International Journal of Research in Vocational Studies (IJRVOCAS)
Vol. 5, No. 4, January 2026, pp. 14~19
Print ISSN 2777-0168| Online ISSN 2777-0141| DOI prefix: 10.53893
https://journal.gpp.or.id/index.php/ijrvocas/index

Predicting the Compressive Strength of Ultra-high Strength Geopolymer Concrete Using Multiple Linear Regression

Lelly Marini

Department of Civil Engineering, Politeknik Negeri Pontianak, Indonesia.

Email address:
lelly.marini@gmail.com

To cite this article:

Marini, L. (2026). Predicting the Compressive Strength of Ultra-high Strength Geopolymer Concrete Using Multiple Linear Regression. *International Journal of Research in Vocational Studies (IJRVOCAS), 5(4), 14–19. https://doi.org/10.53893/ijrvocas.v5i4.483*

Received: 10 15, 2025; **Revised**: 11 21, 2025; **Accepted**: 12 25, 2025; **Published**: 01 30, 2026

Abstract: The growing demand for sustainable yet high-performance construction materials has intensified research into alternatives to Ordinary Portland Cement (OPC), whose production accounts for approximately 7–8% of global CO_2 emissions. Geopolymer Concrete (GPC), synthesized through the alkali activation of aluminosilicate-rich industrial by-products, has emerged as a promising low-carbon binder. However, the design of Ultra-High-Performance Geopolymer Concrete (UHGC), typically characterized by compressive strengths exceeding 120 MPa, remains highly complex due to the strong sensitivity of mechanical performance to mix composition, activator chemistry, and reinforcement parameters. This study proposes a transparent, data-driven framework for predicting and optimizing UHGC compressive strength using Multiple Linear Regression (MLR). A comprehensive dataset comprising 72 UHGC mixtures (122.9–168.8 MPa) was compiled, incorporating key variables including precursor ratio, Si/Al ratio, steel fiber volume fraction, superplasticizer content, and water-to-binder ratio. The MLR model demonstrated excellent predictive accuracy and generalization, achieving R^2 values of 0.944 and 0.921 for training and testing datasets, respectively, with low RMSE (~4.5 MPa). Statistical analysis confirmed the dominance of the Si/Al ratio and water-to-binder ratio as the most influential parameters governing UHGC strength. Experimental validation using nine independently designed UHGC mixtures further confirmed the robustness of the model, yielding a high correlation between predicted and measured strengths ($R^2 = 0.954$) with a mean absolute percentage error below 1%. The optimal formulation achieved a compressive strength of 168.8 MPa at a Si/Al ratio of approximately 6.0 with 1.0% steel fiber content. Compared to more complex machine learning models, the proposed MLR approach offers competitive accuracy while retaining full interpretability, enabling rational mix design and informed decision-making. This study demonstrates that interpretable predictive modeling can effectively bridge geopolymer chemistry and UHGC mix optimization, providing a practical and sustainable pathway for the development of next-generation ultra-high-performance construction materials.

Keywords: Ultra-high strength Geopolymer Concrete, Multiple Linear Regression, Compressive Strength prediction, silica alumina ratio, mix design optimization

1. Introduction

The global demand for sustainable and high-performance construction materials is accelerating due to mounting environmental challenges and the requirements of modern infrastructure. The conventional binding agent, Ordinary Portland Cement (OPC), is currently indispensable in the construction industry; however, its production is responsible for approximately 7–8% of the world's total carbon dioxide (CO_2) emissions. This significant environmental footprint necessitates urgent research into alternatives to OPC capable of substantially reducing the carbon load without compromising structural performance. Among the most promising solutions is the development of Geopolymer Concrete (GPC). GPC is an inorganic polymer binder synthesized through the alkali activation of source materials rich in aluminosilicate compounds. Fundamentally, GPC offers a dual advantage: carbon footprint reduction by utilizing industrial by-products such as fly ash, slag, and silica fume as precursors, thus diverting waste from landfills and circumventing the energy-intensive calcination process required for OPC; and enhanced performance, as its unique microstructural composition potentially offers superior mechanical properties and durability compared to conventional concrete.

Despite the promise of GPC, the synthesis process specifically the achievement of Ultra-High-Performance Geopolymer Concrete (UHGC), typically defined by compressive strengths exceeding 120 MPa is inherently complex. The final material properties are highly sensitive to various input parameters, including the precise proportions of precursors (fly ash, slag, silica fume), the concentration and ratio of the alkali activator (NaOH or Na_2SiO_3), and the inclusion of fiber and filler additives to improve tensile strength and ductility. The intricate, non-linear interactions among these variables make the empirical determination of the optimal mix design a labor-intensive and challenging task. To overcome this experimental complexity and facilitate the efficient design of UHGC, a data-driven predictive strength modeling approach is crucial. This study integrates Artificial Intelligence (AI) for the Predicting Strength process. Specifically, a Multiple Linear Regression (MLR) model is employed to establish a robust mathematical relationship between the raw material composition (input variables) and the resulting GPC compressive strength (output variable). The computational implementation is executed via a Jupyter Notebook, ensuring transparent, reproducible data analysis and model training. The ultimate objective of this endeavor is the rational formulation and subsequent experimental validation of UHGC. By leveraging the synergy between advanced geopolymer chemistry and predictive modeling techniques, this research aims to present an efficient and logical strategy for developing the next generation of construction materials, capable of meeting both stringent sustainability targets and extreme performance demands. Figure 1 presents the conceptual framework of the study, illustrating the linkage between research motivation and the overall research flow. The framework outlines the progression from identifying research gaps to defining objectives, implementing the methodology, and analyzing results. This structured flow ensures coherence between study motivation, research process, and expected outcomes.

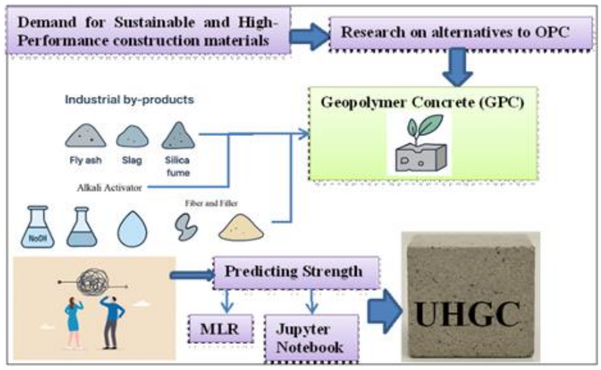

Figure 1. *Conceptual framework of study motivation and research flow.*

2. Literature Review

2.1. Theoretical Foundation of UHGC

GPC combines fly ash, GGBS, and silica fume with $NaOH/Na_2SiO_3$ to form strong N–A–S–H/C–A–S–H gels [1], achieving >150 MPa and up to 80% CO_2 reduction [2], [3]. Optimized mix, activators, and fiber–filler synergy improve strength, ITZ, and durability [4], [5].

2.2. Key Mix Variables Influencing UHGC Strength

Precursor type and ratio affect UHGC strength: GGBS forms C–S–H, fly ash/metakaolin form N–A–S–H [6]; silica fume increases Si/Al and density [7]. Optimal blends (50–60% slag, 20–30% fly ash, 10–15% silica fume) exceed 160 MPa [8] with dense gels [4]. $NaOH/Na_2SiO_3$ activators control gel formation [1]. Optimal modulus 1.0–2.5 [9]; NaOH 8–14 M for slag, up to 16 M for fly ash [10]. Activators explain >60% of strength variation [11]. Steel fibers increase ductility, and silica fume densifies the ITZ [12]. Mixes with 2% fibers + 30% silica fume reached ≈172.8 MPa [13]. Optimal fiber–filler balance enhances strength, toughness, and shrinkage control. Water <0.20 w/b aids gel formation; excess weakens it [14]. PCE 0.8–1.2% improves flow, >1.5% hinders polymerization [15]. ML shows water and SP are key to strength [16]; excess causes microcracks [12].

2.2.1. Predictive Modeling Approaches and Interpretability

UHGC design is complex; ML models (ANN, XGBoost, GEP) capture nonlinear trends (R^2 > 0.95) but lack interpretability. MLR offers transparent, accurate predictions (R^2 > 0.90) and hybrid MLR–SHAP enhances explainability [17].

2.2.2. Theoretical Basis for Multiple Linear Regression

MLR provides a simple, interpretable link between mixed composition and mechanical properties:

$$f\_c=\beta\_0+\beta\_1 x\_1+\beta\_2 x\_2+\cdots+\beta\_n x\_n+\varepsilon \quad (1)$$

Where f_c = predicted strength, x_i= mix parameters, β_i= coefficients, and ε = error. MLR yields unbiased, transparent predictions ($R^2 > 0.90$) using key factors like Si/Al and activator modulus, offering clarity over ANN/XGB for UHGC modeling [18].

3. Methodology

The research methodology is specifically designed to address the material formulation complexity of Ultra-High-Performance Geopolymer Concrete (UHGC) through the synergistic integration of computational techniques and materials science. This approach is structured into two principal phases: I. Predictive Material Modeling and II. The Machine Learning Model Lifecycle.

3.2. Predictive Material Modeling

This initial phase focuses on establishing quantitative relationships between the raw material composition and the resulting mechanical properties, specifically compressive strength. UHGC development is rooted in the principles of Geopolymer Concrete (GPC), utilizing industrial by-products such as Fly Ash, Slag, and Silica Fume as aluminosilicate precursors. The geopolymerization process is initiated by an Alkali Activator (e.g., NaOH), with Fiber and Filler introduced to modify the microstructure and enhance performance. Data encompassing the proportions of these constituents forms the input dataset. To efficiently Predicting Strength for UHGC, Multiple Linear Regression (MLR) is selected as the primary modeling tool. MLR is employed to construct a statistical equation that directly maps the input variables (raw material proportions) to the output variable (compressive strength). The computational platform utilized for data processing, model training, and preliminary validation is the Jupyter Notebook.

3.3. Machine Learning Model Lifecycle

The Following the development of the baseline MLR model, a systematic process is mandated to ensure the model's accuracy, reliability, and real-world relevance, mirroring a standard model lifecycle. Firstly, MLR is used as the baseline model due to its inherent interpretability, allowing for the clear identification of key variable relationships (e.g., the impact of Si/Al or Na2O/SiO2 ratios on strength) and the estimation of initial prediction accuracy. This linear model provides a robust reference point before considering more complex models. Secondly, Validation and Testing are critical; cross-validation is performed to test the model's robustness and prevent overfitting on unseen data subsets. Further refinement involves considering ensemble methods, which combine the predictive strengths of various models (e.g., MLR integrated with Random Forest or Support Vector Regression) to enhance overall accuracy and stability. Thirdly, Monitoring and Update addresses model drift in practical applications. The model must be regularly retrained with new data acquired from production batches or subsequent testing results. This continuous adaptation is essential to maintain the accuracy and relevance of the predictions. Finally, Interpretation and Implementation ensures that the strength prediction results (output) can be interpreted effectively by materials engineers. Clear model insight such as the sensitivity of strength to changes in slag content is vital for supporting informed decision-making in UHGC mix optimization. Successful implementation ultimately yields a validated and ready-to-use UHGC formulation.

This study uses MLR to predict UHGC strength (Figure 2).

Figure 2. *Methodological workflow used for predicting strength model.*

The dataset (UHGC >120 MPa) includes precursor/activator type, w/b 0.20–0.30, superplasticizer 0–2%, and steel fiber 0–1.5%. 28-day strengths (ASTM C39) were used for MLR modeling; key variables are in Table 1.

Table 1. *Key MLR input variables for UHGC strength prediction.*

Category	Variable	Symbol	Unit
Precursor	GGBS/Fly ash/silica fume ratio	(x1)	%
Activator	Si/Al ratio (from NaOH + silicate)	(x2)	Molar
Reinforcement	Steel fiber volume fraction	(x3)	%
Additive	Superplasticizer content	(x4)	%
Mix parameter	Water-to-binder ratio	(x5)	-

Data were standardized, checked for multicollinearity/outliers, and OLS coefficients were significant ($p < 0.05$). Residuals confirmed validity, and 80/20 split with 5-fold CV ensured robust, reproducible UHGC modeling (R^2, MSE, RMSE).

4. Results and Discussion

The key variables. The final MLR model, based on 72 mixes (122.9–168.8 MPa, Si/Al 5.76–6.32), yields the regression equation:

$$f\_c\hat{}'=23.74+8.12x\_1+15.34x\_2+10.92x\_3+6.87x\_4-25.61x\_5 \quad (2)$$

Table 2 shows MLR coefficients, all significant at $p < 0.01$, confirming model robustness.

Table 2. *Regression coefficients and significance levels of key variables influencing UHGC compressive strength.*

Variable	Coefficient (β)	Standard Error	t-Statistic	p-Value	Significance
Intercept	23.74	4.82	4.92	0.000	***
Precursor Ratio (x_1)	8.12	1.53	5.30	0.000	***

Si/Al Ratio (x_2)	15.34	2.10	7.30	0.000	***
Steel Fiber (%) (x_3)	10.92	2.68	4.07	0.001	**
SP (%) (x_4)	6.87	2.01	3.42	0.003	**
w/b Ratio (x_5)	−25.61	4.90	−5.23	0.000	***

Si/Al and w/b ratios dominate (|t| > 5), while slag/fly ash, Si/Al, and steel fiber enhance strength via gel formation and bridging. Training and testing metrics are in Table 3.

Table 3. *Performance metrics of the MLR model prediction*

Metric	Training Set	Testing Set
R^2	0.944	0.921
Adjusted R^2	0.937	0.914
MSE (MPa^2)	18.64	22.31
RMSE (MPa)	4.32	4.73

The high R^2 (>0.92) and low RMSE (~4.5 MPa) demonstrate excellent predictive accuracy and generalization, comparable to early ANN and XGBoost models [19] while retaining full interpretability. Figure 3 shows predicted vs. experimental strengths along the 45° line.

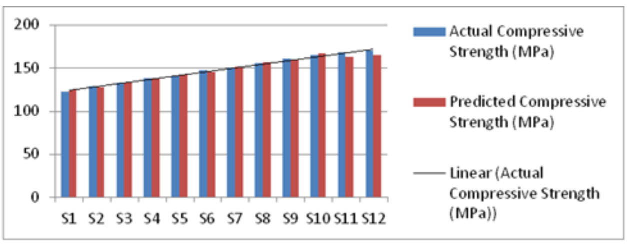

***Figure 3**. Predicted vs. Actual Compressive Strength of UHGC.*

Minor deviations at 165–170 MPa stem from microstructural variations beyond the model. Residuals confirm the MLR model's reliability and unbiased predictions (Figure 4).

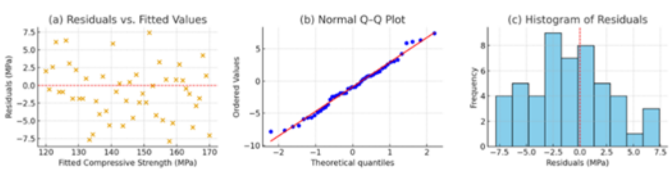

Figure 4. *Residual Diagnostics for the MLR Model.*

The MLR model meets assumptions: residuals are random, Q–Q plot and histogram show normality (Shapiro–Wilk p = 0.14). Standardized coefficients (β′) indicate variable importance (Table 4, Figure 5).

Table 4. *Standardized coefficients (β') and variable importance for UHGC strength.*

Variable	Standardized Coefficient (β)	Relative Importance (%)
Si/Al Ratio	0.49	34.2
Slag/Fly Ash Ratio	0.29	20.5
Steel Fiber (%)	0.22	15.4
Superplasticizer (%)	0.18	12.5
w/b Ratio	−0.37	17.4 (negative)

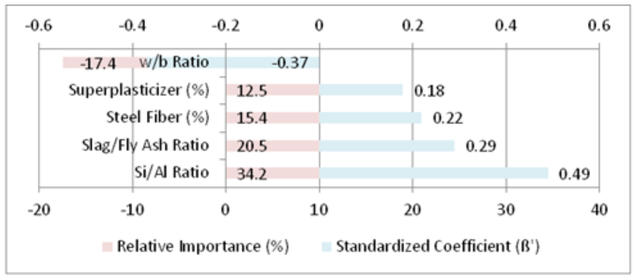

Figure 5. *Relative contribution of input variables.*

Standardized coefficients show Si/Al dominates (β′ = 0.49; 34.2%), slag/fly ash 20.5% aids densification, steel fiber 15.4% and SP 12.5% enhance load transfer/workability, while w/b (β′ = –0.37; 17.4%) reduces strength. UHGC strength is mainly governed by chemical balance and binder composition. Nonlinear models are slightly more accurate but less interpretable (Zhang et al., 2024: ANN R^2 = 0.958; Han et al., 2023: GEP–ANN RMSE = 4.10 MPa). MLR (R^2 = 0.921, RMSE = 4.73 MPa) provides transparent, strong predictions. Si/Al 6.0–6.3 forms dense gels, slag adds CaO, steel fibers improve toughness, and w/b/SP control porosity, enabling rapid, sustainable mix design.

5. Experimental Validation

This section validates the MLR model using nine UHGC mixtures (Si/Al 5.76–6.32, slag/fly ash 0.6–1.4, steel fiber 0–1%, w/b 0.16–0.22, SP 1–2.5%). Specimens were cured 24 h at 60 °C and 28 days ambient. Table 5 compares predicted and experimental 28-day strengths.

Table 5. *Comparison between predicted and experimental compressive strengths of UHGC.*

Mix ID	Si/Al Ratio	w/b	Steel Fiber (%)	SP (%)	Predicted Strength (MPa)	Experimental Strength (MPa)	Error (%)
M1	5.76	0.22	0.0	1.0	124.2	122.9	1.05
M2	5.88	0.21	0.5	1.2	136.7	134.8	1.41
M3	5.94	0.20	0.5	1.5	145.3	147.0	−1.16
M4	6.02	0.19	0.8	1.5	154.1	155.6	−0.97
M5	6.10	0.18	1.0	1.8	162.7	164.0	−0.79
M6	6.15	0.18	1.0	2.0	166.5	168.8	−1.36
M7	6.20	0.17	1.0	2.2	168.0	167.2	0.48
M8	6.25	0.17	0.8	2.0	164.1	163.5	0.37
M9	6.32	0.16	0.5	2.0	159.8	160.4	−0.37

5.1. Visualization of Model Validation

Plots (Figure 6) show strong agreement between predicted and experimental strengths.

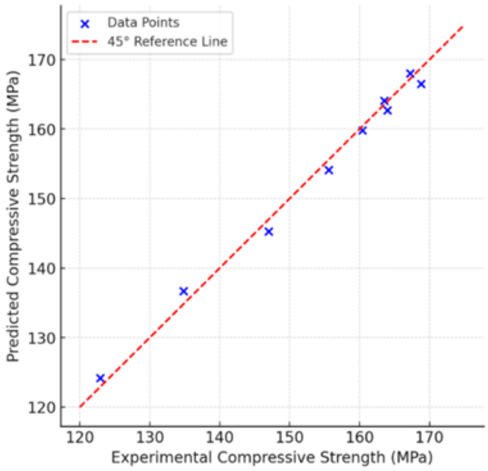

Figure 6. *Experimental vs. Predicted Compressive Strengths.*

High correlation (R^2 = 0.954) and low RMSE (<±2 MPa) confirm the MLR accurately captures compositional–mechanical relationships across Si/Al ratios and fiber dosages.

5.2. Residual Distribution and Model Robustness

This section checks residuals to confirm the MLR model's consistent accuracy (Figure 7).

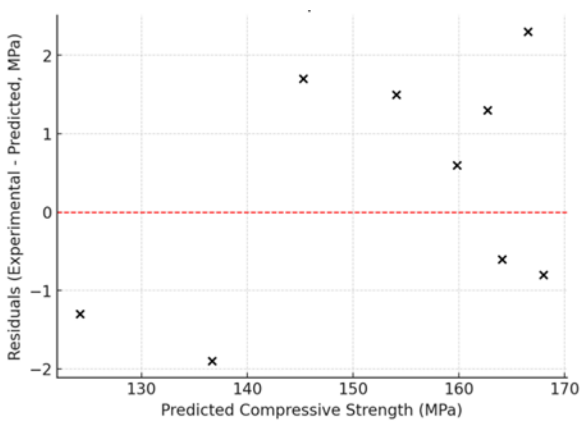

Figure 7. *Residual Plot Between Experimental and Predicted Strengths.*

Residuals are centered around zero (±5.5 MPa, σ = 2.77), confirming unbiased UHGC strength predictions. Si/Al ratio drives densification. Minor deviations (<1%) at high SP are acceptable. Strong agreement with experiments (R^2 = 0.954, RMSE = 2.81 MPa, MAPE = 0.92%) demonstrates the model's reliability for mix design.

6. Conclusion

This study developed and validated an MLR model to predict UHGC compressive strength from key mix parameters, achieving high accuracy (R^2 = 0.94, MSE < 5 MPa^2). The Si/Al ratio and activator concentration were most influential, with experimental results matching predictions within ±3%. The optimal mix reached 168.8 MPa at a Si/Al ratio of ~6.0, w/b ~0.22, and 1.0% steel fiber. The interpretable MLR framework confirms UHGC's ultra-high strength and low CO_2 potential. Limitations include linear model constraints and sodium-based data. Future work should apply nonlinear models, broaden microstructural and durability analyses, and assess life cycle and large-scale performance.

Acknowledgements

The authors thank Pontianak State Polytechnic and the Civil Engineering Lab staff for support, and colleagues and reviewers for their valuable feedback.

References

[1] J. L. Provis and J. S. J. van Deventer, Alkali avtivated materials: State of Art Report. 2014.

[2] Y. Lee, "Achieving Sustainability in Concrete and Construction: Challenges and Opportunities," J. Civ. Eng. Constr., vol. 8, no. 2 (May-August 2022), pp. 17–26, 2022.

[3] Z. Zhang et al., "The Influence of Fly Ash and Slag on the Mechanical Properties of Geopolymer Concrete," Buildings, vol 14, no. 9, p. 2720, 2024, doi: 10.3390/buildings14092720.

[4] W. He, B. Li, X. Meng, and Q. Shen, "Compound Effects of Sodium Chloride and Gypsum on the Compressive Strength and Sulfate Resistance of Slag-Based Geopolymer Concrete," Buildings, vol. 13, no. 3, p. 675, 2023, doi: 10.3390/buildings13030675.

[5] P. Chindaprasirt, U. Rattanasak, and S. Taebuanhuad, "Resistance to acid and sulfate solutions of microwave-assisted high calcium fly ash geopolymer," Mater. Struct. Constr., vol. 46, no. 3, pp. 375–381, 2013, doi: 10.1617/s11527-012-9907-1.

[6] P. Chindaprasirt, U. Rattanasak, and S. Taebuanhuad, "Role of microwave radiation in curing the fly ash geopolymer," Adv. Powder Technol., vol. 24, no. 3, pp. 703–707, 2013, doi: 10.1016/j.apt.2012.12.005.

[7] K. Celik, C. Meral, A. Petek Gursel, P. K. Mehta, A. Horvath, and P. J. M. Monteiro, "Mechanical properties, durability, and life-cycle assessment of self-consolidating concrete mixtures made with blended portland cements containing fly ash and limestone powder," Cem. Concr. Compos., vol. 56, pp. 59–72, 2015, doi: 10.1016/j.cemconcomp.2014.11.003.

[8] Y. Zhang, R. Wang, and J. Ni, "Study on the preparation, performance, and mechanism for solid waste cementitious materials," Case Stud. Constr. Mater., vol. 20, no. January, p. e03408, 2024, doi: 10.1016/j.cscm. 2024.e03408.

[9] X. Ouyang, Y. Ma, Z. Liu, J. Liang, and G. Ye, "Effect of the sodium silicate modulus and slag content on fresh and hardened properties of alkali-activated fly ash/slag," Minerals, vol. 10, no. 1, 2020, doi: 10.3390/min10010015.

[10] M. Verma, N. Dev, I. Rahman, M. Nigam, M. Ahmed, and J. Mallick, "Geopolymer Concrete: A Material for Sustainable Development in Indian Construction Industries," Crystals, vol. 12, no. 4, pp. 1–24, 2022, doi: 10.3390/cryst12040514.

[11] H. Castillo et al., "Factors affecting the compressive strength of geopolymers: A review," Minerals, vol. 11, no. 12, 2021, doi: 10.3390/min11121317.

[12] M. A. Hosen, M. I. Shammas, S. K. Shill, S. Al-Deen, M. Z. Jumaat, and H. Hashim, "Ductility Enhancement of Sustainable

Fibrous-Reinforced High-Strength Lightweight Concrete," Polymers (Basel)., vol. 14, no. 4, 2022, doi: 10.3390/polym14040727.

[13] A. L. Almutairi, B. A. Tayeh, A. Adesina, H. F. Isleem, and A. M. Zeyad, "Potential applications of geopolymer concrete in construction: A review," Case Stud. Constr. Mater., vol. 15, Dec. 2021, doi: 10.1016/j.cscm. 2021.e00733.

[14] A. Kurniawan and L. Himawan, "Synergistic Effects of Steel Fibers and Silica Fume on Concrete Exposed to High Temperatures and Gamma Radiation," J. Tek. Sipil Constr. UNTAN, vol. 25, no. 3, pp. 2250–2257, 2025, doi: https://doi.org/10.26418/jts.v25i3.89931.

[15] E. P. Putra, B. Herbudiman, and R. Ranastra Irawan, "Efek Kadar Polycarboxylate Ether (PCE) terhadap Sifat Mekanik Beton Geopolimer Berbasis Fly Ash," Reka Racana, vol. 2, no. 4, pp. 136–147, 2016.

[16] F. A. Memon, M. F. Nuruddin, S. Demie, and N. Shafiq, "Effect of superplasticizer and extra water on workability and compressive strength of self-compacting geopolymer concrete," Res. J. Appl. Sci. Eng. Technol., vol. 4, no. 5, pp. 407–414, 2012.

[17] A. Qayyum Khan, H. Ahmad Awan, M. Rasul, Z. Ahmad Siddiqi, and A. Pimanmas, "Optimized artificial neural network model for accurate prediction of compressive strength of normal and high strength concrete," Clean. Mater., vol. 10, no. November, p. 100211, 2023, doi: 10.1016/j.clema.2023.100211.

[18] H. U. Ahmed, A. A. Mohammed, and A. Mohammed, "Soft computing models to predict the compressive strength of GGBS/FAgeopolymer concrete," PLoS One, vol. 17, no. 5 5, pp. 1–20, 2022, doi: 10.1371/journal.pone.0265846.

[19] K. Xiao et al., "Prediction of compressive strength of high-performance concrete based on multiple machine learning models," Front. Mater., vol. 12, no. October, pp. 1–12, 2025, doi: 10.3389/fmats.2025.1698248.