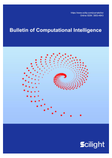
Scilight

Bulletin of Computational Intelligence
https://www.sciltp.com/journals/bci



Article

Explainable Predictive Modelling of Sustainable Slag–Fly Ash Based Geopolymer Concrete with Life Cycle and Carbon-Neutrality Assessment

Saad Shamim Ansari, Mohd Asif Ansari and Syed Muhammad Ibrahim *

Department of Civil Engineering, Zakir Husain College of Engineering and Technology, Aligarh Muslim University, Aligarh 202002, Uttar Pradesh, India
* Correspondence: ibrahim.cv@amu.ac.in

How To Cite: Ansari, S.S.; Ansari, M.A.; Ibrahim, S.M. Explainable Predictive Modelling of Sustainable Slag–Fly Ash Based Geopolymer Concrete with Life Cycle and Carbon-Neutrality Assessment. *Bulletin of Computational Intelligence* **2026**, *2*(1), 54–82. https://doi.org/10.53941/bci.2026.100004

Received: 11 November 2025
Revised: 7 January 2026
Accepted: 9 January 2026
Published: 20 January 2026

Abstract: This study proposes a novel, unified framework integrating scientometric analysis, machine learning (ML), explainable artificial intelligence (XAI), and cradle-to-gate life cycle assessment (LCA) to evaluate and predict the performance of slag-fly ash-based geopolymer concrete (SFGPC). A scientometric review of 441 publications (2009–2025) guided the systematic assembly of a dataset comprising 363 SFGPC mixes. Five ML models were trained to predict compressive strength (f_c), with Gradient Boosting (GB) achieving the highest accuracy, yielding $R^2 = 0.954$, $RMSE = 3.15$ MPa, $MAE = 1.81$ MPa during training, and $R^2 = 0.95$, $RMSE = 3.128$ MPa, $MAE = 2.41$ MPa during testing. Multi-layered XAI analysis identified age, slag content, and alkaline-to-binder ratio as the most influential parameters and revealed governing nonlinear interactions. Sustainability assessment showed that the fly ash-dominant mix exhibited the lowest global warming potential (156 kg CO_2-eq/m^3), the most favourable sustainability index, and the smallest residual emissions after a 25% carbon offset. A user-oriented graphical user interface (GUI) was developed for real-time strength prediction. The novelty of this work lies in introducing an explainable, data-driven, and sustainability-integrated decision-support system for designing transparent and low-carbon geopolymer concretes.

Keywords: geopolymer concrete; scientometric; machine learning; explainable artificial intelligence; life cycle assessment; sustainability

1. Introduction

Concrete stands as the most widely utilized material globally, second only to water, with an approximate consumption of 1 cubic meter of concrete per person per year [1]. The demand for ordinary Portland cement (OPC) closely aligns with the growing requirements for concrete. However, as the global commitment to sustainability intensifies, the conventional use of OPC has come under scrutiny due to its significant environmental impacts [2–5]. The prodigious carbon footprint, resource depletion, and energy-intensive production processes associated with OPC have raised compelling concerns about the industry's role in climate change and ecological degradation. In response, there have been notable advancements in cement manufacturing, resulting in a nearly 30% reduction in emissions over the past few decades [6]. While commendable progress has been made, a critical environmental challenge persists in the form of the de-calcination of limestone during cement production. Even under the hypothetical scenario where cement kilns exclusively utilize 100% renewable energy, the process would still generate 0.78 units of CO_2 per unit of cement [7,8]. This underscores the complexity of achieving a carbon-neutral or low-carbon footprint in cement manufacturing despite strides toward cleaner energy sources [9,10]. Addressing this challenge necessitates a deeper exploration of alternative building materials that not only meet the stringent criteria of sustainability but also surpass the environmental performance of traditional cement.

Geopolymer concrete (GPC) has been emerging as an alternative to Ordinary Portland Cement Concrete (OPCC) [11,12] that not only exhibits mechanical properties comparable to OPCC but also presents a significantly reduced carbon footprint, offering a promising avenue for enhancing the environmental sustainability of the construction industry [13]. Compared to OPCC, GPC exhibits a significant reduction in CO_2 emissions by 40% to 90% [14]. Most GPC compositions provide a noteworthy 80–90% decrease in CO_2 emissions, which includes the carbon emissions produced during the synthesis of alkali activators [15]. This reduction has positioned GPC as a viable pathway for decarbonizing the construction sector, particularly when combined with industrial by-products such as fly ash (FlA) and ground granulated blast furnace slag (GGBFS).

The properties of GPC are intricately affected by various parameters such as the molarity of NaOH, sodium silicate to sodium hydroxide ratio (SS/SH), alkaline liquid-to-binder ratio (Alk./B), curing temperature, and the concentration of Al and Si in precursors, etc. Achieving the desired properties of the GPC mixture requires a considerably larger number of experimental mixes compared to OPCC due to highly non-linear interactions among these parameters. In this context, data-driven and soft computing approaches, particularly machine learning (ML) have become increasingly vital for the accurate and efficient prediction of the properties of the various types of composites [16–23]. Recent experimental and modelling studies have further clarified the role of FlA and slag chemistry on the fresh and hardened behavior of GPC. Study [24] evaluated the compressive strength (f_c) of FlA–slag-based GPC as a function of hydraulic, silica, alumina and lime moduli, as well as SS content, and showed that appropriate modulation of these indices can significantly enhance strength development in hybrid FlA–slag systems. Paruthi et al. [25] reported that incorporating silica fume and alccofine into GGBS-based GPC improves mechanical performance under different curing temperatures, underscoring the sensitivity of slag-rich binders to both reactive silica content and thermal regime. On the modelling side, Ahmed et al. [26] used support vector regression (SVR) combined with grey wolf optimization to predict the f_c of GGBFS-based GPC, confirming that parameters such as water-to-binder ratio, GGBFS fraction, and superplasticizer (Sp) dosage are critical drivers of strength. Collectively, these studies highlight the strong coupling between FlA content, activator chemistry, and mechanical behavior.

However, most ML models used for predicting the properties of GPC function as ‘black-box’ systems, offering limited interpretability despite high accuracy, which restricts their adoption in practical scenarios. To solve this black box issue, Explainable Artificial Intelligence (XAI) techniques, through model agnostic Shapley Additive Explanations (SHAP), help interpret these models by identifying key influencing features and their interactions [27,28]. Although highly relevant, such methods to predict and interpret the predictions in GPC research are scant.

Furthermore, to comprehensively assess the sustainability of GPC, it is essential to perform a Life Cycle Assessment (LCA) alongside mechanical and ML-assisted investigations [29,30]. LCA provides a systematic framework for quantifying the environmental impacts, including carbon emissions, energy consumption, and air pollutants from raw material extraction to production [31]. Integrating LCA with ML and XAI methods can enable a more holistic understanding of GPC’s performance, guiding the development of low-impact, durable, and efficient materials for future applications.

Despite significant progress in predicting GPC properties, existing studies are largely constrained by three major shortcomings: (i) heavy reliance on black-box ML models without transparent interpretability, limiting engineering adoption; (ii) absence of integrated frameworks that combine data-driven prediction with sustainability assessment; and (iii) lack of structured, literature-guided datasets and deployable tools that enable practical implementation. To address these gaps, this study proposes a unified and explainable framework that integrates scientometric analysis for informed dataset development, advanced ML modelling for strength prediction, multi-layer XAI for model transparency, cradle-to-gate LCA for environmental evaluation, and a graphical user interface (GUI) for real-time engineering use. The novelty of this work lies in merging predictive analytics with interpretability and sustainability assessment, thereby advancing state-of-the-art research toward transparent, reliable, and low-carbon GPC design.

1.1. Recent Advances and Limitations in ML-Based Predictions of GPC Strength

Several ML models have been employed to predict the properties of GPC. Random forest (RF), decision tree (DT), and extreme gradient boosting (XGBoost) models were developed from the 110 experimental datasets for predicting the f_c of GPC made from FlA. Among the various models, XGBoost demonstrated superior prediction performance [32]. Ahmed et al. [33] employed ANN, M5-tree, linear regression (LR), and multiple LR (MLR) models to predict the f_c of GGBFS-FlA-based GPC. Utilizing 220 dataset from the literature, input parameters included the contents of GGBFS, FlA, SH, SS, Si/Al of FlA, Si/Ca of GGBFS, binder ratio, and molarity. Sensitivity analysis identified the binder ratio as the most influential parameter. Notably, the ANN model demonstrated superior prediction performance. In another study, the f_c of GGBFS-GPC was predicted using multiple methodologies, including LR, SVR, differential evolution, grey wolf optimization, mantra rays foraging optimization (MRFO), genetic algorithm, and particle swarm optimization. Out of various input parameters, Sp dosage was identified as the most influential in f_c prediction. The SVR-MRFO hybrid model performed better than other models in terms of correlation and errors values [26]. Kina et al. [34] include specimen age, NaOH molarity, natural zeolite and silica fume content, and GGBFS content to predict the f_c of GGBFS-GPC by developing three ensemble ML (EML) models including DT, Bootstrap aggregating, and Least Squares Boosting (LSBoost). Specimen age was identified as the most influential factor. LSBoost exhibited the highest accuracy (98.25%). Similar recent studies were performed by various researchers [35–46] to predict the various properties of GPC.

While the above studies demonstrate the growing use of ML techniques for predicting the properties of GPC, they predominantly rely on black-box models with limited to no interpretability [22,26,32–36,47,48]. Although several ensemble and hybrid approaches have achieved high predictive accuracy, they offer little understanding of how input features influence model outputs. Most evaluations have been restricted to statistical metrics such as coefficient of correlation (R^2), root mean squared error (*RMSE*) and mean absolute error (*MAE*), without incorporating explainable frameworks capable of revealing global or instance-specific feature contributions. While previous studies have successfully employed various ML models for predicting the f_c of GPC, their adoption in practice remains limited due to the absence of deployable tools that facilitate ease of use for engineers and practitioners. This limitation highlights the need for user-friendly interfaces that translate ML outputs into accessible decision-support tools. Without interpretability and accessible deployment, the practical adoption of these ML models remains constrained. This highlights a critical research gap, namely, the need for transparent, explainable, and deployable ML solutions that can support informed decision-making in sustainable concrete design.

1.2. Research Methodology

This study adopts a data-driven approach to predict the f_c of SFGPC and to evaluate its environmental profile through LCA using the following steps.

Scientometric analysis: A scientometric analysis was conducted to review the literature data, identify the research trends related to SFGPC.

Development of dataset: A dataset of 363 samples of SFGPC with different combinations of input features and their corresponding f_c values for developing ML models was prepared. After randomization, the dataset is split into two subsets: one for training (80% of the total data) and one for testing (20% of the total data).

Comparative analysis of ML models (black box models): Five ML models, namely ANN, DT, RF, adaptive boosting (AdB), and gradient boosting (GB), were developed to predict the f_c of SFGPC. The performance of each model is statistically measured using metrics (*RMSE*, *MAE* and R^2), scatter plots for training and testing values, and a Taylor diagram.

Explainable artificial intelligence (XAI): The statistically best ML model was chosen among the five, then uses multi-layered XAI techniques to provide interpretable and understandable explanations of the black box ML model. XAI analysis adopts two layers of explainability: the primary layer, which uses summary plot and force plot to show the feature importance and contribution, and the second layer, which uses individual conditional expectation (ICE) with partial dependence plot (PDP) to show the feature effects and interactions.

Optimization: 3D Surface plots were generated to investigate the effect of two of the most significant factors, as suggested by the multi-layered XAI analysis, namely the Alk./B ratio and the age, on the f_c of SFGPC, for two types of binders: FlA and GGBFS. Patterns and insights revealed were discussed by the plots and their implications for the design and performance of SFGPC.

GUI development: To enable practical deployment and facilitate user interaction with the trained ML model, a prototype GUI was developed (Section S1 of Supplementary Materials). This user-friendly interface allows engineers and researchers to input mix design parameters such as binder content, Alk./B ratio, NaOH molarity, SS/SH ratio, curing age, and admixture dosage, etc. The GUI processes these inputs and outputs the predicted f_c.

1.3. Life Cycle and Carbon Neutrality Assessment

A cradle-to-gate LCA was performed to quantify the environmental impacts of the SFGPC mixes. The LCA evaluated embodied energy (EE), global warming potential (GWP), air emissions (NO_x, SO_x, and particulate matter (PM)), carbon neutrality assessment and normalized indices to reflect their sustainability performance alongside mechanical properties. The LCA results were then used to identify the mixture that successfully balances structural performance with lower environmental impacts. Finally, Carbon neutrality, defined as achieving net-zero GWP, was evaluated by estimating the GWP offset achievable through carbon capture, utilization, and storage (CCUS) and material substitution strategies.

The overall methodology adopted in the study is illustrated in Figure 1.

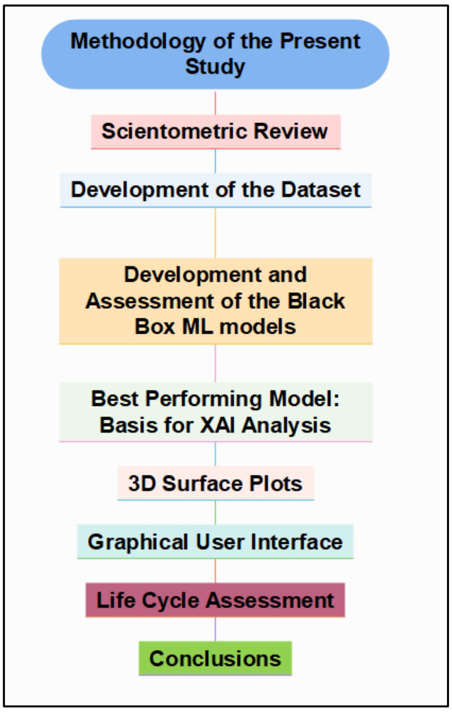


Figure 1. Methodology of the present study.

2. Scientometric Analysis

Scientometric analysis has been conducted to assess literature data, aiming to explore the research trend related to GPC and to collect the data systematically for developing ML models. Employing scientometric analysis ensures a more objective and impartial outcome [49]. Given the prolific output of research articles within the scientific community, it is crucial to identify credible sources for database inclusion. Scopus and Web of Science are acknowledged as the most comprehensive, effective, and unbiased databases for literature searches. Notably, Scopus, with its broader coverage and inclusion of more recent publications, is favoured.

During the extraction of bibliometric data, an initial Scopus search for "Geopolymer Concrete" retrieved 5281 items. To enhance precision, the search was subsequently refined to encompass only original articles, review studies, conference papers, and book chapters published in English, specifically within the domains of Engineering, material science, and environmental science. This refined search produced 3984 results. Further refinement was achieved by introducing the term "Fly Ash" and "Ground Granulated Blast Furnace Slag" to focus explicitly on studies that used these binders for the preparation of GPC. Following the application of these filters in the Scopus database, 441 results were retained. The obtained database was saved in CSV format to facilitate subsequent analysis using suitable software. The analysis was conducted utilizing VOSviewer version 1.6.20 software and the Biblioshiny package in the RStudio [50].

The dataset spanning the years 2009 to 2025, comprising 441 documents from 135 distinct sources. The annual growth rate of 16.59% underscores the significant expansion of scholarly contributions over the specified timespan. The document contents reveal a rich landscape, with 1816 unique keywords plus identifiers and 932 author's keywords, elucidating the diverse thematic focus of the research. Authors engaged in this scholarly endeavor total 1059, with only 13 documents being single-authored. Collaboration emerges as a hallmark, evidenced by an average of 3.36 co-authors per document, with approximately 14.97% of collaborations transcending international boundaries. Notably, the average citations per document stand at 23.15, reflecting a moderate level of scholarly impact. Furthermore, the dataset comprises 328 articles, 17 book chapters, and 96 conference papers, underscoring the multidimensionality of the research output, revealing a dynamic and collaborative research landscape characterized by growth, impact, and internationalisation.

The sematic diagram for most occurring keywords is shown in Figure 2, offering insights into thematic concentrations and their network metrics. The most occurring keywords from the dataset extracted are “fly ash”, “slags”, “geopolymer concrete”, “inorganic polymer” and “compressive strength”. This analysis aids in understanding the interconnected nature of research themes and highlights pivotal concepts shaping the scholarly discourse within the examined dataset.

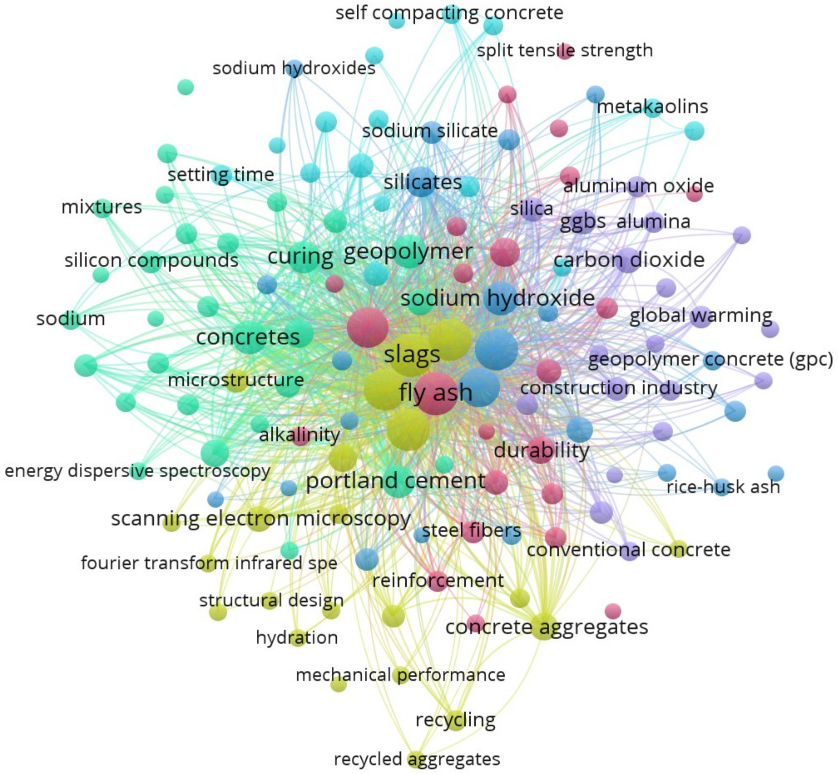


Figure 2. Semantic diagram for the most occurring keywords.

Figure 3 presents the thematic evolution of keywords within the research landscape of GPC over distinct periods. The selection of time brackets for the thematic evolution analysis was guided by publication density patterns observed in the scientometric dataset rather than by uniform temporal segmentation. Analysis of annual publication trends revealed three distinct phases: (i) an initial formative period (2009–2015) characterized by relatively low and sporadic publication output, (ii) a consolidation and rapid growth phase (2016–2022) marked by a sharp increase in annual publications, and (iii) a recent expansion phase (2023–2025) exhibiting sustained high publication density with diversification toward data-driven modelling, and sustainability-oriented research. These breakpoints were identified by examining inflection points in cumulative publication counts and year-wise output, ensuring that each time window contained a statistically meaningful number of studies and represented a distinct stage in the evolution of the research field.

Noteworthy transitions include the shift from ‘concrete’ (2009–2015) to “geopolymer” (2016–2022), suggesting an evolving research focus. The progression from “fly ash” (2009–2015) to “slag” and “circular economy” (2016–2022) signifies a thematic shift towards sustainable construction practices aligned with circular economy principles. Furthermore, the persistence of “fly ash” as a keyword from 2009–2015 to 2016–2022, accompanied by associated terms like “compressive strength”, “alkaline solution”, and “workability”, underscores its enduring significance in the discourse. The evolution from “fly ash” (2016–2022) to diverse theme like “ambient curing” (2023–2025) highlights the expanding scope of geopolymer research within this period. The scientometric keyword co-occurrence analysis revealed that terms such as “ambient curing,” “steam curing,” and “thermal curing” frequently appear in the literature. These terms do not represent thematic research domains; instead, they reflect commonly adopted experimental curing conditions reported across geopolymer studies. Their frequent appearance in the mapping output indicates the diversity of curing regimes investigated in the field rather than distinct conceptual themes. Additionally, the emergence of new keywords such as “Taguchi method” and “metakaolin” in 2023–2025 suggests evolving methodologies and broader conceptual frameworks. This analysis contributes to a nuanced understanding of the dynamic thematic evolution within the studied research domain.

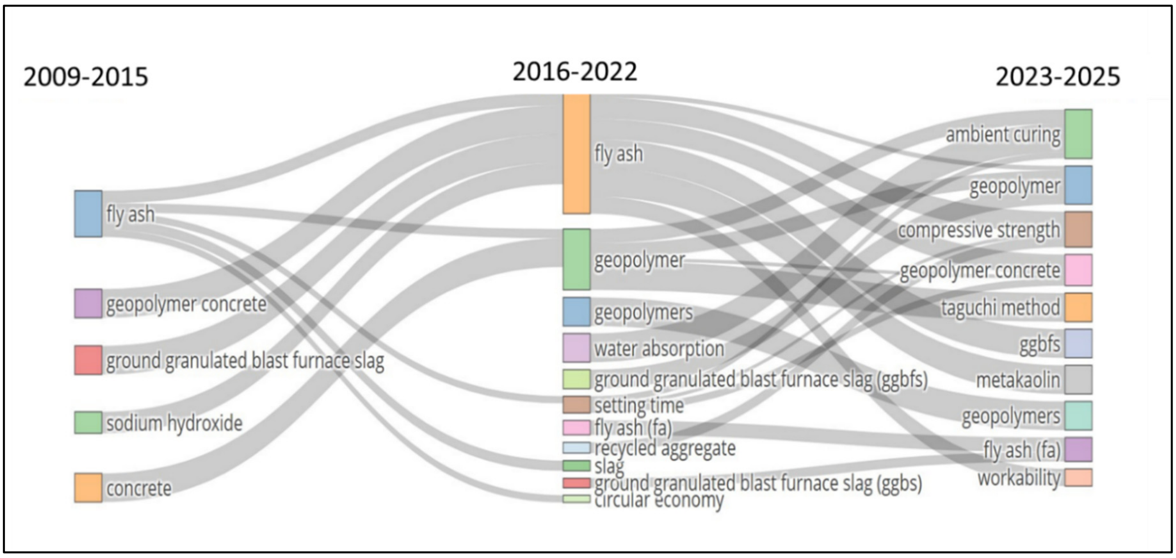


Figure 3. Thematic evolution over time, illustrating the dynamic shifts.

The thematic map, shown in Figure 4, discerns the research landscape into four distinct categories, reflecting the dynamics of emerging, basic, motor, and niche themes. In the emerging theme quadrant, "Ground Granulated Blast Furnace Slag" stands out as a central and evolving focus, indicating a burgeoning research trend. The basic theme quadrant is characterized by the prominence of "Sorptivity, Water Absorption and Bond Strength", suggesting foundational aspects of study within this domain. The motor theme quadrant is dominated by the extensive cluster around "Mechanical Properties", indicating a core and driving force within the research discourse. Finally, the fourth quadrant, featuring the cluster around "Fly Ash", signifies a niche theme that, while well-established, remains a specific and distinctive area of focus within the broader research landscape. This categorization offers a detailed perspective on the evolving and stable thematic concentrations, providing valuable insights into the overarching trends and specialized niches shaping the research discourse.

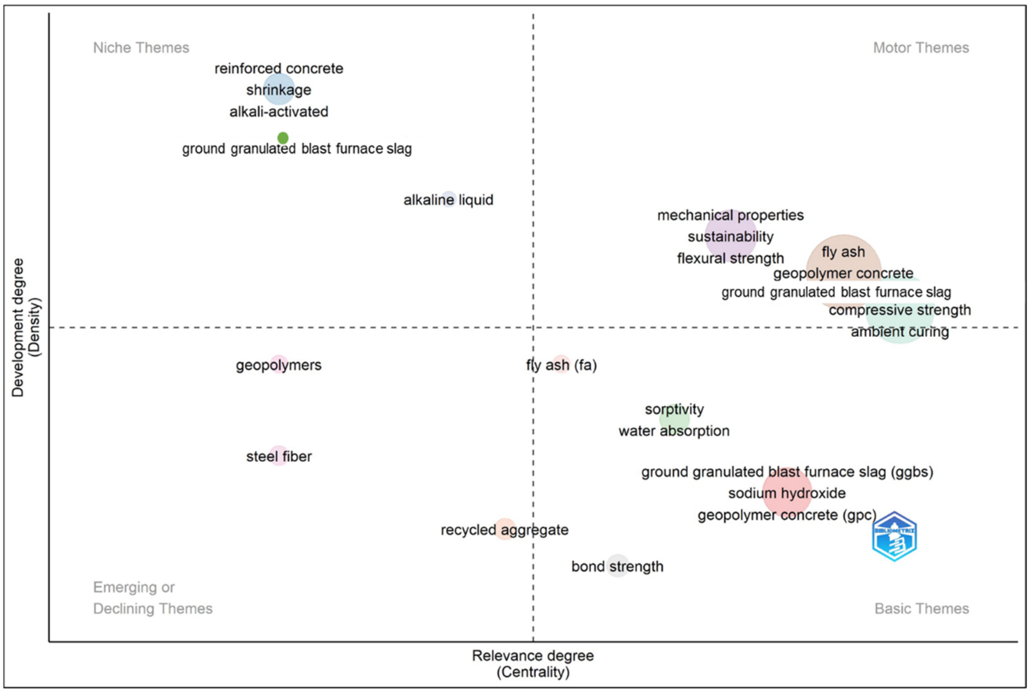


Figure 4. Thematic map illustrating research concentrations.

3. Methods and Materials

3.1. Modelling Techniques

The models, developed for predicting the f_c of SFGPC, encompass both conventional and ensemble approaches. The conventional ML (CML) models consist of the ANN and DT, while the EML models comprise

RF, AdB, and GB. The selection of the CML and EML models used in this study was guided by both the nature of GPC data and evidence from prior research. Notably, CML models operate as standalone algorithms, whereas EML models harness the collective power of multiple algorithms to enhance predictive accuracy. Although the CML and EML models were not applied concurrently, assessing both categories offers complementary insights, such as CML models provide interpretable baseline behaviour while EML models leverage ensemble diversity to enhance accuracy. The theoretical underpinnings of all ML models employed in this study are elaborated below.

3.1.1. Artificial Neural Network

ANNs are nonlinear universal function approximators composed of interconnected processing units (“neurons”) organized in layers. Each neuron computes a weighted sum of its inputs followed by a nonlinear activation function. For a neuron j, the output is expressed as Equation (1):

$$y_j = \sigma\left(\sum_{i=1}^{n} w_{ij} x_i + b_j\right) \tag{1}$$

where w_{ij} and b_j are the learnable weights and biases, respectively, and $\sigma(\cdot)$ denotes the activation function (ReLU in this study). Model training is performed using backpropagation, which minimizes the loss function L by updating weights via gradient descent Equation (2):

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} - \eta \frac{\partial L}{\partial w_{ij}} \tag{2}$$

where η is the learning rate. The ANN is capable of capturing complex nonlinear relationships among mix parameters such as binder composition, activator chemistry, and curing age, making it suitable for modelling the intricate behaviour of GPC.

3.1.2. Decision Tree

DT partitions the feature space into homogeneous regions using recursive binary splits. At each internal node, the algorithm selects the feature x_k and s threshold that maximize impurity reduction. For regression, impurity is commonly measured using variance Equation (3):

$$\mathrm{Var}(S) = \frac{1}{|S|} \sum_{i \in S} (y_i - \bar{y}_S)^2 \tag{3}$$

The optimal split minimizes Equation (4):

$$\Delta = \mathrm{Var}(S) - \left(\frac{|S_L|}{|S|} \mathrm{Var}(S_L) + \frac{|S_R|}{|S|} \mathrm{Var}(S_R) \right) \tag{4}$$

where S_L and S_R are the left and right child nodes. DTs are easy to interpret but prone to overfitting and high variance, which motivates the use of ensemble approaches.

3.1.3. Random Forest

RF is an ensemble of DT constructed using bootstrap aggregation (“bagging”). Each tree is trained on a randomly sampled subset of the data and a random subset of features at each split. For a forest of T trees, the prediction is Equation (5):

$$\hat{y} = \frac{1}{T} \sum_{t=1}^{T} f_t(x) \tag{5}$$

Bagging reduces variance, while feature randomness decorrelates trees, thereby improving generalization. RF is particularly effective for handling nonlinearities and interactions among variables such as alkali ratios, aggregate contents, and binder proportions.

3.1.4. Adaptive Boosting

AdB constructs an ensemble by sequentially training weak learners (typically shallow DTs). Each learner focuses on instances mis-predicted by its predecessor by assigning higher weights to them. For iteration m, the weighted error is Equation (6):

$$\epsilon_m = \sum_{i=1}^{N} w_i^{(m)} \mathbb{I}\left(y_i \neq h_m(x_i)\right) \tag{6}$$

The model weight is computed as Equation (7):

$$\alpha_m = \ln\left(\frac{1-\epsilon_m}{\epsilon_m}\right) \tag{7}$$

and sample weights are updated as Equation (8):

$$w_i^{(m+1)} = w_i^{(m)} \exp\left(\alpha_m \mathbb{I}\left(y_i \neq h_m(x_i)\right)\right) \tag{8}$$

The final prediction aggregates all weak learners Equation (9):

$$\hat{y} = \sum_{m=1}^{M} \alpha_m h_m(x) \tag{9}$$

AdB effectively reduces bias and improves accuracy when underlying patterns are subtle or nonlinear.

3.1.5. Gradient Boosting

GB model in this study, constructs an additive ensemble by fitting new learners to the negative gradients of the loss function. For a model consisting of m learners, the iterative update is Equation (10):

$$F_m(x) = F_{m-1}(x) + \nu h_m(x) \tag{10}$$

where ν is the learning rate and $h_m(x)$ approximates the gradient Equation (11):

$$h_m(x) \approx -\frac{\partial L(y, F(x))}{\partial F(x)} \tag{11}$$

For regression with *MSE* Equation (12):

$$g_i^{(m)} = y_i - F_{m-1}(x_i) \tag{12}$$

representing the residuals that each new tree attempts to model. GB's ability to sequentially minimize residual patterns enables it to capture highly nonlinear interactions, such as the synergistic influence of slag content, activator chemistry, and curing age on compressive strength.

Prior to model training, the dataset was subjected to a structured preprocessing pipeline to enhance numerical stability and improve model performance. All input features were normalized using Min-Max scaling, defined as $x' = \frac{x - x_{min}}{x_{max} - x_{min}}$, where x_{min} and x_{max} represent the minimum and maximum values of each feature in the dataset. This approach preserves distributional characteristics while ensuring that no variable dominates due to differences in scale. Normalization was especially important because the input parameters (e.g., aggregates, binder contents, molarity, and Alk./B ratio) vary significantly in magnitude. After normalization, the dataset was randomly divided into 80% training and 20% testing subsets. All ML models were trained on the normalized dataset to ensure fair comparison and robust convergence.

3.2. Materials

Database Development and Statistical Analysis

The data set used in this study consists of 363 samples of SFGPC with different combinations of these features and their corresponding f_c values [51–65]. The dataset contains eleven features, out of which ten are the inputs and a single output. It is important to note that the dataset includes only the mass-based quantities of fine (Fagg.) and coarse (Cagg.) aggregates in kg/m^3, as reported in the literature. Aggregate size gradations were not included because most publications did not consistently provide detailed size ranges. Therefore, Fagg. and Cagg. in this study refer solely to the reported mass of aggregates and not to specific particle size distributions. Also, in alkali-activated GPC systems, the primary liquid phase is supplied through sodium silicate and NaOH solutions, whose inherent water content is governed by activator concentration and silica modulus. In addition to this inherent water, several studies explicitly reported the use of supplementary water to improve workability or casting feasibility. Such externally added water is represented in the dataset by the parameter "Additional Water (W)" and was included only when clearly documented in the source publications. Therefore, the dataset distinguishes between water intrinsically present in the activator solutions and supplementary water added during mixing, both of which influence fresh and hardened properties. Furthermore, the dataset used in this study was assembled, the preparation, curing, and testing procedures reflect the standards followed in the respective studies. Only those publications that explicitly adhered to recognized standards, such as ASTM C39 for f_c testing, ASTM C192 or equivalent national standards for specimen preparation, and clearly reported curing regimes (ambient, oven, steam), were included in the dataset. The descriptive statistics of the data set are shown in Table 1.

Table 1. Statistical analysis of the data collected.

Features	Abbreviations	Unit	Min.	Max.	Mean	Std. Deviation	Skewness	Kurtosis
Fly Ash	FlA	kg/m^3	0	500	258.3	123.86	−0.495	−0.686
Ground Granulated Blast Furnace Slag	GGBFS	kg/m^3	0	488	151.6	109.96	0.822	0.218
Fine Aggregate	Fagg.	kg/m^3	57.01	707	565.2	138.19	−2.7	7.388
Coarse Aggregate	Cagg.	kg/m^3	132.26	1662	1137.2	287.87	−2.185	5.669
Alkaline soln./Binder ratio	Alk./B	-	0.035	0.719	0.43	0.11	−0.193	1.491
Sodium Silicate/Sodium Hydroxide	SS/SH	-	0	5.006	2.26	0.59	0.496	5.092
Molarity	M	Molar	6	16	10.8	2.35	0.184	−0.832
Superplasticizer	Sp	kg/m^3	0	16	4.50	5.23	0.914	−0.248
Additional Water	W	kg/m^3	0	100.32	9.33	22.58	2.599	5.446
Age	A	Days	3	180	25.39	29.84	3.03	11.465
Compressive Strength	f_c	MPa	0.407	68.602	35.32	14.61	0.259	−0.678

A complete dataset consisting of 363 mix designs and their corresponding compressive strength values, extracted from 15 peer-reviewed publications, has been included as Supplementary Material (Dataset S3). The statistical analysis of the dataset reveals notable trends in the distribution of its constituent materials and influencing parameters. The two primary binders as FlA and GGBFS exhibit contrasting distribution characteristics. FlA ranges from 0 to 500 kg/m^3 with a mean of 258.32 and displays a slightly left-skewed and platykurtic distribution, indicating a relatively balanced spread of values with few extreme highs. In contrast, GGBFS, ranging from 0 to 488 kg/m^3 with a mean of 151.57, is right-skewed and slightly leptokurtic, suggesting a tendency towards lower values with occasional high values.

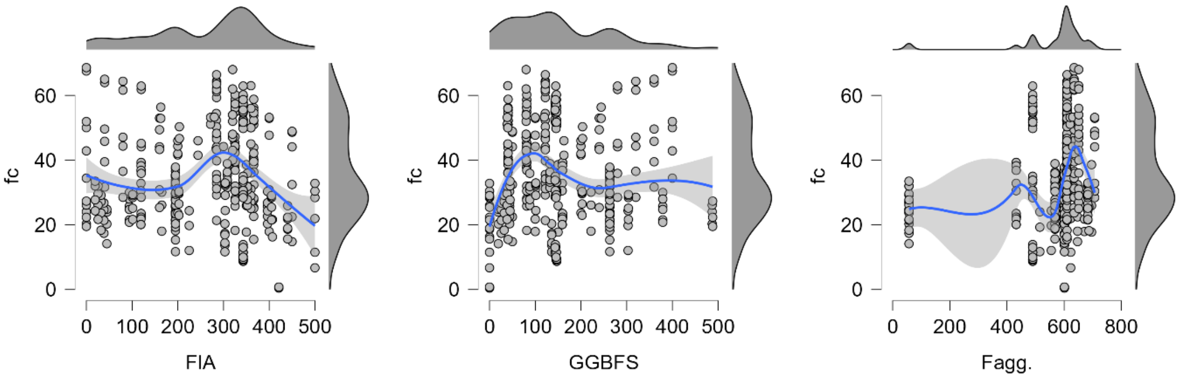
Among the aggregates, both Fagg. and Cagg. show left-skewed and highly leptokurtic distributions, reflecting a concentration of values toward the higher end of their respective ranges (57.01–707 kg/m^3 for Fagg. and 132.26–1662 kg/m^3 for Cagg.) with extreme peaks, indicative of consistency in mix proportions. The Alk./B exhibits a slightly left-skewed, moderately leptokurtic distribution with values between 0.035 and 0.719, showing that most mixes favoured moderate alkaline dosages. The SS/SH ratio is right-skewed and highly leptokurtic (0–5.006, mean = 2.26), highlighting frequent use of moderate SS/SH ratios with some extreme values. The molarity of NaOH (M) varies from 6 to 16 (mean = 10.84), with a near-symmetric but platykurtic distribution, indicating controlled variability in the concentration of the alkaline activator. The Sp and W features exhibit right-skewed distributions, particularly W with a high skewness of 2.599, suggesting that water was added sparingly in most mixes, with few cases requiring significantly higher amounts. The curing age (A) of specimens, ranging from 3 to 180 days with a high positive skewness (3.03) and kurtosis (11.47), indicates that most samples were tested at early ages, with a small number extending to long-term curing.

Finally, the target variable, f_c, spans from 0.407 to 68.60 MPa, with a mean of 35.32 MPa. Its slightly right-skewed and platykurtic distribution suggests a fair spread around the mean with fewer outliers. These insights confirm that the dataset captures a wide variety of SFGPC mix designs, covering a realistic range of material proportions and curing conditions, making it suitable for developing robust and generalizable predictive models.

3.3. Data Visualization

3.3.1. Scatter Plots and Histogram

Scatter plots, along with the violin density plots on the sides, are shown in Figure 5. The scatter plots show that some features have a clear positive or negative correlation with the f_c, such as FlA, GGBFS, and A, while others have a complex or weak correlation, such as Fagg., Cagg., and W. Figure 6 represents the histograms of f_c showing how often different values of the f_c occur in the data set. The histograms also show that the highest f_c is reached at 180 days, with a value of 68.602 MPa.



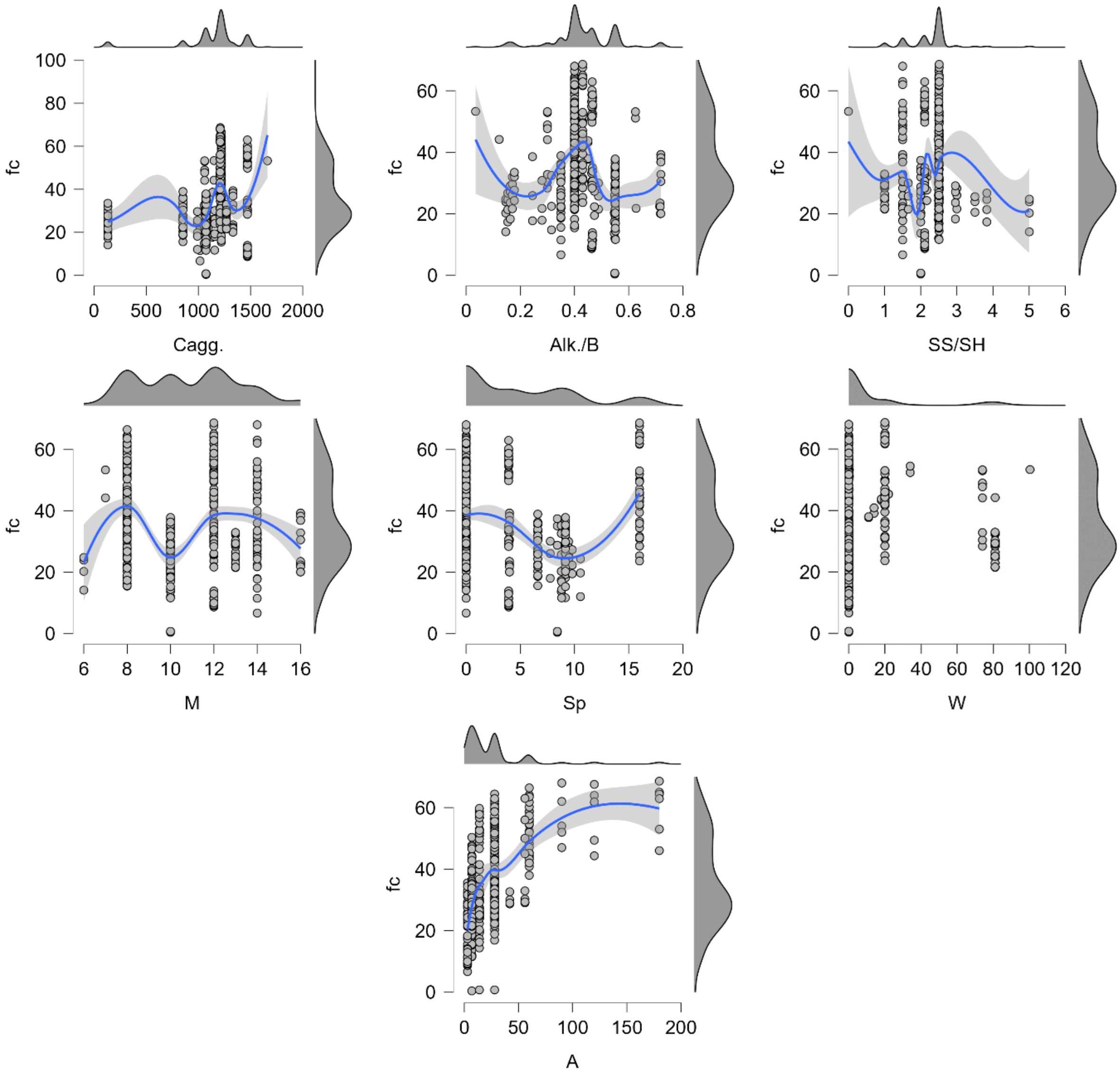


Figure 5. Scatter plots with violin density between input features and f_c of SFGPC.

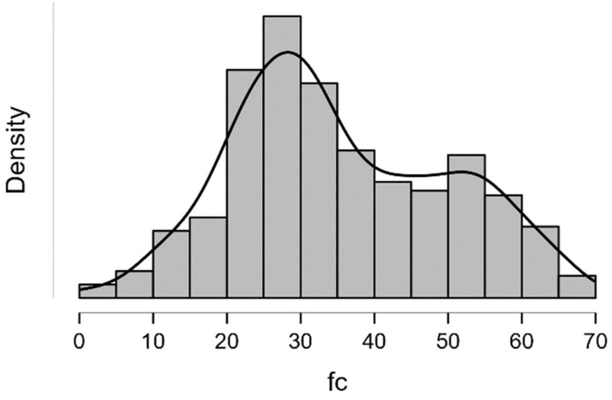


Figure 6. Frequency of f_c values of SFGPC within the dataset.

3.3.2. Mixed Correlation Plot

Figure 7 shows the mixed correlation plot between the f_c of SFGPC and the input features. The plot indicates that the f_c has a positive correlation with the "A", and "aggregates". This means that increasing these features will likely to increase the f_c of GPC. On the other hand, the f_c has a negative correlation with the "M". This means that increasing these features will decrease the f_c of SFGPC. The plot also shows that the f_c has low correlation with the other remaining inputs. This means that changing these features will have a little effect on the f_c of SFGPC.

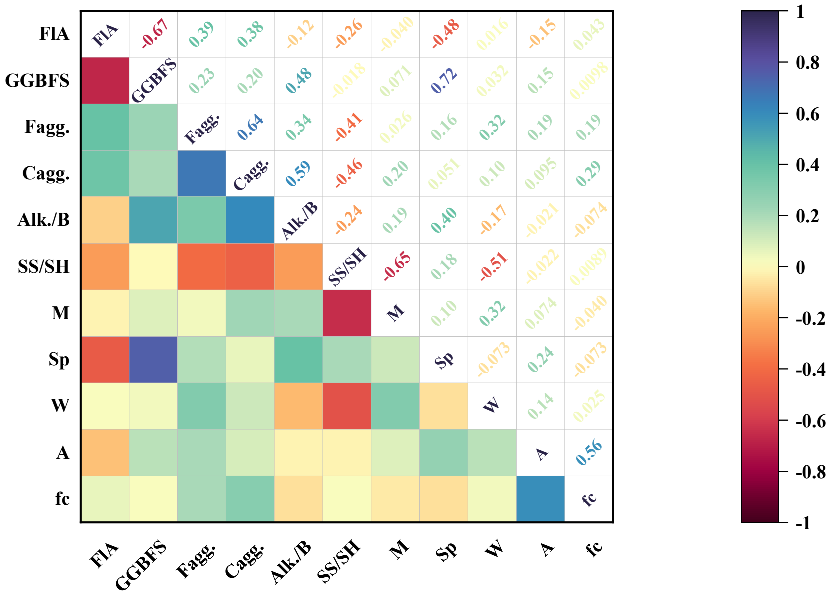


Figure 7. Mixed correlation plot.

3.4. Machine Learning Methodology

Two CML models, namely ANN and DT and three EML, which are RF, AdB and GB have been used for predicting the f_c of sustainable SFGPC. The data set of 363 instances was split into two subsets: one for training (290 instances, 80% of the data) and one for testing (73 instances, 20% of the data). The performance of each model was measured using three metrics: *RMSE*, *MAE* and R^2. The results of the comparison between various ML models developed were also visualised using scatter plots and a Taylor diagram, which show the correlation and dispersion of the predicted and actual values of the f_c for both the training and testing phases. The model that achieved the highest performance in terms of all the metrics and visualizations was chosen as the basis for XAI, which aims to provide interpretable and understandable explanations of the black box ML models. In addition, a prototype GUI was developed with the best-performing model to instantly predict the f_c of the sustainable SFGPC based on user-defined inputs.

4. Results and Discussions

4.1. Hyperparameter Tuning

The hyperparameter tuning process was designed to be objective and non-arbitrary. For each model, the search space was defined based on commonly reported ranges in prior geopolymer ML studies and constrained to avoid excessive model complexity. Preliminary sensitivity checks were conducted to identify parameters with the greatest influence on prediction error, and only these parameters were included in the grid-search to reduce redundancy. Model configurations and weights were evaluated using k-fold cross-validation, and the optimal hyperparameter set was selected based on the minimum average validation error, subject to the additional constraint that training and testing errors remain closely aligned (Section S2 of Supplementary Materials). This ensured that the final model configurations were selected through performance-based criteria rather than subjective preference. As far as the model specific explanation is being concerned, for the GB model, the learning rate, tree depth, and number of estimators were jointly optimized to balance bias and variance. Lower learning rates combined with a moderate number of estimators were preferred to prevent overfitting, while shallow tree depths were selected to limit model complexity. The final configuration was chosen because it provided the lowest cross-validated error and exhibited stable convergence without divergence between training and testing performance. The tuned hyperparameters ensured stable model convergence and minimized the risk of overfitting. Table 2 below shows the tuned hyperparameters and the performance metrics for each model.

Table 2. Tuned parameters for all ML models.

Model	Optimized Hyper-Parameters
ANN	• Number of hidden layers—2 • Number of neurons in hidden layers—12,12 • Activation function—ReLu • Solver—Adam • Regularization—0.0005 • Maximal number of iterations—1,000,000
DT	• Minimum number of instances in leaves—2 • Do not split subsets smaller than—3 • Limit the maximal tree depth to—125
RF	• Number of Trees—50 • Do not split subset smaller than—2 • Limit depth of individual trees—8
AdB	• Base estimator—Tree • Number of estimators—5 • Learning rate—0.6 • Fixed seed for random generator—40 • Regression loss function—exponential
GB	• Number of trees—100 • Learning rate—0.100 • Limit depth of individual trees—5 • Do not split subset smaller than—3

4.2. Model Performance Parameters

The performance of the five ML models, was compared and evaluated using three metrics: *RMSE*, *MAE*, and R^2. Table 3 shows the values of the performance metrics for each model after the training and testing phases, respectively. Figure 8 shows the radar diagrams that illustrate the comparison of the performance metrics visually. The results indicate that the GB model, which is an EML model, achieved the highest performance in terms of all the metrics and visualizations.

Table 3. Performance metrics for various ML models used to predict the f_c of GPC.

S. No.	Model	Phase	*RMSE*	*MAE*	R^2
1.	ANN	Training	4.582	3.424	0.903
		Testing	5.452	4.335	0.848
2.	DT	Training	3.913	2.618	0.929
		Testing	6.747	4.637	0.767
3.	RF	Training	3.718	2.749	0.936
		Testing	4.248	3.461	0.908
4.	AdB	Training	3.249	1.759	0.951
		Testing	4.641	3.401	0.89
5.	GB	Training	3.15	1.809	0.954
		Testing	3.128	2.411	0.95

4.3. Scatter Plots

Scatter plots, as shown in Figure 9, show the comparison of the predicted and actual values of the f_c of SFGPC for the five ML models after the training and testing phases. The plots also include the error lines that represent the deviation of the predicted values from the actual values within a range of ±20%. The results from Figure 9 reveal that the GB model, which is an EML model, has the highest prediction accuracy and the lowest deviation among all the models, followed by the AdB model. The CML models, namely ANN and DT, have lower prediction accuracy and higher deviation than the EML models.

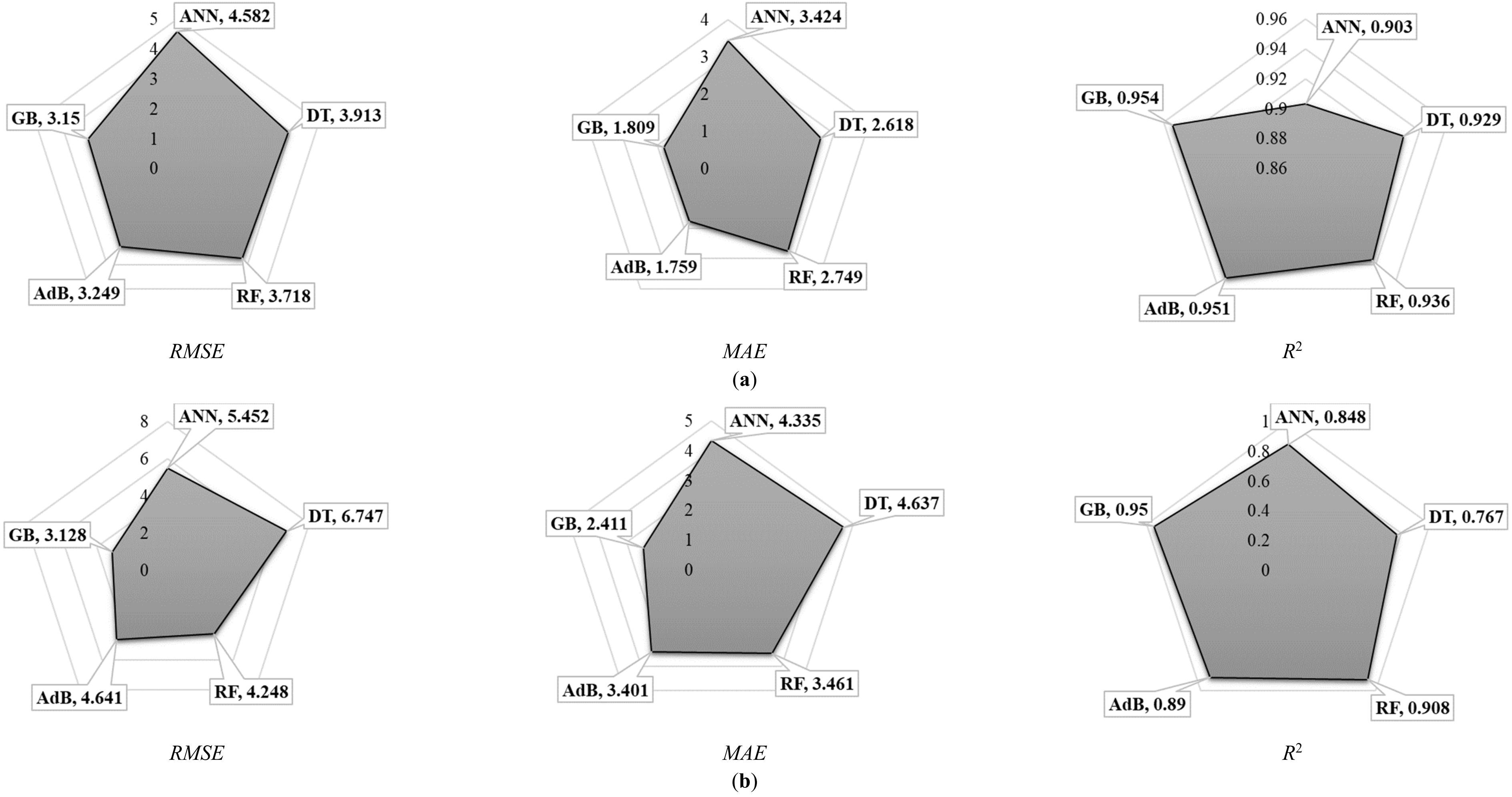


Figure 8. Radar diagrams showing the comparison of performance between various ML models after (**a**) training and (**b**) testing phases.

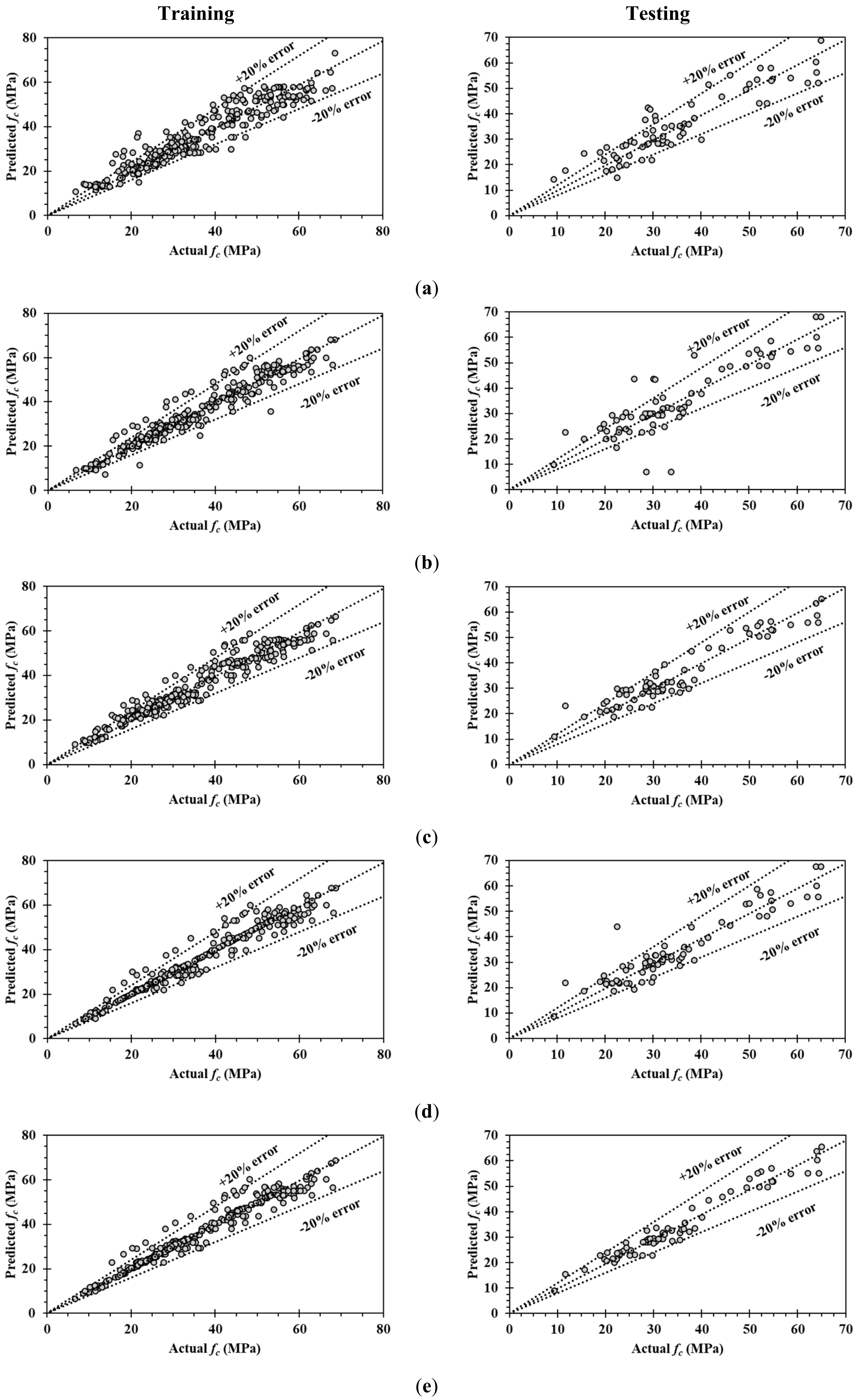
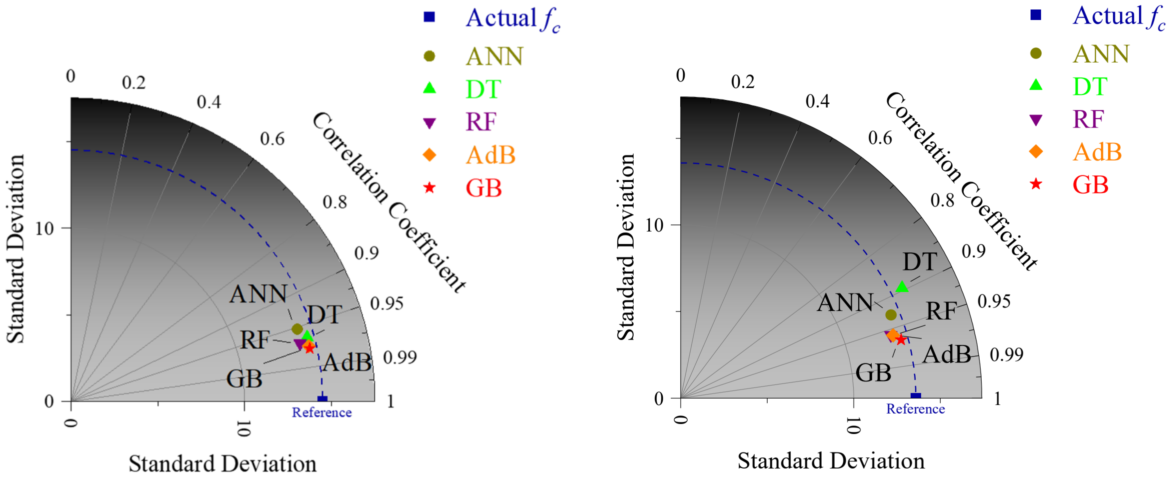


Figure 9. Scatter plots of predicted and actual values of f_c by (**a**) ANN, (**b**) DT, (**c**) RF, (**d**) AdB and (**e**) GB models after training and testing phases, respectively.

4.4. Taylor Diagram

The Taylor diagram in Figure 10 compares the performance of the five black box ML models. The diagram shows that the GB model has the highest correlation coefficient, the lowest error, and the closest standard deviation to the reference standard deviation of the actual f_c. This indicates that the GB model has the best agreement with the actual f_c among all the models. The other EML models, namely RF and AdB, also have high correlation coefficients, low errors, and close standard deviations to the reference standard deviation. The EML models have better performance than the CML models, which have lower R^2, comparatively higher errors, and farther standard deviations from the reference standard deviation.



(**a**) (**b**)

Figure 10. Taylor diagram for various ML models after (**a**) training, and (**b**) testing phase, respectively.

The results from Table 3, Figures 8–10 demonstrate that the GB model outperforms the other four ML models in predicting the f_c of SFGPC for this dataset. The other EML models, namely RF and AdB, also have good performance, but not as good as the GB model. The CML models have comparatively lower performance and higher deviation than the EML models. Hence, the GB model is selected as the best predictive model for the f_c of SFGPC for this dataset, and thus, the GB model is chosen as the basis for multi-layered XAI analysis for providing insights into the feature importance and the decision rules that govern the prediction process.

4.5. Model Explanation Using XAI

4.5.1. Primary Layer of Explainability

The primary layer of explainability provides a summary of the feature importance and effects as suggested by the GB model. The primary layer of explainability consists of two types of plots: a summary plot and a force plot. The summary plot ranks the features from the most important to the least important, based on their performance metric values. The force plot shows the contribution of each input feature to the prediction of the f_c of SFGPC for a specific instance in the dataset. The primary layer of explainability helps to understand the overall and individual predictions of the GB model.

Summary Plot

The summary plot by GB model shown in Figure 11 represents the relative importance of the input features in predicting the f_c of SFGPC. The plot reveals that the most important feature is the age “A” of the concrete specimens, followed by the amount of “GGBFS”, “Alk./B”, “Cagg.”, “FlA”, “Fagg.”, “Sp”, “SS/SH”, “W” and “M”. The plot also shows that some features have a positive impact on the f_c, such as “A”, “GGBFS”, and “Cagg.”, some have a negative impact, such as “Alk./B”, “FlA”, and “M”, while others have a little impact on the f_c within this dataset. The plot provides a useful overview of the feature importance and the direction of their effects on the prediction:

- Age: The age of the concrete specimens is the most important feature, as it has a positive effect on the prediction. This means that the older the specimens, the higher their f_c. This is because the geopolymerization process continues over time, resulting in a denser and stronger matrix [13].

- Ground Granulated Blast Furnace Slag: The amount of GGBFS in the binder material is the second most important feature, as it has a positive effect on the prediction. This means that the more GGBFS, the higher the f_c. This is because GGBFS can enhance the geopolymerization process by providing additional silicon and aluminum, as well as reducing the heat of hydration and improving the durability of the concrete [66].
- Alkaline soln./Binder ratio: The ratio of the alkaline solution to the binder material is the third most important feature, as it has a negative effect on the prediction. This means that the higher the ratio, the lower the f_c. This is because a high ratio can cause an excess of alkaline solution, which affects the geopolymerization process and the durability of the concrete [13].
- Coarse Aggregate: The amount of coarse aggregate in the mixture is the fourth most important feature, as it has a positive effect on the prediction. This means that the more the aggregate, the higher the f_c. This is because coarse aggregate can provide resistance to the applied load and reduce the shrinkage of the concrete. However, the effect of coarse aggregate is also complex and nonlinear, as it also depends on the size and shape of the particles [67].
- Fly Ash: The amount of FlA in the binder material is the fifth most important feature, as it has a negative effect on the prediction. This means that the more FlA, the lower the f_c [68].
- Fine Aggregate: The amount of fine aggregate in the mixture is the sixth most important feature, as it has a positive effect on the prediction. This means that the more the fine aggregate content, the higher the f_c. This is because fine aggregate can fill the voids and improve the density of the concrete. However, the effect of fine aggregate is complex and nonlinear, as it also depends on the grading and shape of the particles [69].
- Superplasticizer: The amount of superplasticizer in the mixture is the seventh most important feature, as it has a positive or negative effect on the prediction, depending on the value. This means that the optimal amount of superplasticizer can increase the f_c, while too much or too little can decrease it. This is because superplasticizer can improve the workability and fluidity of the mixture, but it can also affect the water content and the geopolymerization process [69].
- Sodium Silicate/Sodium Hydroxide: The SS/SH ratio in the alkaline activator solution is the eight most important feature, as it has less effect on the f_c prediction [70].
- Water: The amount of additional water in the mixture is the ninth most important feature, as it has a little or positive effect on the f_c prediction.
- Molarity: The molarity of the alkaline activator solution is the least important feature, as it has a negligible effect on the f_c prediction.

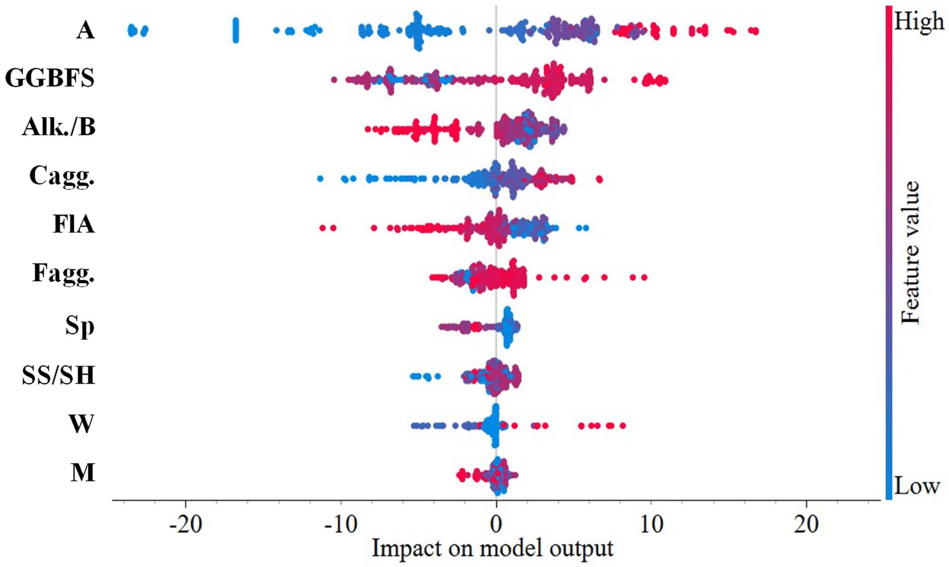


Figure 11. Summary plot showing the global and local feature importance as suggested by the GB model.

Force Plot

The force plot in Figure 12 shows the contribution of each input feature to the prediction of the f_c of SFGPC for a specific instance in the dataset. The base value is the average f_c of the training data, which is 34.88 MPa. The final predicted value is the sum of the base value and the SHAP values of each feature, which is 51.94 MPa. The plot indicates that the force features for this instance are A, Cagg., Alk./B, and Fagg., Sp, SS/SH, W and M.

GGBFS and FlA were the resistant variables for this specific instance. The plot also shows that the features have different effects on the f_c depending on their values. For example, increasing “A” tends to increase the f_c, while increasing FlA tends to decrease the f_c. The plot also reveals that some features have negligible impact on the f_c for this instance, such as Sp, SS/SH, W and M. Therefore, the force plot provides a detailed explanation of how the GB model predicts the f_c of SFGPC for a specific instance, based on the values and importance of each feature.

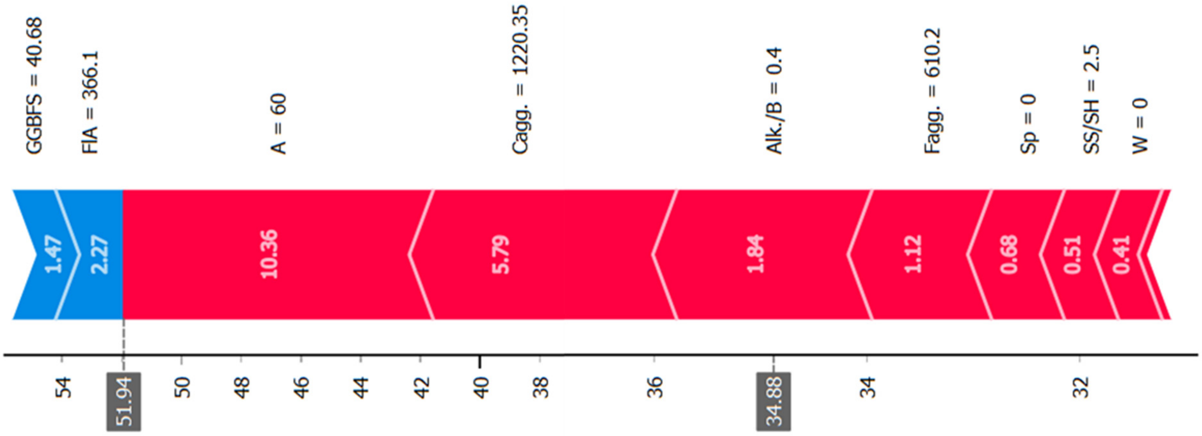


Figure 12. Force plot for a specific instance by GB model.

4.5.2. Second Layer of Explainability

Herein, the second layer of explainability refers to the use of ICE with PDP curves to show how the f_c changes as a function of each feature, while holding the other features constant. This layer provides more granular and individualized information than the first layer, which uses summary plot and force plot to show the overall feature importance and contribution. This helps in gaining more insights into the relationship between the features and the target variable, as well as the variability and interactions among the features.

ICE with PDP Curves

The ICE with PDP curves in Figure 13 shows the effect of varying each of the most significant features on the f_c of SFGPC, while keeping the other features fixed at their average values. The curves reveal that some features have a linear or nonlinear relationship with the f_c, such as A, GGBFS, and Alk./B, while others have a comparative weak relationship, such as FlA. The curves also show that some features have a positive or negative impact on the f_c, depending on their values. For example, increasing A tends to increase the f_c, while increasing Alk./B tends to decrease the f_c. By this the ICE with PDP curves provides a deeper understanding of how the GB model predicts the f_c of SFGPC, based on the values and effects of each feature.

4.5.3. 3D Surface Plots

To investigate the effect of two of the significant factors, namely the Alk./B ratio and the age, on the f_c, 3D Surface plots were generated using the GB model as shown in Figure 14. The plots show the f_c values for different combinations of Alk./B and age, for two types of binders: FlA and GGBFS independently and the synergistic behaviour too. The plots reveal some interesting patterns and insights, as discussed. One of the patterns that can be observed from the plots is that the f_c decreases with the increase of Alk./B ratio, regardless of the type of binder. This implies that higher alkalinity is unfavourable for the geopolymerization process and the formation of a denser and stronger matrix. Another pattern is that the f_c increases with the increase of age, especially for the GPC made with GGBFS. This indicates that the GPC made with GGBFS has a higher potential for strength development over time. A third pattern is that the f_c of GPC made with GGBFS is consistently higher than the f_c of GPC made with FlA, for all values of Alk./B and age. This suggests that the GGBFS is a more suitable binder for GPC than the FlA, as it provides better mechanical properties. These patterns and insights can help optimize the design and performance of SFGPC for various applications.

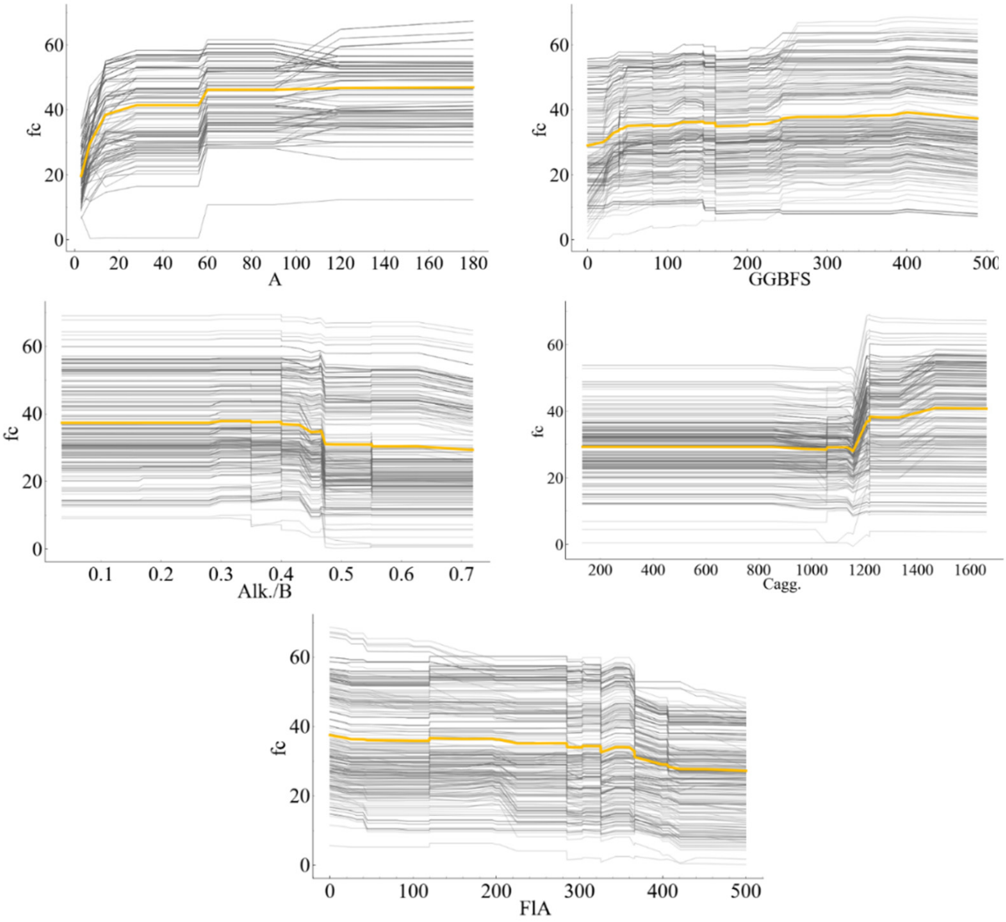
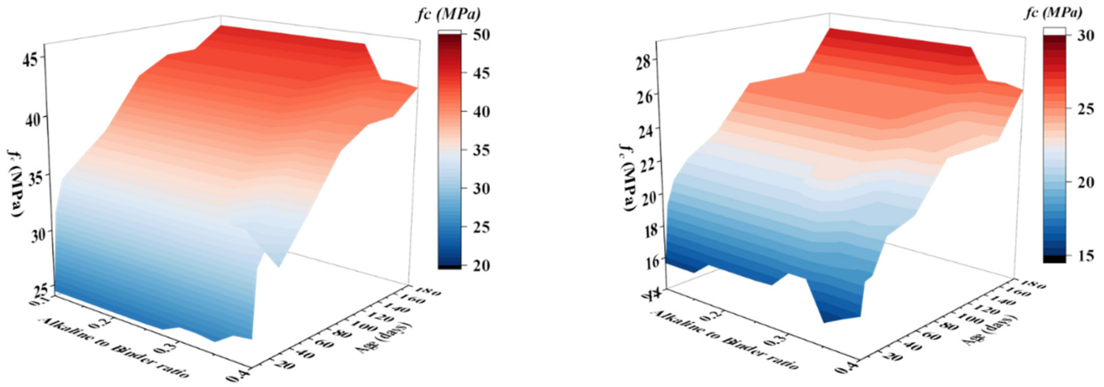
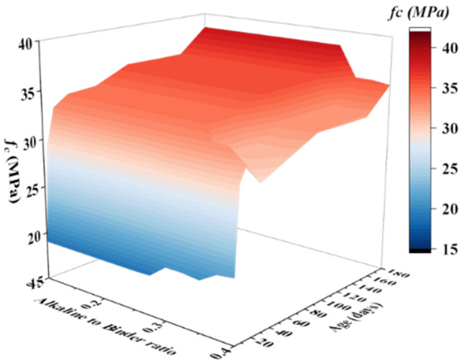


Figure 13. ICE with PDP curves for the most significant features, showing how the f_c of SFGPC varies with each feature value.



GGBFS as binder FlA as binder



GGBFS and FlA both as binder

Figure 14. Effect of Alk./B ratio with Age on the f_c using 3D surface plots when GGBFS or/and FlA was/were used as binder.

5. Sustainability Analysis of GPC Mixes

A cradle-to-gate LCA was conducted to evaluate the environmental performance of three GPC mixes with comparable f_c (~27.3–28.4 MPa): Mix 1 (GGBFS-Dominant), Mix 2 (Hybrid or synergistic), and Mix 3 (FlA-Dominant). The mixes composition used for LCA are shown in Table 4. The analysis focused on EE, GWP, carbon neutrality and air emissions, including Nitrogen Oxides (NO_x), Sulfur Oxides (SO_x), and PM. To account for mechanical performance alongside environmental effects, two normalized sustainability indices were defined: Sustainability Index (S_i) and Energy Efficiency Index (EE_i). Carbon neutrality potential was assessed through an assumed 25% GWP offset, based on carbonation (10%), CO_2 injection in aggregates (5%), and material substitution (10%). Although intentionally limited in scope, this assessment provides a comparative understanding of how the key parameters identified through XAI, particularly binder composition and activator dosage, influence both mechanical performance and environmental outcomes, which is important for sustainable mix design in structural applications.

Table 4. Three mix designs for LCA.

Parameter	Mix 1	Mix 2	Mix 3
FlA (kg/m^3)	0	94	405
GGBFS (kg/m^3)	488	379	0
Fagg. (kg/m^3)	601	601	683
Cagg. (kg/m^3)	1073	1073	1269
Alk./B	0.4734	0.5497	0.4
SS/SH	2	2.5135	1
Molarity (M)	10	10	13
Sp (kg/m^3)	9.79	9.49	0
Additional Water (kg/m^3)	0	0	81
Curing Age (days)	28	28	28
f_c (MPa)	27.313	27.96	28.357

5.1. LCA Assumptions

To streamline and standardize the LCA, a set of key assumptions were made. The system boundary was defined as cradle to gate, excluding curing, use, and end-of-life phases. Environmental coefficients were sourced from ICE v3.0, Ecoinvent 3.8, and [71,72]. The life cycle inventory coefficients (LCI) used in this study are summarized in Table 5.

Table 5. LCI coefficients used for LCA.

Material/Component	EE (MJ/kg)	GWP (kg CO_2-eq/kg)	NO_x (kg/kg)	SO_x (kg/kg)	PM (kg/kg)
FlA	0.12	0.009	0.00001	0.00002	0.000005
GGBFS	1.00	0.07	0.00010	0.00015	0.00003
Fagg.	0.10	0.005	0.000005	0.00001	0.000002
Cagg.	0.10	0.005	0.000005	0.00001	0.000002
SH	20.0	1.50	0.0010	0.0020	0.0005
SS	11.0	1.20	0.0008	0.0015	0.0004
Sp	25.0	2.00	0.0012	0.0025	0.0006
Additional Water	0.01	0.00	0.0000	0.0000	0.0000

All raw materials were assumed to be transported 100 km by diesel truck, while sodium silicate and sodium hydroxide were produced without recycling. The energy for mixing was kept constant at 5 MJ/m^3, with associated emissions of 0.4 kg CO_2-eq/m^3. The functional unit for all calculations was 1 m^3 of concrete, and the S_i and EEi were normalized by f_c to enable a fair comparison across mixes. These assumptions reflect industry conditions and allow for a reasonable and consistent comparison of environmental impacts.

5.2. Mix Compositions and Environmental Performance

The three GPC mixes differ in binder composition, activator content, and water usage, influencing their mechanical and environmental performance. Table 6 represents the calculation details for the sustainability indicators. Table 7 summarizes the sustainability metrics as GWP, EE, Air Emissions, S_i, and EE_i.

- ***Sustainability Indices***:

  $S_i = GWP/f_c$.

  $EE_i = EE/f_c$.

Table 6. Calculation details.

Parameter	**Mix M1 (GGBFS Dominant)**	**Mix M2 (Hybrid or Synergistic)**	**Mix M3 (FlA Dominant)**
Activator Mass	Alk./B = 0.4734, SS/SH = 2 → 231.03 kg/m^3 (154.02 kg/m^3 SS, 77.01 kg/m^3 SH)	Alk./B = 0.5497, SS/SH = 2.5135 → 258.80 kg/m^3 (186.34 kg/m^3 SS, 72.46 kg/m^3 SH)	Alk./B = 0.4, SS/SH = 1 → 162 kg/m^3 (81 kg/m^3 SS, 81 kg/m^3 SH)
GWP (kg CO_2-eq/m^3)	(488 × 0.07) + (601 × 0.005) + (1073 × 0.005) + (154.02 × 1.2) + (77.01 × 1.5) + (9.79 × 2.0) + (0 × 0.0) + [(488 + 601 + 1073 + 231.03 + 9.79 + 0) × 0.01] + 0.4 = **302.64**	**287.05**	**156.00**
EE (MJ/m^3)	(488 × 1.0) + (601 × 0.1) + (1073 × 0.1) + (154.02 × 11) + (77.01 × 20) + (9.79 × 25) + (0 × 0.01) + [(488 + 601 + 1073 + 231.03 + 9.79 + 0) × 0.1] + 5 = **2693.35**	**2347.44**	**1966.56**
NO_x (kg/m^3)	(488 × 0.0001) + (601 × 0.000005) + (1073 × 0.000005) + (154.02 × 0.0008) + (77.01 × 0.001) + (9.79 × 0.0012) + (0 × 0.0) + [(488 + 601 + 1073 + 231.03 + 9.79 + 0) × 0.00002] + 0.00001 = **0.190**	**0.182**	**0.103**
SO_x (kg/m^3)	(488 × 0.00015) + (601 × 0.00001) + (1073 × 0.00001) + (154.02 × 0.0015) + (77.01 × 0.002) + (9.79 × 0.0025) + (0 × 0.0) + [(488 + 601 + 1073 + 231.03 + 9.79 + 0) × 0.00003] + 0.00002 = **0.340**	**0.324**	**0.184**
PM (kg/m^3)	(488×0.00003) + (601×0.000002) + (1073×0.000002) + (154.02×0.0004) + (77.01×0.0005) + (9.79×0.0006) + (0×0.0) + [(488+601+1073+231.03+9.79+0) × 0.00001] + 0.000005 = **0.083**	**0.079**	**0.047**

Table 7. Sustainability metrics for the three GPC mixes.

Parameter	**Mix 1**	**Mix 2**	**Mix 3**
GWP (kg CO_2-eq/m^3)	302.64	287.05	156.00
EE (MJ/m^3)	2693.35	2347.44	1966.56
S_i (kg CO_2/MPa)	11.08	10.27	5.50
EE_i (MJ/MPa)	98.62	83.96	69.36
NO_x Emissions (kg/m^3)	0.19	0.182	0.103
SO_x Emissions (kg/m^3)	0.34	0.324	0.184
PM Emissions (kg/m^3)	0.083	0.079	0.047

5.3. Comparative Sustainability Analysis

Among the three GPC mixes with comparable f_c, Tables 6 and 7 demonstrates that the Mix 1 demonstrated the highest environmental footprint across all evaluated indicators. This elevated impact is primarily attributed to its high GGBFS content (488 kg/m^3) and substantial alkali activator demand (231.03 kg/m^3), particularly with a SS/SH ratio of 2.0. These components are known to be energy-intensive and carbon-intensive, significantly increasing both EE and GWP.

Mix 2, incorporating a moderate quantity of both GGBFS (379 kg/m^3) and FlA (94 kg/m^3), also required a high activator dosage (258.80 kg/m^3, SS/SH = 2.5135). Although its environmental indicators were marginally lower than Mix 1, the overall impacts remained considerably higher than those of Mix 3. The elevated SS/SH ratio and high cumulative alkali content contributed notably to its environmental burden, particularly in terms of EE and air emissions.

In contrast, Mix 3, which exclusively used FlA (405 kg/m^3) as the binder and operated at a significantly lower activator dosage (162 kg/m^3) with an SS/SH ratio of 1.0, exhibited the most favourable environmental performance. It achieved the lowest GWP (156.00 kg CO_2-eq/m^3), EE (1966.56 MJ/m^3), and emissions (NO_x, SO_x, and PM) among all mixes, while maintaining a comparable f_c of 28.36 MPa. The high reactivity of FlA, combined with optimized alkali usage, reduced the environmental burden significantly.

The effectiveness of Mix 3 is further reinforced by the normalized sustainability indices. Its S_i = 5.50 kg CO_2-eq/MPa is approximately 50% lower than that of Mix 1 (S_i = 11.08), and its EE_i = 69.36 MJ/MPa) is around 30%

better than Mix 1 (EE$_i$ = 98.62). These normalized values underscore that, despite identical strength requirements, Mix 3 delivers a more resource-efficient and climate-conscious solution.

The comparative results are presented in Figures 15–17. Figure 15 illustrates the normalized sustainability metrics (S$_i$ and EE$_i$) relative to Mix 3. It is evident that Mix 3 consistently performs best across all criteria, reflecting superior carbon and energy efficiency. Figure 16 depicts the contributions of individual material components to GWP. It clearly shows that Mix 1 exhibits the highest greenhouse gas emissions, largely due to its high GGBFS and activator content, while Mix 3 records the lowest emissions, primarily owing to its FlA-based binder and reduced alkali dosage. Figure 17 presents a bubble chart plotting f_c against the S$_i$. It reveals that Mix 3 achieves the most favourable balance, maintaining strength while substantially lowering carbon intensity. Collectively, these graphical representations reaffirm the environmental and mechanical superiority of the FlA-dominant Mix 3, highlighting its potential as a highly sustainable and technically sound solution for low-carbon GPC applications.

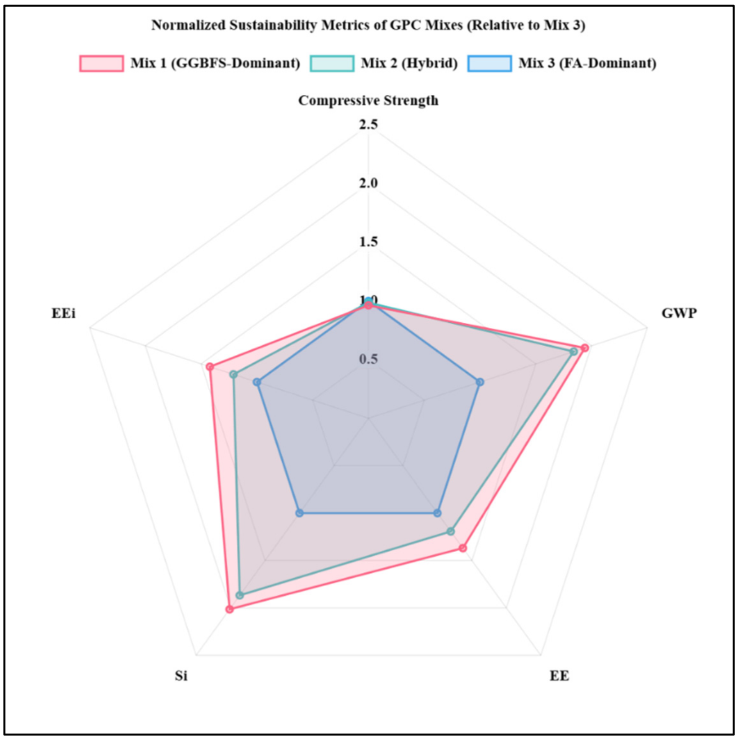


Figure 15. Radar diagram showing the normalized sustainability metrics of GPC mixes (relative to Mix 3).

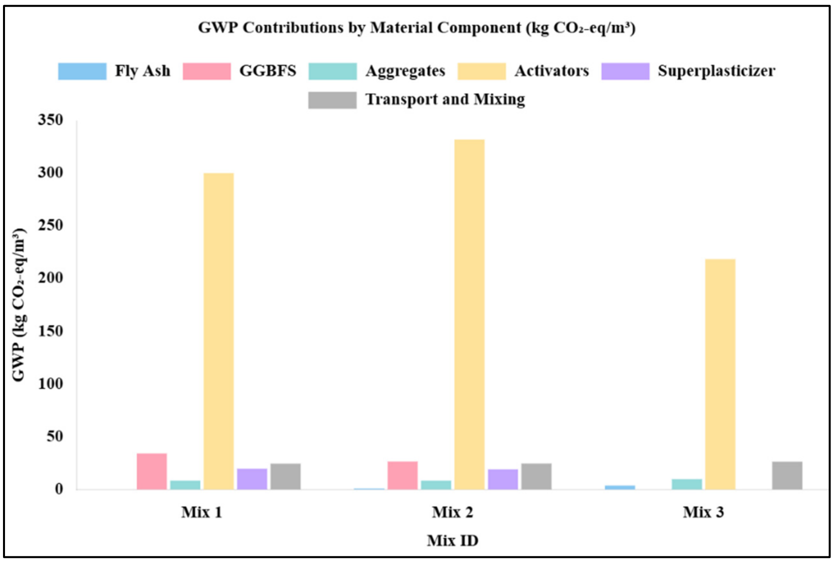


Figure 16. GHG contributions by material components.

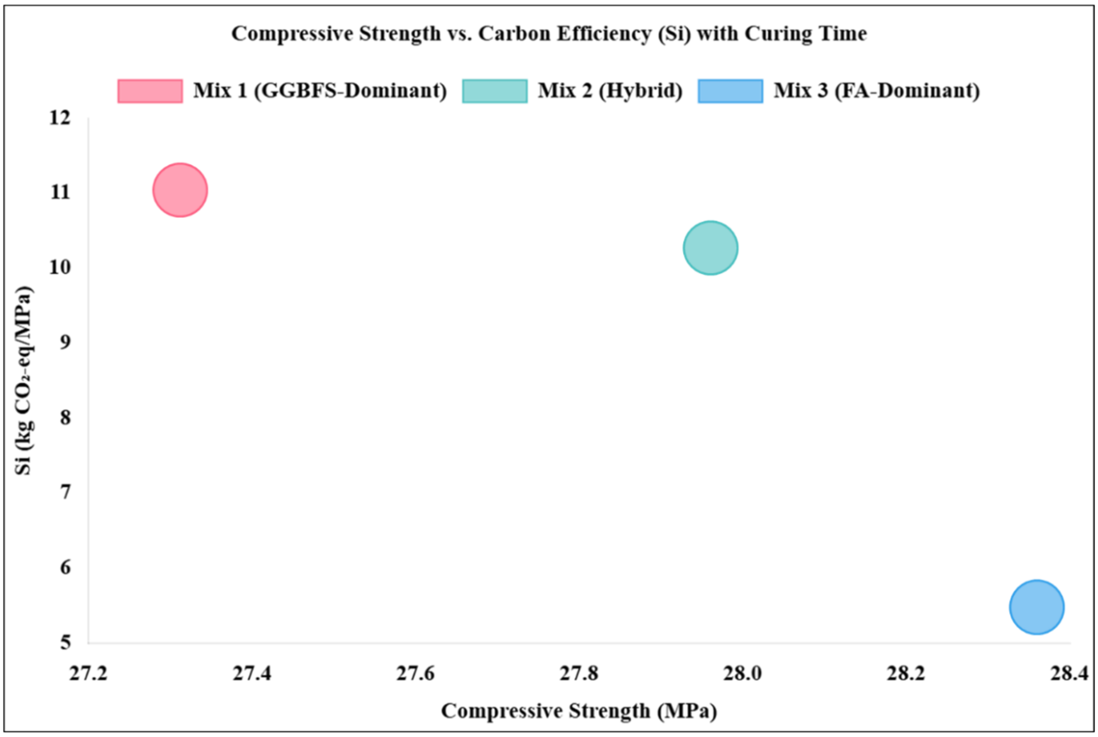


Figure 17. Bubble chart showing the comparison between f_c vs. S_i with curing time.

5.4. Air Emissions (NO_x, SO_x, and PM) Assessment

To further clarify these trends, Table 7, Figure 18 present the air emissions due to the three GPC mixes. Mix 1 emits the highest NO_x (0.190 kg/m^3), SO_x (0.340 kg/m^3) and PM (0.083 kg/m^3), predominantly due to its GGBFS content and large amount of SS and SH. Mix 2 showed moderately reduced emissions. Mix 3 again outperformed both, owing to its minimal activator dosage and reliance on low-impact FlA.

This underscores the opportunity to minimize air emissions by optimizing material selections and mixture proportions, without compromising mechanical properties. Furthermore, this highlights the role of Ecoinvent 3.8 coefficients in identifying key contributors to emissions.

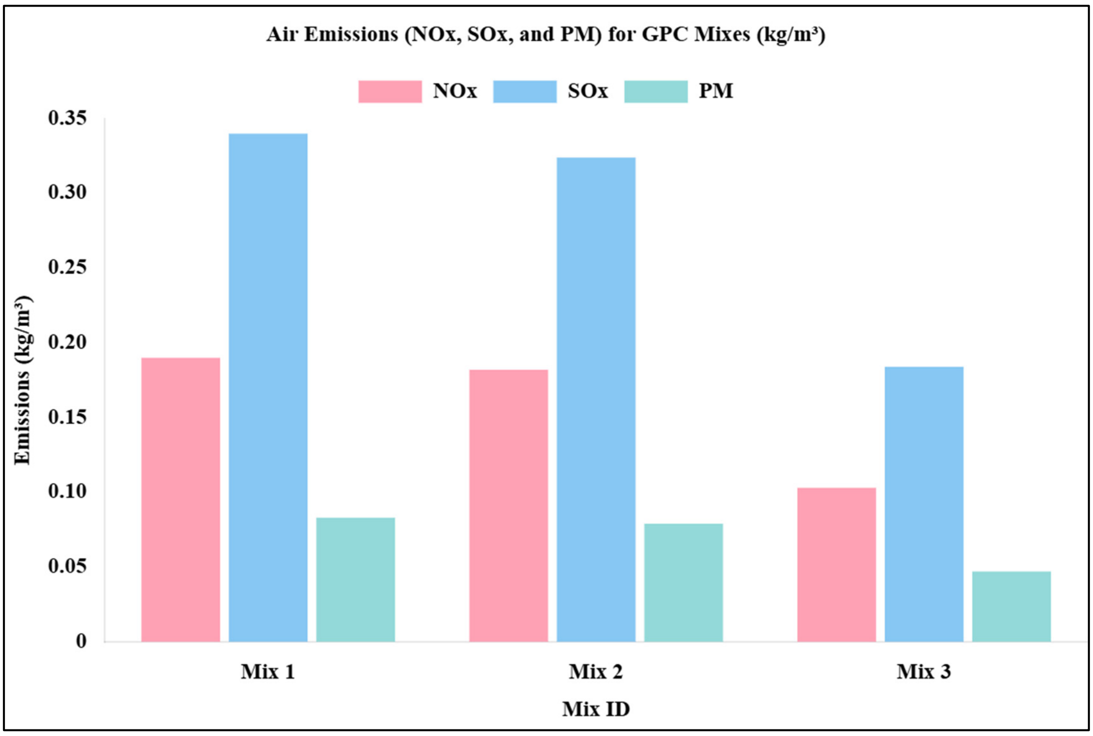


Figure 18. Air emissions (NO_x, SO_x and PM) for the three GPC mixes.

From the sustainability analysis, it is evident that alkali activator content plays a key role in determining the environmental impact of GPC. Mix 1, characterized by a high activator dosage (231.03 kg/m^3) and SS/SH ratio of 2.0, exhibits the highest GWP and air emissions, primarily due to the intensive energy and carbon footprint of SS and SH production. In contrast, Mix 3, with a significantly lower activator mass (162 kg/m^3) and balanced SS/SH ratio of 1.0, demonstrates a 30–45% reduction in environmental indicators such as GWP, EE, NO_x, SO_x, and PM, showcasing the importance of activator optimization. Unlike the hybrid or synergistic approach used in the earlier mixes, the binder in Mix 3 consists solely of FlA, which, despite its slower reactivity compared to GGBFS, performs well due to the optimized activator formulation and curing regime. This mix reflects the principles of resource circularity and emissions minimization, achieving good mechanical strength with minimal environmental burden. Water content also influences both mechanical and sustainability outcomes. Mix 3 includes 81 kg/m^3 of water, which slightly dilutes strength but aids in workability and reaction kinetics, without the need for Sp. This highlights the potential benefit of incorporating water-reducing admixtures in FlA-rich mixes to further enhance performance. Mix 2, despite its balanced binder composition, uses the highest activator mass (258.80 kg/m^3) and also suffers from elevated SS/SH ratios (2.5135), resulting in a relatively high environmental footprint, albeit marginally better than Mix 1.

Overall, Mix 3 emerges as the most environmentally and technically viable option, delivering low GWP (156.00 kg/m^3), moderate energy demand (1966.56 MJ/m^3), and the best normalized indices (S_i = 5.50 kg/MPa, EE_i = 69.36 MJ/MPa). Mix 2 may be improved by lowering SS/SH ratios or employing blended activators, while Mix 1 is only recommended in cases where GGBFS is locally available and environmental regulations are less stringent.

5.5. Carbon Neutrality Assessment

Carbon neutrality, defined as achieving net-zero GWP, was evaluated by estimating the GWP offset achievable through carbon capture, utilization, and storage (CCUS) and material substitution strategies. A 10% GWP reduction was assumed via carbonation, where alkaline binders absorb atmospheric CO_2 during the curing process, as supported by findings from [73]. An additional 5% reduction was considered from CO_2 injection into aggregate systems, based on the benchmark offset reported by [74]. A further 10% reduction was attributed to material substitution, such as partial replacement of GGBFS with FlA or the use of low-carbon activators. Together, these mechanisms amount to a total assumed GWP offset of 25% for all mixes.

Mix 1:

An initial GWP of 302.64 kg CO_2-eq/m^3 was reduced by 75.66 kg/m^3 (25%), yielding a residual GWP of 226.98 kg/m^3. Achieving full carbon neutrality would require offsetting the remaining 226.98 kg/m^3, equivalent to planting approximately 23 trees per cubic meter (assuming 10 kg CO_2/tree/year over 10 years) or employing high-cost CCUS technologies, rendering this mix the least favourable in terms of neutrality feasibility.

Mix 2:

The initial GWP of 287.05 kg CO_2-eq/m^3 was reduced by 71.76 kg/m^3, leaving a residual GWP of 215.29 kg/m^3, which would require around 22 trees/m^3 for complete neutrality. While slightly more efficient than Mix 1, its elevated activator content and SS/SH ratio still pose sustainability challenges.

Mix 3:

With the lowest GWP of 156.00 kg CO_2-eq/m^3, a 25% offset removes 39.00 kg/m^3, resulting in a residual GWP of 117.00 kg/m^3—the lowest among all mixes. This would require the equivalent of only 12 trees/m^3 to reach neutrality. Owing to its reduced reliance on high-impact activators and efficient binder system, Mix 3 emerges as the most promising solution for carbon-neutral applications.

As illustrated in Figure 19, the normalized carbon neutrality metrics, including residual GWP, total GWP offset, and the Carbon Neutrality Index (CNI), highlight the comparative viability of the three GPC mixes, with Mix 3 demonstrating the most feasible pathway toward net-zero emissions.

The Carbon Neutrality Index (CNI = Offset GWP/Initial GWP) is 0.25 for all mixes, as offset strategies were uniformly applied. Higher CNI values (closer to 1) would require advanced CCUS or renewable energy in production, which are currently limited for alkali activators. Mix 3's balance of strength and low residual GWP positions it as the most viable for carbon-neutral concrete.

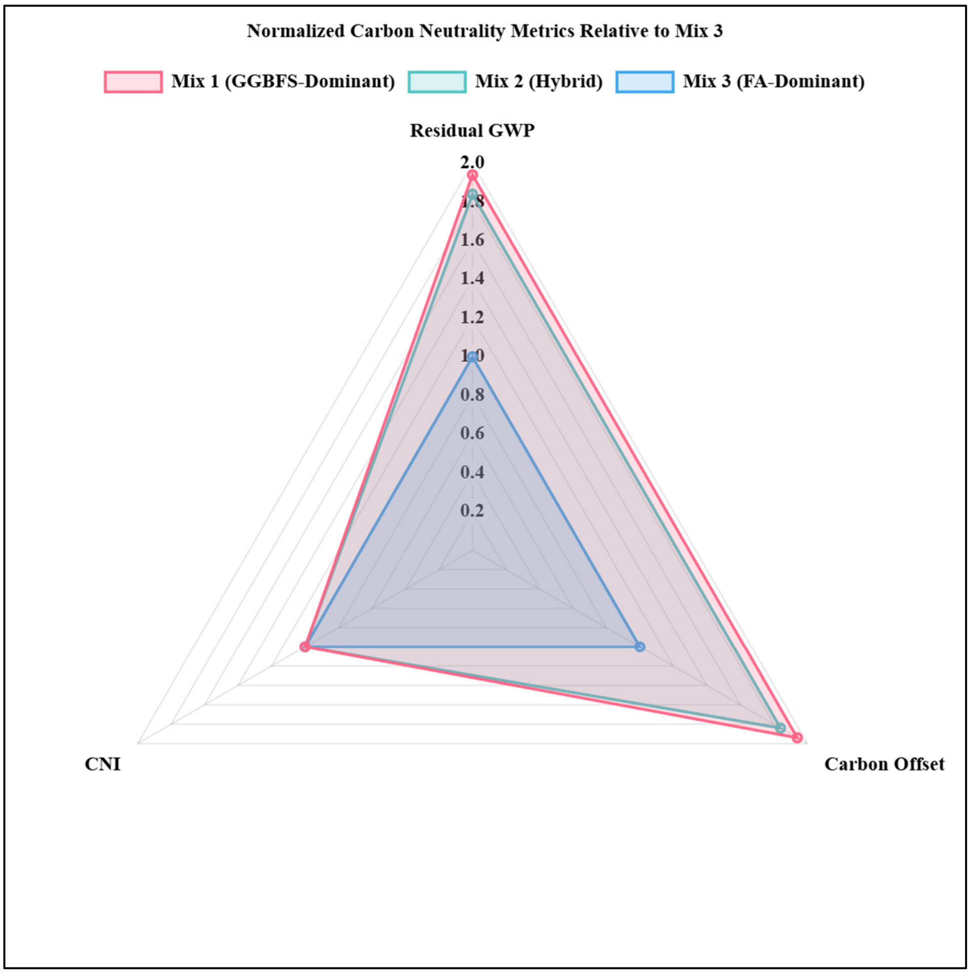


Figure 19. Normalized carbon neutrality metrics relative to Mix 3.

6. Conclusions

This study presents a comprehensive, explainable, and sustainability-driven framework for modelling the compressive strength of SFGPC. A curated dataset of 363 mixes was developed from literature, guided by a scientometric mapping of 441 publications. Five ML models were trained, among which GB demonstrated superior predictive performance, achieving $R^2 = 0.954$, $RMSE = 3.15$ MPa, $MAE = 1.81$ MPa during training and $R^2 = 0.95$, $RMSE = 3.128$ MPa, $MAE = 2.41$ MPa during testing. Multi-layered XAI revealed that curing age, GGBFS content, and alkaline-to-binder ratio are the dominant features governing strength development, with clear nonlinear trends. LCA confirmed that the fly ash–dominant mix achieved the lowest global warming potential (156 kg CO_2-eq/m^3), highest resource efficiency, and smallest post-offset residual GWP. The GUI developed in this work enables practitioners to instantly predict strength using user-provided mix parameters. The novelty of this study lies in its integration of ML, XAI, GUI deployment, and LCA into a single transparent and engineering-ready framework for designing next-generation sustainable geopolymer concretes.

6.1. Key Findings

The study conducts a scientometric analysis to review the literature data, identify the research trends related to GPC, and collect the data systematically for developing ML models.

- Develops five ML models, two CML models as ANN, DT, and three EML models as RF, AdB, and GB, to predict the f_c of SFGPC based on different input features.
- Compares and evaluates the performance of the ML models using metrics, scatter plots, and Taylor diagram. The results indicate that the GB model outperforms the other models in terms of prediction accuracy and agreement with the actual f_c.
- The explanations from the multi-layered XAI techniques revealed that the “A”, “GGBFS”, and “Alk./B” were the most influential features in predicting the f_c of SFGPC. The explanations also showed how the features varied and interacted with each other, and how they affected the prediction outcome.

- 3D Surface plots were made for showing the effect on f_c of SFGPC with variation in Alk./B and the age (A), for two types of binders: FlA and GGBFS independently and the synergistic behaviour too. The main findings are:
    - Higher Alk./B ratios tend to reduce the f_c of SFGPC.
    - Increased curing age significantly enhances strength, particularly in GGBFS-based mixes.
    - GGBFS was found to be a more effective binder than FlA in achieving higher f_c.
- To bridge the gap between model development and real-world application, a user-friendly GUI was developed. A prototype GUI demonstrating the practical deployment of the trained ML model is provided as Supplementary Material. The GUI allows practitioners to input mix parameters and obtain real-time f_c predictions using the trained GB model.
- LCA results clearly indicate that the FlA-dominant GPC mix offers the most favourable balance between mechanical strength and environmental performance. With the lowest GWP, EE, and air emissions, along with a feasible path toward carbon neutrality, Mix 3 (FlA-dominant) demonstrates superior sustainability without compromising f_c. Its optimized use of FlA and reduced alkali activator dosage minimize environmental burdens while delivering good mechanical performance. These findings underscore Mix 3's potential as a green, durable, and resource-efficient alternative for sustainable construction applications.

6.2. Research Limitations

- The dataset, although larger than many prior GPC ML studies, is still dependent on published literature and may not cover the full design space of SFGPC mixtures.
- Certain experimental variables, such as curing humidity, reaction kinetics, and microstructural indices, were unavailable in the literature dataset and therefore were not included as predictive features.
- The compressive strength modelling does not incorporate long-term durability behaviour (e.g., chloride ingress, shrinkage, carbonation).
- LCA results are based on cradle-to-gate boundaries; transportation and end-of-life aspects were beyond the study scope.

6.3. Recommendations for Future Research

- Generate larger experimental datasets through laboratory testing to overcome literature bias and expand the design domain.
- Incorporate durability-related outputs (e.g., permeability, chloride resistance) into future ML-XAI modelling frameworks.
- Integrate uncertainty quantification or Bayesian approaches to capture variability in geopolymer chemistry and curing conditions.
- Extend the LCA from cradle-to-gate to cradle-to-grave, including demolition, recycling, and carbon sequestration potentials.
- Explore hybrid ML–mechanistic modelling to link XAI insights with geopolymer reaction chemistry.

Supplementary Materials

The additional data and information can be downloaded at: https://media.sciltp.com/articles/others/2601201604223206/BCI-25110071-supplement.pdf. Figure S1: Prototype GUI interface for predicting the f_c of sustainable SFGPC. Table S1: ANN architecture details. Table S2: ANN training hyperparameters. Table S3: Optimized DT hyperparameters. Table S4: Optimized RF hyperparameters. Table S5: Optimized AdB hyperparameters. Table S6: Optimized AdB hyperparameters.

Author Contributions

S.S.A.: Writing—original draft, Visualization, Validation, Software, Investigation, Data curation, Conceptualization. M.A.A.: Writing—original draft, Visualization, Software, Data curation, Conceptualization. S.M.I.: Writing—review & editing, Supervision, Project administration, Investigation, Formal analysis, Conceptualization. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The data used in this study were assembled from published literature sources. The compiled dataset and all relevant information required to reproduce the results are provided in the Supplementary Materials. Additional details can be obtained from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

During the preparation of this work, the authors used AI-based tools to assist with language editing. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

1. Ansari, S.S.; Shariq, M.; Mohammad, Z.; et al. Effect of elevated temperature on the structural performance of reinforced high volume fly ash concrete. *Structures* **2023**, *57*, 105168. https://doi.org/10.1016/j.istruc.2023.105168.
2. Zhang, B.; Feng, Y.; Zhou, X.; et al Dynamic mechanical behaviour and life cycle assessment of rubberised solid waste-based geopolymer concrete. *J. Clean. Prod.* **2025**, *501*, 145247. https://doi.org/10.1016/j.jclepro.2025.145247.
3. Guo, S.; Hu, J.; Dai, Q. A critical review on the performance of portland cement concrete with recycled organic components. *J. Clean. Prod.* **2018**, *188*, 92–112. https://doi.org/10.1016/j.jclepro.2018.03.244.
4. de Brito, J.; Kurda, R. The past and future of sustainable concrete: A critical review and new strategies on cement-based materials. *J. Clean. Prod.* **2021**, *281*, 123558. https://doi.org/10.1016/j.jclepro.2020.123558.
5. Ansari, S.S.; Ibrahim, S.M.; Hasan, S.D.; et al. Effect of perforated fly ash cenospheres and nano-silica on the thermo-structural behaviour of sustainable concrete. *Structures* **2025**, *80*, 109800. https://doi.org/10.1016/j.istruc.2025.109800.
6. Juenger, M.C.G.; Winnefeld, F.; Provis, J.L.; et al. Advances in alternative cementitious binders. *Cem. Concr. Res.* **2011**, *41*, 1232–1243. https://doi.org/10.1016/j.cemconres.2010.11.012.
7. Degefu, D.M.; Liao, Z.; Berardi, U.; et al. Geopolymer concrete for net-zero buildings: Correlating paste chemistry with monolith hygrothermal performance. *Resour. Conserv. Recycl.* **2023**, *189*, 106743. https://doi.org/10.1016/j.resconrec.2022.106743.
8. Wang, D.; Noguchi, T.; Nanao, M.; et al. Parametric investigation and scale testing on accelerated CO_2 sequestration of cement-based materials by utilizing industrial waste heat. *J. Clean. Prod.* **2024**, *475*, 143716. https://doi.org/10.1016/j.jclepro.2024.143716.
9. Dong, W.; Li, W.; Tao, Z. A comprehensive review on performance of cementitious and geopolymeric concretes with recycled waste glass as powder, sand or cullet. *Resour. Conserv. Recycl.* **2021**, *172*, 105664. https://doi.org/10.1016/j.resconrec.2021.105664.
10. Martínez, A.; Miller, S.A. A review of drivers for implementing geopolymers in construction: Codes and constructability. *Resour. Conserv. Recycl.* **2023**, *199*, 107238. https://doi.org/10.1016/j.resconrec.2023.107238.
11. Wang, Y.; Liu, H.; Nie, Z.; et al. Experimental investigation into the influence of calcium aluminate cement on the micro- and macro-mechanical properties of the interfacial transition zone in geopolymer concrete. *Dev. Built Environ.* **2025**, *21*, 100618. https://doi.org/10.1016/j.dibe.2025.100618.
12. Umer, M.; Ahmad, J.; Mukhtar, H. Innovative valorization of biomass waste-derived sodium silicate for geopolymer concrete synthesis: Sustainability assessment and circular economy potential. *J. Clean. Prod.* **2024**, *452*, 142181. https://doi.org/10.1016/j.jclepro.2024.142181.
13. Wang, M.-R.; Jia, D.-C.; He, P.-G.; et al. Microstructural and mechanical characterization of fly ash cenosphere/metakaolin-based geopolymeric composites. *Ceram. Int.* **2011**, *37*, 1661–1666. https://doi.org/10.1016/j.ceramint.2011.02.010.
14. Zhang, P.; Gao, Z.; Wang, J.; et al. Properties of fresh and hardened fly ash/slag based geopolymer concrete: A review. *J. Clean. Prod.* **2020**, *270*, 122389. https://doi.org/10.1016/j.jclepro.2020.122389.

15. Valente, M.; Sambucci, M.; Chougan, M.; et al. Reducing the emission of climate-altering substances in cementitious materials: A comparison between alkali-activated materials and Portland cement-based composites incorporating recycled tire rubber. *J. Clean. Prod.* **2022**, *333*, 130013. https://doi.org/10.1016/j.jclepro.2021.130013.
16. Silewu, K.; Kahanji, C.; Simwanda, L.; et al. Intelligent Data Driven Ensemble Approaches for Bending Strength Prediction of Ultra-High Performance Concrete Beams. *Bull. Comput. Intell.* **2025**, *1*, 31–52. https://doi.org/10.53941/bci.2025.100003.
17. Benzaamia, A.; Ghrici, M.; Rbouh, R.; et al. Prediction of Chloride Resistance Level in Concrete Using Optimized Tree-Based Machine Learning Models. *Bull. Comput. Intell.* **2025**, *1*, 104–117. https://doi.org/10.53941/bci.2025.100007.
18. Asteris, P.G.; Skentou, A.D.; Bardhan, A.; et al. Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models. *Cem. Concr. Res.* **2021**, *145*, 106449. https://doi.org/10.1016/j.cemconres.2021.106449.
19. Le, T.-T.; Skentou, A.D.; Mamou, A.; et al. Correlating the Unconfined Compressive Strength of Rock with the Compressional Wave Velocity Effective Porosity and Schmidt Hammer Rebound Number Using Artificial Neural Networks. *Rock. Mech. Rock. Eng.* **2022**, *55*, 6805–6840. https://doi.org/10.1007/s00603-022-02992-8.
20. Asteris, P.G.; Lemonis, M.E.; Nguyen, T.-A.; et al. Soft computing-based estimation of ultimate axial load of rectangular concrete-filled steel tubes. *Steel Compos. Struct.* **2021**, *39*, 471–491.
21. Mahmood, W.; Mohammed, A.S.; Asteris, P.G.; et al. Modeling Flexural and Compressive Strengths Behaviour of Cement-Grouted Sands Modified with Water Reducer Polymer. *Appl. Sci.* **2022**, *12*, 1016. https://doi.org/10.3390/app12031016.
22. Shamim Ansari, S.; Muhammad Ibrahim, S.; Danish Hasan, S. Conventional and Ensemble Machine Learning Models to Predict the Compressive Strength of Fly Ash Based Geopolymer Concrete. *Mater. Today Proc.* **2023**, *in press*. https://doi.org/10.1016/j.matpr.2023.04.393.
23. Ansari, S.S.; Ibrahim, S.M.; Hasan, S.D. Interpretable Machine-Learning Models to Predict the Flexural Strength of Fiber-Reinforced SCM-Blended Concrete Composites. *J. Struct. Des. Constr. Pract.* **2025**, *30*. https://doi.org/10.1061/JSDCCC.SCENG-1496.
24. Baqer, B.T.; Mohammed, A.S. Evaluating the compressive strength of fly ash-slag-based geopolymer concrete: Impact of hydraulic, silica, alumina, and lime moduli, and sodium silicate using various predictive models. *Innov. Infrastruct. Solut.* **2025**, *10*, 191. https://doi.org/10.1007/s41062-025-01973-4.
25. Paruthi, S.; Khan, A.H.; Isleem, H.F.; et al. Influence of silica fume and alccofine on the mechanical performance of GGBS-based geopolymer concrete under varying curing temperatures. *J. Struct. Integrity Maint.* **2025**, *10*, 2447661. https://doi.org/10.1080/24705314.2024.2447661.
26. Ahmed, H.U.; Mostafa, R.R.; Mohammed, A.; et al. Support vector regression (SVR) and grey wolf optimization (GWO) to predict the compressive strength of GGBFS-based geopolymer concrete. *Neural Comput. Appl.* **2023**, *35*, 2909–2926. https://doi.org/10.1007/s00521-022-07724-1.
27. Ibrahim, S.M.; Ansari, S.S.; Hasan, S.D. Towards white box modeling of compressive strength of sustainable ternary cement concrete using explainable artificial intelligence (XAI). *Appl. Soft Comput.* **2023**, *149*, 110997. https://doi.org/10.1016/j.asoc.2023.110997.
28. Ansari, S.S.; Ibrahim, S.M.; Hasan, S.D.; et al. Experiments and predictive modelling on sustainable cementitious mortar with perforated fly ash cenospheres and effectively dispersed nano silica. *Mater. Today Commun.* **2025**, *47*, 113020. https://doi.org/10.1016/j.mtcomm.2025.113020.
29. Bianco, I.; Ap Dafydd Tomos, B.; Vinai, R. Analysis of the environmental impacts of alkali-activated concrete produced with waste glass-derived silicate activator—A LCA study. *J. Clean. Prod.* **2021**, *316*, 128383. https://doi.org/10.1016/j.jclepro.2021.128383.
30. Ma, X.; Hu, H.; Luo, Y.; et al. A carbon footprint assessment for usage of recycled aggregate and supplementary cementitious materials for sustainable concrete: A life-cycle perspective in China. *J. Clean. Prod.* **2025**, *490*, 144772. https://doi.org/10.1016/j.jclepro.2025.144772.
31. Martínez, A.; Miller, S.A. Life cycle assessment and production cost of geopolymer concrete: A meta-analysis. *Resour. Conserv. Recycl.* **2025**, *215*, 108018. https://doi.org/10.1016/j.resconrec.2024.108018.
32. Nguyen, M.H.; Mai, H.-V.T.; Trinh, S.H.; et al. A comparative assessment of tree-based predictive models to estimate geopolymer concrete compressive strength. *Neural Comput. Appl.* **2023**, *35*, 6569–6588. https://doi.org/10.1007/s00521-022-08042-2.
33. Ahmed, H.U.; Mohammed, A.A.; Mohammed, A. Soft computing models to predict the compressive strength of GGBS/FA- geopolymer concrete. *PLoS ONE* **2022**, *17*, e0265846. https://doi.org/10.1371/journal.pone.0265846.
34. Kina, C.; Tanyildizi, H.; Turk, K. Forecasting the compressive strength of GGBFS-based geopolymer concrete via ensemble predictive models. *Constr. Build. Mater.* **2023**, *405*, 133299. https://doi.org/10.1016/j.conbuildmat.2023.133299.
35. Shahmansouri, A.A.; Yazdani, M.; Ghanbari, S.; et al. Artificial neural network model to predict the compressive strength of eco-friendly geopolymer concrete incorporating silica fume and natural zeolite. *J. Clean. Prod.* **2021**, *279*, 123697. https://doi.org/10.1016/j.jclepro.2020.123697.

36. Jin, J.; Wu, T.; Zhang, Z.; et al. Application of optimization-based estimation analysis for predicting the compressive strength of eco-friendly modified geopolymer concrete. *Struct. Concr.* **2022**, *25*, 1489. https://doi.org/10.1002/suco.202201051.
37. Onyelowe, K.C.; Ebid, A.M.; Awoyera, P.; et al. Prediction and validation of mechanical properties of self-compacting geopolymer concrete using combined machine learning methods a comparative and suitability assessment of the best analysis. *Sci. Rep.* **2025**, *15*, 6361. https://doi.org/10.1038/s41598-025-90468-4.
38. Bypour, M.; Yekrangnia, M.; Kioumarsi, M. Machine learning-driven optimization for predicting compressive strength in fly ash geopolymer concrete. *Clean. Eng. Technol.* **2025**, *25*, 100899. https://doi.org/10.1016/j.clet.2025.100899.
39. Diksha, S.; Dev, N.; Goyal, P.K. Utilizing an enhanced machine learning approach for geopolymer concrete analysis. *Nondestruct. Test. Eval.* **2025**, *40*, 904–931. https://doi.org/10.1080/10589759.2024.2334434.
40. Yang, H.; Li, H.; Jiang, J. Predictive modeling of compressive strength of geopolymer concrete before and after high temperature applying machine learning algorithms. *Struct. Concr.* **2025**, *26*, 1699–1732. https://doi.org/10.1002/suco.202400552.
41. Chhetri Sapkota, S.; Dahal, D.; Yadav, A.; et al. Analyzing the Behavior of Geopolymer Concrete with Different Novel Machine-Learning Algorithms. *J. Struct. Des. Constr. Pract.* **2025**, *30*. https://doi.org/10.1061/JSDCCC.SCENG-1724.
42. Wang, X.; Zhang, H.; Yin, L.; et al. Frost resistance of steel fiber geopolymer concrete studied via the machine learning method. *Structures* **2025**, *73*, 108444. https://doi.org/10.1016/j.istruc.2025.108444.
43. Ulloa, N.; Morales León, M.A.; Silva Palmay, L.F.; et al. Evaluating the compressive strength of industrial wastes-based geopolymer concrete with machine learning models. *Constr. Build. Mater.* **2025**, *472*, 140891. https://doi.org/10.1016/j.conbuildmat.2025.140891.
44. Nguyen, H.A.T.; Pham, D.H.; Le, A.T.; et al. Transfer learning framework for modelling the compressive strength of ultra-high performance geopolymer concrete. *Constr. Build. Mater.* **2025**, *459*, 139746. https://doi.org/10.1016/j.conbuildmat.2024.139746.
45. Hassan, A.; Saleh, R.A.A.; Al-Sameai, H.; et al. Novel hybrid machine learning framework for high-fidelity prediction of fly ash-based geopolymer concrete strength. *Compos. Struct.* **2026**, *378*, 119906. https://doi.org/10.1016/j.compstruct.2025.119906.
46. Abdellatief, M.; Hassan, Y.M.; Elnabwy, M.T.; et al. Investigation of machine learning models in predicting compressive strength for ultra-high-performance geopolymer concrete: A comparative study. *Constr. Build. Mater.* **2024**, *436*, 136884. https://doi.org/10.1016/j.conbuildmat.2024.136884.
47. Ling, Y.; Wang, K.; Wang, X.; et al. Prediction of engineering properties of fly ash-based geopolymer using artificial neural networks. *Neural Comput. Appl.* **2021**, *33*, 85–105. https://doi.org/10.1007/s00521-019-04662-3.
48. Toufigh, V.; Jafari, A. Developing a comprehensive prediction model for compressive strength of fly ash-based geopolymer concrete (FAGC). *Constr. Build. Mater.* **2021**, *277*, 122241. https://doi.org/10.1016/j.conbuildmat.2021.122241.
49. Ravikumar, S.; Agrahari, A.; Singh, S.N. Mapping the intellectual structure of scientometrics: A co-word analysis of the journal Scientometrics (2005–2010). *Scientometrics* **2015**, *102*, 929–955. https://doi.org/10.1007/s11192-014-1402-8.
50. Aria, M.; Cuccurullo, C. bibliometrix: An R-tool for comprehensive science mapping analysis. *J. Inf. Informetr.* **2017**, *11*, 959–975. https://doi.org/10.1016/j.joi.2017.08.007.
51. Mustakim, S.M.; Das, S.K.; Mishra, J.; et al. Improvement in Fresh, Mechanical and Microstructural Properties of Fly Ash- Blast Furnace Slag Based Geopolymer Concrete by Addition of Nano and Micro Silica. *Silicon* **2021**, *13*, 2415–2428. https://doi.org/10.1007/s12633-020-00593-0.
52. Bellum, R.R.; Nerella, R.; Madduru, S.R.C.; et al. Mix Design and Mechanical Properties of Fly Ash and GGBFS-Synthesized Alkali-Activated Concrete (AAC). *Infrastructures* **2019**, *4*, 20. https://doi.org/10.3390/infrastructures4020020.
53. Suresh, G.V.; Karthikeyan, J. Performance enhancement of green concrete. *Proc. Inst. Civ. Eng. Eng. Sustain.* **2018**, *171*, 191–203. https://doi.org/10.1680/jensu.15.00066.
54. Pasupathy, K.; Berndt, M.; Castel, A.; et al. Carbonation of a blended slag-fly ash geopolymer concrete in field conditions after 8 years. *Constr. Build. Mater.* **2016**, *125*, 661–669. https://doi.org/10.1016/j.conbuildmat.2016.08.078.
55. Deb, P.S.; Nath, P.; Sarker, P.K. Drying Shrinkage of Slag Blended Fly Ash Geopolymer Concrete Cured at Room Temperature. *Procedia Eng.* **2015**, *125*, 594–600. https://doi.org/10.1016/j.proeng.2015.11.066.
56. Khan, M.S.H.; Castel, A.; Akbarnezhad, A.; et al. Utilisation of steel furnace slag coarse aggregate in a low calcium fly ash geopolymer concrete. *Cem. Concr. Res.* **2016**, *89*, 220–229. https://doi.org/10.1016/j.cemconres.2016.09.001.
57. Bellum, R.R.; Muniraj, K.; Madduru, S.R.C. Exploration of mechanical and durability characteristics of fly ash-GGBFS based green geopolymer concrete. *Discov. Appl. Sci.* **2020**, *2*, 919. https://doi.org/10.1007/s42452-020-2720-5.
58. Gopalakrishnan, R.; Chinnaraju, K. Durability Characteristics of Slag/Fly Ash Based Alumina Silicate Concrete at Ambient Temperature. *Silicon* **2021**, *13*, 3619–3628. https://doi.org/10.1007/s12633-020-00889-1.
59. Bellum, R.R.; Muniraj, K.; Indukuri, C.S.R.; et al. Investigation on Performance Enhancement of Fly ash-GGBFS Based Graphene Geopolymer Concrete. *J. Build. Eng.* **2020**, *32*, 101659. https://doi.org/10.1016/j.jobe.2020.101659.
60. Mathew, G.; Issac, B.M. Effect of molarity of sodium hydroxide on the aluminosilicate content in laterite aggregate of laterised geopolymer concrete. *J. Build. Eng.* **2020**, *32*, 101486. https://doi.org/10.1016/j.jobe.2020.101486.

61. Bellum, R.R.; Muniraj, K.; Madduru, S.R.C. Characteristic Evaluation of Geopolymer Concrete for the Development of Road Network: Sustainable Infrastructure. *Innov. Infrastruct. Solut.* **2020**, *5*, 91. https://doi.org/10.1007/s41062-020-00344-5.
62. Abbass, M.; Singh, D.; Singh, G. Properties of hybrid geopolymer concrete prepared using rice husk ash, fly ash and GGBS with coconut fiber. *Mater. Today Proc.* **2021**, *45*, 4964–4970. https://doi.org/10.1016/j.matpr.2021.01.390.
63. Abdul Sani, M.F.A.; Muhamad, R.; Mo, K.H. Effect of Ground Granulated Blast Furnace Slag as Partial Replacement in Fly Ash-Based Geopolymer Concrete. In Proceedings of the 3rd National Conference on Wind & Earthquake Engineering and International Seminar On Sustainable Construction Engineering, Kuala Lumpur, Malaysia, 12–13 July 2019.
64. Verma, M.; Dev, N. Effect of ground granulated blast furnace slag and fly ash ratio and the curing conditions on the mechanical properties of geopolymer concrete. *Struct. Concr.* **2022**, *23*, 2015–2029. https://doi.org/10.1002/suco.202000536.
65. Das, S.K.; Shrivastava, S. Siliceous fly ash and blast furnace slag based geopolymer concrete under ambient temperature curing condition. *Struct. Concr.* **2021**, *22*, E341–E351. https://doi.org/10.1002/suco.201900201.
66. Prusty, J.K.; Pradhan, B. Multi-response optimization using Taguchi-Grey relational analysis for composition of fly ash-ground granulated blast furnace slag based geopolymer concrete. *Constr. Build. Mater.* **2020**, *241*, 118049. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2020.118049.
67. Toniolo, N.; Boccaccini, A.R. Fly ash-based geopolymers containing added silicate waste. A review. *Ceram. Int.* **2017**, *43*, 14545–14551. https://doi.org/10.1016/j.ceramint.2017.07.221.
68. Dave, S.V.; Bhogayata, A. The strength oriented mix design for geopolymer concrete using Taguchi method and Indian concrete mix design code. *Constr. Build. Mater.* **2020**, *262*, 120853. https://doi.org/10.1016/j.conbuildmat.2020.120853.
69. Hajimohammadi, A.; Ngo, T.; Vongsvivut, J. Interfacial chemistry of a fly ash geopolymer and aggregates. *J. Clean. Prod.* **2019**, *231*, 980–989. https://doi.org/10.1016/j.jclepro.2019.05.249.
70. Phoo-ngernkham, T.; Maegawa, A.; Mishima, N.; et al. Effects of sodium hydroxide and sodium silicate solutions on compressive and shear bond strengths of FA–GBFS geopolymer. *Constr. Build. Mater.* **2015**, *91*, 1–8. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2015.05.001.
71. Weil, M.; Dombrowski, K.; Buchwald, A. Life-cycle analysis of geopolymers. In *Geopolymers*; Elsevier: Amsterdam, The Netherlands, 2009; pp. 194–210. https://doi.org/10.1533/9781845696382.2.194.
72. Weil, M.; Buchwald, A.; Dombrowski-Daube, K. *How to Assess the Environmental Sustainability of Geopolymers? A Live Cycle Perspective*; Trans Tech Publications Ltd.: Baech, Switzerland, 2010; pp. 186–191. https://doi.org/10.4028/www.scientific.net/AST.69.186.
73. Shi, C.; Jiménez, A.F.; Palomo, A. New cements for the 21st century: The pursuit of an alternative to Portland cement. *Cem. Concr. Res.* **2011**, *41*, 750–763. https://doi.org/10.1016/j.cemconres.2011.03.016.
74. Shivaprasad, K.N.; Yang, H.-M.; Singh, J.K. A path to carbon neutrality in construction: An overview of recent progress in recycled cement usage. *J. CO2 Util.* **2024**, *83*, 102816. https://doi.org/10.1016/j.jcou.2024.102816.