Green Technologies and Sustainability 4 (2026) 100275

Contents lists available at ScienceDirect

Green Technologies and Sustainability

journal homepage:
https://www.keaipublishing.com/en/journals/green-technologies-and-sustainability/

Full-length article

Prediction of flexural and split tensile strength of waste glass-concrete composite using machine learning algorithms

Derrick Mirindi [a,b,*], David Sinkhonde [c], Tajebe Bezabih [c], Frederic Mirindi [e], Oluwakemi Oshineye [a], Patrice Mirindi [d]

[a] School of Architecture and Planning, Morgan State University, Baltimore, United States of America
[b] School of Engineering and Applied Science, University of Pennsylvania, PA, United States of America
[c] Department of Civil Engineering, Pan Africa University Institute for Basic Sciences and Technology, Nairobi, Kenya
[d] Department of Agricultural Economics, Extension and Rural Development, University of Pretoria, Pretoria, South Africa
[e] Department of Economics, University of Manitoba, Winnipeg, Canada

HIGHLIGHTS

- Machine learning models predict the mechanical properties of concrete-glass composite.
- Characteristics of glass.
- Mechanical properties of concrete-glass composite.
- Methodological innovation for robust machine learning models to optimize materials for sustainable construction.

ARTICLE INFO

Keywords:
Waste glass concrete
Machine learning
AdaBoost
XGBoost
LightGBM
Gaussian process

ABSTRACT

Waste material, including glass, presents significant environmental challenges due to its non-biodegradable nature and low global recycling rates. Incorporating waste glass into concrete offers a sustainable solution, but predicting its effects on mechanical properties, particularly flexural (f_b) and split tensile (f_t) strengths, remains complex. This study utilizes machine learning (ML) algorithms (decision tree (DT), extreme gradient boosting (XGBoost), adaptive boosting (AdaBoost), light gradient boosting machine (LightGBM), support vector regression (SVR), and gaussian process (GP)) to predict f_b and f_t strengths based on compressive strength (f_c), concrete age, and glass replacement percentage of glass-concrete composites. Thirteen experimental studies were utilized using secondary data. Results demonstrate that Pearson correlation analysis reveals strong interdependence among mechanical properties (f_c-f_b: 0.809-0.876, f_c-f_t: 0.927-0.948, f_b-f_t: 0.943-0.970), with negligible influence of glass type and moderate positive impact of replacement percentage. The ML algorithms each offer unique predictive strengths—most notably, XGBoost training model achieves near-perfect accuracy (with R^2 equal to 0.9991). However, k-fold cross-validation revealed overfitting concerns limiting applicability to conventional concrete compositions. Non-parametric analyses reveal moderate f_c-f_b correlations (Spearman's ρ = 0.5879, p=0.0739) and statistically significant f_c-f_t relationships (ρ = 0.6364, p=0.0479), while ML models achieve high predictive accuracy by exploiting multi-feature interactions beyond simple pairwise correlations. These ML models enable optimized mix designs, advancing sustainable construction through efficient waste glass utilization as a partial aggregate replacement.

1. Introduction

The management of waste materials has become a critical environmental challenge in recent years [1–4]. This issue is driven by several factors, including the growing scarcity of landfill space, the depletion of raw materials, and the increasingly urgent need to reduce the carbon footprint of the construction industry [5]. Among these waste materials, glass waste stands out because of its non-biodegradable nature and extremely slow decomposition. It can take millennia to break down. The large volumes of glass discarded each year further intensify the problem. The production and use of glass have a rich history that dates back to ancient civilizations [6]. Glass has traditionally been valued for its optical properties and durability. Today, it has become

* Corresponding author at: School of Architecture and Planning, Morgan State University, Baltimore, United States of America.
E-mail address: demir1@morgan.edu (D. Mirindi).

https://doi.org/10.1016/j.grets.2025.100275
Received 6 July 2025; Received in revised form 4 September 2025; Accepted 14 September 2025
Available online 17 September 2025

ubiquitous in modern society. Glass is widely used in packaging, architecture, automotive components, high-tech devices, and renewable energy systems [7].

This application diversification has come about mainly due to innovations in processing and manufacturing techniques. Specialized varieties of glass have been developed with tailored properties. For example, sodium-calcium, lead, barium, aluminosilicate, borosilicate, and luminescent glasses all meet specific industrial and architectural requirements [8]. Nonetheless, this production growth has a significant downside. Large amounts of waste glass are generated during production, use, and disposal. This waste adds to the growing problem of solid waste management.

The manufacturing process for glass, from raw material extraction to final product formation, is inherently energy-intensive and emits a substantial amount of carbon dioxide. Furthermore, waste glass disposal poses a serious environmental challenge [9]. This applies whether the waste comes from industrial production or from post-consumption sources. The non-biodegradable nature of this material means it accumulates in landfills, consuming valuable space and adding to the growing waste crisis [10]. Recent data shows the scale of this problem. In the United States alone, an estimated 10.4 million metric tons of waste glass are generated annually. Of this amount, nearly 60% is disposed of in landfills, while only 26% is recovered for recycling. The rest was routed to combustion facilities [11]. Furthermore, global annual waste glass production is approximately 100 million tons, yet the recycling rate remains alarmingly low at around 26%. China, the USA, and the European Union collectively account for nearly two-thirds of this total, contributing about 33, 32, and 20 Mt, respectively [12]. This low rate reflects both the production scale and the persistent challenge of effectively recycling this material.

Meanwhile, global glass production is expected to grow, reflecting industrialization and population growth. The packaging industry, for instance, dominates this expansion, consuming nearly 60% of total production, while the architectural and automotive sectors collectively account for about 30% [13]. Specialized and renewable sectors, which employ advanced glasses for solar panels, smartphones, optical fibers, and protective screens, form the rest [14]. Naturally, this upward trend signals a growing need for innovative solutions to reduce waste and find valuable applications for discarded glass. One promising approach to addressing this challenge lies in the reuse of waste glass in the production of cement and concrete [15]. Specifically, waste glass can be ground into small particles and incorporated into a cement matrix as a partial replacement for sand or gravel [16]. The addition of waste glass reduces the consumption of virgin resources and energy. It can also enhance specific mechanical properties of the composite, yielding a more durable and environmentally friendly form of concrete.

Some specialized waste glasses, for example, liquid crystal display (LCD) glass, waste window glass, and cathode ray tube (CRT) glass, possess unique physical, chemical, and pozzolanic properties [17]. These properties can affect their behavior when incorporated into cement and concrete [15]. The eventual performance of the composite is a complex phenomenon governed by numerous factors, including the type of waste glass, its particle size, its dosage, and its interaction with the cement hydration products [18]. Furthermore, these properties can affect not only the compressive strength (f_c) but also the flexural strength (f_b) and split tensile strength (f_t), which are crucial for understanding a material's ability to resist bending and tension and, therefore, for its structural applications [19].

To maximize the benefits of adding waste glass to cement and to control its effects on mechanical properties, a deep understanding of these relationships is required. Traditionally, this understanding has been developed through extensive trial-and-error methods [20]. This approach involves preparing numerous batches with different proportions of waste glass, testing their mechanical properties, and attempting to draw empirical relationships from the results. However, this approach is time-intensive, expensive, and often unreliable, particularly when attempting to extrapolate results to a wider range of conditions [21].

Given these challenges and the limitations of conventional trial-and-error approaches, more efficient methods are needed. This is where machine learning (ML) algorithms become particularly powerful [22–25]. ML algorithms, a branch of artificial intelligence (AI), can capture complex and non-obvious relationships within multivariate data. This ability allows for accurate predictions of material properties, such as f_c, f_b, and f_t [26]. The predictions are made using a set of easily measurable variables. Furthermore, using conventional statistical techniques, ML methods can quantify the relative influence of different factors and uncover patterns that may be hard to observe. Recent investigations have demonstrated the potential of ML to successfully predict the mechanical properties of concrete with reasonable accuracy [27]. Nevertheless, these studies predominantly focus on f_c, underexploring the ability to predict other key properties, such as f_b and f_t, based on the f_c. These properties govern a material's resistance to bending and tension, influencing its structural performance in many applications.

Despite extensive research into glass-concrete composites, a significant gap exists in comparative analyses of mechanical properties among varieties of waste glasses in concrete. With the rapid advance of AI, the current literature lacks comprehensive studies that evaluate and compare the cross mechanical properties prediction of these materials using advanced ML and DL algorithms. Importantly, no comprehensive study to date has addressed these issues by directly comparing multiple ML algorithms [23]. Therefore, this study aims to conduct a comprehensive review of the mechanical properties (compressive strength (f_c), flexural strength (f_b), and split tensile (f_t)) of various glass-concrete composites by predicting the f_b and f_t strengths based on the percentage of waste glass as a partial replacement of aggregate, the age of the concrete, and f_c. strength. These algorithms include decision tree (DT), extreme gradient boosting (XGBoost), adaptive boosting (AdaBoost), light gradient boosting machine (LightGBM), support vector regression (SVR), and Gaussian process (GP). By predicting the mechanical properties of concrete made with glasses, this study will revolutionize the understanding and application of glass in concrete technology, potentially leading to more sustainable, cost-effective, time-efficient, and high-performance concrete for structural applications.

2. Data and methods

Our comprehensive data collection strategy encompassed an approach that gathers relevant research on glasses in concrete applications across diverse scholarly repositories. Beyond Google Scholar and ResearchGate, we expanded our search to include Web of Science, Scopus, ScienceDirect, IEEE Xplore, ProQuest, JSTOR, Microsoft Academic, and Semantic Scholar to ensure comprehensive coverage of systematic literature. Based on our title, we employed a hierarchical search strategy with increasing specificity across three levels. In Search Level 1, we utilized the search query "glass" AND "concrete"; Level 2 narrowed the focus with "glass" AND "concrete" AND "mechanical properties", while Level 3 further refined our search with "glass" AND "concrete" AND "mechanical properties" AND "machine learning".

For instance, Fig. 1 presents a citation map by country for Search Level 1 search results from 2016 to 2025, extracted from DimensionsAI and visualized using VOSviewer. The data reveal a precise global distribution of research activity and impact in the field of glass-enhanced concrete and related topics. Furthermore, Fig. 2 illustrates the network of countries based on the number of documents published. China leads by a significant margin, with 724 documents and 7189 citations, indicating both a high volume of research output and substantial influence within the academic community. The United States, while producing fewer documents (207), demonstrates a high citation count (4881), suggesting that its publications are highly impactful. Other notable contributors include Saudi Arabia (197 documents, 3638 citations),

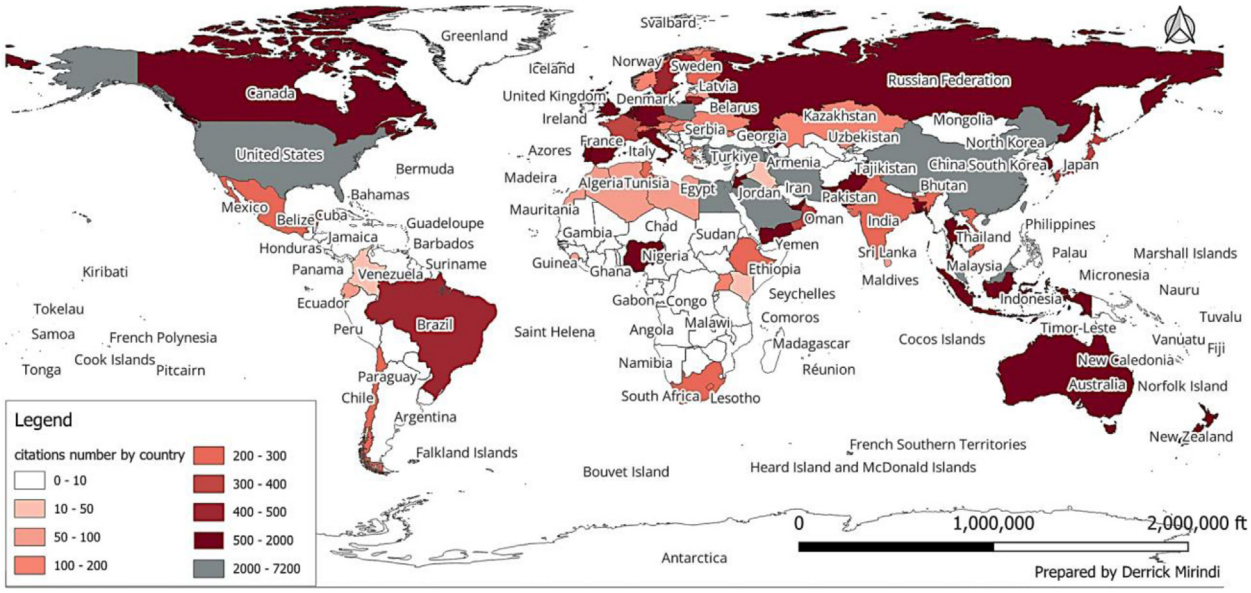

Fig. 1. Citation map by country of the Search Level 1 from 2016–2025 - extracted from DimensionsAI with VOSviewer.

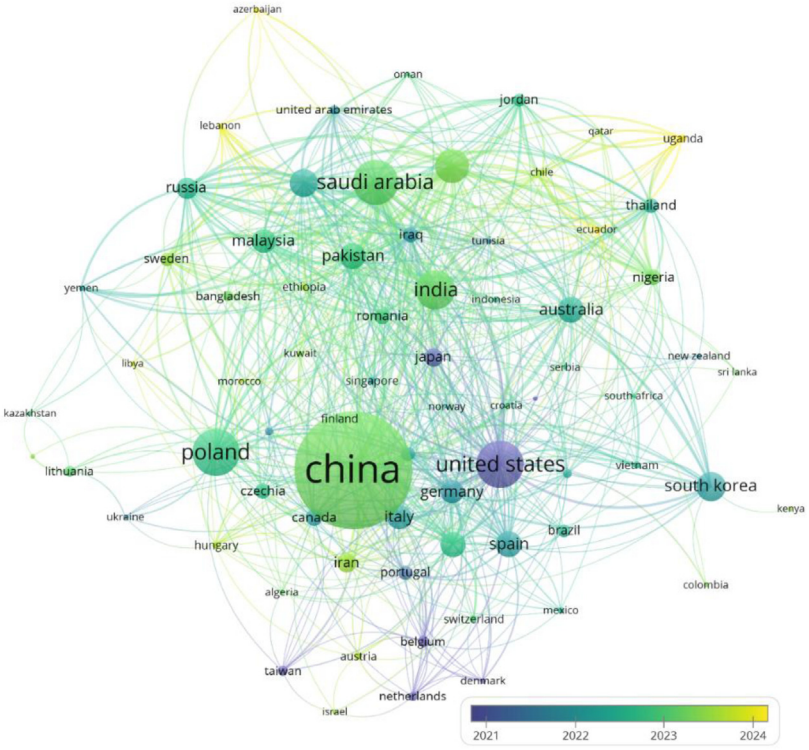

Fig. 2. Network of countries of the Search Level 1 from 2016–2025 - extracted from DimensionsAI with VOSviewer.

Iceland (161 documents, 4203 citations), and Poland (207 documents, 2117 citations). Countries such as Egypt, South Korea, and Türkiye also make meaningful contributions, although their overall document and citation counts are lower.

Fig. 3 displays the publication trends from 2022 to 2025 for the top seven research fields associated with glass-enhanced concrete (Search level 1), as determined by the total number of publications over this period. Indeed, Engineering leads all fields with a total of 439 publications, followed by *Civil Engineering* (372), *Built Environment and Design* (138), *Building* (135), *Chemical Sciences* (97), *Macromolecular and Materials Chemistry* (72), and *Materials Engineering* (57). This pattern highlights the central role of *Engineering* and *Civil Engineering* in advancing glass-enhanced concrete research while also demonstrating the interdisciplinary nature of the topic, as evidenced by sustained contributions from fields such as Chemical Sciences and Materials Engineering.

Furthermore, the network map in Fig. 3 highlights the centrality and interconnectedness of leading institutions in this field. Prince Sattam

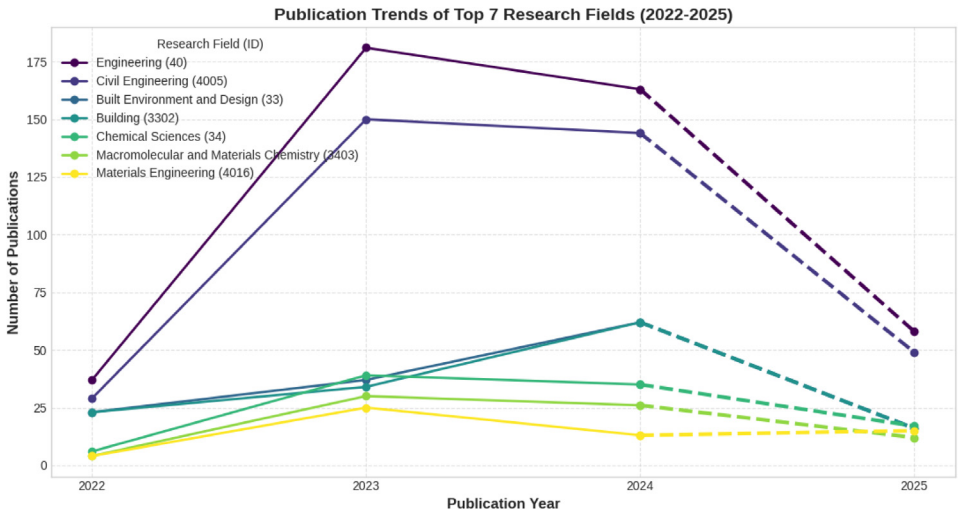

Fig. 3. Timeline of Publications of the Search Level 1 from 2014–2024 in 5 research fields - extracted from Dimensions with VOSviewer.

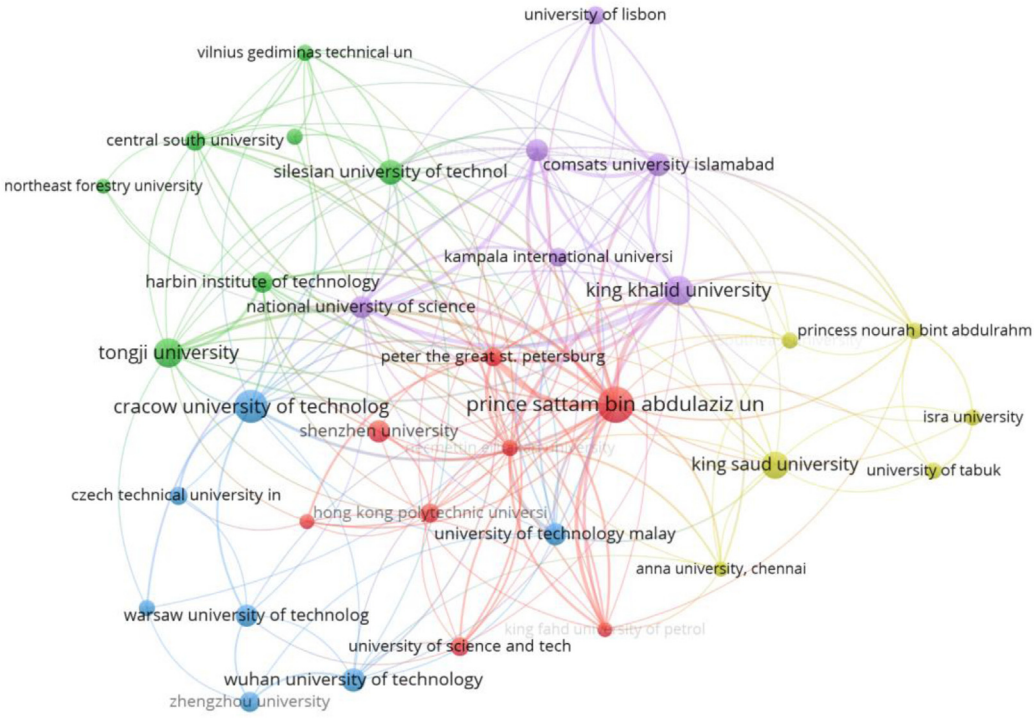

Fig. 4. Network for institutions for the Search Level 1 from 2016–2025 - extracted from Dimensions with VOSviewer.

bin Abdulaziz University emerges as the most prominent institution, with the highest number of documents (38), citations (1688), and total link strength (171), reflecting its important role in both research output and collaborative influence. Other major hubs include Cracow University of Technology, King Khalid University, Tongji University, and King Saud University, each contributing substantial publication counts and forming strong collaborative ties. The clusters indicate regional and thematic collaboration patterns, with institutions from China, Saudi Arabia, Poland, and Malaysia forming distinct but interconnected groups. Notably, the network demonstrates a high degree of international collaboration, as evidenced by the dense web of links between institutions across Asia, Europe, and the Middle East (see Fig. 4).

Fig. 5 presents the author collaboration network for the same period, revealing the key individual contributors and their collaborative dynamics. The network is characterized by several prominent clusters, each centered around highly productive and influential authors. Yasin Onuralp Özçelik, Qudeer Hussain, Sergey A. Stel'makh, and Jawad Ahmad stand out as leading figures, each with a substantial number of publications and citations. Özçelik and Ahmad, in particular, are at the core of dense clusters, indicating frequent co-authorship and strong collaborative ties with a wide array of researchers. Authors such as Kennedy C. Onyelowe and Evgenii M. Shcherban further contribute to the network's connectivity, linking different research groups and facilitating the exchange of ideas. The diversity of author affiliations and the presence of both established and emerging scholars reflect a vibrant and evolving research community.

Fig. 6 depicts the journal publication network for research on glass-enhanced concrete from 2016 to 2025. The journal *Materials* stands out as the central publication hub, with the highest metrics: 845

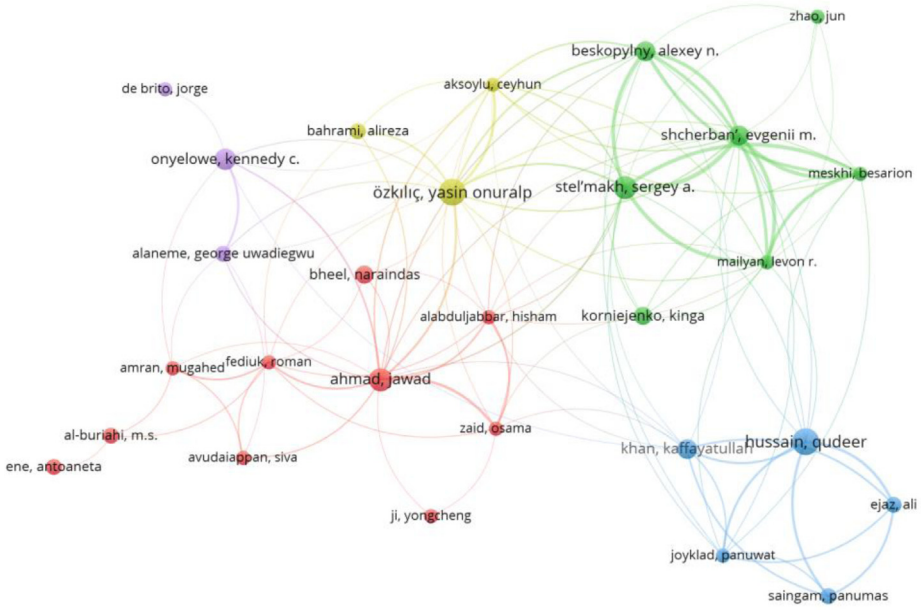

Fig. 5. Network by authors of the Search Level 1 from 2016–2025 - extracted from DimensionsAI with VOSviewer.

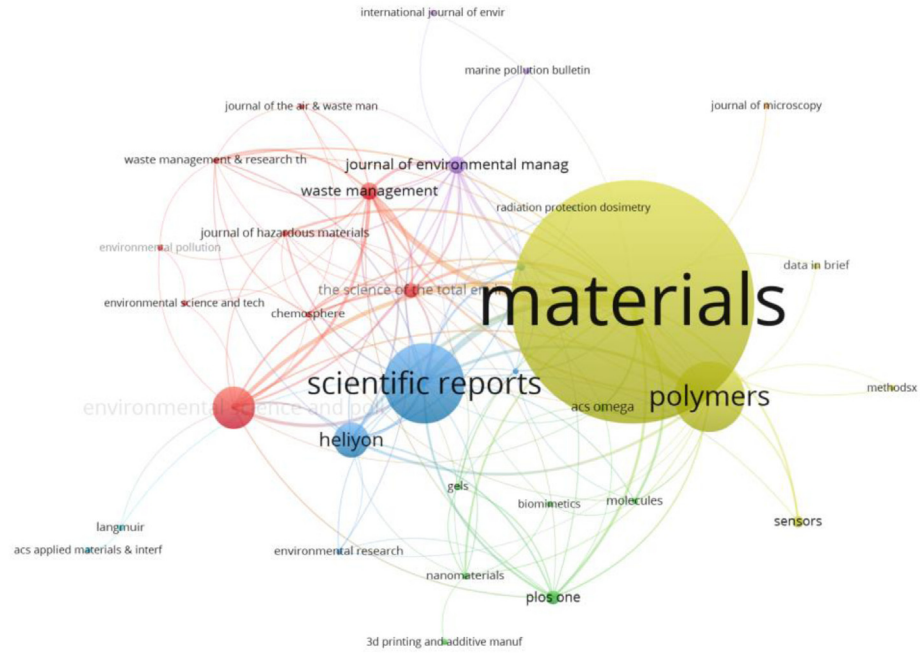

Fig. 6. Network by journals publication of the Search Level 1 from 2016–2025 - extracted from DimensionsAI with VOSviewer.

documents and 10,674 citations. This dominance shows *Materials* as the primary journal publication for both the volume and impact of research in glass-enhanced concrete, reflecting its broad scope and appeal to materials science researchers. Other significant journals include *Scientific Reports* and *Polymers*, with 281 and 245 documents, respectively, and substantial citation counts (2079 and 3996). The network further reveals important clusters around environmental and sustainability-focused journals such as *Environmental Science and Pollution Research*, *Journal of Environmental Management*, and *Waste Management*. These journals, while publishing fewer total documents, contribute significant citation influence and serve as bridges between materials science and environmental research communities. The presence of multidisciplinary and open-access journals such as *Heliyon* and *PLOS ONE* facilitates broader dissemination and cross-disciplinary collaboration.

3. Glass characteristics

Waste glass (WG) demonstrates significant potential as a fine aggregate substitute in concrete applications, particularly when processed to smaller particle sizes such as 75 μm, as illustrated in Fig. 7(a) [28]. The chemical composition analysis presented in Table 1 reveals that glass contains high silica content (72.86% SiO_2) along with substantial calcium oxide (11.58% CaO) and sodium oxide (10.10% Na_2O) [28]. These properties indicate favorable pozzolanic properties for concrete applications [29].

However, microscopic examination of glass particles reveals critical structural limitations, as shown in Fig. 8, where larger glass particles exhibit significant micro-cracks that can compromise concrete performance [28]. These inherent defects create pores and provide pathways

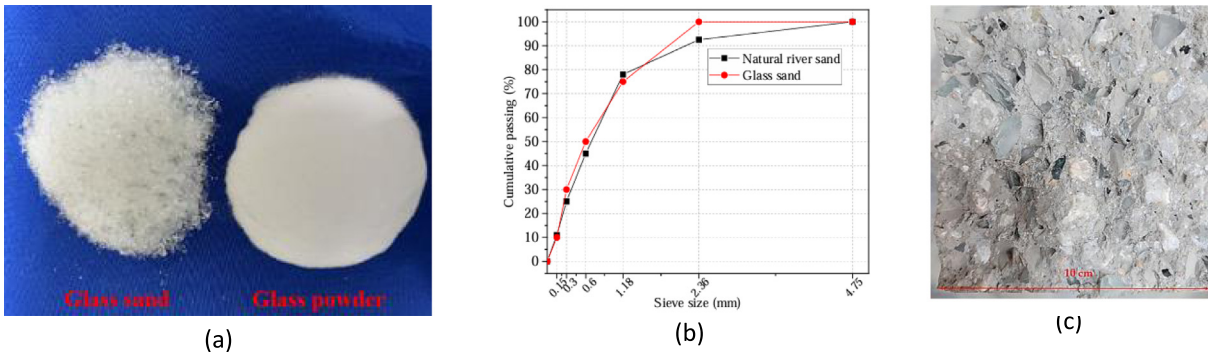

Fig. 7. (a) Physical appearance of glass sand and glass powder, (b) Grain size distribution of fine aggregate [28], (c) recycled glass in concrete [31].

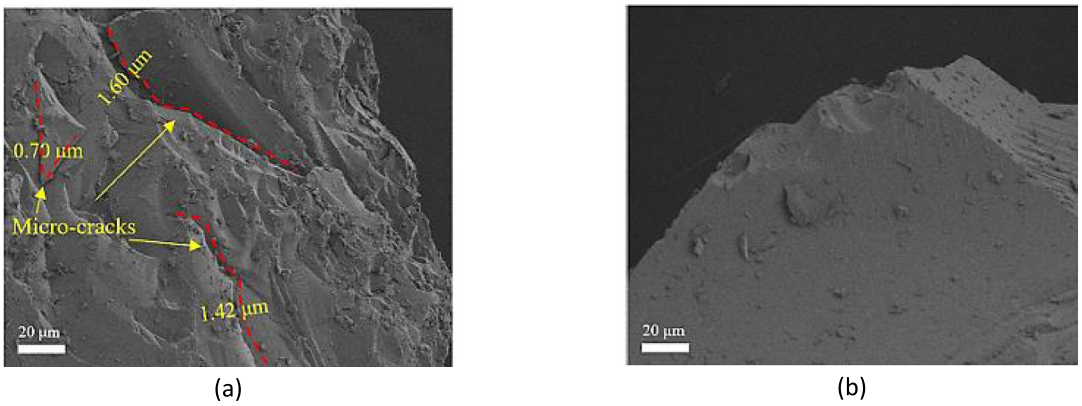

Fig. 8. Microscopic characteristics of glass: (a) glass sand and (b) glass powder [28].

Table 1
Chemical composition of waste glass [28].

Material	SiO_2	CaO	Al_2O_3	Fe_2O_3	MgO	SO_3	Na_2O	K_2O
Glass	72.86	11.58	1.84	0.84	1.92	0.34	10.10	0.45

for deleterious reactions, potentially increasing alkali-silica reaction (ASR) susceptibility [29]. To mitigate these concerns, the glass sand particle size range can be strategically reduced to 0–2.36 mm, as illustrated in Fig. 7 (b) [28], as research demonstrates that finer glass particles below certain thresholds significantly reduce ASR risk while enhancing pozzolanic activity [29,30].

Additionally, as depicted in Fig. 9, we have classified glass materials based on their characteristics in six groups as follows:

I. Barium-based glasses are characterized by high density, approximately ranging between 4.492 and 4.942 g/cm^3 for a glass chemical composition of $65Bi_2O_3$-xBaO-(35-x)B_2O_3, where the ratio composition may vary between bismuth oxide (Bi_2O_3) (65%), barium oxide (BaO) (2% to 10%) and boron oxide (B_2O_3) (25% to 35%) [32]. With the glass composition of xBaO–5ZnO–5MgO–14Na_2O–1Li_2O–(75–x)B_2O_3 (namely; barium oxide (BaO), zinc oxide (ZnO), magnesium oxide (MgO), sodium oxide (Na_2O), lithium oxide (Li_2O), boron oxide (B_2O_3)), the refractive index can rise with increasing BaO content, with values ranging approximately from about 2.20 to 2.45 as BaO concentration increases from 0 to 50 mol% [33]. The melting point of borosilicate glass generally falls between 820 °C and 1,100 °C [34]. Recent findings show that increasing Ba^{2+} ion concentration leads to a decrease in polaron radius and inter-ionic distance, while the oxygen packing density increases with higher BaO content, indicating a more compact glass structure [33]. Indeed, the addition of BaO also enhances the formation of non-bridging oxygens, which improves the glass's ability to absorb gamma and X-rays, making these glasses suitable for use as transparent protective materials against X-ray and gamma radiation. These properties are promising for optical, optoelectronic, and radiation shielding applications.

II. Sodium-calcium glasses, or soda-lime glasses, are the most widely used glass type due to their cost-effectiveness and versatile properties. Generally, their composition which can vary, primarily consists of silica (SiO_2), sodium oxide (Na_2O), and calcium oxide (CaO), ranging between 50% and 70%, 15% and 22.5%, and 15% and 22.5%, respectively [35]. Physical properties of soda-lime glass include a density of approximately 2.5 g/cm^3, a Mohs hardness of 5–6, a thermal conductivity of approximately 1.0 W/m K, a thermal expansion coefficient reaching almost 9×10^{-6}/K, and a softening point around 720 °C. The density and mass of glass significantly influence its stress distribution and structural suitability [36].

III. Lead-infused glasses, specifically pseudo-binary (100–x)$Li_2B_4O_7$ –xPb_3O_4 chemical composition (with lithium tetraborate ($Li_2B_4O_7$) and lead tetroxide (Pb_3O_4)) where PbO content ranging from 0 to 70 mol%, exhibit increasing density and decreasing molar volume as Pb_3O_4 content rises [37]. The density increases linearly with PbO content, while the glass transition temperature (Tg) and molar volume decrease non-linearly. These glasses are valued for their high density and refractive index, making them suitable for optical and radiation shielding applications [38]. PbO acts as both a network modifier at low concentrations and a network former at higher concentrations, influencing the glass structure and properties. The addition of Pb_3O_4 leads to the formation of strong B–O–Pb linkages, resulting in a more compact glass structure [39,40].

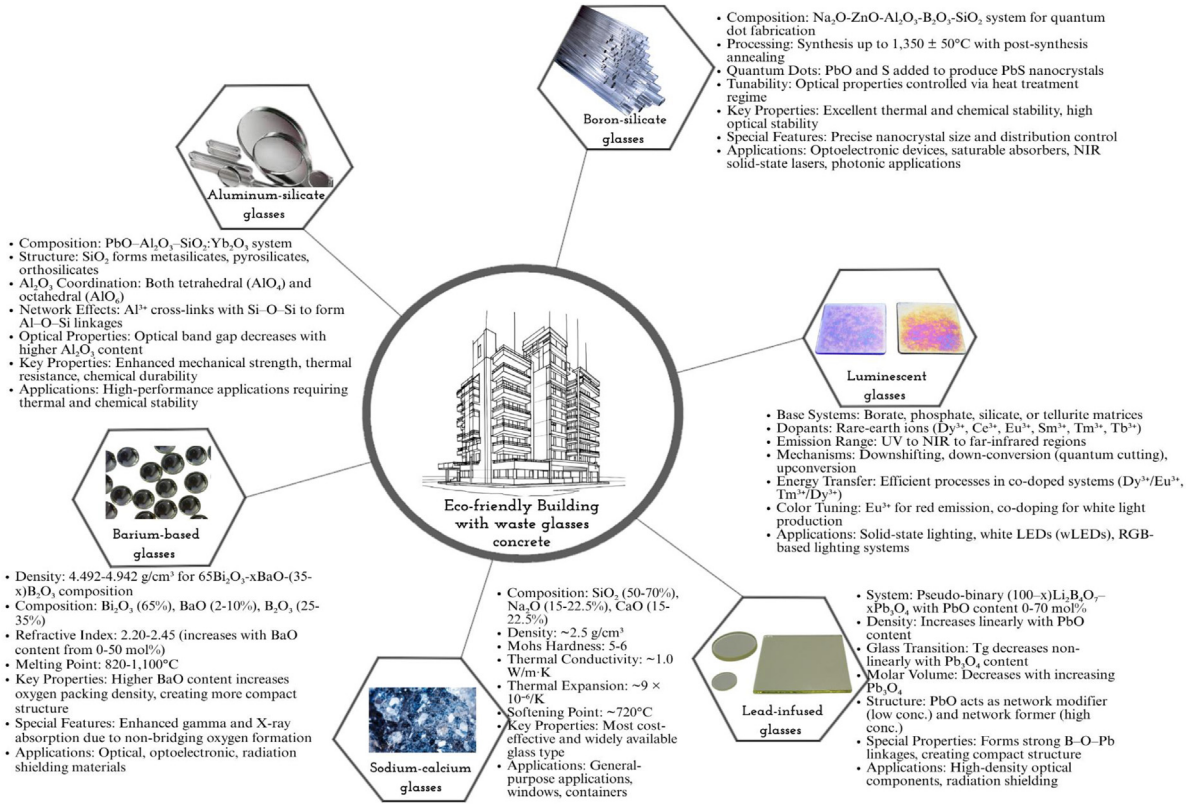

Fig. 9. Glass classification based on their characteristics (own elaboration).

IV. Aluminum-silicate glasses are engineered for enhanced mechanical strength, thermal resistance, and chemical durability. In the $PbO–Al_2O_3–SiO_2:Yb_2O_3$ glass composition (namely; lead oxide (PbO), aluminum oxide (Al_2O_3), silica (SiO_2), and ytterbium oxide (Yb_2O_3)), SiO_2 acts as the primary glass former, resulting in the formation of metasilicates, pyrosilicates, and orthosilicates [41], while Al_2O_3 is incorporated in both tetrahedral (AlO_4) and octahedral (AlO_6) coordination [42]. Increasing Al_2O_3 content leads to structural depolymerization of the glass network, as indicated by infrared spectral analysis, with aluminum ions (Al^{3+}) cross-linking with Si–O–Si units to form Al–O–Si linkages. This modification results in changes to the vibrational bands and a decrease in the optical band gap with higher Al_2O_3 concentration, confirming network disruption. The presence of AlO_6 units at higher Al_2O_3 concentrations also causes local structural distortions, affecting the optical properties and Stark splitting of Yb3+ ions. These structural features contribute to the glass's suitability for high-performance applications requiring thermal and chemical stability [43].

V. Boron-silicate glasses are notable for their excellent thermal and chemical stability. In the context of quantum dot (QD) fabrication, boron-silicate glasses are typically prepared using a $Na_2O-ZnO-Al_2O_3-B_2O_3-SiO_2$ system (namely; sodium oxide (Na_2O), zinc oxide (ZnO), aluminum oxide (Al_2O_3), boron oxide (B_2O_3)–silica (SiO_2)), with PbO and S added to produce PbS nanocrystals within the glass composition [44]. Indeed, these glasses are synthesized at high temperatures (up to 1350 ± 50 °C) and subsequently annealed, which allows precise control over the nanocrystal size and distribution through post-synthesis thermal treatment. The resulting materials exhibit high optical and thermal stability, and their optical properties can be finely tuned by adjusting the heat treatment regime. This tunability, combined with the inherent durability of boron-silicate glasses, makes them promising for optoelectronic devices, including saturable absorbers for near-infrared solid-state lasers and other photonic applications.

VI. Luminescent glasses are materials that incorporate rare-earth (lanthanide) ions or other luminescent centers into a glass matrix, enabling a wide range of photoluminescent properties. The base composition of these glasses can vary, commonly including borate, phosphate, silicate, or tellurite systems, with the addition of rare-earth ions such as Dy^{3+} (dysprosium), Ce^{3+} (cerium), Eu^{3+} (europium), Sm^{3+} (samarium), Tm^{3+} (thulium), or Tb^{3+} (terbium), among others [45]. These rare-earth ions enable emissions spanning from the ultraviolet (UV) to the near-infrared (NIR) and even far-infrared regions, depending on the dopant and excitation source.

The photoluminescence mechanisms in these materials include downshifting, down-conversion (quantum cutting), and up-conversion, allowing for tailored emission characteristics for applications such as solid-state lighting and white light-emitting diodes (wLEDs). Energy transfer processes are important in co-doped systems (e.g., Dy^{3+}/Eu^{3+}, Tm^{3+}/Dy^{3+}), enabling efficient photon conversion and color tuning. For example, trivalent Eu^{3+} ions are widely used for red emission in red, green, and blue (RGB)-based wLEDs [46–48], while co-doping with Dy^{3+} or Tm^{3+} can produce white light through controlled energy transfer and emission balancing [45].

4. Machine learning algorithms

ML algorithms have undergone significant transformation since their foundational establishment in 1950, particularly in materials science and engineering applications (Fig. 10) [49,50]. Indeed, the emergence of fuzzy logic and neural networks in 1980 marked an important transition from traditional rule-based systems to more adaptive learning mechanisms. These early neural network architectures laid

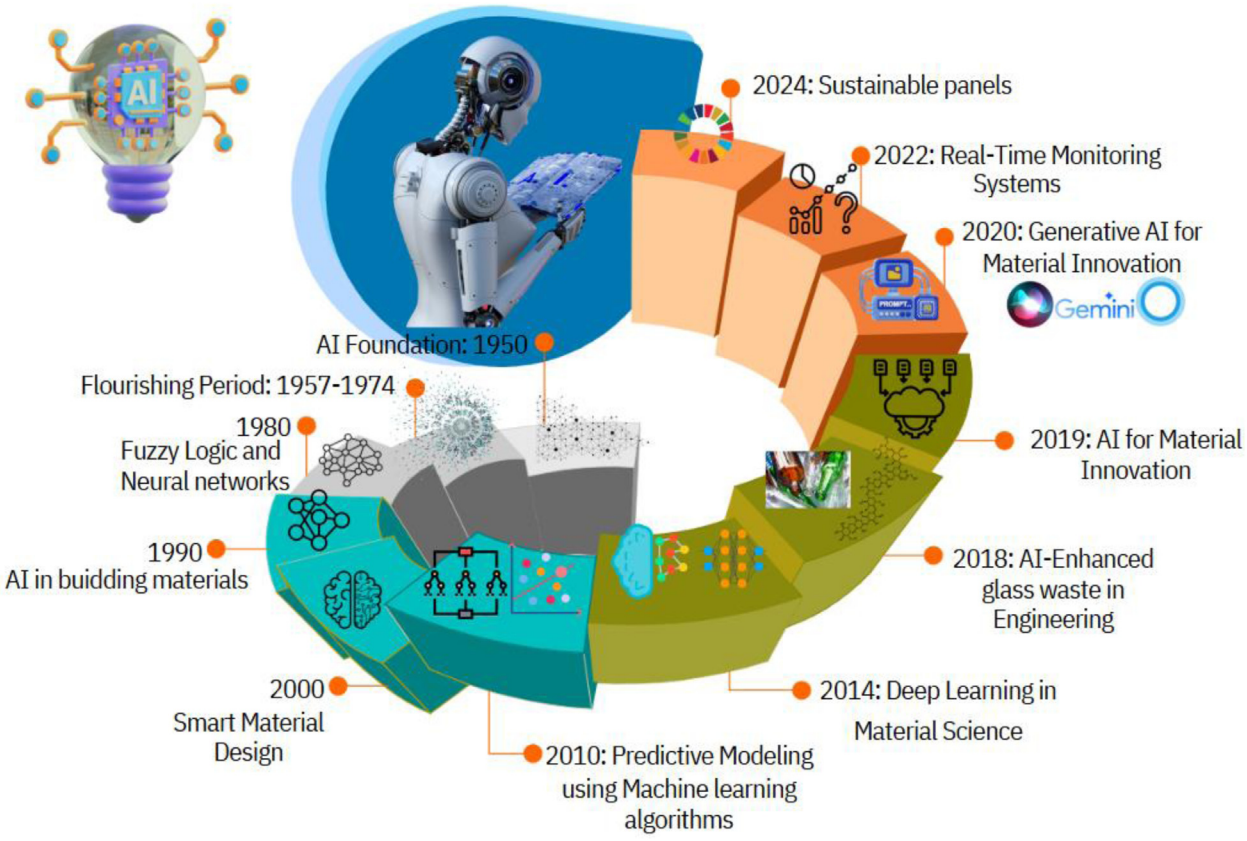

Fig. 10. Historical development of AI, ML, and DL in materials engineering.

the groundwork for contemporary deep learning applications, enabling computers to process complex, non-linear relationships in material properties and behaviors. The integration of AI in building materials during the 1990s illustrates the practical application of supervised learning algorithms, where labeled datasets of material properties were used to train models for predicting structural performance and durability characteristics [51,52].

In addition, the achievement of predictive modeling using ML algorithms in 2010 represented a paradigm shift toward data-driven material design. This period marked the adoption of various algorithms, such as the support vector machine (SVM) and decision tree (DT) methods, for analyzing material composition properties. These algorithms enabled researchers to model complex dependencies between chemical composition, processing parameters, and resulting material characteristics with unprecedented accuracy.

Also, the introduction of DL algorithms in material science in 2014 revolutionized the field by enabling the processing of vast datasets containing structural, compositional, and performance data. Advanced neural network architectures, including convolutional neural network (CNN) and recurrent neural network (RNN) algorithms, facilitated the analysis of complex material microstructures and time-dependent behaviors. The subsequent development of generative artificial intelligence (GenAI) [53] for material innovation in 2020 demonstrated the capability of algorithms to not only predict material properties but also design novel compositions with desired characteristics.

4.1. Pearson correlation coefficient

The Pearson correlation coefficient (r) algorithm, also known as the Pearson product-moment correlation coefficient (PPMCC), measures the strength and direction of a linear relationship between two continuous variables (e.g., f_c_treatment and f_b_control). Its value ranges from −1 to +1 [54,55], and it is calculated using the following formula:

$$min\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\left(\xi_i + \xi_i^*\right) \tag{1}$$

$$r = \frac{\sum_{i=1}^{n}\left(x_i - x'\right)\left(y_i - y'\right)}{\sqrt{\sum_{i=1}^{n}(x_i - x')^2 \sum_{i=1}^{n}(y_i - y')^2}} \tag{2}$$

Alternatively, it can also be expressed as:

$$r = \frac{S_{xy}}{\sqrt{S_{xx} \times S_{yy}}} \tag{3}$$

where n is the number of data points or observations, x_i is the value of the first variable for the *i-th* observation, y_i is the value of the second variable for the *i-th* observation, x' is the mean of the first variable, y' is the mean of the second variable, S_{xy} is the sum of the products of the deviations from the means (covariance), S_{xx} is the sum of the squared deviations from the mean for variable X, and S_{yy} is the sum of the squared deviations from the mean for variable Y.

4.2. Decision tree

Decision tree (DT) is a supervised learning algorithm that models' relationships between input features (f_c) and target variables (f_b and f_t) through hierarchical, interpretable splits [56–58]. For regression tasks, the algorithm partitions the feature space into regions R_m to minimize prediction error, where each leaf node represents a constant value c_m:

$$y' = \sum_{m=1}^{M} c_m \cdot I(x_{f_c} \in R_m) \tag{4}$$

DTs are supervised learning algorithms that model relationships through recursive binary splits of the feature space. For predicting concrete mechanical properties (f_b and f_t from f_c), the regression tree minimizes

the variance at each split using the mean squared error (MSE), the mean absolute error (MAE), and the root mean squared error (RMSE) criteria following Eqs. (5), (6), and (7), respectively:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - y_i')^2 \tag{5}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - y_i'| \tag{6}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n}(y_i - y_i')} \tag{7}$$

where n is the total number of data points, y_i is the actual (observed) value, and y_i' is the predicted value.

For classification tasks, Gini Impurity (G) or Entropy (H) guides splits:

$$G = 1 - \sum_{i=1}^{C} p_i^2 \tag{8}$$

$$H = -\sum_{i=1}^{C} p_i \log_2 p_i \tag{9}$$

where p_i is the proportion of class i in a node. Information Gain (IG) quantifies purity improvement:

$$IG = H_{parent} - \sum_{children} \frac{N_{child}}{N_{parent}} H_{child} \tag{10}$$

The DT models for predicting f_b and f_t from c_s were configured with three key hyperparameters, namely a maximum tree depth of 3 (max_depth=3) to capture non-linear relationships while preventing overfitting, a minimum sample split threshold of 2 (min_samples_split= 2) to enable granular partitioning of the limited dataset (n_f_b = 8, n_f_t = 7), and *random_state* equal to 42 to ensure reproducible splits.

4.3. Extreme Gradient Boosting

The Extreme Gradient Boosting (XGBoost) algorithm operates through an ensemble of regression trees (f_k) that sequentially correct residual errors in predictions [59–61]. For the concrete-glass composite strength prediction task (f_b and f_t from f_c), the model at iteration t can be computed as:

$$y_i^{(t)} = \sum_{k=1}^{t} f_k(x_{f_c}^{(i)}) \tag{11}$$

where each tree f_k belongs to the space of regression trees $\digamma$. The optimization objective combines prediction accuracy and model complexity as follows:

$$\xi^{(t)} = \sum_{i=1}^{n}(y_i - y_i'^{(t-1)} - f_t(x_{f_c}^{(i)}))^2 + \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T} w_j^2 \tag{12}$$

where $\xi(y_i, y_i')$ is the squared error loss $(y_i - y_i')^2$ for regression, $\Omega(f)$ is the regularization term controlling model complexity, T is the number of leaves in tree, w_j is the weight scores on leaves, γ is the minimum loss reduction for partition (complexity control), and λ is L2 regularization on leaf weights.

The sequential tree-building process specifically optimizes for residual errors in strength predictions, where each new tree f_t learns as follows:

$$f_t(x_{f_c}) \approx y - \sum_{k=1}^{t-1} f_k(x_{f_c}) \tag{13}$$

This iterative refinement enables precise modeling of complex relationships between compressive strength and other mechanical properties while maintaining computational efficiency through parallel tree construction.

The XGBoost model was configured with a learning rate (η = 0.1) to balance individual tree contributions and prevent overfitting while maintaining stable convergence for modeling mechanical property relationships. A total of 100 regression trees (*n_estimators*) provided an optimal trade-off between computational efficiency and predictive capacity for the moderate-sized dataset (n_f_b = 8, n_f_t = 7). Tree complexity was constrained to a maximum depth of 3 to capture non-linear interactions between f_c strength and other properties without overfitting. Full dataset sampling (subsample = 1.0) preserved data patterns given limited samples, while a minimum child weight of 1 enabled detection of subtle strength relationships through leaf node splitting. The quadratic loss function (*objective* = *'reg:squarederror'*) optimized continuous strength predictions, with *random_state* equal to 42, ensuring reproducible tree structures.

4.4. Adaptive boosting

Adaptive boosting (AdaBoost) is an ensemble algorithm that sequentially combines weak regression models (maximum depth 3 decision trees in this study) and can be used to predict concrete [62,63]. Indeed, for regression, the algorithm minimizes the linear loss function $L(y, y') = |y - y'|$, updating sample weights iteratively to focus on residuals from prior models. The AdaBoost algorithm consists of the initialization (by assigning equal weights to all N training samples), the weak learner training, the weak learner weight (the weight or "amount of say" (α_t) of the weak learner h_t in the final prediction is calculated based on its error rate ϵ_t), the weight update (by increasing weights for samples with larger residuals), and the final prediction (by combining all T weak learners).

$$y_i - (w^T\phi(x_{f_c}^{(i)}) + b) \le \epsilon + \xi_i/\xi_i^* \tag{14}$$

$$W_0(i) = \frac{1}{N},\ for\ i = 1, \ldots, N \tag{15}$$

where W_o is the initial weight of the ith training sample, and N is the total number of training samples.

$$\epsilon_t = \sum_{i=1}^{N} W_t(i) \times \mathbb{I}\left(y_i \neq h_t\left(x_{f_c}^i\right)\right) \tag{16}$$

where $\mathbb{I}\left(y_i \neq h_t\left(x_i\right)\right)$ is an indicator function that is 1 if the prediction of the weak learner $h_t(x_i)$ is incorrect for the ith sample xi with true label y_i, and 0 otherwise.

$$\alpha_t = \frac{1}{2}\log\frac{1-\epsilon_t}{\epsilon_t} \tag{17}$$

$$W_{t+1}(i) = W_t(i) \times e^{\alpha_t},\ \text{if the i-th sample is misclassified by } h_t \tag{18}$$

$$W_{t+1}(i) = W_t(i) \times e^{-\alpha_t},\ \text{if the i-th sample is correctly classified by } h_t \tag{19}$$

$$H\left(x_{f_c}\right) = \text{sign}\left(\sum_{t=1}^{T}\alpha_t h_t\left(x_{f_c}\right)\right) \tag{20}$$

where $H(xf_c)$ is the final prediction for input x, T is the total number of iterations, α_t is the weight of weak learner h_t, and $h_t(x_{cs})$ is the prediction of weak learner h_t for input x.

The AdaBoost models were configured with a base estimator of DecisionTreeRegressor (*max_depth*=3) to capture non-linear relationships between f_c and target properties f_b and f_t while preventing overfitting. A total of 50 sequential weak learners (*n_estimators*=50) were combined to balance computational efficiency and predictive refinement. The learning rate was set to 1.0 to fully integrate each learner's corrections without dampening which is critical given the limited sample sizes (n_f_b = 8, n_f_t = 7). The linear loss function prioritized minimization of absolute prediction errors ($L = |y - y'|$), aligning with material engineering requirements for strength estimation accuracy. Reproducibility was ensured through *random_state* equal to 42, fixing the initialization of sample weights and tree structures across training runs.

4.5. Light Gradient Boosting Machine

Light Gradient Boosting Machine (LightGBM) is an open-source gradient boosting framework developed by Microsoft that focuses on speed and efficiency, especially for large datasets and high-dimensional data [64,65]. It is widely used in supervised learning tasks such as classification, regression, and ranking. It can optimize predictive performance through leaf-wise tree growth and histogram-based splitting, which is particularly efficient for modeling relationships between f_c, f_b, and f_t. The algorithm minimizes an objective function combining prediction errors and regularization and can be expressed mathematically as follows:

$$\xi(t) = \sum_{i=1}^{N}(y_i - y_i^{'(t-1)} - f_t(x^i_{f_c}))^2 + \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T} w_j^2 \tag{21}$$

with $L2loos = (y_i - y_i^{'(t-1)} - f_t(x^i_{f_c}))^2$ where f_t is the new tree, T the number of leaves, w_j leaf weights, and γ/λ regularization terms.

Unlike XGBoost, which uses a level-wise tree growth strategy, LightGBM employs leaf-wise expansion, splitting the node with maximum gain:

$$Gain = \frac{(\sum x_{f_c\in L} g_i)^2}{\sum x_{f_c\in L} h_i + \lambda} + \frac{(\sum x_{f_c\in R} g_i)^2}{\sum x_{f_c\in R} h_i + \lambda} + \frac{(\sum x_{f_c\in P} g_i)^2}{\sum x_{f_c\in P} h_i + \lambda} \tag{22}$$

where g_i and h_i are first- and second-order gradients of the loss, and L, R, and P denote left, right, and parent nodes.

4.6. Gaussian Process

A Gaussian Process (GP) algorithm is a probabilistic model that defines a distribution over functions [66]. Instead of predicting a single function, a GP provides a probability distribution over possible functions that could fit the observed data [67,68]. This is particularly useful for quantifying uncertainty in predictions. GP regression models the relationship between f_c and mechanical properties (f_b and f_t) as a probabilistic distribution over functions, characterized by a mean function $m(x)$ and a covariance kernel $k(x, x')$. For this study, the Radial Basis Function (RBF) kernel with additive noise governed predictions:

$$k(x_i, x_j) = \sigma_f^2 exp - \frac{\|x_i - x_j\|}{2l^2} + \sigma_n^2\delta_{ij} \tag{23}$$

where σ_f^2 (signal variance) scales function variability, l (length scale) controls input correlation smoothness, and σ_n^2 (noise variance) captures observational uncertainty. The posterior predictive distribution for a new f_c strength value x_* is Gaussian:

$$p(y_* \mid x_*, D) \sim N(k_*^T(K + \sigma_n^2 I)^{-1} y, k_{**} - k_*^T(K + \sigma_n^2 I)^{-1} k_*) \tag{24}$$

GP hyperparameter consists of *n_restarts_optimizer* equal to 15 to mitigate the local minima risks during kernel parameter tuning, alpha equal to 1e-2 to regularize the matrix inversion to stabilize numerical computations, *normalize_y* equal to True to standardized the target variables (y_{fb}, and y_{ft}) to zero mean and unit variance for improved kernel scaling, and *random_state* equal to 42 to ensure the reproducible initialization of hyperparameter optimization.

4.7. Support vector regression

Support vector regression (SVR) is a regression algorithm based on the principles of support vector machine (SVM), which were originally designed for classification [69,70]. While SVM algorithm aims to find a hyperplane that separates different classes with the maximum margin [71], SVR algorithm seeks to find a function that best fits the data points within a specified margin of tolerance [72]. Instead of minimizing the squared error like traditional regression methods, SVR minimizes the error within a predefined ε-insensitive tube. This means that errors that fall within the tube are ignored, and only errors that fall outside the tube are penalized. The goal is to find a function (represented by a hyperplane in a potentially high-dimensional feature space) that has the largest possible margin and keeps as many training points as possible within the ε-tube. SVR aims to minimize the following objective function:

$$min\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{25}$$

$$\begin{cases} y_i - (W^T\phi(X_{cs}^{(i)}) + b) \le \epsilon + \xi_i \\ (W^T\phi(X_{cs}^{(i)}) + b) - y_i \le \epsilon + \xi_i \\ \xi_i, \xi_i^* \ge 0 \end{cases} \tag{26}$$

where $\phi(x_i)$ is a kernel function that maps the input x_i to a higher-dimensional feature space, w is the weight vector, b the bias term, ξ_i/ξ_i^* slack variables for excess positive/negative errors, ϵ is the width of the ε-insensitive tube, defining the tolerance for errors, and C is a regularization parameter that controls the trade-off between maximizing the margin and minimizing the error. The RBF kernel:

$$\phi(x_i, x_j) = exp(-\gamma\|x_i - x_j\|^2) \tag{27}$$

with γ='scale' (automatically set to $1/(n_{features}\cdot Var(x_{f_c}))$) adapts to the input variance, while ϵ equal to 0.1 defines the error tolerance margin. For the limited dataset, this configuration prioritizes generalization over strict fitting, leveraging support vectors at the ϵ-tube boundaries to capture nonlinear relationships between f_c strength and mechanical properties. The regularization parameter C, which equals 100, balances margin width and error penalization, accommodating sparse data while mitigating overfitting.

The ML algorithms used in this study and their corresponding hyperparameters and descriptions are depicted in Table 2.

4.8. Non-parametric correlation coefficients of Spearman and Kendall

Non-parametric correlation methods evaluate monotonic relationships between variables without assuming linearity or normality, including Spearman's rank correlation coefficient (ρ) and Kendall's tau (τ). These techniques are ideal for non-Gaussian data or ordinal measurements, relying on rank transformations to assess associations [73].

Spearman's ρ quantifies the strength and direction of monotonic relationships by applying Pearson's formula to rank-transformed data. For paired observations (X_i, Y_i), ranks $R[X_i]$ and $R[Y_i]$ replace raw values. The coefficient is computed as:

$$\rho = \frac{cov(R[X], R[Y])}{\sigma R[X]\sigma R[Y]} \tag{28}$$

where *R[X]* and *R[Y]* represent the ranks of variables X and Y, respectively. The coefficient ranges from −1 (perfect negative monotonic association) to +1 (perfect positive monotonic association), with 0 indicating no association.

Furthermore, Kendall's τ measures association by comparing concordant and discordant observation pairs. For any two pairs (X_i, Y_i) and (X_j, Y_j), the concordant pair is ($X_i > X_j$) and ($Y_i > Y_j$), or ($X_i < X_j$) and ($Y_i < Y_j$); while the discordant pair is ($X_i > X_j$) and ($Y_i < Y_j$), or ($X_i < X_j$) and ($Yi > Yj$). The τ-b variant (adjusted for ties) is defined as:

$$\tau_b = \frac{P - Q}{\sqrt{(P + Q + T)(P + Q + U)}} \tag{29}$$

where P is the number of concordant pairs, Q the discordant pairs, T ties in X, and U ties in Y. τ ranges from −1 (all pairs discordant) to +1 (all pairs concordant), with 0 implying independence.

4.9. k-fold cross-validation

Cross-validation is widely used to estimate prediction error and model generalization capability [74]. Therefore, we can assume that 5-fold or 10-fold cross-validation can provide an optimal balance between

Table 2
Hyperparameters and description of machine learning algorithms.

Algorithm	Hyperparameter	Value	Description
Decision Tree	max_depth	3	Limits tree depth to capture non-linear relationships while preventing overfitting
	min_samples_split	2	Allows node splitting with ≥2 samples
	random_state	42	Ensures reproducible splits
XGBoost	learning_rate	0.1	Controls the contribution of individual trees
	n_estimators	100	Number of gradient-boosted trees
	max_depth	3	Tree depth limit for non-linear interactions
	subsample	1.0	Uses 100% of the data for each tree
	min_child_weight	1	Allows splits with minimal Hessian sum
	objective	reg:squarederror	Quadratic loss function for regression
	random_state	42	Seed for reproducible tree structures
AdaBoost	estimator	DT (max_depth=3)	Base learner with constrained complexity
	n_estimators	50	Number of sequentially boosted weak learners
	learning_rate	1.0	Full weighting of learner contributions
	loss	linear	Absolute error minimization (L1 loss)
	random_state	42	Controls weight initialization
LightGBM	boosting_type	gbdt	Traditional gradient boosting framework
	num_leaves	31	Maximum leaves per tree (constrained by max_depth=3)
	max_depth	3	Depth limit for computational efficiency
	learning_rate	0.1	Step size shrinkage during boosting
	n_estimators	100	Total number of boosting iterations
	min_child_samples	1	Minimum samples required in leaf nodes
	subsample	1.0	Uses the full dataset for each iteration
	colsample_bytree	1.0	Uses 100% of features per tree
	reg_alpha	0.0	L1 regularization term (disabled)
	reg_lambda	0.0	L2 regularization term (disabled)
	random_state	42	Seed for histogram binning and splitting
Gaussian Process	kernel	1.13^2×RBF + WhiteKernel	RBF for smooth trends + WhiteKernel for noise
	n_restarts_optimizer	15	Optimization attempts to avoid local minima
	alpha	1e−2	Added to kernel matrix diagonal for numerical stability
	normalize_y	True	Standardizes target variables
	random_state	42	Controls kernel parameter initialization
SVR	kernel	rbf	Radial basis function for non-linear mapping
	C	100	Regularization strength (balances margin vs. error)
	gamma	scale	Auto-set to $1/(n_{features} \times \mathrm{Var}(x))$
	epsilon	0.1	Width of ϵ-insensitive tube (ignores errors <0.1 MPa)

computational efficiency and performance assessment for any dataset. Additionally, k-fold cross-validation methodology partitions the dataset into k equally-sized folds. It can be computed mathematically as illustrated in Eq. (30) as follows:

$$CV_{error} = \frac{1}{k}\sum_{i=1}^{k} E_i \tag{30}$$

where E_i represents the error metric (R^2, MSE, or MAE) for the ith fold. Each fold is the test set once, while the remaining k-1 folds form the training set.

The k-fold cross-validation is set up with 5-fold cross-validation with stratified sampling across available data points. Also, we set the random state to a value equal to 42 for reproducibility. This validation strategy ensures each sample appears in the test set exactly once during cross-validation. Also, it is helpful to note that for each fold is calculated based on R^2 as follows:

$$R^2_{fold} = 1 - \frac{\sum(y_{true} - y_{pred})^2}{\sum(y_{true} - y')^2} \tag{31}$$

where $(y_{true} - y_{pred})^2$ is the sum of squared residuals that measures the prediction error, $(y_{true} - y')^2$ measures the total variance in actual data, and y' is the mean of true values.

5. Results and discussion

The incorporation of various types of waste glasses, such as liquid crystal display glass (LCD) or cathode ray tube glass (CRT) as a partial replacement for fine aggregate in concrete has been widely studied for its effects on mechanical properties such as compressive (f_c), flexural (f_b), and split tensile (f_t) strengths. As shown in Figs. 11 and 12, the use of WG, particularly window glass, often results in notable improvements in f_c and f_b strengths. For instance, as depicted in Fig. 11, Dębska et al. [75] obtained control and treatment f_c strengths of 94.73–96.65 MPa and f_b strengths of 24.3–27.73 MPa at 28 days by mixing 14.64% waste glass as sand replacement. In their conclusion, they highlighted that mortars with waste glass can increase performance, with minimal absorbability and no significant density change, and that optimized mixes can simultaneously maximize strength and minimize absorption. Ismail et al. [76] recycled waste glass obtained from containers (bottles, jars) and flat glass (windows), while CRT (cathode ray tube) glass was excluded due to hazardous metals, as a partial sand replacement ranging from 10 to 20%. At 28 days, the 20% recycled glass mix achieved the highest strengths. Indeed, fc and fb strengths increase from 44.0 to 45.9 MPa and from 5.89 to 6.55 MPa, respectively. Abdallah et al. [77] used waste glass as a partial fine-aggregate replacement in concrete at 5%, 15%, and 20%. Similarly, at 28 days, the 20% mix showed f_c strength 32.49 to 34.22 MPa, and f_b strength 4.9 to 5.3 MPa, indicating beneficial late-age effects likely linked to pozzolanic reactions.

When examining f_t strength in Fig. 12, the trends are more nuanced. Indeed, researchers such as Wang et al. [78] mixed ultrahigh-performance shotcrete (UHPS) using recycled waste flat glass sand (100% replacement) and found, at 28 days, peak f_c strengths around 135.2 MPa, and f_t strengths about 19.75 MPa. On the other hand, Tong et al. [30] obtained 28-day control–treatment ranges of 39.8–53.99 MPa f_c and 2.99–364 MPa f_t, respectively, by incorporating up to 20% waste glass in concrete. Similarly, Batayneh et al. [79] obtained control f_c and f_t strengths of approximately 31 MPa and 5 MPa, respectively, while treatments with up to 20% crushed waste glass achieved higher values across both properties. They concluded that replacing fine aggregate with crushed glass (0%–20%) increases mechanical properties compared with normal concrete. In contrast, Wang [80] obtained

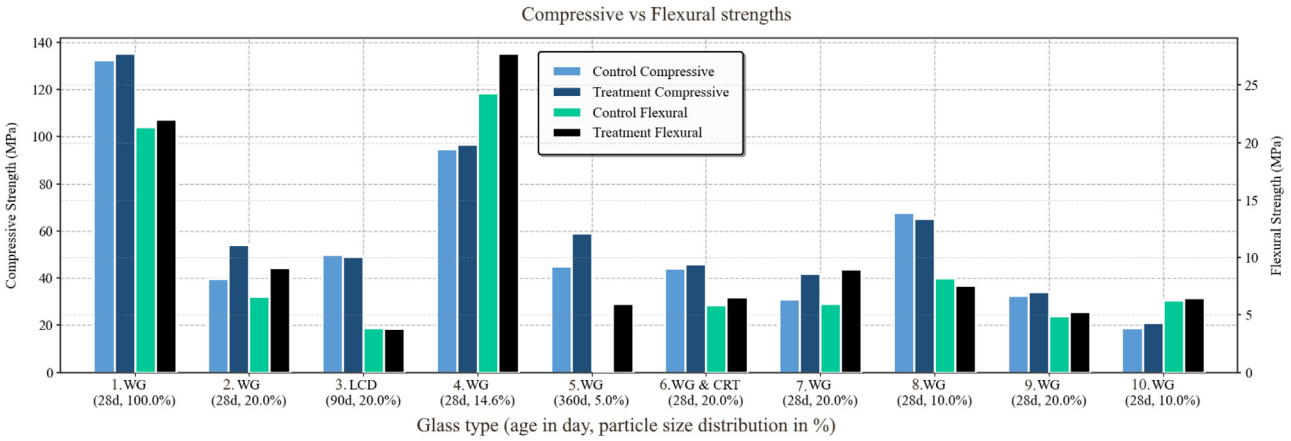

Fig. 11. Compressive and flexural strength of concrete made with glass with 1. [78], 2. [30], 3. [80], 4. [75], 5. [81], 6. [76] 7. [79] 8. [82] 9. [77], 10. [31].

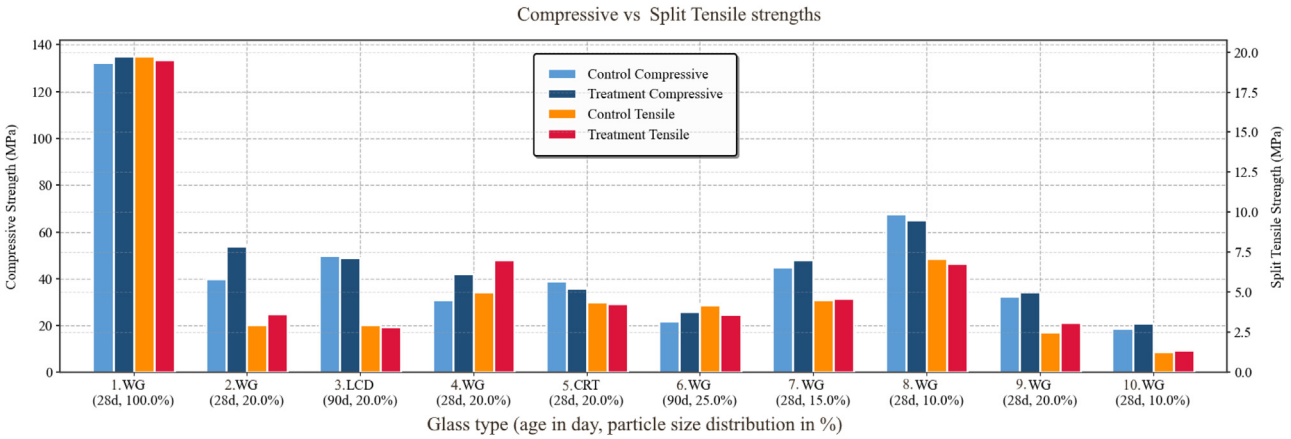

Fig. 12. Compressive and split tensile strength of concrete made with glass with 1. [78], 2. [30], 3. [80], 4. [79], 5. [83], 6. [84], 7. [85], 8. [82], 9. [77], 10. [31].

control strength values slightly higher than treatment values. Indeed, using 20% LCD glass replacement at 90 days reduced f_c strength from approximately 50 to 49 MPa while also lowering f_t strength from almost 2.98 to 2.83 MPa. In line with the article's findings, he highlighted that the concrete performance generally decreases as glass sand replacement increases.

Therefore, most of the literature suggests that waste glass can be beneficial for concrete performance. However, optimal replacement levels and particle sizes are critical for maximizing mechanical properties, as excessive glass content may introduce negative effects such as alkali-silica reaction or reduced bond strength. Therefore, predicting the strength of one property based on the other using ML algorithms can provide valuable insights for optimizing mix designs, enabling more precise control over concrete performance when incorporating waste glass.

For instance, the Pearson correlation matrix reveals strong positive relationships between all mechanical strength properties of glass concrete compared to other variables such as the type of glass, the age of the concrete, or the size of glass replacement in concrete. Fig. 13 reveals that the choice of glass type does not significantly affect concrete strength. All correlations with f_c, f_b, and f_t are near zero or weakly negative. However, the glass replacement percentage in concrete presents significantly different results. It shows a moderate positive correlation to concrete performance. For instance, the correlation with f_t strength is much stronger (0.932 and 925 for control and treatment, respectively), with higher replacement levels improving f_t strength significantly. Concrete age shows weak correlations with strength properties. Most values range from −0.022 to −0.290. Therefore, older age, ranging from 28 days, does not substantially impact glass concrete performance.

Results also indicate coherent behavior across f_c, f_b, and f_t strengths under control and treatment conditions. The most remarkable finding is the near-perfect correlation between control and treatment conditions within each strength property, with f_c, $f_{b,}$ and f_t strength showing a correlation of 0.983, 0.987, and 0.990, respectively. These results suggest that the incorporation of waste glass maintains consistent relative mechanical performance patterns across different concrete mixes, regardless of treatment application. Furthermore, cross-property correlations demonstrate significant interdependence among mechanical characteristics. For example, f_c strength shows strong correlations with f_b strength, ranging from 0.809 to 0.876, indicating that improvements in one property generally correspond to enhancements in the other. Similarly, the relationship between f_b and f_t strengths is notably higher, with correlations ranging from 0.943 to 0.970, with $f_{b\_control}$ and $f_{t\_control}$ showing the highest cross-property correlation at 0.970. Also, the correlations between f_c and f_t, while still substantial at 0.927 to 0.948, are relatively higher than those between f_c and f_b property pairs.

Using the same analysis approach, the DT models accurately predict f_b and f_t strengths of glass concrete using f_c strength, age, and replacement percentage. Fig. 14 shows the f_b model reaches R^2 equal to 0.9911, while Fig. 15 depicts the f_t model gives R^2 equal to 0.9523. Indeed, the DT splits first at $f_c \leq 80.925$ MPa, separating 13 samples with a mean f_b of 9.55 MPa (Fig. 14). In the next stage, the DT model divides the tree into two with 11 samples by further splitting at an age

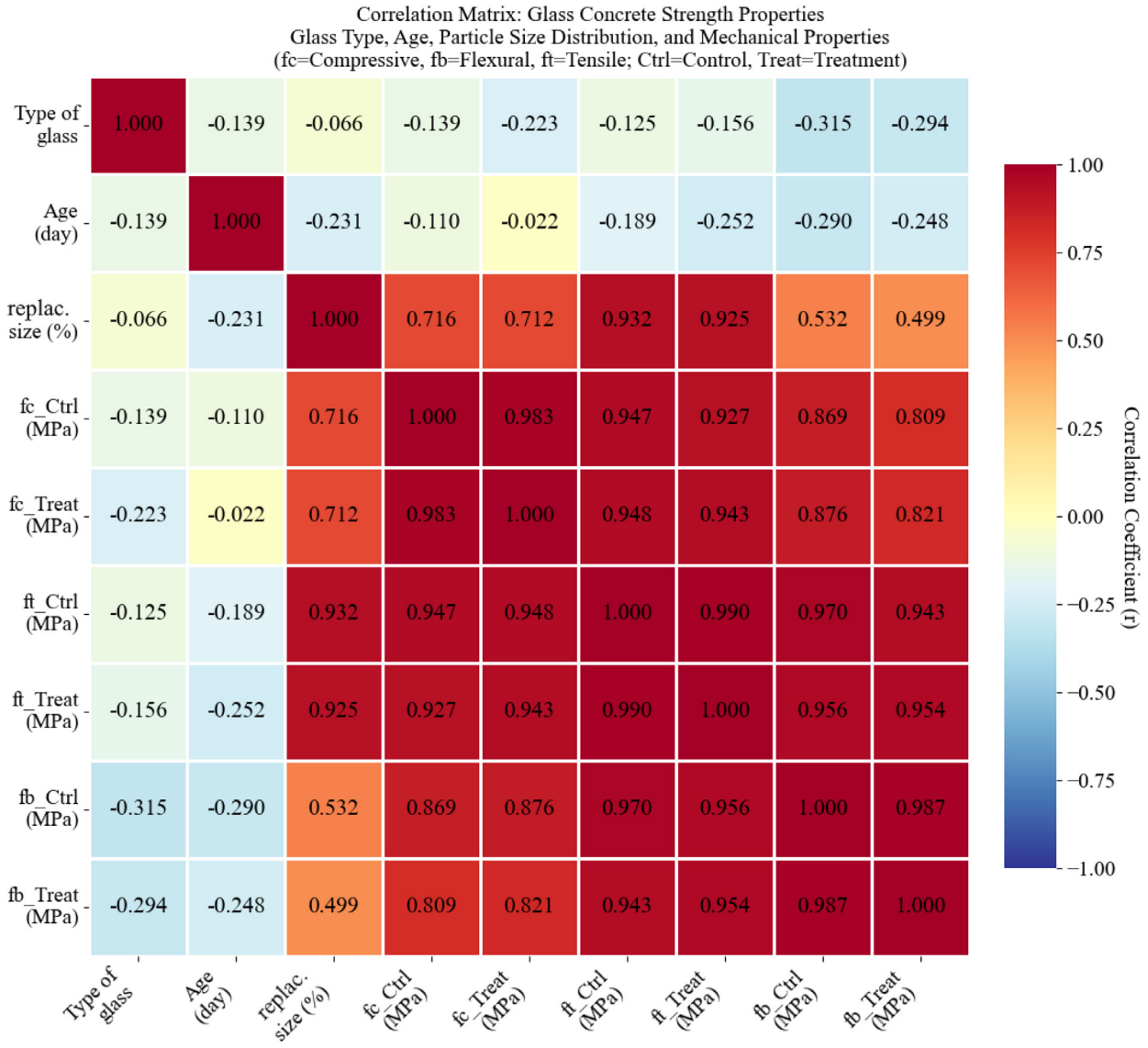

Fig. 13. Pearson correlation matrix of glass-concrete mechanical properties.

of concrete less than or equal to 59.0 days, leading to an average f_b of 6.76 MPa. This results in two branches at f_c equal to 38.11 MPa and 54.0 MPa, narrowing the f_b prediction value between 3.8 and 8.092 MPa.

Similarly, Fig. 15 illustrates the DT for f_t strength, using f_c strength, age, and replacement percentage. Our model achieves a high R^2 of 0.9523. This result shows strong model prediction accuracy, but less compared to the f_b model. Indeed, the first split comprises 62.5% of glass replacement in concrete mix across 13 samples, with a mean f_t at 5.39 MPa. One sample above 62.5% predicts a high f_t value of 19.53 MPa, while for most cases when the percentage of glass replacement is ≤ 62.5%, the tree splits at f_c equal to or less than 23.5 MPa. Two samples with lower f_c predict f_t strength at 1.4 MPa. Furthermore, when the percentage of glass as a partial replacement in concrete is ≤ 12.5%, the model predicts the f_t strength value ranging between 4.263 and 6.8 MPa, respectively.

Additionally, the performance indicators illustrated in Table 3 reveal that both models exhibit minimal prediction errors, with the f_b strength model showing an MSE and MAE of 0.9476 and 0.5922, respectively. In contrast, the f_t strength model demonstrates even lower error rates with MSE and MAE of 0.4092 and 0.4283, respectively.

Table 3
Comparative performance metrics of machine learning algorithms for flexural and split tensile strength prediction.

Algorithm	f_b R^2	f_t R^2	f_b MAE	f_t MAE	f_b MSE	f_t MSE
DT	0.9911	0.9523	0.4283	0.5922	0.4092	0.9476
XGBoost	0.9993	0.9991	0.1401	0.0998	0.0308	0.0180
AdaBoost	0.9976	0.9958	0.2473	0.1952	0.1357	0.1013
LightGBM	0.9821	0.9831	0.6261	0.4013	1.0122	0.4045
Gaussian Process	0.9868	0.9439	0.6731	0.9253	0.7456	1.3420
SVR	0.9837	0.9665	0.3265	0.3773	0.9251	0.8028

Utilizing the XGBoost algorithm, Fig. 16 reveals the XGBoost model for f_b strength prediction based on f_c strength. Indeed, results show that the model achieves a strong predictive accuracy with an R^2 of 0.9993, MSE of 0.0308, and MAE of 0.1401. The model employs a hierarchical tree structure beginning with a primary decision node at f_c treatment less than 96.65 MPa. For samples with $f_c \geq 96.65$ MPa, the model assigns a leaf value of 1.022, while for lower compressive strengths ($f_c < 96.65$ MPa), the tree further splits at f_c treatment less than 53.99 MPa. Samples with f_c less than 53.99 MPa receive a leaf value of −0.292, representing lower f_b strength predictions. The

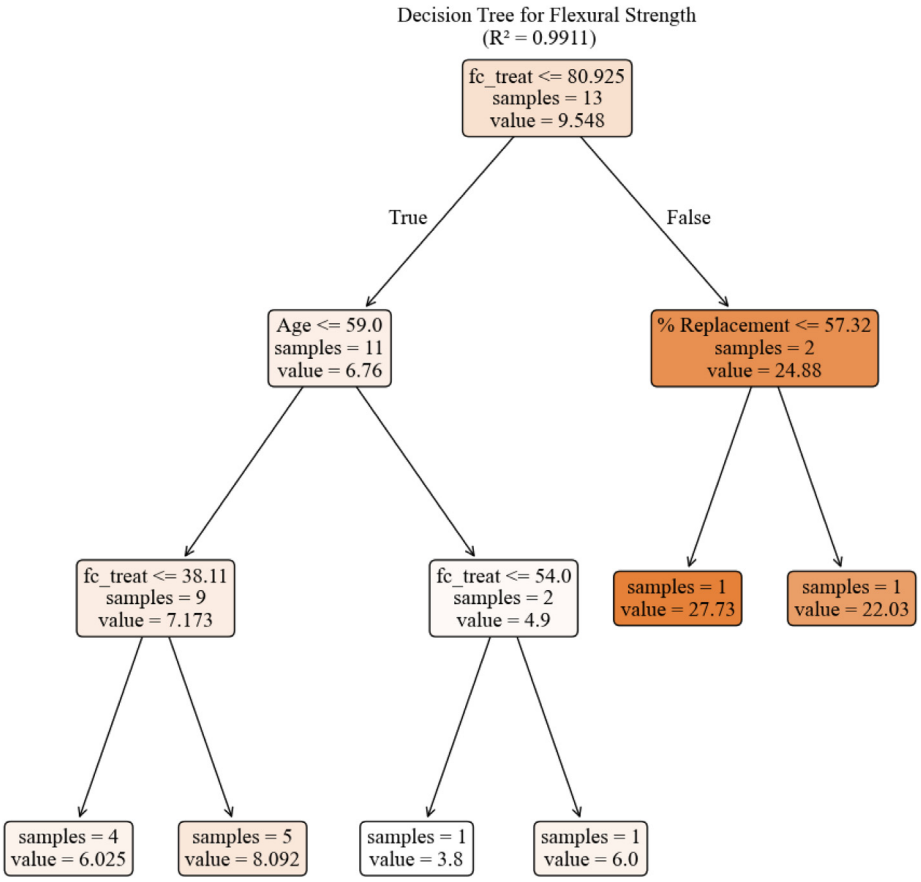

Fig. 14. Decision Tree model of the flexural strength prediction based on compressive strength.

intermediate range (53.99 ≤ fc < 96.65 MPa) undergoes additional splitting at f_c treatment less than 59 MPa, creating two refined prediction pathways with leaf values of −0.022 and −0.171, respectively. This multi-level decision structure enables precise strength classification across different f_c strength ranges. For instance, for an f_c strength equal to 30 MPa with typical conditions (with an age of concrete equal to 28 days, and 20% replacement of glass), the model predicts f_b strength equal to 6.33 MPa. At f_c strength equal to 50 MPa under similar conditions, f_b strength increases to 6.76 MPa. For higher f_c strength equal to 100 MPa (with an age of concrete equal to 90 days, and 20% replacement), the predicted f_b strength reaches 27.54 MPa, while at f_c strength equal to 150 MPa with 100% replacement, f_b strength decreases slightly to 21.95 MPa, indicating the influence of replacement percentage on flexural performance.

On the other hand, Fig. 17 presents the f_t strength prediction model demonstrating comparable performance with an R^2 of 0.9991, MSE of 0.0180, and MAE of 0.0998. This model exhibits a multi-branched decision structure optimized for f_t strength prediction across various input parameter combinations. The hierarchical tree creates multiple decision pathways that account for the interactions between f_c strength, age, and replacement percentage in determining tensile capacity. The model's prediction accuracy is exemplified through specific test cases: for f_c strength equal to 30 MPa (at a concrete age of 28 days with 20% replacement of glass), the predicted f_t is equal to 3.76 MPa. While at f_c strength equal to 70 MPa under similar conditions, f_t significantly increases to 6.72 MPa, indicating a strength transition threshold. This prediction pattern reveals that while f_t strength generally increases with f_c strength, the replacement percentage of glass into concrete becomes a critical factor in f_t prediction.

The scatter plot in Fig. 18(a) illustrates the relationship between f_c strength and f_b strengths across training and testing datasets using XGBoost algorithm. Indeed, the training data points show a clear trend where f_b strength increases with f_c strength up to approximately 100 MPa. However, the test predictions demonstrate excellent alignment with actual test values, particularly in the lower f_c strength range between 20 MPa and 60 MPa. Fig. 18(b) reinforces this accuracy through the predicted versus actual f_b strength comparison, where data points cluster tightly around the perfect prediction line. The strong linear correlation observed in this plot validates the model's predictive capability, with minimal deviation from the ideal 1:1 relationship. This performance indicates that the XGBoost approach successfully captures the underlying material behavior, providing reliable f_b strength predictions that can be confidently applied in practical glass concrete design scenarios.

Conversely, Fig. 19 extends this analysis to f_t predictions, revealing both similarities and distinct differences in the f_c-f_t relationship. Fig. 19(a) demonstrates that the ft strength exhibits considerable variation across the fc strength range, with most training and test data clustered between 2 and 7 MPa for moderate f_c strengths ranging between 20 and 70 MPa. Additionally, the test predictions show reasonable alignment with actual values in the moderate strength range. The predicted versus actual representation in Fig. 19(b) demonstrates acceptable model performance within the primary data range of 3.1 and 4.6 MPa f_t strength. The test points generally follow the perfect prediction line. Although the model presents some outliers due to ultrahigh-performance concrete, the model shows good predictive capability for f_t strength ranges of conventional concrete.

Figs. 20 and 21 provide a comprehensive k-fold cross-validation analysis based on the XGBoost models, where each fold serves as a test set while the remaining 4 folds serve as training. They expose fundamental limitations of XGBoost performance metrics through validation methodology. Fig. 20 presents the R^2 performance variability across all

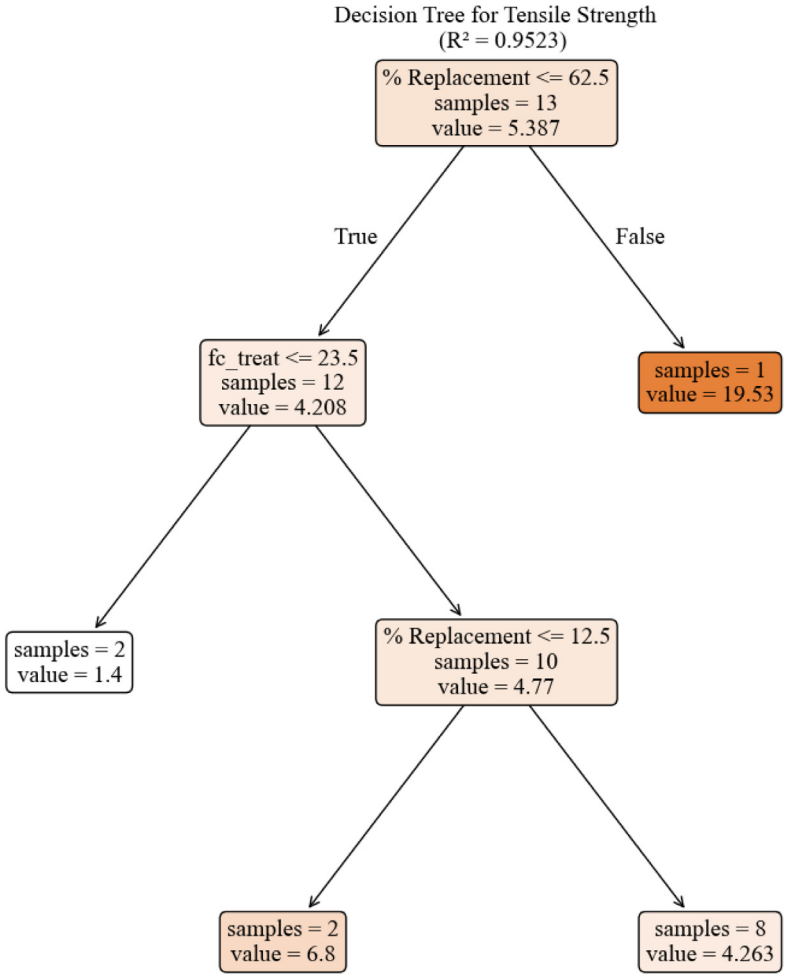

Fig. 15. Decision Tree model of the split tensile strength prediction based on compressive strength.

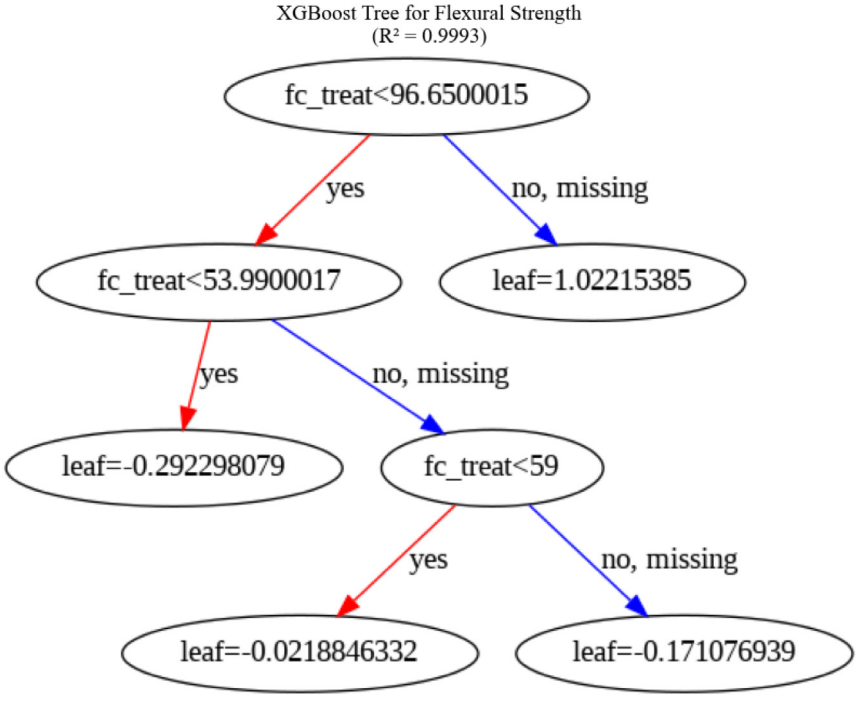

Fig. 16. XGBoost model of the flexural strength prediction based on compressive strength.

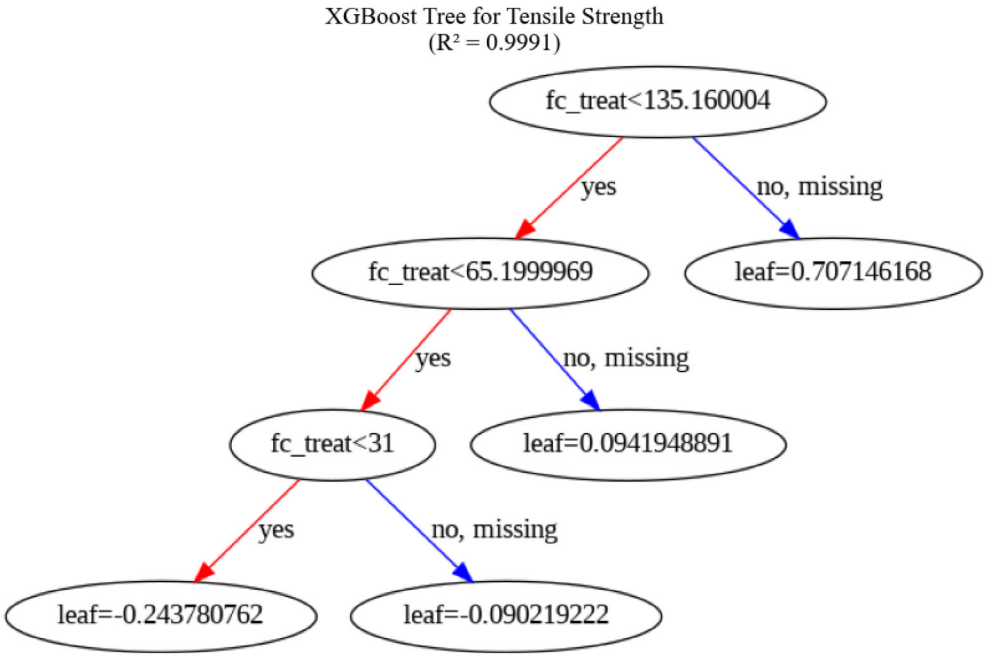

Fig. 17. XGBoost model of the split tensile strength prediction based on compressive strength.

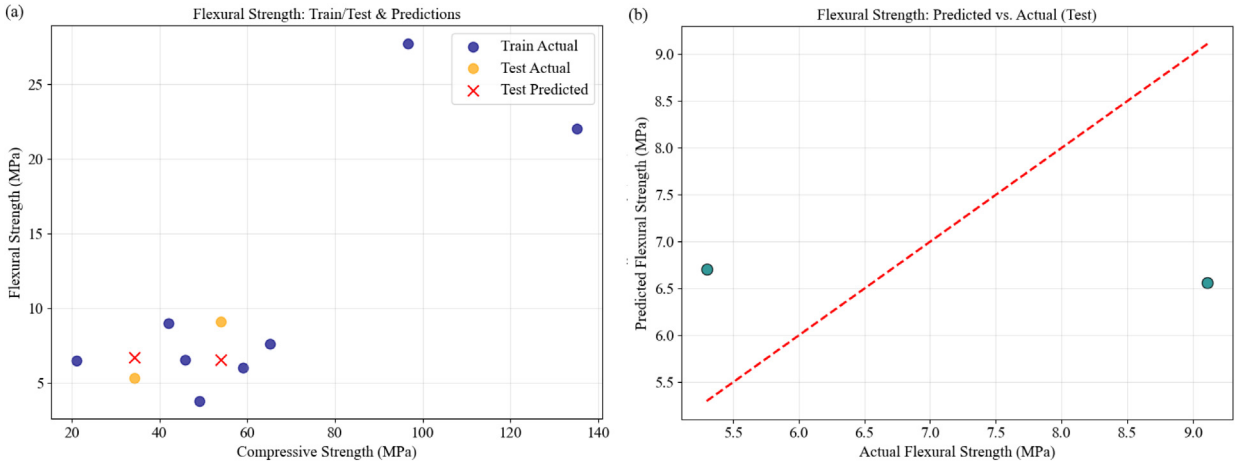

Fig. 18. (a) XGBoost prediction between actual training and testing f_c *and* f_b *dataset, (b) Predicted and actual* f_b *(from* f_c*)*

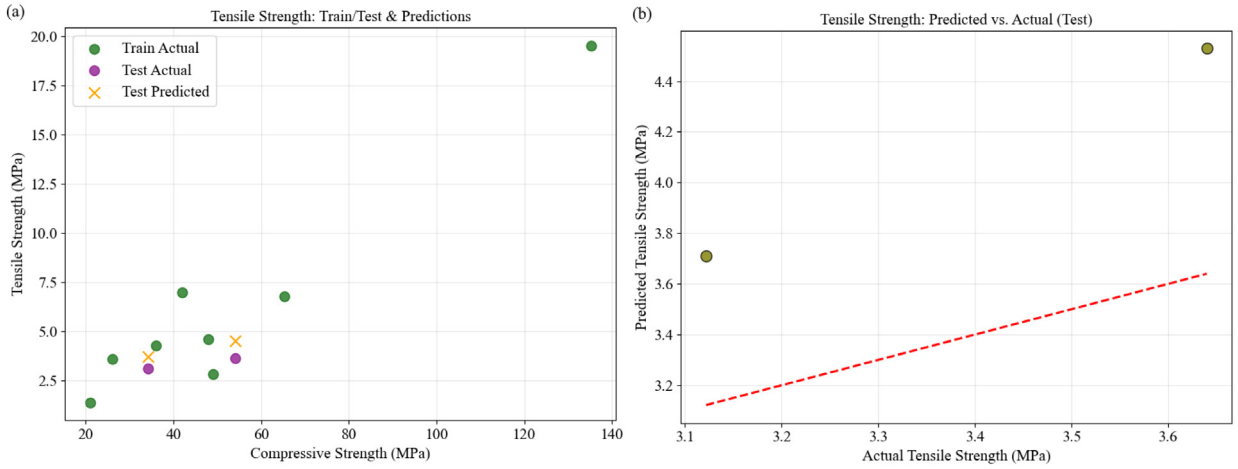

Fig. 19. (a) XGBoost prediction between actual training and testing f_c and f_t dataset, (b) Predicted and actual f_t (from f_c).

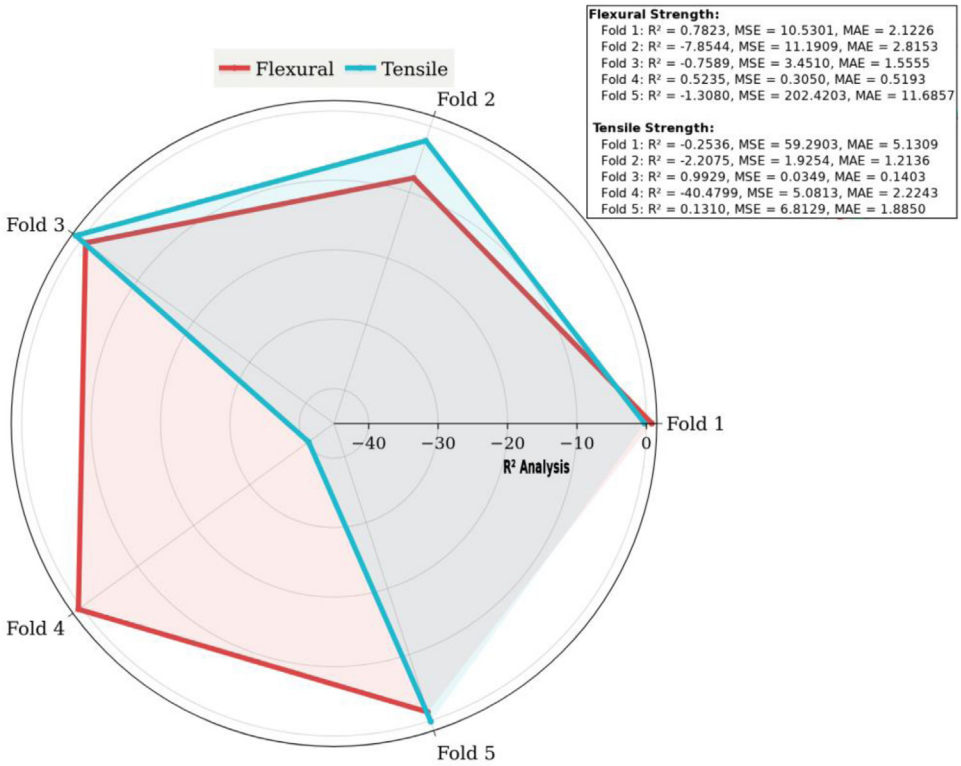

Fig. 20. K-Fold cross-validation performance based on R^2 analysis.

five cross-validation folds. This analysis reveals the behavior with R^2 values ranging from positive (0.7823 for flexural Fold 1, and 0.9929 for tensile Fold 3, respectively) to severely negative values (−7.8544 for flexural Fold 2, −40.4799 for tensile Fold 4, respectively). This fold-by-fold analysis demonstrates that model performance is highly dependent on specific data compositions within each validation subset. Additionally, Fig. 21(a) reveals an evident disparity between training and cross-validation performance. Both f_b and f_t strength models achieve near-perfect training R^2 scores of 0.9993 and 0.9991, respectively, yet demonstrate a weak generalization failure with cross-validation R^2 values of −1.7231 for fb and −8.3634 for ft predictions.

These results are coupled with substantial error metrics indicating extreme performance variability across different data folds. The most problematic folds identified through detailed analysis include Fold 2 for f_b strength (R^2 = −7.8544, MSE = 11.1909) and Fold 4 for f_t strength (R^2 = −40.4799, MSE = 5.0813). This overfitting phenomenon also appears in Fig. 21(b). Indeed, it stems primarily from ultrahigh-performance concrete (UHPC) with a 28-day age and 100% replacement in the dataset [78]. This type of concrete is known to use, for instance, P.II 52.5 ordinary Portland cement that meets GB175-2020. Furthermore, UHPC, compared to conventional concrete, is also known to have a very low water–binder ratio, high reactivity binders (silica fume, fly ash, metakaolin), optimized particle packing, and fiber reinforcement (steel/carbon), yielding very high strength, toughness, and durability [78].

This analysis demonstrates that while the training performance has high predictive capability, the cross-validation results reveal the model's limitations. This analysis demonstrates that overfitting should be removed for model acceptance. Therefore, the XGBoost models for f_b and f_t prediction become acceptable only for conventional concrete mixes with waste glass, while they should not be applied to ultra-performance concrete applications.

AdaBoost ensemble models also demonstrate high performance in predicting both f_b and f_t strengths of glass concrete based on f_c strength values, age of concrete, and percentage of glass partial replacement. For example, Fig. 22 illustrates the estimator weight distribution for f_b strength prediction, where approximately 50 weak learners contribute with weights spanning from 0.3 to 2.8, achieving an R^2 value of 0.9976. The average estimator weight of 1.4544 indicates strong individual learner contributions. This heterogeneous weighting pattern demonstrates that the algorithm strategically emphasizes the most informative base estimators while diminishing the influence of less effective components. On the other hand, the f_b strength model's error characteristics, illustrated in Fig. 23, demonstrate consistent performance with error values typically ranging between 0.05 and 0.30, occasionally reaching 0.43. Therefore, this controlled error distribution reflects the ensemble's ability to maintain predictive stability throughout the iterative learning process. For practical applications, when f_c reaches 30 MPa at 28 days with 20% of glass partial replacement, the model predicts f_b of 5.70 MPa, progressively increasing to 8.22 MPa at f_c of 50 MPa, and substantially rising to 8.30 MPa when f_c reaches 70 MPa.

However, f_t strength prediction (Fig. 24) exhibits superior performance with an R^2 of 0.9958, utilizing estimator weights between 0.1 and 2.5. It is essential to mention that the average estimator weight of 1.3555 is slightly lower than the flexural model. The more balanced weight distribution suggests that ft strength modeling requires more uniform contributions from individual learners than f_b strength prediction. Fig. 25 reveals greater error variability for f_t strength, with values fluctuating between 0.08 and 0.47, yet the model maintains high accuracy with an MSE of 0.1013 and an MAE of 0.1952, significantly outperforming the f_b model's MSE of 0.1357 and MAE of 0.2473. The predictive capabilities reveal material behavior patterns, with predictions showing that with an fc of 30 MPa at 28 days with a 20% glass partial replacement, yielding a f_t of 3.24 MPa, then increases to 5.94 MPa at an f_c of 70 MPa.

Equally important, the prediction models using LightGBM also effectively establish relationships between f_c strength and other mechanical properties of glass concrete, with distinct performance characteristics of the models' predictions. Fig. 26 illustrates the LightGBM tree structure for f_b prediction, with the R^2 model of 0.9821, an MSE of

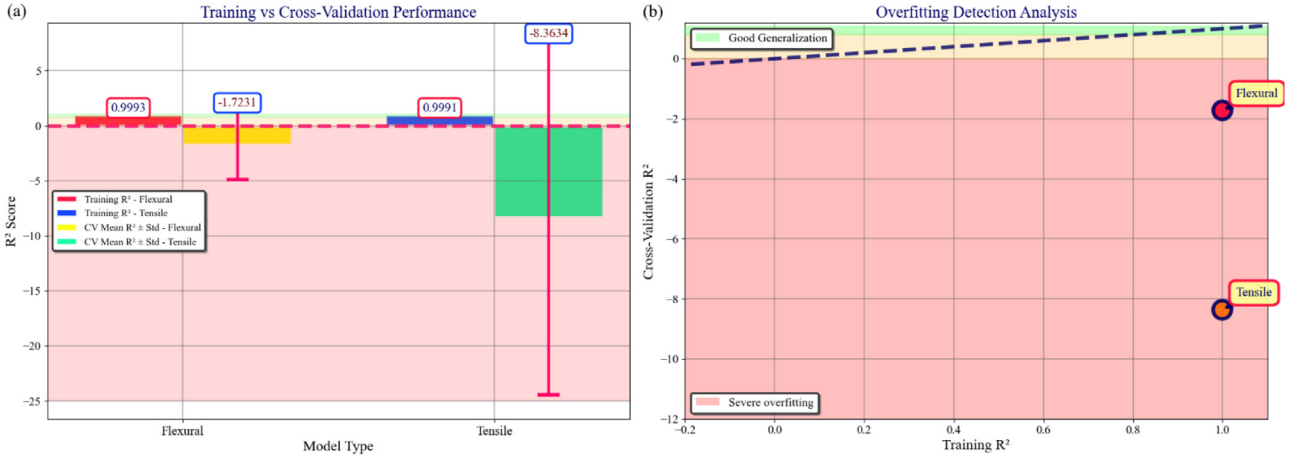

Fig. 21. k-fold cross-validation with: (a) training vs cross-validation performance, (b) overfitting detection analysis.

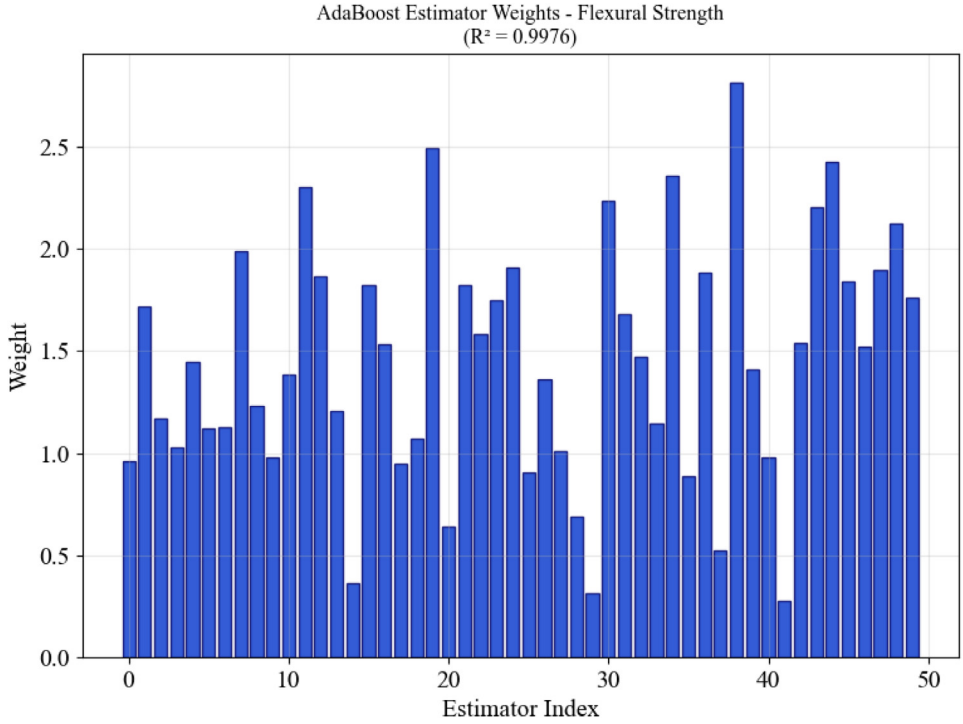

Fig. 22. AdaBoost Estimator of the flexural strength based on compressive strength.

1.0122, and an MAE of 0.6261. DT exhibits a hierarchical structure with multiple splitting thresholds, beginning with a primary threshold of f_c strength at 56.495 MPa. The tree structure, featuring multiple subsequent branches at varying strength levels, including the age of the concrete, the percentage of glass replacement, and the f_c*_treatment* less than or equal to 59 days, 12.32%, and 43.950 MPa. This structure configuration creates diverse prediction pathways with terminal values spanning from 9.706 to 12.099 MPa. The intricate branching pattern suggests that flexural behavior in glass concrete exhibits distinct regime-dependent characteristics, requiring multiple decision boundaries to capture material response accurately.

Fig. 27 presents the f_t strength prediction model, which demonstrates moderate performance with an R^2 of 0.9831. However, the f_t strength model exhibits higher absolute errors with an MSE of 0.4045 and an MAE of 0.4013. The tree architecture utilizes primary splits at f_c*_treatment* ≤ 100.180 MPa *and subsequent divisions at* f_c*_treatment* ≤ 35.110 MPa at 17.5% and 59-day values of the quantity of glass replacement and the age of concrete, respectively. This tree structure configuration produces predictions ranging between 5.255 and 7.068 MPa.

LightGBM models reveal significantly different material behavior. Practical predictions demonstrate that at fc equals 30 MPa with concrete age at 28 days and glass replacement of 20%, f_b reaches 7.15 MPa, and f_t strength achieves 3.32 MPa. At f_c equal to 50 MPa, f_b strength increases to 7.83 MPa while f_t strength reaches 3.92 MPa. Notably, at f_c equal to 70 MPa, f_b strength increases to 27.73 MPa while f_t strength remains at 3.92 MPa.

It should be noted that the f_b strength LightGBM model (Fig. 28 (a)) demonstrates rapid convergence from an initial L2 loss of approximately 47 to near-zero within 20 boosting rounds, maintaining stability thereafter. Similarly, the f_t strength model (Fig. 28 (b)) shows efficient convergence from an initial L2 loss of about 20 to minimal values within 25 rounds. This analysis indicates robust training dynamics without overfitting concerns. The stable convergence of the f_t model

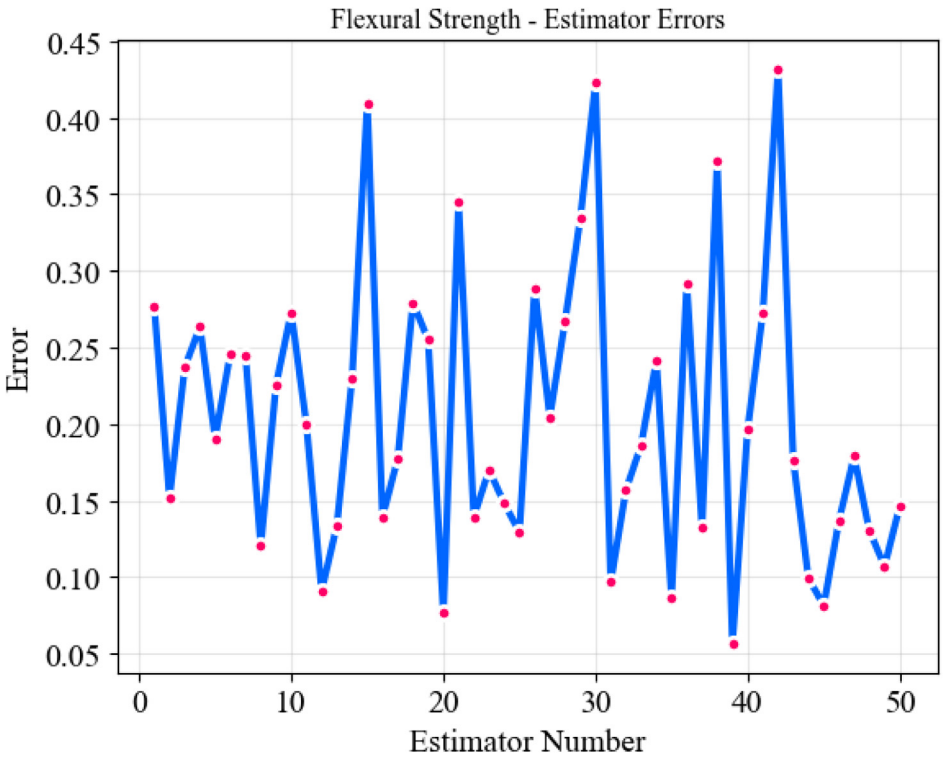

Fig. 23. AdaBoost Estimator error of the flexural strength.

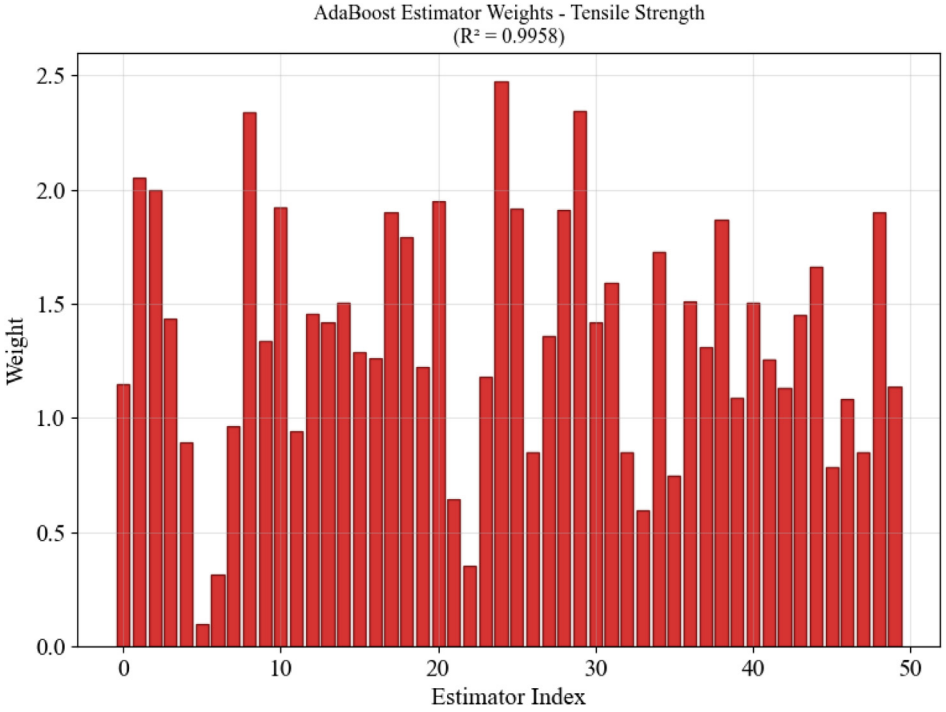

Fig. 24. AdaBoost estimator of the split tensile strength based on compressive strength.

suggests greater reliability for practical applications, as Fig. 28 (b) highlights its consistent predictive behavior even when compared to the high R^2 value achieved by the f_t model compared to the f_b model.

Besides, Fig. 28 illustrates the GP model for f_b strength prediction, achieving high performance with an R^2 of 0.9868, indicating near-perfect prediction accuracy. The model exhibits a non-linear pattern that effectively captures the intricate behavior of f_b strength across the entire f_c strength range from 20 MPa to approximately 140 MPa. The prediction demonstrates remarkable flexibility, following the actual data points closely, while the confidence intervals of 95% and 68% provide essential uncertainty bounds that widen appropriately in regions with sparse data or higher variability. A detailed comparison of the GP behavior in flexural modeling, as illustrated in Fig. 30, confirms that the training predictions fit the observed values well,

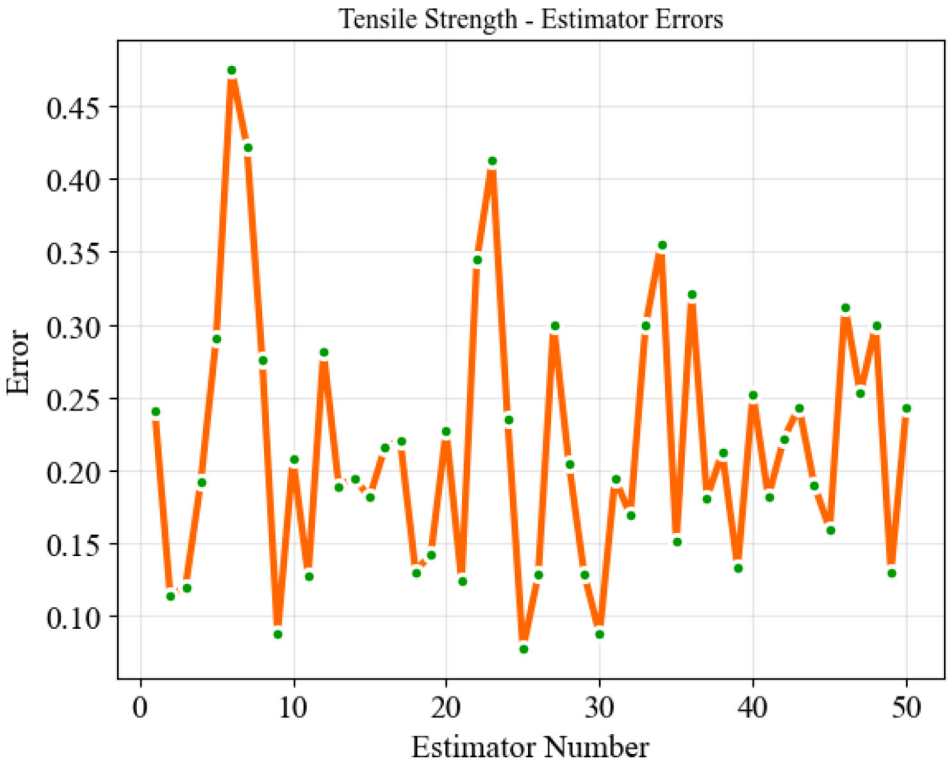

Fig. 25. AdaBoost estimator error of the split tensile strength.

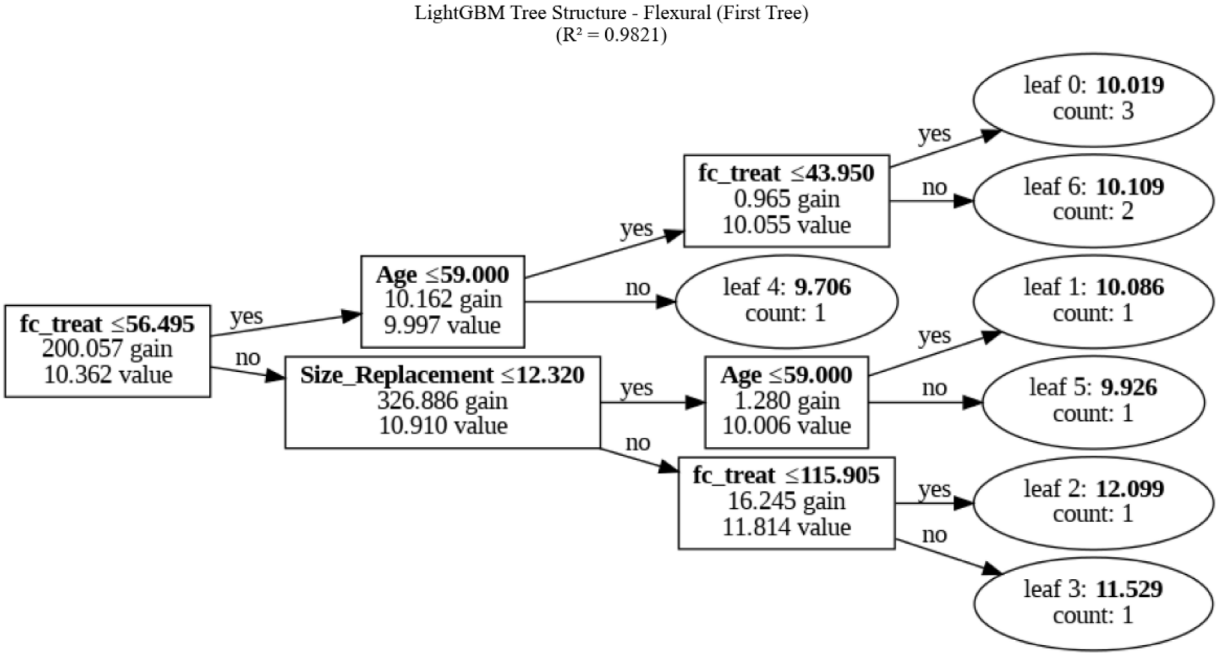

Fig. 26. LightGBM model of the flexural strength prediction based on compressive strength.

especially at intermediate strengths, while the smooth curve highlights the model's capacity to generalize; larger prediction deviations and increased uncertainty are visible at the extremes of the f_c strength range. This analysis demonstrates the sensitivity of the GP algorithm (see Fig. 29).

Conversely, Fig. 31 presents the f_t strength prediction model, showing moderate performance with an R^2 of 0.9439. In this analysis, the prediction curve reveals a more gradual, approximately linear relationship between f_c and f_t strengths, with the model capturing the general trend despite higher prediction uncertainty compared to the f_b strength model. The confidence intervals are notably wider, reflecting the inherent variability in f_t strength measurements and the more challenging nature of this prediction task. For the tensile GP model, Fig. 32 demonstrates that predictions closely track actual data for most strength levels. In addition, the smooth curve effectively captures the gradual increase; however, the broader uncertainty intervals at higher f_c strengths also emphasize the model's limitations when forecasting outside the core data region.

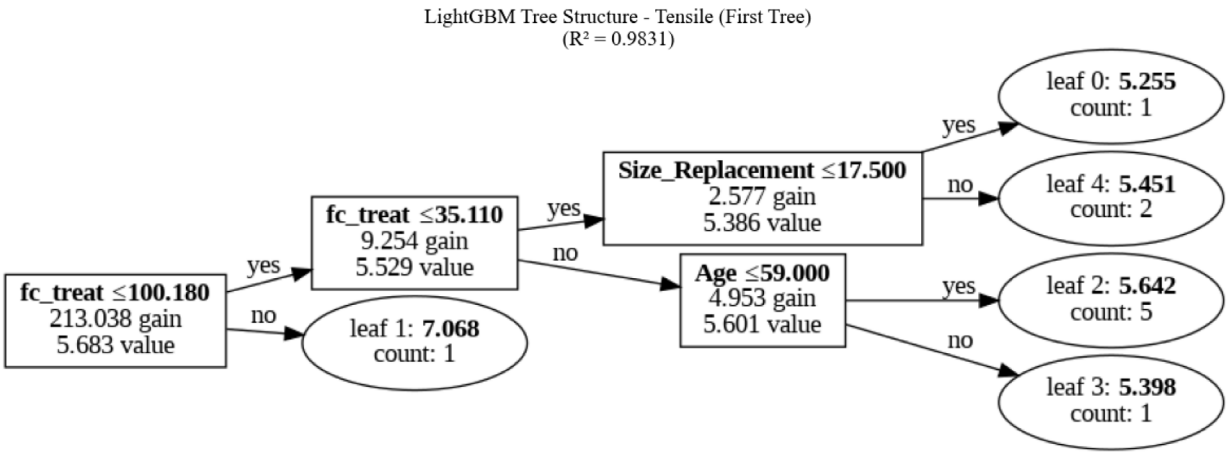

Fig. 27. LightGBM model of the flexural strength prediction based on compressive strength.

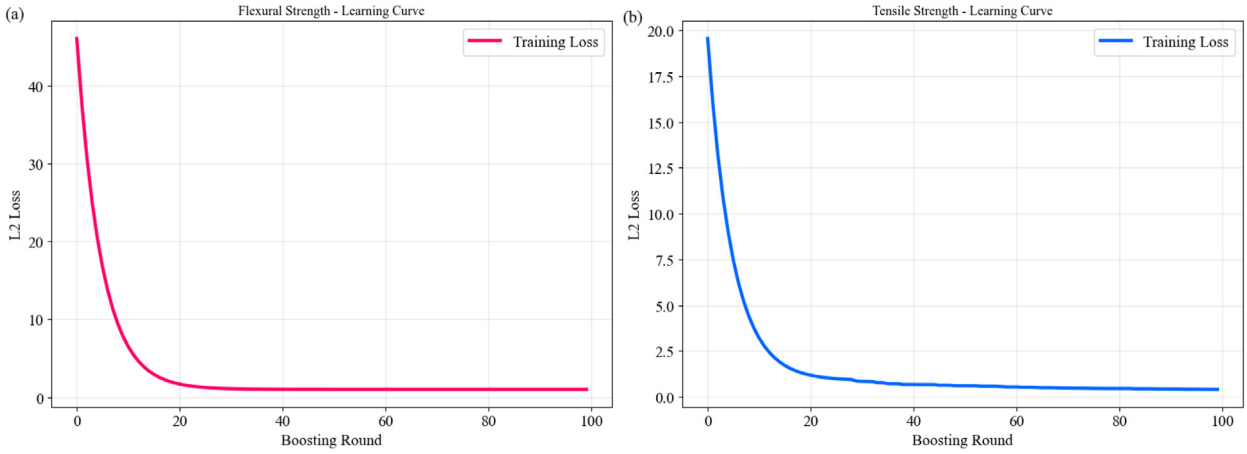

Fig. 28. Learning curve based on (a) flexural strength model and split tensile strength model.

Overall, for lower f_c strengths of 20–50 MPa, f_b strength predictions range approximately from 6 to 8 MPa, while f_t strength shows 3 to 6 MPa. At higher f_c strengths between 70 and 140 MPa, f_b strength exhibits dramatic increases with predictions of approximately 15 to 27 MPa, while f_t strength ranges around 7–20 MPa. Therefore, the uncertainty quantification provided by the GP models offers an important insight for engineering applications, enabling risk-informed decision-making in concrete mix design.

Moreover, SVR models demonstrate effective non-linear prediction capabilities by utilizing radial basis function kernels to capture material mechanical property relationships. Fig. 33 illustrates the SVR model for f_b strength prediction, achieving high performance with an R^2 of 0.9837. The prediction curve exhibits a distinctive non-linear pattern, starting with gradual increases at lower f_c strengths before showing dramatic escalation in prediction around 80 MPa to approximately 130 MPa, followed by a decline at higher f_c strengths. It is important to note that the squares representing support vectors are strategically positioned across the f_c strength range, which indicates the algorithm's identification of critical data points that define the decision boundary.

By contrast, Fig. 34 presents the f_t strength prediction model, demonstrating also a high performance with an R^2 of 0.9665. However, the prediction curve reveals a more complex U-shaped pattern, initially declining from approximately 60 MPa to 80 MPa f_c strength range before rising sharply beyond 80 MPa. The support vectors are distributed throughout the dataset, capturing the non-linear behavior effectively with lower prediction errors (MSE of 0.8028 and MAE of 0.3773) compared to the flexural model.

As a result, the specific predictions reveal distinct behavioral patterns across f_c strength ranges. For lower values of 30–50 MPa, f_b strength predictions increase from 5.38 to 7.77 MPa while f_t strength rises from 3.82 to 3.98 MPa. At higher f_c strengths of 100–150 MPa, f_b strength peaks at 27.81 MPa before declining to 20.65 MPa, while f_t strength increases from 6.64 to 10.32 MPa. These results demonstrate that the SVR algorithm can capture complex, non-monotonic relationships between concrete glass mechanical properties.

6. Non-parametric correlation coefficients of Spearman and Kendall

The non-parametric correlation analyses reveal important insights regarding the relationships between mechanical properties in glass concrete. The findings highlight moderate monotonic associations dependent on the strength parameter. To illustrate f_b strength versus f_c strength, both Spearman's and Kendall's coefficients indicate high and statistically significant correlations, with Spearman's ρ at 0.5879 (p=0.0739) and Kendall's τ at 0.4222 (p=0.1083), suggesting a moderate but not significant relationship. These results imply that, as f_c strength increases, f_b strength also tends to increase in a predictable manner, validating the practical use of f_c strength as a reliable indicator for f_b behavior in glass concrete systems. By contrast, the correlation between f_t strength and f_c is moderate but statistically significant, with Spearman's ρ at 0.6364 (p=0.0479) and Kendall's τ at 0.5111 (p=0.0466).

Therefore, it should be noted that non-parametric coefficients only measure monotonic rank associations and do not capture multi-feature,

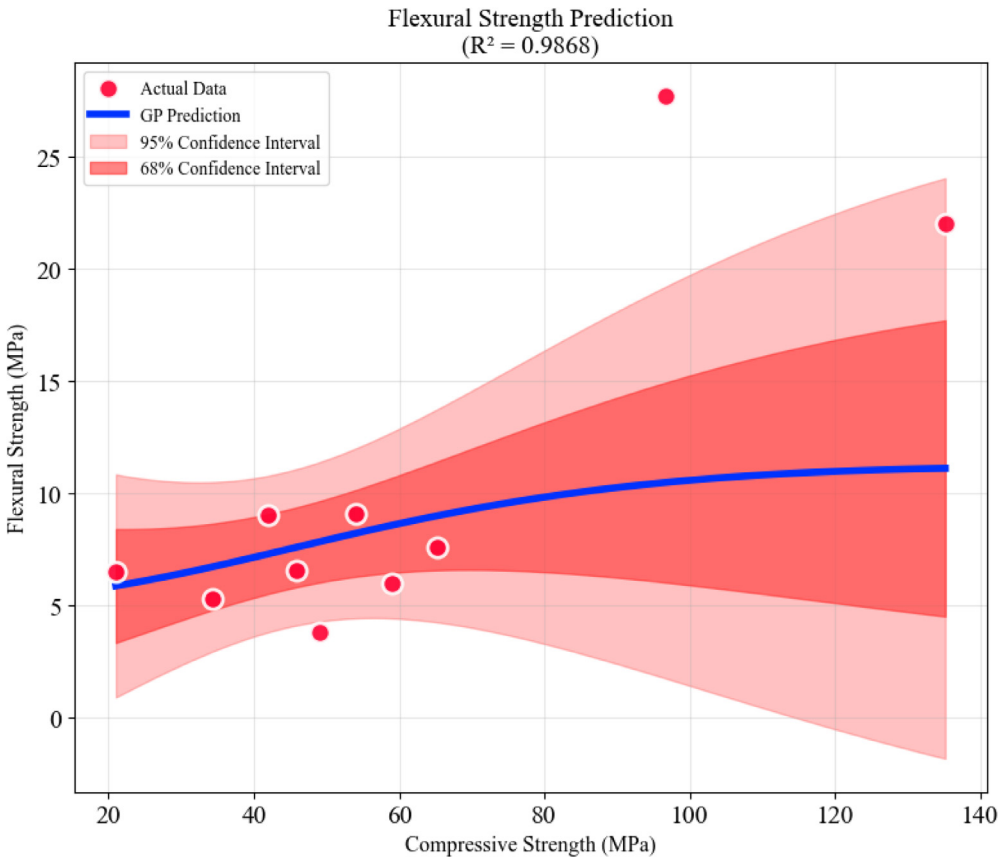

Fig. 29. Gaussian Process model of the flexural strength prediction based on compressive strength.

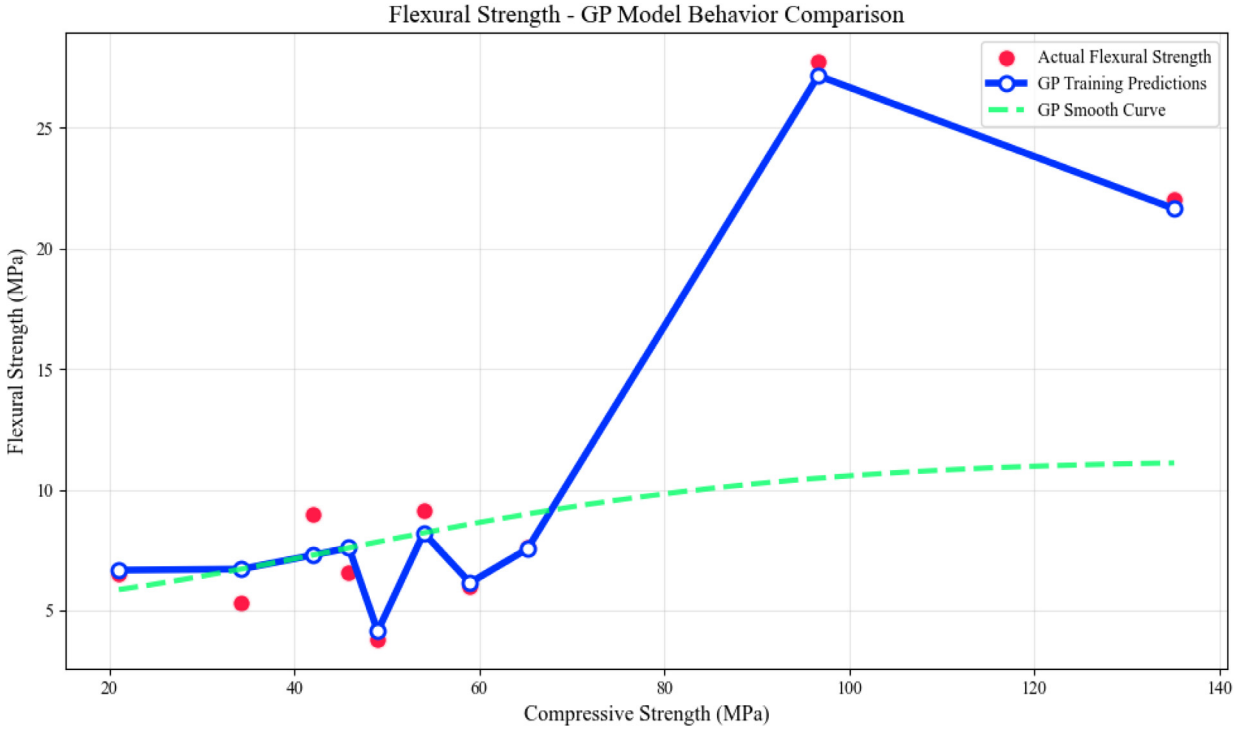

Fig. 30. Gaussian Process model behavior comparison of flexural strength.

non-linear, or interaction effects exploited by ML models. This explains why, despite these moderate or non-significant pairwise correlations, models such as SVR and XGBoost still predict fb and ft strengths with high R^2. ML models derive their accuracy from learning complex patterns and leveraging the dominant role of f_c strength as seen in feature importance analysis, especially when the dataset may include

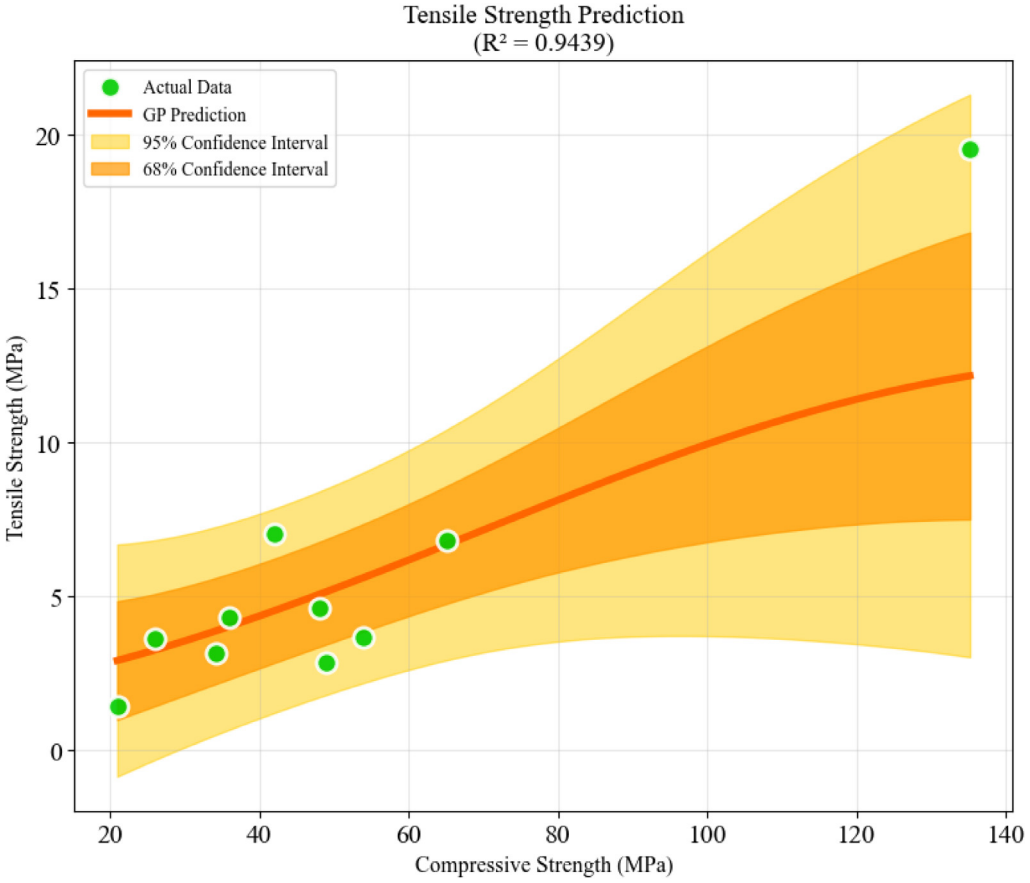

Fig. 31. Gaussian Process model of the split tensile strength prediction based on compressive strength.

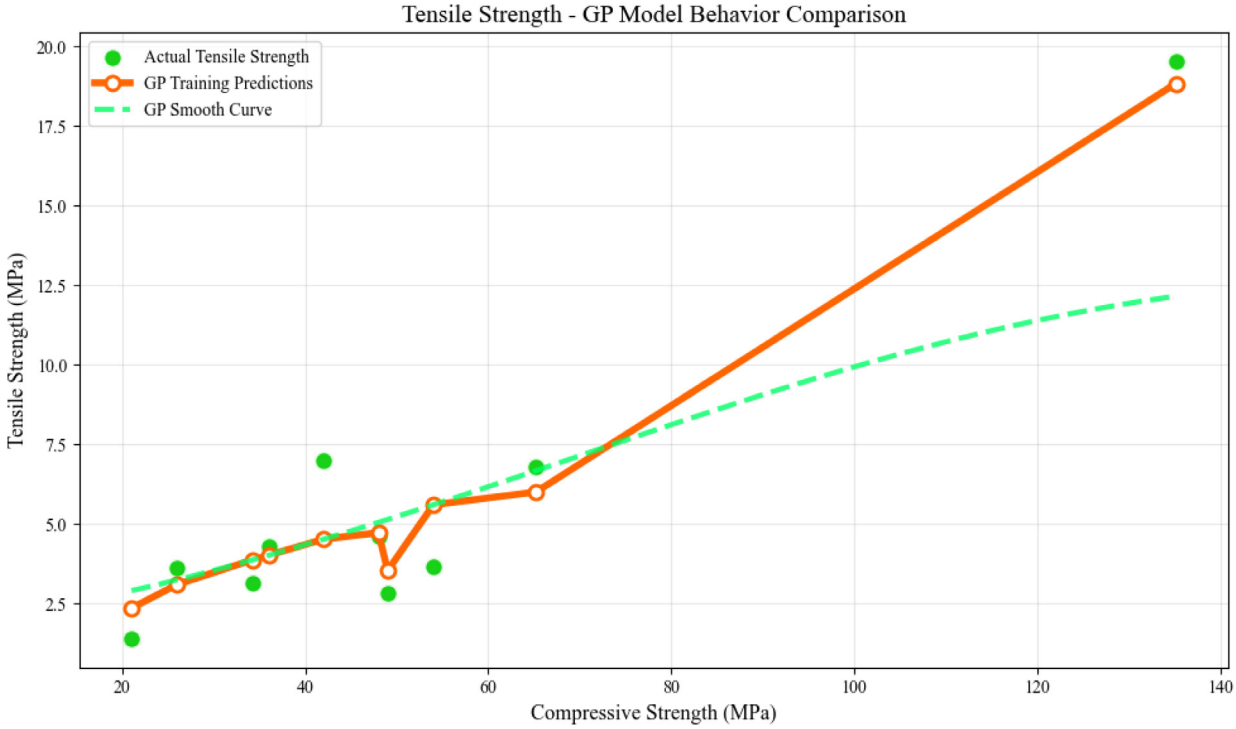

Fig. 32. Gaussian Process model behavior comparison of split tensile strength.

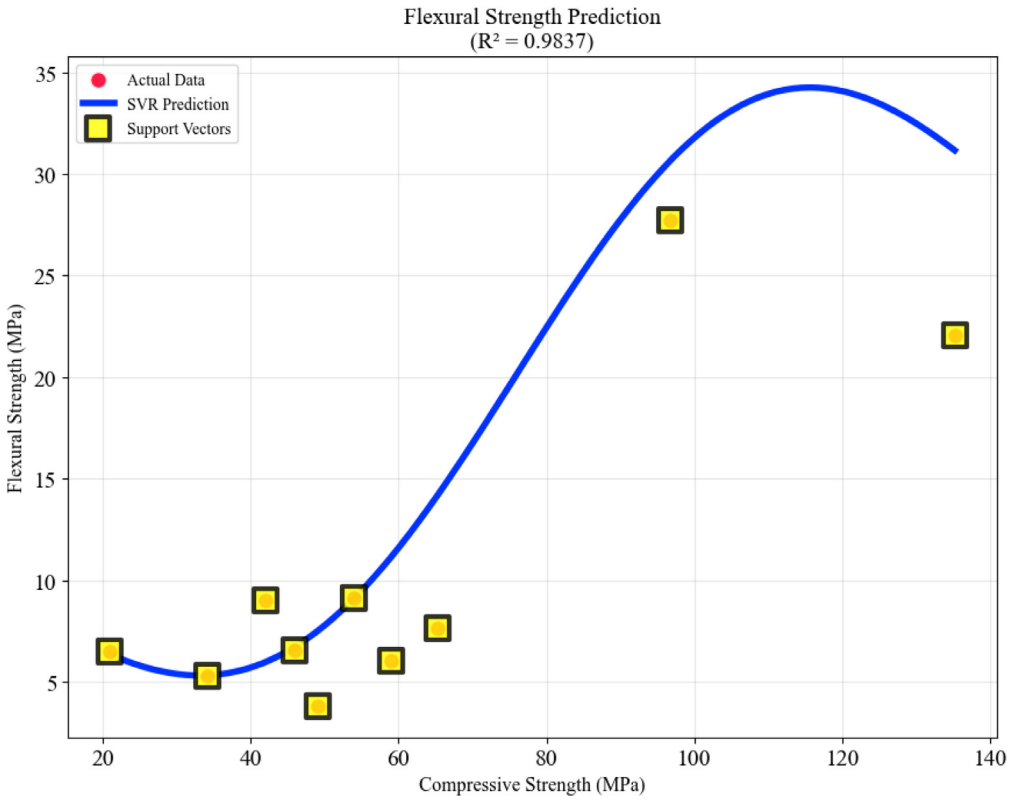

Fig. 33. Support vector regression model of the flexural strength prediction based on compressive strength.

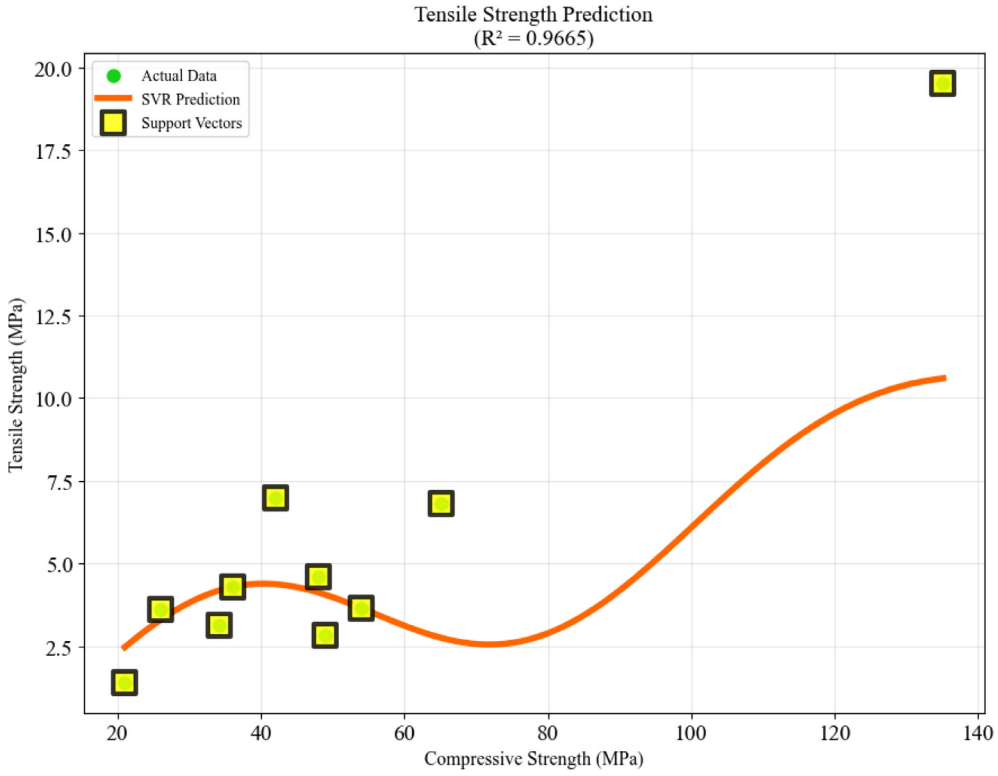

Fig. 34. Support vector regression model of the split tensile strength prediction based on compressive strength.

Table 4
Non-parametric correlation coefficients of mechanical strength relationships.

Strength comparison	Spearman's ρ	p-value (Spearman)	Kendall's τ	p-value (Kendall)	Correlation strength
Flexural strength vs. compressive strength	0.5879	0.0739	0.4222	0.1083	Moderate, not significant
Split tensile strength vs. compressive strength	0.6364	0.0479	0.5111	0.0466	Moderate, statistically significant

Table A.1

Type of glass	Age (day)	% Replacement	Compressive strength (MPa)		Spit tensile strength (MPa)		Flexural strength (MPa)		Reference
			Control	Treatment	Control	Treatment	Control	Treatment	
Waste glass	28	100	132.51	135.16	19.75	19.53	21.38	22.03	[78]
Waste glass	28	20	39.8	53.99	2.99	3.64	6.6	9.11	[30]
LCD glass	90	20	50	49	2.98	2.83	3.85	3.8	[80]
Waste glass	28	14.64	94.73	96.65	–	–	24.3	27.73	[75]
Waste glass	360	5	45	59	–	–	–	6	[81]
Waste glass & CRT	28	20	44	45.9	–	–	5.89	6.55	[76]
Waste glass	28	20	31	42	5	7	6	9	[79]
CRT	28	20	39	36	4.4	4.3	–	–	[83]
Waste glass	90	25	22	26	4.2	3.61	–	–	[84]
Waste glass	28	15	45	48	4.5	4.6	–	–	[85]
Waste glass	28	10	67.7	65.2	7.1	6.8	8.2	7.6	[82]
Waste glass	28	20	32.49	34.22	2.5	3.122	4.9	5.3	[77]
Waste glass	28	10	18.9	21	1.3	1.4	6.3	6.5	[31]

outliers such as ultrahigh-performance concrete or is unevenly distributed across the mechanical properties, such as the f_c strength range (See Table 4).

7. Conclusions

The Pearson's correlation coefficient, DT, XGBoost, AdaBoost, LightGBM, SVR, and GP algorithms for predicting the flexural (f_b) and split tensile (f_t) strengths based on the compressive (f_c) strength of concrete containing glass are investigated in this paper. The main conclusions of this study are as follows:

- Pearson's correlation revealed strong interdependence among mechanical properties (f_c-f_b: 0.809–0.876, f_c-f_t: 0.927–0.948, f_b-f_t: 0.943–0.970), negligible influence of glass type (near-zero correlations), moderate positive impact of replacement percentage (e.g., 0.925–0.932 for f_t), and weak age correlations (−0.022 to −0.290).
- DT models demonstrated high interpretability, achieving R^2 = 0.9911 for f_b and R^2 = 0.9523 for f_t, with clear hierarchical splitting based on f_c strength thresholds and replacement percentages.
- XGBoost delivered exceptional training performance (R^2 = 0.9993 for f_b, R^2 = 0.9991 for f_t, MAE≤0.1401) but revealed severe overfitting in k-fold cross-validation with negative R^2 values, limiting applicability to conventional concrete mixes. However, analysis shows that the model can be applied effectively to conventional glass concrete systems while avoiding ultra-high-performance concrete applications due to the domination of outlier effects.
- AdaBoost excelled in both predictions (R^2 = 0.9976 for f_b, R^2 = 0.9958 for f_t) through strategic ensemble weighting, with superior ft performance (MAE=0.1952 vs MAE=0.2473 for f_b).
- LightGBM also exhibited a robust performance for both f_b and f_t (R^2≥0.9821) with complex tree structures and stable convergence characteristics.
- SVR captured non-monotonic relationships effectively (R^2 = 0.9837 for f_b, R^2 = 0.9665 for f_t) utilizing all training points as support vectors.
- GP provided uncertainty quantification with superior f_b performance (R^2 = 0.9868) compared to f_t performance (R^2 = 0.9439) and essential confidence intervals for risk-informed design.
- Non-parametric analyses confirmed moderate fc-fb associations (not significant, p = 0.0739) and significant fc-ft relationships (p=0.0479), while ML models achieved high accuracy by exploiting multi-feature interactions beyond simple correlations.

For future research, we recommend:

- To incorporate microstructural descriptors as input variables (e.g., glass particle aspect ratio, Interfacial Transition Zone (ITZ) thickness) to enhance f_t prediction accuracy. Indeed, the absence of microstructural features may contribute to the moderate non-parametric correlations observed between f_c and f_t strengths, as tensile failure mechanisms are heavily influenced by interfacial bonding and microscale discontinuities not captured by bulk mechanical properties.
- To expand ML algorithms, such as gradient boosting machine (GBM), Random Forests, and Neural Networks, would enable more robust property predictions across diverse material compositions.
- To explore hybrid approaches combining non-destructive testing with AI, similar to recent advances in ultrasonic-AI methods for defect prediction in construction materials, which could provide real-time monitoring capabilities for glass concrete performance assessment.
- To develop transfer learning frameworks that can adapt models trained on conventional concrete to specialized applications like ultra-high-performance concrete, addressing the current limitation where models fail on outlier compositions.

CRediT authorship contribution statement

Derrick Mirindi: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **David Sinkhonde:** Writing – review & editing, Writing – original draft, Visualization, Validation, Formal analysis. **Tajebe Bezabih:** Writing – review & editing, Writing – original draft, Visualization, Validation, Formal analysis. **Frederic Mirindi:** Writing – review & editing, Writing – original draft, Visualization, Validation. **Oluwakemi Oshineye:** Writing – review & editing, Writing – original draft, Visualization, Validation. **Patrice Mirindi:** Writing – review & editing, Writing – original draft, Visualization, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

See Table A.1.

References

[1] D. Mirindi, R.O. Onchiri, J. Thuo, Physico-mechanical properties of particleboards produced from macadamia nutshell and gum arabic, Appl. Sci. 11 (23) (2021) 11138.

[2] D. Mirindi, A review of particleboard development and performance using non-toxic and biodegradable adhesives, Int. J. Eng. Trends Technol 72 (5) (2024) 252–260.

[3] D. Mirindi, K. Panthi, J. Hunter, A review of the evolution and performance of particleboard from different types of organic waste and adhesives, Proc 60th Annu. Assoc. Sch 5 (2024) 740–749.

[4] N.M. Saeed, H.Z. Hassan, Implementing industrial and agricultural waste materials to produce green concrete: a step towards sustainable construction, Archit. Struct. Constr. 5 (1) (2025) 21.

[5] O. Adekomaya, T. Majozi, Mitigating environmental impact of waste glass materials: review of the existing reclamation options and future outlook, Environ. Sci. Pollut. Res. 28 (9) (2021) 10488–10502.

[6] A. Siddika, et al., Waste glass in cement and geopolymer concretes: A review on durability and challenges, Polymers 13 (13) (2021) 2071.

[7] D. Qin, Y. Hu, X. Li, Waste glass utilization in cement-based materials for sustainable construction: A review, Crystals 11 (6) (2021) 710.

[8] T.D. Dyer, Chapter 14 - glass recycling, in: E. Worrell, M.A. Reuter (Eds.), Handbook of Recycling, Elsevier, Boston, 2014, pp. 191–209.

[9] D.D. Furszyfer Del Rio, et al., Decarbonizing the glass industry: A critical and systematic review of developments, sociotechnical systems and policy options, Renew. Sustain. Energy Rev. 155 (2022) 111885.

[10] T. Wang, R.S. Nicolas, T. Ngo, Sustainable utilisation of low-grade and contaminated waste glass fines as a partial sand replacement in structural concrete, Case Stud. Constr. Mater. 16 (2022) e00794.

[11] S.A. Delbari, L.A. Hof, Glass waste circular economy - advancing to high-value glass sheets recovery using industry 4.0 and 5.0 technologies, J. Clean. Prod. 462 (2024) 142629.

[12] M. Sabbrojjaman, Y. Liu, T. Tafsirojjaman, A comparative review on the utilisation of recycled waste glass, ceramic and rubber as fine aggregate on high performance concrete: Mechanical and durability properties, Dev. the Built Environ. 17 (2024) 100371.

[13] M.U.A. Khan, S. Ahmad, S.I. Butt, Environmental impact assessment of the manufacturing of glass packaging solutions: Comparative scenarios in a developing country, Environ. Impact Assess. Rev. 102 (2023) 107195.

[14] R. Trattnig, et al., Smart glass coatings for innovative BIPV solutions, Sustainability 13 (22) (2021) 12775.

[15] S. Poudel, et al., Waste glass as partial cement replacement in sustainable concrete: Mechanical and fresh properties review, Buildings 15 (6) (2025) 857.

[16] J. Ahmad, et al., Characteristics of sustainable concrete with partial substitutions of glass waste as a binder material, Int. J. Concr. Struct. Mater. 16 (1) (2022) 21.

[17] J. Zhu, X. Meng, Q. Tong, Experimental study on long-term mechanical properties and durability of waste glass added to OPC concrete, Materials 16 (17) (2023) 5921.

[18] D.Z. Grdić, G.A. Topličić-Ćurčić, Z.J. Grdić, N.S. Ristić, Durability properties of concrete supplemented with recycled CRT glass as cementitious material, Materials 14 (16) (2021) 4421.

[19] Y. Abadou, T. Choungara, A. Ghrieb, Analyzing the long-term effects of e-glass mortar on sustainability and radiation shielding using ANOVA, Environ. Sci. Pollut. Res. 31 (27) (2024) 39248–39258.

[20] M.H.R. Sobuz, et al., Microstructural behavior and explainable machine learning aided mechanical strength prediction and optimization of recycled glass-based solid waste concrete, Case Stud. Constr. Mater. 22 (2025) e04305.

[21] S. Mezaouri, et al., Optimization and characterization of cementitious composites combining maximum amounts of waste glass powder and treated glass aggregates, Int. J. Concr. Struct. Mater. 18 (1) (2024) 25.

[22] M.R. Raju, et al., Machine learning for predicting strength properties of waste iron slag concrete, Heliyon 11 (3) (2025).

[23] M. Derrick, S. David, B. Tajebe, A review on aerospace-AI, with ethics and implications, J. Civ. Constr. Environ. Eng. 10 (2) (2025) 60–74.

[24] D. Sinkhonde, et al., Ensemble machine learning algorithms for efficient prediction of compressive strength of concrete containing tyre rubber and brick powder, Clean. Waste Syst. 10 (2025) 100236.

[25] D. Sinkhonde, D. Mashava, D. Mirindi, Predicting the compressive strength of rubberized concrete incorporating brick powder based on MLP and RBF neural networks, Waste Manag. Bull. 3 (1) (2025) 219–233.

[26] J.F. Vargas, et al., Machine-learning-based predictive models for compressive strength, flexural strength, and slump of concrete, Appl. Sci. 14 (11) (2024) 4426.

[27] H.A. Shah, et al., Application of machine learning techniques for predicting compressive, splitting tensile, and flexural strengths of concrete with metakaolin, Materials 15 (15) (2022) 5435.

[28] B. Wen, et al., The synergistic utilization of glass aggregates and glass powder on the thermal and mechanical properties of concrete, Materials 18 (10) (2025) 2405.

[29] J. Ahmad, et al., A step towards concrete with partial substitution of waste glass (WG) in concrete: A review, Materials 15 (7) (2022) 2525.

[30] G. Tong, et al., Response tests on the effects of particle size of waste glass sand and glass powder on the mechanical and durability performance of concrete, Sci. Rep. 14 (1) (2024) 25445.

[31] Ö. Zeybek, et al., Influence of replacing cement with waste glass on mechanical properties of concrete, Materials 15 (21) (2022) 7513.

[32] P. Kaur, et al., Investigation of bismuth borate glass system modified with barium for structural and gamma-ray shielding properties, Spectrochim. Acta Part A: Mol. Biomol. Spectrosc. 206 (2019) 367–377.

[33] A. Saeed, Y. Elbashar, S. El Khameesy, A novel barium borate glasses for optical applications, Silicon 10 (2018) 569–574.

[34] Boyitechnology, Exploring Glass Melting Points: What You Need to Know. 2024; Available from: https://www.boyiprototyping.com/materials-guide/glass-melting-point/.

[35] Z. Jin, et al., From window panes to bone regeneration: structure, viscosity and bioactivity of soda lime silicate glasses., J. Mater. Chem. B (2025).

[36] Continentaltrade, What is soda lime glass made of? 2025; Available from: https://www.continentaltrade.com.pl/en/our-offer/technical-glass/types-of-materials/soda-lime-glass.

[37] I. Kashif, et al., The role of lead oxide on structural and physical properties of lithium diborate glasses, J. Alloys Compd. 539 (2012) 124–128.

[38] N. Ghoneim, H. ElBatal, I. Ali, Shielding behavior of V2O5 doped lead borate glasses towards gamma irradiation, J. Alloys Compd. 509 (24) (2011) 6913–6919.

[39] F. Fayon, C. Bessada, D. Massiot, High-resolution double-quantum 31P MAS NMR study of the intermediate-range order in crystalline and glass lead phosphates, Inorg. Chem. 38 (23) (1999) 5212–5218.

[40] M. Ganguli, K. Rao, Structural role of PbO in Li2O–PbO–B2O3 glasses, J. Solid State Chem. 145 (1) (1999) 65–76.

[41] K.J. Rao, Structural Chemistry of Glasses, Elsevier, 2002.

[42] D. Müller, et al., Solid-state high-resolution 27Al nuclear magnetic resonance studies of the structure of CaO-Al2O3-P2O5 glasses, 1983.

[43] M.S. Rao, et al., The role of ligand coordination on the spectral features of Yb3+ ions in lead aluminum silicate glasses, J. Mol. Struct. 1007 (2012) 185–190.

[44] G. Rachkovskaya, G. Zakharevich, V. Golubkov, Formation of the pbs quantum dots in boron-silicate glass matrix, in: Physics, Chemistry and Application of Nanostructures, Reviews and Short Notes, World Scientific, 2009, pp. 136–139.

[45] E. Erol, et al., Recent progress in lanthanide-doped luminescent glasses for solid-state lighting applications—A review, J. Phys.: Condens. Matter. 33 (48) (2021) 483001.

[46] H. You, M. Nogami, Optical properties and local structure of Eu3+ ions in sol-gel TiO2- SiO2 glasses, J. Phys. Chem. B 108 (32) (2004) 12003–12008.

[47] K. Jha, M. Jayasimhadri, Structural and emission properties of Eu3+-doped alkaline earth zinc-phosphate glasses for white LED applications, J. Am. Ceram. Soc. 100 (4) (2017) 1402–1411.

[48] G. Lakshminarayana, et al., Eu3+-doped fluoro-telluroborate glasses as red-emitting components for W-LEDs application, Opt. Mater. 99 (2020) 109555.

[49] D. Mirindi, T.N. Sanders, J. Hunter, Integration of artificial intelligence and smart technologies in offsite construction: a comprehensive review, Transform. Constr. Off-Site Methods Technol. (2024).

[50] D. Mirindi, A review of the advances in artificial intelligence in transportation system development, J. Civ. Constr. Environ. Eng. 9 (3) (2024) 72–83.

[51] D. Mirindi, D. Sinkhonde, T. Bezabih, A review on aerospace-AI, with ethics and implications, 2025.

[52] D. Mirindi, et al., The Role of Artificial Intelligence in Building Information Modeling, in: Proceedings of the 2025 Computers and People Research Conference, 2025.

[53] D. Mirindi, D. Sinkhonde, T. Bezabith, Advanced evaluation of BIM-GenAI using OpenAI o1 and ethical considerations, in: Proceedings of the 2025 Computers and People Research Conference, 2025.

[54] P. Sedgwick, Pearson's correlation coefficient, Bmj 345 (2012).

[55] J. Benesty, J. Chen, I. Cohen, Pearson correlation coefficient, in: Noise Reduction in Speech Processing, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 1–4.

[56] D. Sinkhonde, et al., Applications of machine learning algorithms on the compressive strength of laterite blocks made with metakaolin-based geopolymer and sugarcane molasses, Waste Manag. Bull. (2025) 100212.
[57] J. Mingers, An empirical comparison of selection measures for decision-tree induction, Mach. Learn. 3 (1989) 319–342.
[58] W. Buntine, Learning classification trees, in: Artificial Intelligence Frontiers in Statistics, Chapman and Hall/CRC, 2020, pp. 182–201.
[59] T. Chen, et al., Xgboost: extreme gradient boosting, R Packag. Version 0.4-2 1 (4) (2015) 1–4.
[60] Z.A. Ali, et al., Extreme gradient boosting algorithm with machine learning: A review, Acad. J. Nawroz Univ. 12 (2) (2023) 320–334.
[61] Y.-C. Chang, K.-H. Chang, G.-J. Wu, Application of extreme gradient boosting trees in the construction of credit risk assessment models for financial institutions, Appl. Soft Comput. 73 (2018) 914–920.
[62] A. Taherkhani, G. Cosma, T.M. McGinnity, AdaBoost-CNN: An adaptive boosting algorithm for convolutional neural networks to classify multi-class imbalanced datasets using transfer learning, Neurocomputing 404 (2020) 351–366.
[63] D.-C. Feng, et al., Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach, Constr. Build. Mater. 230 (2020) 117000.
[64] J. Fan, et al., Light gradient boosting machine: An efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data, Agricult. Water. Manag. 225 (2019) 105758.
[65] J. Guo, et al., Prediction of heating and cooling loads based on light gradient boosting machine algorithms, Build. Environ. 236 (2023) 110252.
[66] E. Schulz, M. Speekenbrink, A. Krause, A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions, J. Math. Psych. 85 (2018) 1–16.
[67] J. Wang, An intuitive tutorial to Gaussian process regression, Comput. Sci. Eng. 25 (4) (2023) 4–11.
[68] F. Bachoc, F. Gamboa, N. Venet, A Gaussian process regression model for distribution inputs, IEEE Trans. Inform. Theory 64 (10) (2018) 6620–6637.
[69] D. Mirindi, A. Khang, F. Mirindi, Artificial intelligence (AI) and automation for driving green transportation systems: A comprehensive review, in: Driving Green Transportation System Through Artificial Intelligence and Automation: Approaches, Technologies and Applications, 2025, pp. 1–19.
[70] F. Zhang, L.J. O'Donnell, Chapter 7 - support vector regression, in: A. Mechelli, S. Vieira (Eds.), Machine Learning, Academic Press, 2020, pp. 123–140.
[71] H. Cevikalp, Best fitting hyperplanes for classification, IEEE Trans. Pattern Anal. Mach. Intell. 39 (6) (2017) 1076–1088.
[72] M.V. Suganyadevi, C.K. Babulal, Support vector regression model for the prediction of loadability margin of a power system, Appl. Soft Comput. 24 (2014) 304–315.
[73] D. Mirindi, et al., Performance of machine learning algorithms to evaluate the physico-mechanical properties of nanoparticle panels, Green Technol. Sustain. 3 (4) (2025) 100235.
[74] T. Fushiki, Estimation of prediction error by using K-fold cross-validation, Stat. Comput. 21 (2) (2011) 137–146.
[75] B. Dębska, L. Lichołai, G.J.B. Silva, Effects of waste glass as aggregate on the properties of resin composites, Constr. Build. Mater. 258 (2020) 119632.
[76] Z.Z. Ismail, E.A. Al-Hashmi, Recycling of waste glass as a partial replacement for fine aggregate in concrete, Waste Manage. 29 (2) (2009) 655–659.
[77] S. Abdallah, M. Fan, Characteristics of concrete with waste glass as fine aggregate replacement, Int. J. Eng. Tech. Res 2 (6) (2014) 11–17.
[78] J. Wang, S. Wu, X. Qian, Study on the mechanics and self-sensing properties of ultrahigh-performance shotcrete containing waste glass aggregates, Sci. Rep. 15 (1) (2025) 2936.
[79] M. Batayneh, I. Marie, I. Asi, Use of selected waste materials in concrete mixes, Waste Manage. 27 (12) (2007) 1870–1876.
[80] H.-Y. Wang, A study of the effects of LCD glass sand on the properties of concrete, Waste Manage. 29 (1) (2009) 335–341.
[81] M.C. Limbachiya, Bulk engineering and durability properties of washed glass sand concrete, Constr. Build. Mater. 23 (2) (2009) 1078–1083.
[82] E.E. Ali, S.H. Al-Tersawy, Recycled glass as a partial replacement for fine aggregate in self compacting concrete, Constr. Build. Mater. 35 (2012) 785–791.
[83] W. Song, et al., Effects of recycled CRT glass fine aggregate size and content on mechanical and damping properties of concrete, Constr. Build. Mater. 202 (2019) 332–340.
[84] O.M. Olofinnade, et al., Strength and microstructure of eco-concrete produced using waste glass as partial and complete replacement for sand, Cogent Eng. 5 (1) (2018) 1483860.
[85] Z. Steyn, A. Babafemi, R. Combrinck, Concrete containing waste recycled glass, plastic and rubber as sand replacement, Constr. Build. Mater. 269 (2021) 121242.