligns with previous research findings. In those studies, researchers noted that 80% of the impurities from soaking the bio-aggregate are dissolved in the first cycle [16]. Furthermore, as no color change is observed between the fourth and fifth washing cycles, the washing process is deemed sufficient up to the fourth cycle.



(a) Hot water treatment (b) Alkali treatment

Fig. 4 Color change of the wastewater of BF in each cycle

This observation is further supported by the weight loss of BF after HWA treatment, as shown in Table 4, where the dry weight and water absorption of BF after treatment decrease by 12.48% and 31.88%, respectively, compared with untreated BF.

Table 4 Water absorption of untreated BF and HWA-treated BF

Type of BF	Dry weight (g)	Wet weight (g)	WA (%)
Untreated BF	21.55	99.30	360.79
HWA-treated BF	18.86	65.21	245.76

3.2. Peak Analysis

Changes in the FTIR spectra and wavenumber can reveal alterations in the chemical composition of the BF after treatment, as presented in Table 5, while the FTIR spectra are shown in Fig. 5. Notable broad peaks are observed at 3341 cm^{-1} and 3338 cm^{-1}, attributed to the hydrogen-bonded O–H stretching vibrations of hydroxyl groups in α-cellulose. The peak shifts from 3341 cm^{-1} to 3338 cm^{-1} with increased intensity, indicating greater exposure of hydroxyl groups on the cellulose surface following the removal of amorphous components. This enhancement in hydroxyl group accessibility is critical for improving fiber-matrix bonding in composite applications. The absorption peaks at 2895 cm^{-1} correspond to the symmetric and asymmetric C–H stretching vibrations of methyl and methylene groups present in hemicellulose, lignin, and cellulose. The peak of the aliphatic CH-group from lipid and wax also shifted from 2324 cm^{-1} to 2322 cm^{-1}.

Table 5 Wavenumber of untreated and HWA-treated BF

Wavenumber (cm^{-1})		Possible chemical content	Reference
Untreated	HWA-Treated		
3341	3338	Stretching vibration -OH from α-cellulose.	[26]
2895	2895	Stretching vibration C-H -CH2 from hemicellulose, lignin, and cellulose	[27]
2324	2322	Stretching vibration of aliphatic CH- from lipid and wax	[27]
1727	-	Stretching vibration C=O from hemicellulose	[26]
1603	1600	Stretching vibration aromatic structural vibration of benzene within lignin	[27]
1424	1420	Stretching vibration CO	[27]
1240	1267	Stretching vibration C-O / O-H flexing from cellulose	[27]
1034	1030	Stretching vibration C-O / C-C	[27]
833	829	β-glycosidic linkage	[27]

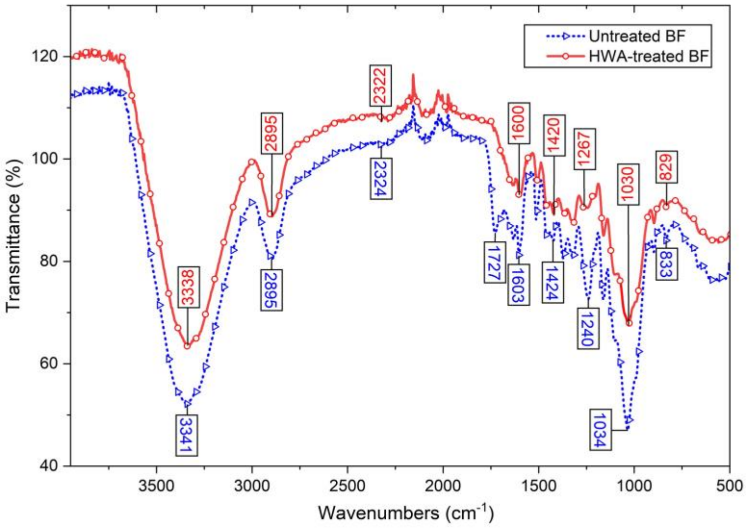


Fig. 5 FTIR spectra of HWA-treated BF and untreated BF

The most notable finding is the disappearance of the CO= Peak associated with hemicellulose in the HWA-Treated BF, indicating that the treatment effectively degraded the ester linkages associated with hemicellulose and lignin. In the FTIR spectrum of untreated BF, the characteristic peak at 1727 cm^{-1} is clearly visible. In contrast, the peak in HWA-treated BF is significantly diminished or nearly absent, confirming substantial removal of hemicellulose during HWA treatment. This observation is consistent with previous studies showing that alkali treatment can remove up to 90% of acetyl groups from bagasse.[26].

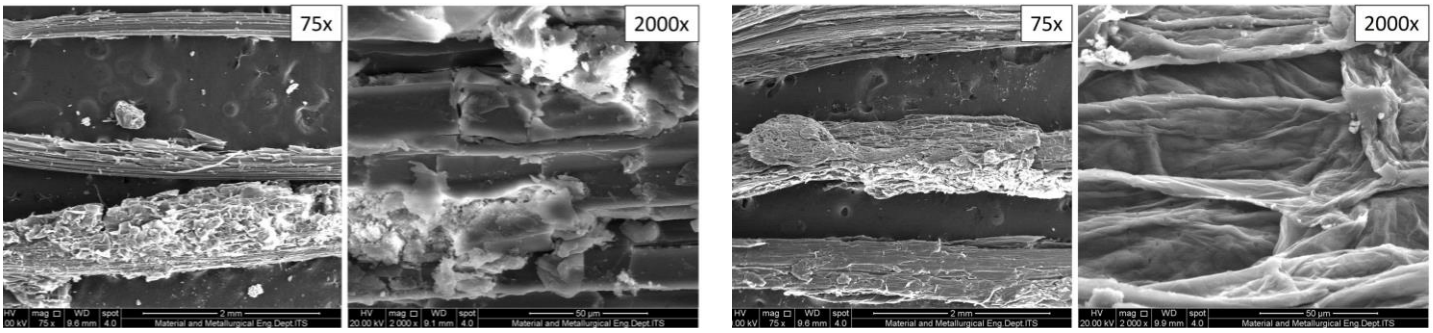
3.3. Morphological changes

The SEM images illustrate the morphological characteristics of the BF, revealing irregular and relatively coarse surface features. As shown in Fig. 6, the fiber diameters are not uniform, and the diameters are approximatelyin the range of 200 μm to 400 μm. The outer layer of the fibers is mainly composed of lignin, hemicellulose, and wax, which are commonly identified as surface-adhering components.

At 75x magnification, untreated BF appears as tightly packed and densely bound fiber bundles with minimal separation. The fibers maintain a relatively uniform and compact structure with minimal fibrillation. Individual filaments remain cemented together by lignin and intercellular substances, presenting a monolithic appearance. In contrast, significant morphological changes are observed after HWA treatment. The HWA-treated BF bundles exhibit clear fibrillation, characterized by the splitting and separation of individual fiber filaments into smaller subunits. This fibrillation is attributed to the dissolution and removal of hemicellulose that binds the elementary fibers together. Consequently, the HWA-treated BF appears more dispersed and less densely packed compared to the untreated BF.

At a magnification of 2000x, cellulose microfibrils in the HWA-treated BF are clearly visible, with an average diameter of approximately 8 μm, whereas the surface of the untreated BF remains covered by impurities. The surface of HWA-treated BF exhibits a highly textured, wrinkled, and corrugated morphology with pronounced elongated groove patterns running along the fiber axis. These grooves and ridges represent exposed cellulose microfibrils that are previously embedded beneath the non-cellulosic layer. The surface roughness is markedly enhanced due to the removal of the encasing impurity layer. These observations are consistent with previous findings [28].

SEM analysis demonstrates that HWA treatment induces pronounced morphological changes in BF. The transition from tightly bound, impurity-covered fiber bundles to fibrillated, clean-rich structures with pronounced surface texture and exposed microfibrils fundamentally alters the fiber's suitability for composite applications. The enhanced surface roughness, increased effective surface area from fibrillation, and removal of impurities collectively result in superior fiber-matrix bonding and improved load transfer efficiency in composite materials.



(a) Untreated BF (b) HWA-treated BF

Fig. 6 SEM image for untreated BF and HWA-treated BF

3.4. Density

As shown in Fig. 7, an increase in the BF dosage leads to a reduction in density. The incorporation of BF at dosages ranging from 3% to 5% results in a density decrease of approximately 8.6% to 17.8% compared with the plain BFG (NS). In contrast, for a given BF dosage, the HWA-treated BFG exhibits higher density than the untreated BFG, with increases of 7.49%, 7.28%, and 6.74% at the dosage of 3%, 4%, and 5%, respectively.

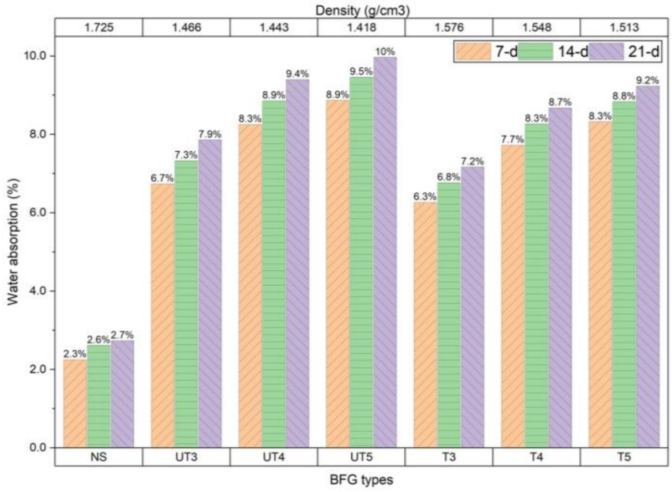


Fig. 7 Density and water absorption of BFG with various BF dosages and types at 7, 14, and 28 days

3.5. Water absorption

The water absorption of the BFG is presented in Fig. 7. The results indicate that increasing the BF dosage leads to higher water absorption for all BFG mixtures. This behavior is attributed to the intrinsic moisture uptake of BF and the formation of additional air voids within the geopolymer matrix. During mixing, BF absorbs water, which subsequently evaporates during curing, resulting in increased porosity and a corresponding reduction in density. The effectiveness of the HWA treatment in mitigating water absorption is clearly demonstrated. As shown in Fig. 7, BFG incorporating HWA-treated BF exhibits reductions in water absorption of 7.03%, 6.38%, and 6.17% at dosages of 3%, 4%, and 5%, respectively, at 7 days of age.

3.6. Mechanical properties of the BFG

Fig. 8 shows the compressive strength and split tensile strength of test specimens, while representative compression failure modes of specimens T5 and UT4 are shown in Fig. 9. The plain BFG (NS), used as the control specimen, exhibits the highest compressive strength and split tensile strength. This establishes the baseline mechanical performance of the unreinforced geopolymer matrix. The relatively high compressive strength reflects the inherent load-bearing capacity of the fully polymerized geopolymer binder. Meanwhile, the substantially lower tensile strength highlights the intrinsically brittle nature of cementitious materials.

It can be observed that the compressive strength of the composite decreases as the BF content in the BFG increases. For untreated BF (UT3–UT5), compressive strength reductions relative to plain BFG (NS) are 60.73%, 62.23%, and 64.30% at BF dosages of 3%, 4%, and 5%, respectively. On the other hand, HWA-treated BF (T3–T5) exhibits smaller reductions of 45.47%, 49.46%, and 52.98% at the same dosages. The decrease in the compressive strength is primarily attributed to pores introduced in the mixture. This finding aligns with previous studies indicated that natural fibers have limited bond compatibility with Portland cement and geopolymers [29-30]. Specifically, the layers of lignin, hemicellulose, and waxes on untreated fiber surfaces hinder effective fiber-matrix bonding, creating weak interfaces where stress concentration can initiate premature failure [28].

Examining the split tensile to compressive strength ratio (*T/C ratio*) reveals that fiber-reinforced specimens exhibit a higher tensile-compressive ratio than the plain BFG. This suggests that fibers contribute proportionally more to tensile reinforcement than to compressive strength. Such behavior is consistent with established fiber reinforcement mechanisms: fibers bridge cracks under tension and help maintain post-peak load capacity, while offering minimal reinforcement in the brittle compressive failure mode of geopolymers [17].

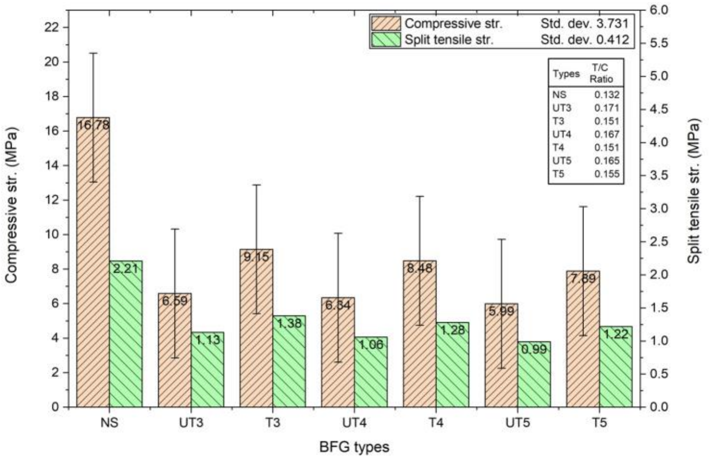


Fig. 8 Compressive and split tensile strength of BFG

Fig. 9 Typical compressive failures of BFG (T5-UT4)

Despite the overall strength reductions compared to plain BFG (NS), HWA-treated BFG (T3-T5) exhibits significantly superior properties compared to untreated BFG (UT3-UT5). The compressive strength of treated specimens is increased by 38.85%, 33.75%, and 31.72% relative to untreated BF composites at dosages of 3%, 4%, and 5%, respectively. Similarly, the incorporation of HWA-treated BF improves split tensile strength by 22.12%, 20.75%, and 23.23% compared to the corresponding untreated BF specimens.

Fig. 10 shows the compressive stress-strain behavior of the bagasse fiber-reinforced geopolymer composites. Specimens T3–T5 display both higher peak compressive strength and greater deformability than UT3–UT5. At higher fiber dosages, the composites exhibit a transition to sigmoidal stress–strain curves, indicating enhanced ductility.

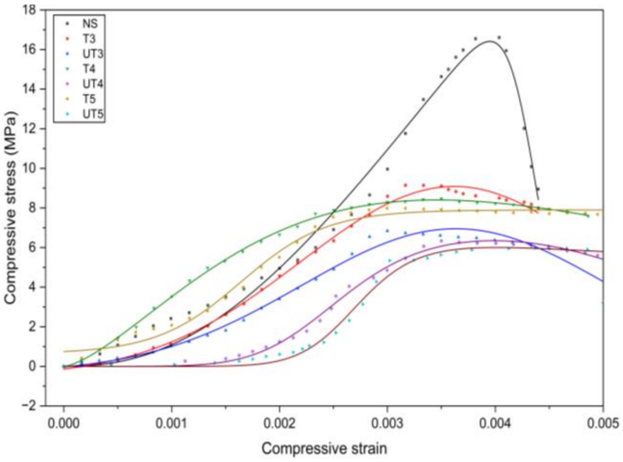


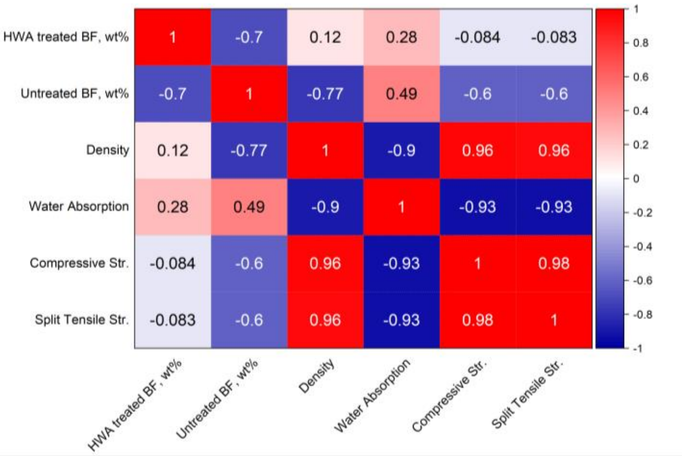
Fig. 10 Compressive Stress-Strain Relationship

For plain BFG (NS), the sharp vertical drop after peak stress indicated abrupt loss of load-bearing capacity and minimal energy absorption, which is characteristic of unreinforced cementitious matrices. In contrast, the HWA-treated BFG (T3–T5) exhibits extended plateau regions after peak stress, where the stress remains relatively high as strain increases. This pseudo-ductile response demonstrates that bridging fibers actively preserve load transfer across cracks even beyond matrix failure. Although peak strength decreases systematically with higher fiber dosage, the strain capacity remains largely unchanged, suggesting that the fiber bridging mechanism is not significantly affected by matrix dilution.

Ultimate failure in treated-fiber composites results from tensile rupture or fiber pull-out, leading to a more ductile failure mode and enhanced energy absorption. These results confirm that hwa treatment is an effective approach for improving natural fiber performance. Consistent with previous studies, alkali treatment increases fiber tensile strength and stiffness [15] [26]. In addition, alkali treatment improves the matrix compatibility through the reduction of water-soluble compounds [16]. It is also noted that ambient-temperature curing tends to reduce overall geopolymer performance.

3.7. *Correlation and regression model*

The correlation between the compressive strength, water absorption, and density is analyzed using the Pearson correlation, as shown in Fig. 11. The results indicate a significant negative correlation between the water absorption and both the mechanical properties and density of the specimens, suggesting that a higher WA ratio adversely affects the performance and density. This is likely due to the increased porosity. The results also indicate a negative correlation between the untreated BF and the density, likely due to agglomeration and increased porosity.



	HWA treated BF, wt%	Untreated BF, wt%	Density	Water Absorption	Compressive Str.	Split Tensile Str.
HWA treated BF, wt%	1	-0.7	0.12	0.28	-0.084	-0.083
Untreated BF, wt%	-0.7	1	-0.77	0.49	-0.6	-0.6
Density	0.12	-0.77	1	-0.9	0.96	0.96
Water Absorption	0.28	0.49	-0.9	1	-0.93	-0.93
Compressive Str.	-0.084	-0.6	0.96	-0.93	1	0.98
Split Tensile Str.	-0.083	-0.6	0.96	-0.93	0.98	1

Fig. 11 Pearson correlation coefficient matrix

Since density, compressive strength, and split tensile strength have a higher r-value and significant positive correlations, it is reasonable to develop a regression model between density and compressive strength, as well as between compressive strength and split tensile strength. This study proposes polynomial regression models to characterize the relationships between density and compressive strength, and between compressive strength and split tensile strength. The corresponding regression equations are presented as follows:

$$f_c' = 179.9 - 251\rho + 91.14\rho^2 \tag{1}$$

$$f_t = 0.122f_c' + 0.3403 \tag{2}$$

where f_c' is the compressive strength, f_t is the split tensile strength, and ρ is the density. Eq. (1) yields standard error (S), R^2, and adjusted R^2 values of 0.33670, 99.19%, and 99.10%, respectively. Furthermore, Eq. (2) has an S, R^2, and adjusted R^2 values of 0.0767235, 96.59%, and 96.59%, respectively.

3.8. *Data-driven modeling of split tensile strength*

(1) Construction data sets and parameter input

The experimental dataset is presented in Table 6. Since MK, AA, alkali concentration, and curing temperature are consistently applied across all specimens, these features are excluded from the analysis. Hyperparameters for the ML algorithms are optimized using GridSearchCV in Scikit-learn, as summarized in Table 7. The dataset is divided into training and testing subsets at an 80:20 split ratio, with 80% used for model training and 20% for performance evaluation. This approach ensures a robust assessment of the model's generalization capability.

Table 6 Experiment dataset

Specimens	HWA-treated BF, wt%	Untreated BF, wt%	Density, g/cm3	Water Absorption	Compressive Str., MPa	Split Tensile Str., MPa
NS_1	0%	0%	1.713	2.16%	16.50	1.98
NS_2	0%	0%	1.726	2.27%	16.71	2.28
NS_3	0%	0%	1.735	2.31%	17.19	2.37
UT3_1	0%	3%	1.455	6.65%	6.15	1.07
UT3_2	0%	3%	1.468	6.74%	6.69	1.11
UT3_3	0%	3%	1.476	6.82%	6.93	1.21
UT4_1	0%	4%	1.432	8.15%	5.97	0.99
UT4_2	0%	4%	1.445	8.25%	6.41	1.05
UT4_3	0%	4%	1.452	8.35%	6.64	1.14
UT5_1	0%	5%	1.413	8.82%	5.75	0.86
UT5_2	0%	5%	1.417	8.87%	5.87	0.97
UT5_3	0%	5%	1.423	8.91%	6.35	1.05
T3_1	3%	0%	1.569	6.23%	8.93	1.26
T3_2	3%	0%	1.578	6.27%	9.18	1.41
T3_3	3%	0%	1.582	6.29%	9.34	1.47
T4_1	4%	0%	1.520	7.65%	8.21	1.19
T4_2	4%	0%	1.547	7.73%	8.41	1.27
T4_3	4%	0%	1.578	7.79%	8.82	1.35
T5_1	5%	0%	1.499	8.21%	7.67	1.18
T5_2	5%	0%	1.517	8.35%	7.83	1.22
T5_3	5%	0%	1.522	8.40%	8.17	1.29

Table 7 Parameters of the ML Algorithms

ML Algorithm	Regressor Package	Parameter
Random Forest (RF)	RandomForestRegressor( )	n_estimators=200, min_samples_split=2, min_samples_leaf=1, random_state=42
Adaptive Boosting (ADA)	AdaBoostRegressor( )	n_estimators=1800, learning_rate=1, loss= "linear", random_state=42
eXtreme Gradient Boosting (XGB)	xgb.XGBRegressor( )	n_estimators=1800, learning_rate=0.01, booster="gblinear", objective='reg:squarederror', random_state=42

(2) Evaluation model and data analysis

The three models are evaluated using statistical metrics, including R^2, MAE, and RMSE. The evaluation results are shown in Table 8 and Fig. 12. As Table 8 indicates, XGB outperforms the other algorithms, achieving an R^2 value of 0.925 and a low MAE of 0.086. Train RMSE of 0.047 and Test RMSE of 0.121 suggest a good balance between fitting the training data and generalizing to the test data, with minimal overfitting. XGB, using the GBLinear booster, effectively captures the underlying trend while avoiding overfitting due to its linear approximation approach.

The RF metrics are close to those of XGB, confirming RF as a robust model on this dataset. In contrast, the ADA produces the highest test error and the lowest R^2 and a low Train RMSE of 0.010 compared to Test RMSE, suggesting potential overfitting. Fig. 12 presents the confidence bands of all models where the true regression relationship lies with 95% confidence. The XGB model exhibits markedly narrow confidence bands (orange/yellow), reflecting substantially lower prediction uncertainty and correspondingly higher model confidence in predictions. Conversely, the ADA model demonstrates considerably wider confidence bands (purple/light purple), suggesting elevated prediction uncertainty and diminished reliability.

Table 8 Model performance evaluation for predicting the split tensile strength.

ML Algorithm	Index			
	MAE	R^2	Train RMSE	Test RMSE
RF Regressor	0.09	0.921	0.102	0.124
ADA Regressor	0.123	0.860	0.010	0.165
XGB Regressor	0.086	0.925	0.047	0.121

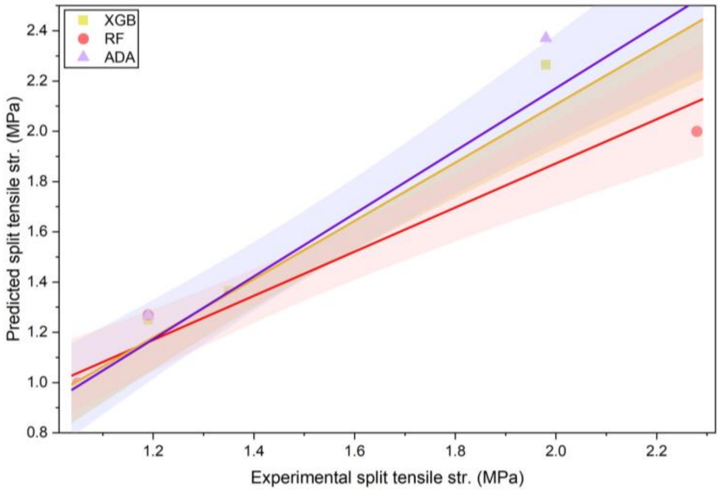


Fig. 12 Split tensile: actual vs XGB, ADA, RF predicted results

The RF model occupies an intermediate position, with moderately wide confidence bands (red/pink), reflecting acceptable yet suboptimal confidence levels in predictive estimation. Fig. 13 presents a boxplot comparison of experimental and predicted split tensile strength across all ML models. The XGB model has the highest median value of 1.25 MPa, a mean of 1.48 MPa, and the most compact interquartile range (IQR) of approximately 1.05-2.26 MPa. The relatively tight box indicates consistent predictions with minimal variability, demonstrating reliable and stable model performance. The narrow prediction range and close alignment with the experimental median of 1.21 MPa highlight strong generalization capability. XGB and RF demonstrate approximately symmetric distributions with evenly extended whiskers, indicating normally distributed prediction errors with few extreme outliers. The ADA exhibits an asymmetric distribution with a notably extended upper whisker to 2.37 MPa, suggesting occasional high outlier predictions and right-skewed prediction errors.

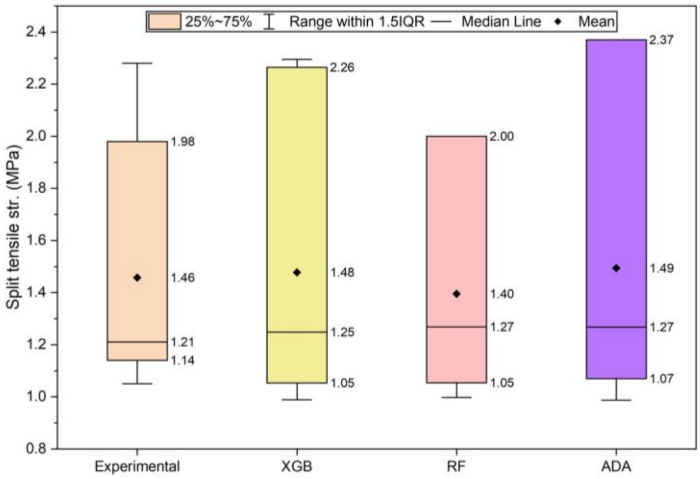


Fig. 13 Boxplot comparison of experimental and predicted split tensile strength

The feature importance in the prediction results of each algorithm is illustrated in Fig. 14. The larger the absolute mean SHAP value, the greater the feature's contribution to the overall prediction. Compressive strength, water absorption, and density are identified as the most influential features in the prediction for all algorithms.

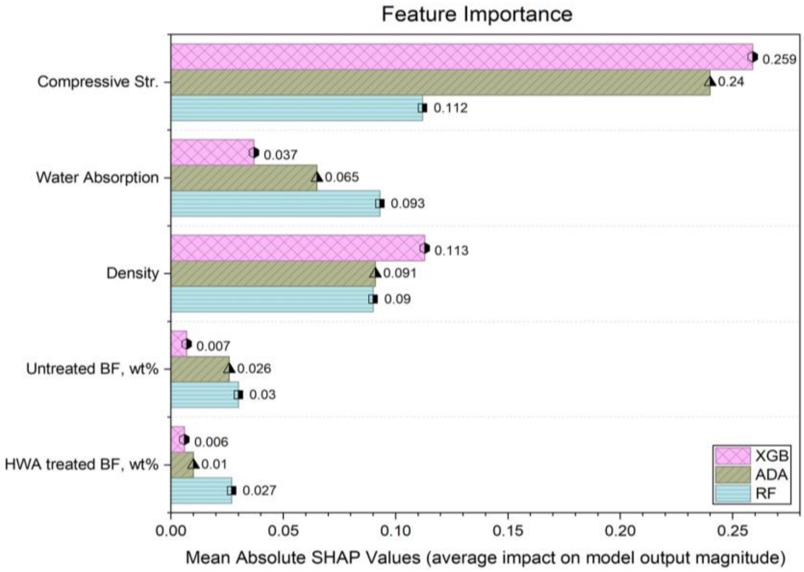


Fig. 14 Feature importance factor in the prediction models

Compressive strength emerges as the most influential predictor across all three models, exhibiting substantially higher mean absolute SHAP values compared to all other features. For the XGB, the compressive strength demonstrates a mean SHAP value of 0.259, representing approximately 2.3 times the influence of the second-ranked feature. The ADA shows a similarly high value of 0.24, while the RF registers 0.112. This pronounced dominance indicates that compressive strength is the primary determinant of split tensile strength predictions. This is consistent with the established mechanical relationships between compressive and tensile properties in cementitious composites.

Water absorption and density constitute secondary yet significant features in the prediction framework, indicating a moderate influence on model predictions. The relatively higher importance in the RF model suggests that the RF may place greater emphasis on porosity-related characteristics, as water absorption directly reflects the pore structure and connectivity within the geopolymer matrix. Density shows comparable influence across models, with the consistency of density contributions across algorithms indicating a reliable role as a predictor of mechanical properties. Both water absorption and density are intrinsically linked to the quality of the geopolymer matrix, degree of polymerization, and the presence of microcracks or voids. All of these factors have a direct impact on tensile strength capacity.

The results indicate that all models can effectively capture the relationships among features in the dataset, with the overall complexity classified as moderate. This also highlights the importance of the dataset size and feature variability for ensuring model stability and predictive accuracy.

3.9. Digital microscope inspection

Fig. 15 shows the visualization of the BFG specimens after the compressive test. In specimen UT4, some unreacted particles are observed within the gel, and the BF bridging gaps inside the matrix is clearly visible. This indicates that a higher BF tensile strength corresponds to a higher BFG tensile strength. Therefore, the BFG using the HWA-treated BF has a higher strength compared to the untreated BF [15], as validated by the compressive test results. Specimen UT5 shows the agglomeration of BF. This agglomeration increases the porosity of the BFG and may result in a lower performance of the BFG.

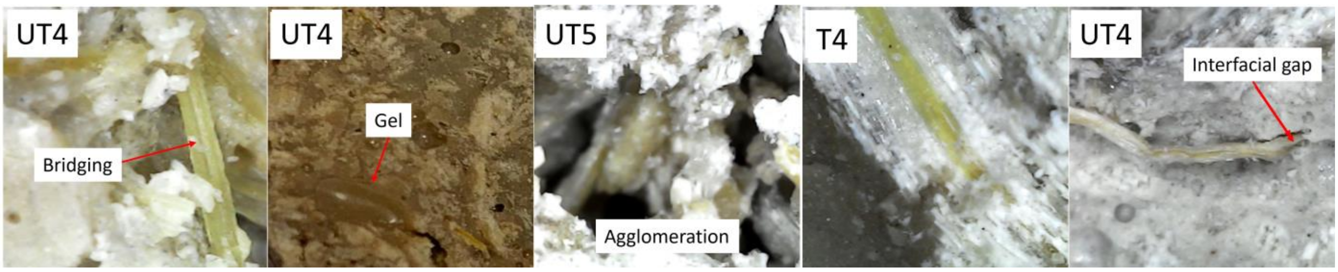


Fig. 15 Visual inspection of the BFG specimens

In specimen T4, the formation of crystalline structures around the BF is visible, indicating that HWA-treated BF exhibits stronger bonding compared to the untreated BF in specimen UT4. The loosened bonding in specimen UT4 may be due to the drying of the trapped water inside the BFG after 28 days of curing.

4. Conclusions

HWA treatment was applied to sugarcane BF, which was incorporated into BFG. Fiber morphology, composite mechanical properties, and predictive performance were evaluated through experimental testing, microstructural analysis, and machine learning models. The main findings are summarized as follows:

(1) HWA treatment effectively removed hemicellulose and surface impurities, reduced water absorption, and improved fiber surface uniformity, leading to stronger fiber–matrix bonding.
(2) Geopolymers reinforced with HWA-treated BF exhibited higher compressive strength, improved ductility, and enhanced post-peak energy absorption compared to those with untreated fibers.
(3) Among the predictive models, XGBoost demonstrated the highest accuracy and reliability, with compressive strength, water absorption, and density identified as the most influential parameters.
(4) HWA treatment promoted natural fiber compatibility in geopolymer systems, while data-driven modeling offers a robust tool for predicting and optimizing mechanical properties.

Future work may focus on expanding the dataset and exploring additional fiber treatments to further improve model accuracy and composite performance.

Acknowledgments

This research was made possible due to financial support from the Directorate of Research, Technology, and Public Services, Ministry of Education and Culture, Research, and Technology of the Republic of Indonesia, and the Directorate of Research and Public Services, Institut Teknologi Sepuluh Nopember (ITS) through Research Grant PDD with Contract No: 1730/PKS/ITS/2024, and LPDP Scholarship for supporting, scholarship registration number 202402210100645. The authors gratefully acknowledge financial support from the Institut Teknologi Sepuluh Nopember for this work, under the project scheme of the Publication Writing and IPR Incentive Program (PPHKI) 2025.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] H. H. Ahmad and Tavio, “Experimental Study of Cold-Bonded Artificial Lightweight Aggregate Concrete,” AIP Conference Proceedings, vol. 1977, article no. 030011, 2018.

[2] D. Christianto, Tavio, and D. Kurniadi, “Effect of Steel Fiber on the Shear Strength of Reactive Powder Concrete,” IOP Conference Series: Materials Science and Engineering, vol. 508, article no. 012006, 2019.

[3] Z. Zhang, L. Zhang, H. Liu, and J. Yin, “Effects of Metakaolin and Sodium Silicate Treatment on Highwater Content Dredged Clay for Improved Construction Fill Performance,” Construction and Building Materials, vol. 411, article no. 134196, 2024.

[4] G. Sharmila and R. Jeyalakshmi, “Unravelling the Stability of Solvated Silicate Species in Sodium Silicate Activator Solution of Different Silica Modulus by FTIR and NMR Studies and Their Chemical Reactivity on Metakaolin and Fly Ash Geopolymerisation,” Sustainable Chemistry and Pharmacy, vol. 39, article no. 101578, 2024.

[5] G. M. Amantino, N. P. Hasparyk, F. Tiecher, and R. D. Toledo Filho, “Assessment of Bio-Aggregate Concretes’ Properties with Rice Residue,” Journal of Building Engineering, vol. 52, article no. 104348, 2022.

[6] H. M. Hamada, A. Al-attar, S. Beddu, M. Kamal, S. T. Yousif, and A. Majdi, “Impact of Rice Husk Ash on Geopolymer Concrete: A Literature Review and Future Directions,” Case Studies in Construction Materials, vol. 22, article no. e04476, 2025.

[7] B. Kozub, S. Gądek, B. Tyliszczak, L. Wojnar, and K. Korniejenko, “Leveraging 3D Printing Capability for Geopolymer Composites Based on Fly Ash with Cotton Fibers Addition”, International Journal of Engineering and Technology Innovation, vol. 14, no. 3, pp. 231-243, 2024.

[8] A. Workiye and E. Woldesenbet, “Porous Maize Stalk Cellulose Fiber-Reinforced Geopolymer Composites for Heat Insulation at the Bottom Side of a Local Electric Stove”, International Journal of Engineering and Technology Innovation, vol. 22, pp. 20-29, 2022.

[9] R. N. Yanou, R. C. KAZE, A. Adesina, J. G. D. Nemaleu, S. B. K. Jiofack, and J. N. Y. Djobo, “Performance of Laterite-Based Geopolymers Reinforced with Sugarcane Bagasse Fibers,” Case Studies in Construction Materials, vol. 15, article no. e00762, 2021.

[10] R. Y. Nkwaju, J. N. Y. Djobo, J. N. F. Nouping, P. W. M. Huisken, J. G. N. Deutou, and L. Courard, “Iron-Rich Laterite-Bagasse Fibers Based Geopolymer Composite: Mechanical, Durability and Insulating Properties,” Applied Clay Science, vol. 183, article no. 105333, 2019.

[11] P. Rovnaník, “Effect of Curing Temperature on the Development of Hard Structure of Metakaolin-Based Geopolymer,” Construction and Building Materials, vol. 24, no. 7, pp. 1176-1183, 2010.

[12] C. D. Atiş, E. B. Görür, O. Karahan, C. Bilim, S. Ilkentapar, and E. Luga, “Very High Strength (120 MPa) Class F Fly Ash Geopolymer Mortar Activated at Different NaOH Amount, Heat Curing Temperature and Heat Curing Duration,” Construction and Building Materials, vol. 96, pp. 673-678, 2015.

[13] L. K. Turner and F. G. Collins, “Carbon Dioxide Equivalent (CO2-e) Emissions: A Comparison Between Geopolymer and OPC Cement Concrete,” Construction and Building Materials, vol. 43, pp. 125-130, 2013.

[14] L. Prasad, S. Kumar, R. V. Patel, A. Yadav, V. Kumar, and J. Winczek, “Physical and Mechanical Behaviour of Sugarcane Bagasse Fibre-Reinforced Epoxy Bio-Composites,” Materials, vol. 13, no. 23, article no. 5387, 2020.

[15] M. Djalal, M. Nafissa, R. Mansour, M. Jawaid, M. Hocine, and B. Lamia, “Effect of Alkali Treatment on New Lignocellulosic Fibres from the Stem of the Aster Squamatus Plant,” Journal of Materials Research and Technology, vol. 32, pp. 2882-2890, 2024.

[16] M. Y. R. da Gloria, V. M. Andreola, D. O. J. dos Santos, M. Pepe, and R. D. Toledo Filho, “A Comprehensive Approach for Designing Workable Bio-Based Cementitious Composites,” Journal of Building Engineering, vol. 34, article no. 101696, 2021.

[17] J. Cai, J. Pan, J. Han, and X. Wang, “Mechanical Behaviors of Metakaolin-Based Engineered Geopolymer Composite under Ambient Curing Condition,” Journal of Materials in Civil Engineering, vol. 34, no. 7, pp. 1-12, 2022.

[18] L. Breiman, "Random Forests," Machine Learning, vol. 45, pp. 5-32, 2001.

[19] P. F. Zhang, M. Iqbal, D. Zhang, X. L. Zhao, and Q. Zhao, "Bond Strength Prediction of FRP Bars to Seawater Sea Sand Concrete Based on Ensemble Learning Models," Engineering Structures, vol. 302, article no. 117382, 2024.

[20] R. E. Schapire and Y. Freund, "Bodomosting: Foundations and Algorithms," Kybernetes, vol. 42, no. 1, pp. 164-166, 2013.

[21] M. H. R. Sobuz, M. H. Khan, M. K. I. Kabbo, A. H. Alhamami, F. S. Aditto, M. S. Sajib, et al., "Assessment of Mechanical Properties with Machine Learning Modeling and Durability, and Microstructural Characteristics of a Biochar-Cement Mortar Composite," Construction and Building Materials, vol. 411, article no. 134281, 2024.

[22] Y. Zeng, Y. Chen, Y. Liu, T. Wu, Y. Zhao, D. Jin, et al., "Prediction of Compressive and Flexural Strength of Coal Gangue-Based Geopolymer Using Machine Learning Method," Materials Today Communications, vol. 44, article no. 112076, 2025.

[23] W. Huo, Z. Zhu, H. Sun, B. Ma, and L. Yang, "Development of Machine Learning Models for the Prediction of the Compressive Strength of Calcium-Based Geopolymers," Journal of Cleaner Production, vol. 380, part 2, article no. 135159, 2022.

[24] S. M. Lundberg and S. I. Lee, "A Unified Approach to Interpreting Model Predictions," Proceedings of Neural Information Processing Systems (NIPS'17), ACM Press, pp. 4768-4777, 2017.

[25] X. Zhao, P. F. Zhang, Q. Zhao, D. Zhang, Y. Tuerxunmaimaiti, and H. Cao, "A SHAP Algorithm-Based Prediction of the Interlaminar Shear Strength Degradation of G/BFRP Bars Embedded in Concrete Exposed to Marine Environment," Case Studies in Construction Materials, vol. 22, article no. e04770, 2025.

[26] A. Bartos et al., "Alkali Treatment of Lignocellulosic Fibers Extracted from Sugarcane Bagasse: Composition, Structure, Properties," Polymer Testing, vol. 88, article no. 106549, 2020.

[27] Fahmida-E-Karim, M. R. Islam, P. Roy, M. Hasan, S. Islam, and S. Haque, "Analysis of Water Absorbency and Surface Behavior of Kapok and Bagasse Fibers After Alkali Treatment," Cleaner Waste Systems, vol. 12, article no. 100357, 2025.

[28] A. Workiye and E. Woldesenbet, "Flexural Strength and Porosity of NaOH-Treated Maize Stalk Cellulose-Fibers-Reinforced Geopolymer Composites", International Journal of Engineering and Technology Innovation, vol. 25, pp. 44-53, 2023.

[29] H. M. Hamada, A. Al-Attar, M. Kamal, S. Beddu, and A. Majdi, "Advancing the Sustainability of Fiber-Reinforced Geopolymer Concrete Using Natural Plant Fibers: A Comprehensive Review of Properties and Impacts," Structures, vol. 77, article no. 109201, 2025.

[30] B. Poletanovic, J. Dragas, I. Ignjatovic, M. Komljenovic, and I. Merta, "Physical and Mechanical Properties of Hemp Fibre Reinforced Alkali-Activated Fly Ash and Fly Ash/Slag Mortars," Construction and Building Materials, vol. 259, article no. 119677, 2020.