Computer Modeling in Engineering & Sciences

Tech Science Press

https://doi.org/10.32604/cmes.2026.086781

ARTICLE

Data-Driven Design and Optimization of Sustainable and Low-Carbon Calcium Sulfoaluminate Cement Blends Incorporating Blast Furnace Slag

Ahmad Nawaz[1,*], Hilal Khan[2], Salamat Ullah[3], Hamad Almujibah[4,5], Ali E. A. Elshekh[5] and Maaz Osman Bashir[5]

[1]Guangdong Provincial Key Laboratory of Durability for Marine Civil Engineering, College of Civil and Transportation Engineering, Shenzhen University (SZU), Shenzhen, China
[2]NUST Institute of Civil Engineering (NICE), School of Civil and Environmental Engineering (SCEE), National University of Sciences and Technology (NUST), Sector H-12, Islamabad, Pakistan
[3]Center for Mechanics under Extreme Environments, Ningbo University, Ningbo, China
[4]Research Center of Basic Sciences, Engineering and High Altitude, Taif University, Taif, Saudi Arabia
[5]Department of Civil Engineering, College of Engineering, Taif University, Taif, Saudi Arabia

*Corresponding Author: Ahmad Nawaz. Email: ahmadnawaz@szu.edu.cn

Received: 05 June 2026; Accepted: 30 July 2026; Published: 28 August 2026

ABSTRACT: Calcium sulfoaluminate (CSA) cement is considered a promising low-carbon alternative to ordinary Portland cement owing to its lower clinkerization temperature and reduced CO_2 emissions. The incorporation of blast furnace slag can further enhance the sustainability of CSA-based binders by lowering clinker content, reducing cost and embodied carbon emissions, while maintaining satisfactory mechanical performance. However, optimizing CSA-slag systems remains challenging due to the complex interactions among binder composition, clinker mineralogy, and slag replacement levels. This study therefore aims to predict the compressive strength of CSA-slag binders, identify the mixture parameters governing it, and optimize mixture proportions for balanced mechanical, environmental, and economic performance, using a machine learning (ML)-based framework trained on 232 experimental samples compiled from the literature. Three ML models, Decision Tree (DT), Random Forest (RF), and Extreme Gradient Boosting (XGB), were optimized using a Genetic Algorithm and integrated through stacked ensemble learning with a multilayer perceptron meta-learner. Among the developed models, the DT-XGB ensemble achieved the highest predictive accuracy for compressive strength (R^2 = 0.965, RMSE = 3.168 MPa). Shapley additive explanations (SHAP) analysis identified the water-to-cement ratio as the most influential parameter governing compressive-strength prediction, followed by CSA clinker content, curing age, ye'elimite content, M (sulfate-to-ye'elimite ratio), slag content, and belite content. Furthermore, a multi-objective optimization framework based on NSGA-II was employed to simultaneously optimize compressive strength, embodied CO_2 emissions, and cost. The optimum mixture contained 30% CSA clinker and 70% slag at a water-to-cement ratio of 0.4, achieving 57.9 MPa compressive strength, 379 kg/m^3 embodied CO_2 emissions, and a cost of 91.8 USD/m^3. These findings provide an effective strategy for designing sustainable, high-performance, and low-carbon CSA-based cementitious materials.

KEYWORDS: Machine learning; calcium sulfoaluminate cement; blast furnace slag; multi-objective optimization; sustainable cement

1 Introduction

During the past decade, cement consumption has increased dramatically due to rising demand for housing and infrastructure in developing countries such as China [1,2]. In 2024, global cement production

was approximately 4.0 billion metric tons, of which China accounted for about 1.9 billion metric tons, representing nearly half of the global output. However, ordinary Portland cement (PC) often comes with high environmental cost due to high energy consumption and CO_2 gas emission during its production [3]. Various alternative binders have been proposed to reduce CO_2 emissions and develop an eco-friendly cementitious binder [4]. Examples include alkali-activated binders/geopolymers [5,6], limestone calcined clay cement (LC_3) [7], magnesium-based cements [8], and calcium aluminate cement [9], which have been associated with lower clinker demand, rapid hardening, high early strength, improved durability, or reduced embodied CO_2 depending on their chemistry and application [5,7,10,11]. Among various alternative binders, calcium sulfoaluminate (CSA) cement is one of the promising alternatives as it requires lower $CaCO_3$ input during production, have lower calcination temperature, and requires less amount of grinding energy compared to PC [5,10]. These factors therefore significantly reduce the CO_2 footprint and the energy requirements associated with the production, and a theoretical study found that CSA cement production generates 30% lower CO_2 emissions compared to PC [12]. Compared to others, CSA cement has various advantages including rapid hardening, high early strength, low shrinkage, and resistance to harmful ions. CSA cement can be used in self-leveling floors, filling grouts, rapid repairs, and latent heat storage material where use of PC would not be optimal [13,14]. However, the limited availability of raw materials as well as low demand for these types of cement makes the product expensive and can only be utilized for special applications.

To further reduce the environmental impact and production cost of CSA cement, the incorporation of supplementary cementitious materials (SCMs) has emerged as one of the most effective and practical pathways [15]. Partial replacement of CSA cement with supplementary cementitious material can be desired from both economic and environmental perspectives [4]. Supplementary cementitious materials (SCMs) such as fly ash, silica fume, metakaolin, rice husk ash, and blast furnace slag have been widely used to partially replace cement clinker, improve pore structure, enhance later-age strength, and reduce environmental impact [16–18]. Among different SCMs, blast furnace slag (BFS) is a promising candidate because it not only reduces clinker consumption and embodied carbon emissions but also promotes the utilization of industrial by-products [18], and its replacement level has been reported to govern the mechanical and durability performance of blended cement mortars [19]. However, the incorporation of slag significantly alters the hydration kinetics, phase assemblage, and strength development behavior of CSA systems. Previous studies on CSA-slag blended binders have shown that slag replacement can reduce CSA clinker demand, embodied CO_2 emissions, and material cost while supporting the utilization of industrial by-products. However, the reported mechanical performance strongly depends on slag dosage, sulfate availability, clinker mineralogy, water-to-cement (W/C) ratio, and curing age [20,21]. Moderate slag replacement has been reported to maintain or improve later-age strength through continued hydration and microstructural refinement, whereas excessive replacement may reduce early-age strength due to clinker dilution and slower slag activation [22,23]. Therefore, designing CSA-slag blended binders with optimized mechanical performance and environmental sustainability remains a complex challenge due to the strong nonlinear interactions among clinker composition, slag dosage, and mixture proportions.

In designing cement-based materials, compressive strength is a critical performance indicator, as it directly affects structural serviceability and safety. Laboratory testing provides the most accurate strength measurements, it becomes time-consuming and costly when dealing with extensive mix designs [24]. Therefore, developing reliable models to predict the compressive strength of SCM-incorporated concrete is essential. However, designing a CSA cement mixture that simultaneously optimizes compressive strength, environmental and economic sustainability remains a complex task due to the strong interdependence among mixture composition [25]. Therefore, the application of soft computing techniques is essential for

rapid and reliable prediction of compressive strength and for the efficient optimization of CSA cement formulations.

In recent years, machine learning has emerged as a powerful tool for modeling complex nonlinear relationships in cementitious materials. Artificial intelligence-based material design has emerged as a powerful tool for accelerating the development of low-carbon binders, where mix proportioning often involves complex interactions among multiple variables such as molar ratios, SCM content, curing regime, and additive dosage [26,27]. Traditional experimental optimization of these parameters is highly time-consuming, resource-intensive, and often insufficient for capturing nonlinear dependencies inherent in sustainable cementitious systems. Machine learning (ML) methods such as artificial neural networks (ANN), support vector regression (SVR), decision trees (DT), and advanced ensemble algorithms including Random Forest (RF), Gradient Boosting Regression (GBR), and Extreme Gradient Boosting (XGB) have therefore gained significant attention for their ability to model high-dimensional, nonlinear datasets and accurately forecast key mechanical and durability properties of low-carbon binders [28–31]. These models have demonstrated exceptional predictive capability across various cementitious composites, including alkali-activated geopolymers [32], CSA cement [25], magnesium phosphate cements [24], and magnesium oxychloride cement systems [30], achieving high correlation coefficients in predicting compressive strength, chloride penetration, shrinkage resistance, and overall durability. Despite these advantages, the performance of individual ML algorithms can vary significantly depending on the dataset characteristics and problem complexity, and no single model consistently provides optimal results [26]. In theory, integrating multiple models through ensemble learning offers a more robust solution by leveraging the complementary strengths of different algorithms. Therefore, this study incorporates high-performing ensemble techniques within a stacking framework to further enhance prediction accuracy and provide a reliable data-driven design tool for sustainable low-carbon binder systems.

In parallel, the growing emphasis on net-zero emissions and sustainable construction demands not only accurate property prediction but also multi-objective optimization of binder formulation. Low-carbon material design must simultaneously consider strength, durability, carbon footprint, and cost [26,27]. Since improvements in one objective often compromise another, the mix design becomes a constrained multi-objective optimization problem. Among different optimization approaches, the Non-dominated Sorting Genetic Algorithm II (NSGA-II) has been widely adopted because of its capability to efficiently generate Pareto-optimal solutions within complex multidimensional search spaces [33,34].

Although previous studies have applied ML methods to predict the performance of cementitious materials, studies integrating interpretable stacked ensemble learning with multi-objective optimization for low-carbon CSA-slag systems remain limited. Therefore, this study proposes an integrated machine learning and optimization framework for sustainable CSA-slag cement systems. The specific objectives are as follows:

- *Development of optimized base learners:*

Random Forest (RF), Decision Tree (DT), and Extreme Gradient Boosting (XGB) algorithms are adopted as base learners within the stacking ensemble framework. Genetic Algorithm (GA) is employed to optimize the hyperparameters of each base learner, ensuring improved predictive performance and generalization.

- *Construction of stacking ensemble models:*

Various combinations of optimized base learners are integrated using a Multi-Layer Perceptron (MLP) as the meta-learner. The stacked models are trained to predict compressive strength, and the best-performing model based on predictive accuracy is selected to construct the objective function for subsequent multi-objective optimization.

- *Model interpretability analysis:*

Feature importance and interpretability analyses are conducted to quantify the contributions of key mixture parameters to strength development. This provides mechanistic insight into how concrete constituents influence the predicted outcomes, enhancing transparency in ML material design.

- *Development of a multi-objective optimization framework:*

Finally, the best-performing ensemble model was coupled with the NSGA-II algorithm to perform multi-objective optimization targeting compressive strength, material cost, and embodied CO_2 emissions. The proposed framework aims to provide an efficient and intelligent strategy for designing sustainable low-carbon CSA cement systems with balanced mechanical, environmental, and economic performance. The overall workflow of the proposed framework, showing the sequential connection among database preparation, machine-learning modeling, model interpretation, and multi-objective optimization, is illustrated in Fig. 1.

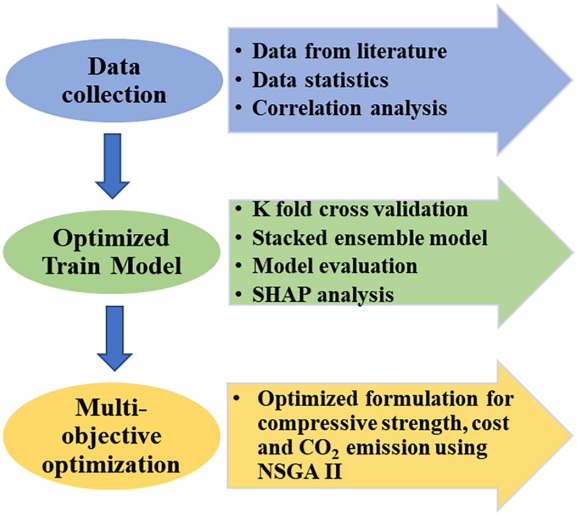

Figure 1: Flowchart of this study.

2 Database Description

2.1 Feature Parameters Configuration

The raw dataset consists of 232 data points collected from peer-reviewed studies on CSA and CSA-slag cement systems published between 2018 and 2024. The experimental data were compiled from the main literature sources [18,21,22,35], with additional data obtained from related studies [36–38]. The dataset compiled for this study features nine input variables designed to predict the compressive strength (CS) of CSA cement, which serves as the output variable. The compressive strength values compiled in the database represent the mean values reported in the source literature. These input variables were selected to comprehensively represent the critical factors influencing CSA cement performance, including the mixture type, its precise mineralogical composition, mix proportions, and curing conditions. A key aspect of the study is the incorporation of Slag (%) as a partial replacement for CSA clinker, a strategy aimed at reducing the clinker factor and developing more sustainable, low-carbon cement formulations. Importantly, in belite-bearing CSA clinkers, slag does not remain inert at elevated replacement levels: belite hydration supplies C-S-H and modest portlandite, while the sulfate-rich environment (governed by M) promotes sulfate activation of slag and the formation of stratlingite (C_2ASH_8), consistent with prior CSA-slag studies [18,21,22,35]. Nonetheless, the effective activator supply diminishes as CSA clinker content decreases, which constrains

strength development at high slag dosages. The initial categorical variable, Type, distinguishes between two mixture forms: P for paste and M for mortar. The mineralogical composition, specifically the content of the key hydraulic phases ye'elimite (%) and belite (%), along with the molar ratio (M) of sulfate to ye'elimite (which is critical for optimal phase formation and varies from 0.75 to 2.39 in this dataset), are fundamental drivers of hydration and strength development. When these mineralogical compositions were not directly reported in the source literature, they were calculated from the oxide compositions using an adapted Bogue's equation (Eqs. (1)–(3)) [25,39], and no missing-value imputation technique was applied in this study.

$$\text{Ye'elimite}\ (\%) = 1.995\,(Al_2O_3\%) - 1.273\,(Fe_2O_3\%) \tag{1}$$

$$\text{Belite}\ (\%) = 2.867\,(SiO_2\%) \tag{2}$$

$$\text{Calcium Sulfate}\ (\%) = 1.700\,(SO_3\%) - 0.445\,(Al_2O_3\%) + 0.284\,(Fe_2O_3\%) \tag{3}$$

The CSA clinker (%) content defines the primary binder, while the water-to-cement (W/C) ratio, which spans from 0.3 to 0.8, is the essential mix parameter governing workability and microstructure. The curing time (Age) is included to capture the effect of hydration duration on strength gain. The curing-age variable ranged from 1 to 360 days, with a mean value of 41.23 days and a standard deviation of 73.23 days. This indicates a right-skewed distribution, where most data points are concentrated at early and intermediate curing ages, while relatively few observations are available at later ages. Such imbalance is typical of literature-derived cementitious-material datasets because early-age and standard-age strengths are more frequently reported than long-term strengths. The dataset encompasses a wide range of values for these features, reflecting data aggregated from controlled experimental programs. For instance, the CSA clinker content varies from 100% down to 30%, allowing for the analysis of slag substitution effects. This diversity is a key advantage of data-driven modeling, as it incorporates a breadth of conditions that facilitate the development of robust predictive models. The output variable, CS, demonstrates a correspondingly wide range of values, from approximately 5 MPa to over 100 MPa, directly reflecting the systematic variation of the input parameters.

Overall, the dataset includes 9 input features, encompassing mix design parameters, raw material characteristics, and curing conditions. These variables include mixture type, CSA clinker content, ye'elimite content, belite content, M-value, slag content, sand-to-binder ratio, water-to-cement ratio, and curing age. Among these variables, Type is a categorical input feature used to distinguish paste specimens (P) from mortar specimens (M), and it was encoded for machine-learning model development. Since Type is categorical, it is not included in the numerical statistical summary shown in Table 1. Compressive strength is the sole target variable predicted by the ML models, and the measurements cover curing ages ranging from 1 to 360 days. Material cost and embodied CO_2 emissions were not treated as machine-learning target variables; instead, they were calculated from the mixture proportions using the unit cost and embodied-carbon factors of the constituent materials presented in Table 2. These calculated values were subsequently used, together with the predicted compressive strength, as objective functions in the NSGA-II multi-objective optimization framework. Furthermore, to ensure model robustness, it is essential to standardize data formats extracted from different literature sources and harmonize the representation of mix designs across various sample types (e.g., paste and mortar) and material systems. For this purpose, an additional mixture parameter, namely the sand-to-binder ratio, was employed to reflect differences between paste and mortar specimens. Furthermore, the distribution characteristics of the numerical dataset variables are illustrated in Fig. 2 using histograms. While Table 2 reports the exact statistical values, Fig. 2 visually shows the frequency distribution and clustering of the variables. The histograms indicate that the dataset covers a broad range of CSA clinker content, slag replacement level, mineralogical composition, W/C ratio, curing age, and compressive strength, with some clustering around commonly used mixture proportions and testing ages.

Table 1: Statistical summary of dataset.

Features	Mean	Standard Deviation	Minimum	Maximum
CSA clinker	79.220	23.083	30	100
Ye'elimite%	53.922	14.731	27.5	69.4
Belite	22.754	8.092	18	48.9
M	1.572	0.602	0.75	2.39
Slag%	16.883	22.013	0	70
Sand to binder ratio	0.584	1.190	0	3
W/C	0.5112	0.098	0.3	0.8
Age	41.233	73.236	1	360
CS (Target feature)	32.589	16.661	5	104

Table 2: The cost and embodied carbon values of raw materials [27,40,41].

Materials	Market Price (USD/ton)	Embodied Carbon (kg/ton)	Unit Weight (kg/m^3)
CSA cement	90	620	2850
Slag	60	143	2800
Sand	28	4.8	2650
water	0.9	0	1000

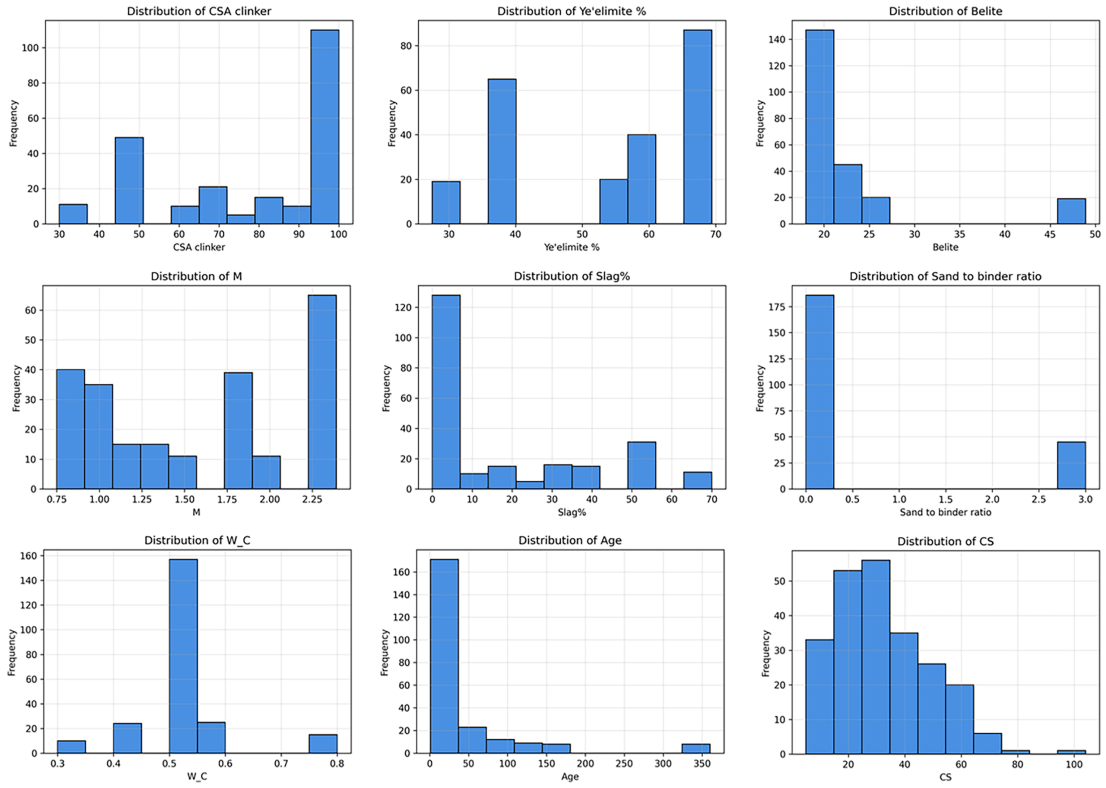

Figure 2: Data distribution histogram.

2.2 Correlation Analysis

Fig. 3 presents a Pearson correlation matrix mapping the linear dependencies among nine key variables governing the compressive strength (CS) of CSA cement. CS exhibits a strong positive correlation with CSA clinker content (0.65), confirming its important role in strength development, while negative correlations are observed with slag content (−0.57) and W/C ratio (−0.33). This indicates that higher slag replacement and increased water content tend to reduce strength within the compiled dataset. A notable trend is the strong negative correlation between CSA clinker and slag content (−0.82), which reflects the direct replacement relationship between these two binder components. Therefore, mixtures with a higher CSA clinker-to-slag ratio generally show higher compressive strength, although this effect is also influenced by other variables such as M-value, W/C ratio, and curing age. Mild positive associations are observed with the molar ratio (M) (0.19) and belite (0.25), indicating that higher sulfate levels and belite content contribute modestly to mechanical performance. Other variables display negligible linear interactions with CS, suggesting their influence is either non-linear or overshadowed by the dominant factors identified above.

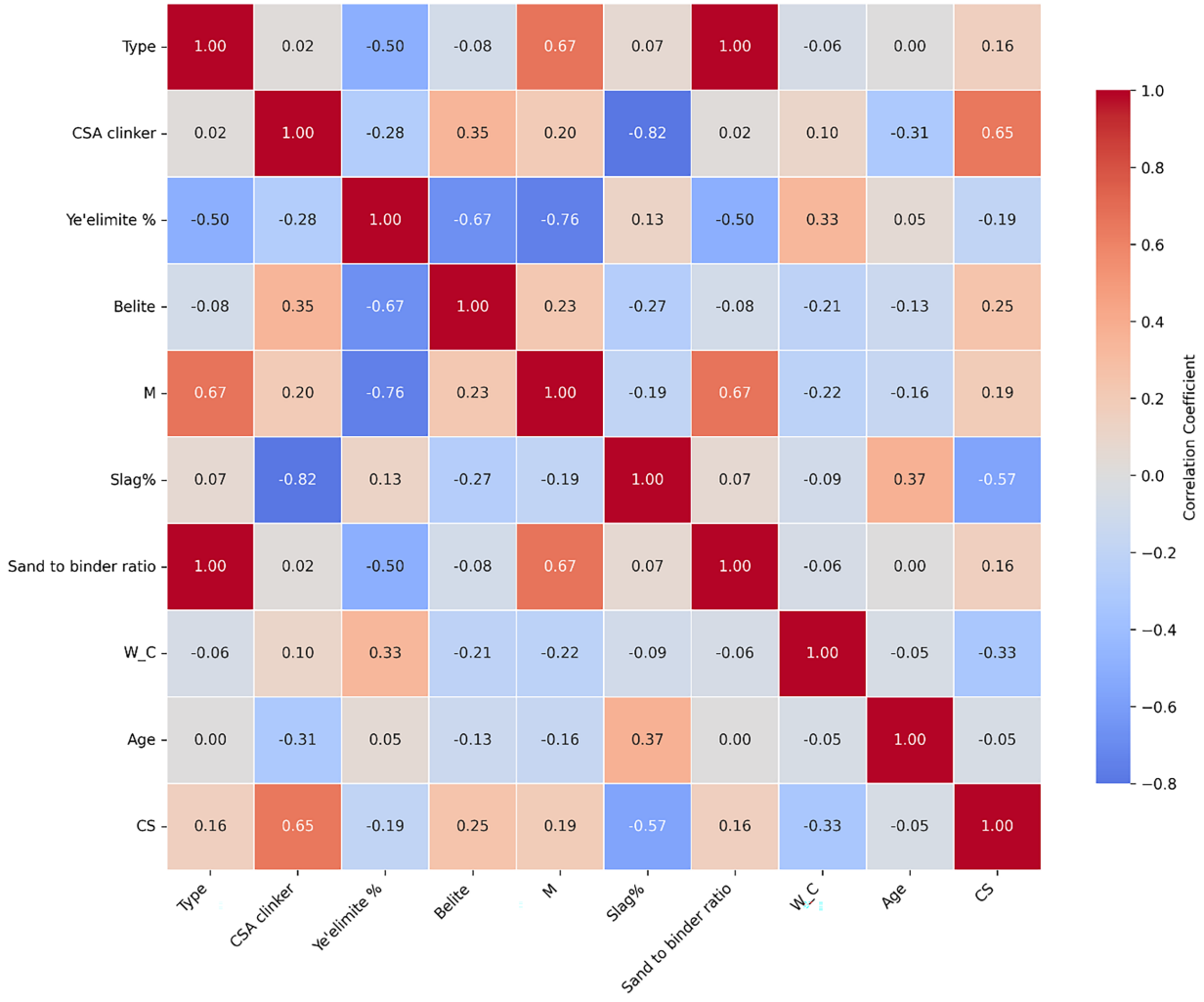

Figure 3: Input and output correlation and heat intensity map.

3 Research Methodology

3.1 Machine Learning Algorithms

Prior to this investigation, various machine learning paradigms had been used for predicting mechanical and durability properties of cementitious materials. Building on established precedents [29,31,32,42], this study employs three tree-based learning algorithms, including Decision Tree (DT), Random Forest (RF), and Extreme Gradient Boosting (XGB). Among these, DT represents a single base learner, whereas RF and XGB are ensemble algorithm methods that combine multiple decision trees through bagging and boosting strategies, respectively. These models are effective for capturing complex nonlinear relationships commonly observed in cementitious-material datasets, particularly when the available experimental data are limited. Because all three use decision trees as their building blocks, they are capable of identifying underlying nonlinear patterns and feature interactions within the dataset. The use of these models enhances the accuracy, reliability, and interpretability of the predictive analysis. The generalized workflow is depicted in Fig. 4.

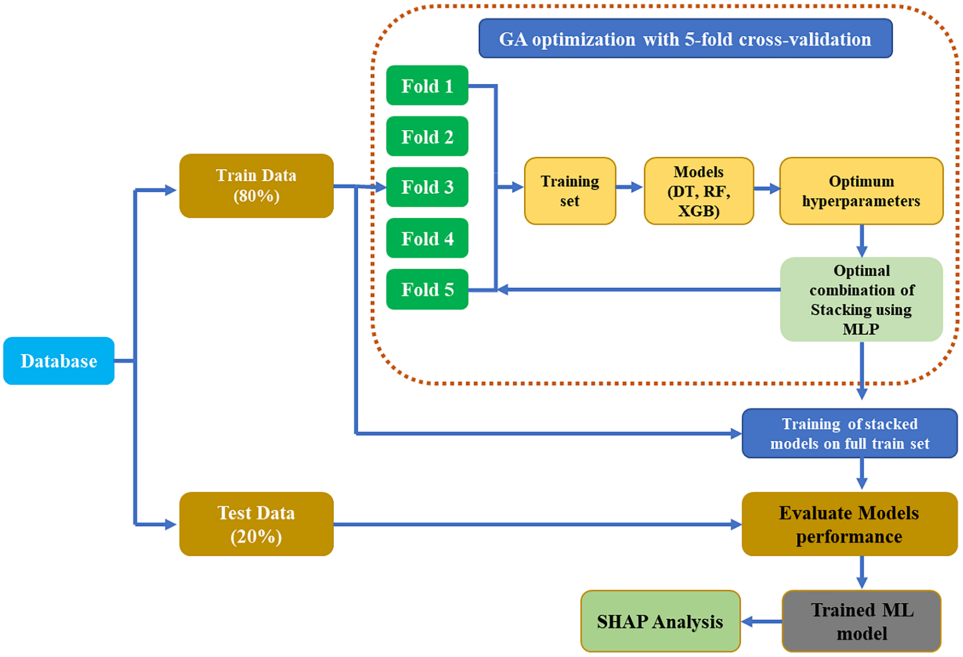

Figure 4: Establishment of stacked ML models.

To ensure reproducibility, the compiled dataset was first checked for consistency before model development. Duplicate and incomplete entries were reviewed, units and mixture-design variables were harmonized, and the categorical variable Type was encoded for machine-learning analysis. The final dataset was then randomly divided into training and testing subsets at an [80/20] ratio, where the training subset was used for model development, and the testing subset was held out entirely and reserved solely for independent performance evaluation, ensuring that no test data influenced model training or hyperparameter tuning. During training, five-fold cross-validation was applied exclusively to the training subset, in which the training data were divided into five subsets. In each iteration, four subsets were used for training, and the remaining subset was used for validation. This process was repeated five times so that each subset served once as the validation fold, and the average validation performance was used to guide the Genetic Algorithm–based hyperparameter optimization of the base learners. Consequently, all hyperparameters, including the

decision-tree depth, were selected on the basis of cross-validated validation performance rather than training fit, which mitigates the risk of overfitting despite the limited dataset size.

3.1.1 Decision Tree

The DT regression algorithm predicts continuous output variables by recursively partitioning the dataset into smaller, more homogeneous subsets. At each stage, the algorithm identifies the most informative feature and uses it to split the data, forming a tree-like structure. This splitting process continues for each resulting subset until predefined stopping criteria are met, producing terminal nodes. Each leaf node represents a local prediction, typically given by the mean value of the target variable within that subset. The overall regression output is obtained by aggregating the predictions from the terminal nodes according to the structure of the tree.

3.1.2 Random Forest

RF [43] is a powerful ensemble learning method built upon decision trees, specifically designed to manage non-linear relationships, complex feature interactions, and outliers effectively. To perform regression, the algorithm constructs numerous individual decision trees during training. The process begins by randomly dividing the training dataset into multiple subsets, each containing a portion of the samples and features. Subsequently, a separate decision tree model is developed for each subset, where the feature selected for splitting at each node is chosen randomly from the available features within that subset. Once trained, each tree generates a prediction result when presented with the test dataset. The final output of the Random Forest model is determined by calculating the average or weighted average of all individual tree predictions, ensuring a robust and stable regression outcome.

3.1.3 Extreme Gradient Boost

XGB [44] is a machine learning technique grounded in the gradient boosting tree framework, designed to address both classification and regression problems. Its core principle centers on enhancing prediction accuracy by sequentially training multiple decision tree models and subsequently integrating their outputs. A defining characteristic of XGB is the inclusion of additional regularization terms for every weak learner, which actively constrains model complexity to optimize generalization performance while effectively curbing overfitting attributes that markedly elevate its efficacy and predictive capability. As an optimized evolution of Gradient Boosting Regression, XGB combines built-in regularization, high computational efficiency, and robust generalization capacity, making it a widely utilized tool for predictions in cementitious materials research.

Stacking is an ensemble learning strategy in which multiple base models are combined to enhance prediction accuracy. The core idea is to train several base learners on the same dataset and then use their predictions as inputs to a higher-level model referred to as the meta-learner, which generates the final output. When the base models exhibit complementary strengths, the meta-learner can effectively learn how to balance their individual weaknesses, resulting in improved generalization performance.

To enhance the nonlinear modeling capability of the stack, a Multilayer Perceptron (MLP) is employed as the meta-learner. The MLP architecture consisting of an input layer, one or more hidden layers, and an output layer enables the model to capture complex, nonlinear relationships among the base-model predictions. This makes it particularly suitable for stacking frameworks where interactions between base learners may not be linearly separable.

3.2 Hyperparameter Optimization Using Genetic Algorithm

Hyperparameter optimization is essential for maximizing the predictive performance and generalization ability of machine learning models. In this study, the Genetic Algorithm (GA) [45] was selected as the optimization method for tuning the hyperparameters of the base learners within the ensemble framework. GA is an evolutionary optimization technique that iteratively improves a population of candidate solutions through selection, crossover, and mutation. Individuals with higher fitness values are preferentially selected as parents, crossover recombines genetic information to generate diverse offspring, and mutation introduces random perturbations to prevent premature convergence. The update rules applied in this study are detailed in Eqs. (4)–(6):

$$p_i = \frac{f_i}{\sum_{j=1}^{n} f_j} \tag{4}$$

$$\begin{cases} C_{kl} = C_{kl}(1-b) + C_{jl}b \\ C_{jl} = C_{jl}(1-b) + C_{kl}b \end{cases} \tag{5}$$

$$\acute{C}_{kl} = C_{kl} + N(o, \alpha) \tag{6}$$

p_i denotes the probability that the i-th individual is selected from a population containing n individuals. The variables f_i and f_j represent the fitness values associated with individuals i and j, respectively. In addition, C_{kl} and C_{jl} indicate the genes located at the l-th position of the k-th and j-th chromosomes. The parameter b is a random variable generated within the range [0, 1], which controls the proportion of genetic information exchanged between two parent chromosomes. The parameter α represents the standard deviation of the Gaussian perturbation introduced during the mutation process. GA was chosen because of its strong global search capability and robustness in navigating complex, nonlinear, and high-dimensional parameter spaces. These properties make GA particularly suitable for machine-learning hyperparameter optimization, ensuring that the stacking ensemble achieves improved stability, predictive performance, and generalization.

3.3 Evaluation Indices

To quantitatively assess the predictive capabilities of the developed models, four standard statistical metrics were employed: The Coefficient of Determination (R^2), Pearson Correlation Coefficient (PCC), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). These metrics collectively evaluate the model's goodness-of-fit and prediction accuracy. While R^2 and PCC measure the strength and direction of the linear correlation between predicted and actual values, with values approaching 1 indicating superior fit, MAE and RMSE quantify the magnitude of prediction errors. Specifically, MAE and RMSE represent absolute errors, with lower values signifying higher precision. Together, these indicators provide a comprehensive validation of the model's reliability and efficacy. The mathematical formulations for R^2, PCC, MAE, and RMSE are defined in Eqs. (7)–(10), respectively, where N denotes the total number of samples.

$$PCC = \frac{\sum_{i=1}^{N}(y_i - \overline{y})(\acute{y}_i - \overline{\acute{y}})}{\sqrt{\sum_{i=1}^{N}(y_i - \overline{y})^2} \cdot \sqrt{\sum_{i=1}^{N}(\acute{y}_i - \overline{\acute{y}})^2}} \tag{7}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \acute{y}_i)^2}{\sum_{i=1}^{N}(y_i - \overline{y})^2} \tag{8}$$

$$MAE = \frac{1}{N}\left(\sum_{i=1}^{N}|y_i - \acute{y}_i|\right) \tag{9}$$

$$RMSE = \sqrt{\frac{1}{N}\left(\sum_{i=1}^{N}(y_i - \acute{y}_i)^2\right)} \tag{10}$$

3.4 Shapely Additive Explanations (SHAP)

Machine learning models often exhibit complex internal structures that are difficult to interpret and are therefore commonly regarded as "black box" models. To improve model transparency, SHAP analysis was employed to quantify the contribution of individual input variables to the model predictions. SHAP assigns each feature a Shapley value, which represents its marginal contribution to the prediction based on local model behavior.

The Shapley value for feature j is computed as:

$$\varphi_j(f,x) = \sum_{s\subseteq J/j}\frac{|S|!\,(|J-S|-1)!}{|J|!}\left[f\left(x_s \cup x_j\right) - f\left(x_s\right)\right] \tag{11}$$

where f denotes the trained predictive model, x represents the input sample to be explained, J is the full set of input features, and S is a subset of features excluding j. The weighting term accounts for the contribution of feature j across all possible feature permutations.

SHAP analysis was applied to interpret both the global behavior of the model and the predictions of individual samples. Global explanations were obtained by averaging SHAP values across the dataset, while local explanations were used to assess feature influence relative to a reference prediction. The SHAP analysis was performed using the finalized dataset of 232 experimental data points, ensuring consistent interpretation of the model predictions.

3.5 Multi-Objective Optimization Framework Using NSGA-II

To achieve a balance between mechanical performance, economic feasibility, and environmental sustainability, the Non-dominated Sorting Genetic Algorithm II (NSGA-II) [46] was employed for multi-objective optimization. NSGA-II is an evolutionary optimization algorithm capable of simultaneously optimizing multiple conflicting objectives and generating Pareto-optimal solutions. In this study, compressive strength was maximized, whereas material cost and embodied CO_2 emissions were minimized using the predictive models developed from the experimental dataset. The overall workflow of the optimized prescriptive framework is illustrated in Fig. 5.

The optimization process involved population initialization, non-dominated sorting, crossover, and mutation operations to iteratively identify optimal mixture compositions. Non-dominated sorting was used to rank solutions according to Pareto dominance, while crowding distance was employed to preserve diversity among the generated solutions.

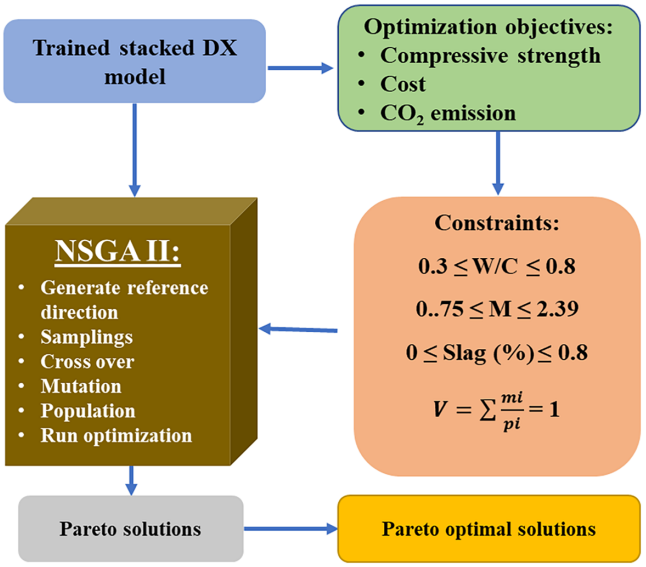

Figure 5: Workflow of optimized ensemble DX model using NSGA II.

The obtained Pareto-optimal solutions provide feasible mixture designs with balanced strength, cost, and environmental performance, thereby supporting the development of sustainable low-carbon CSA cement systems. The decision variables included the water-to-cement ratio (W/C), slag replacement level, curing age, and the sulfate-to-ye'elimite molar ratio represented by M. The curing age was fixed at 56 days during optimization to ensure a consistent comparison among all candidate mixtures. The W/C ratio was constrained between 0.3 and 0.8, while the slag replacement level and M value varied from 0%–70% and 0.75–2.39, respectively. These limits were selected based on the experimental ranges and practical mixture-design conditions represented in the compiled dataset. Therefore, the constraints prevent the optimization algorithm from generating unrealistic mixtures or extrapolating beyond the domain of the trained machine-learning model. In addition, the sum of the volumetric fractions of CSA clinker, slag, water, and sand was constrained to unity to ensure realistic and physically meaningful mixture compositions. These constraints, expressed in Eqs. (12)–(15), were established based on practical mix design considerations and the experimental ranges used in this study.

$$0.3 \leq \frac{W}{C} \leq 0.8 \tag{12}$$

$$0.75 \leq M \leq 2.39 \tag{13}$$

$$0 \leq Slag\,(\%) \leq 70 \tag{14}$$

$$V = \frac{CSA\ clinker}{\rho_{CSA}} + \frac{Slag}{\rho_{slag}} + \frac{Water}{\rho_w} + \frac{Sand}{\rho_s} = 1 \tag{15}$$

The obtained Pareto-optimal solutions provide feasible mixture designs with balanced strength, cost, and environmental performance, thereby supporting the development of sustainable low-carbon CSA cement systems.

4 Results and Discussion

4.1 Optimization of Hyperparameters of Base Learners

This study uses DT, RF, and XGB as base learners and applies a GA for hyperparameter tuning to achieve a consistent and systematic optimization procedure across all models. GA-based tuning is adopted to improve generalization by identifying hyperparameter combinations that control model complexity while reducing prediction error. The optimal hyperparameters selected for each learner are reported in Table 3.

Table 3: Optimal parameters for the ML model used in this study.

Algorithm	Parameter	Optimal Value
DT	max_depth	13
	min_samples_leaf	1
	min_samples_split	3
RF	max_depth	19
	min_samples_leaf	1
	min_samples_split	2
	n_estimators	57
XGB	max_depth	5
	n_estimators	228
	learning_rate	0.118
	gamma	0

The optimized configurations indicate that each algorithm attains its best performance under a different complexity regime, consistent with its learning principle. For DT, the optimal setting (max_depth = 13 max, min_samples_split = 3 min, min_samples_leaf = 1 min) suggests that a moderately deep tree is required to capture nonlinear relationships in the dataset, while split constraints limit excessive partitioning. For RF, GA selected n_estimators = 57 with (max_depth = 19 max, min_samples_split = 2 min, min_samples_leaf = 1), implying that deeper base trees can be effectively leveraged because bootstrap aggregation reduces variance and improves stability. For XGB, the optimal configuration (max_depth = 5, n_estimators = 228, learning rate = 0.118) favors shallower trees with more boosting iterations and regularized splitting, reflecting the boosting strategy of incremental error correction under controlled structural complexity.

The predictive performance of the optimized single learners is summarized in Table 4, which reports both training and testing metrics. The training results show that all three models captured the nonlinear relationships between the input variables and compressive strength with high fitting accuracy. Among them, XGB exhibited the strongest training performance, indicating its high learning capacity, whereas DT also showed strong fitting ability without complete memorization. However, the testing results provide a more practical measure of model generalization. On the testing dataset, DT achieved the highest goodness-of-fit among the single learners, with R^2 = 0.935, PCC ≈ 0.968, and the lowest RMSE of 4.301 MPa. RF delivered comparable correlation performance, with PCC = 0.969, slightly lower R^2 = 0.928, and a higher error level, as reflected by RMSE = 4.558 MPa and MAE = 3.604 MPa. XGB showed the lowest testing R^2 of 0.897 and the highest RMSE of 5.528 MPa, despite its strong training performance, indicating a larger train–test performance gap. However, its MAE of 3.267 MPa remained close to that of DT, suggesting that a limited number of higher-deviation predictions contributed disproportionately to the squared-error metric. Overall, the comparison between training and testing performance indicates that the optimized single learners

exhibit different fitting and generalization characteristics. These differentiated error patterns support their subsequent integration within the stacking framework, where the complementary strengths of the base learners can be combined to improve prediction stability and generalization performance.

Table 4: Performance of base ML models used in this study.

Models	R^2		PCC		MAE		RMSE	
	Train	Test	Train	Test	Train	Test	Train	Test
DT	0.977	0.935	0.988	0.968	0.982	3.278	2.455	4.301
RF	0.982	0.928	0.992	0.969	1.344	3.604	2.189	4.558
XGB	0.998	0.897	0.999	0.955	0.401	3.267	0.535	5.528

4.2 Prediction Results of Stacked Ensemble Models

Considering the differences in learning behavior among the base learners, four stacking combinations were evaluated: DT-RF (DR), DT-XGB (DX), RF-XGB (RX), and DT-RF-XGB (DRX). As shown in Table 5, all ensembles achieved near-perfect training fit ($R^2 \approx 0.999$) together with strong predictive agreement with the testing dataset ($R^2 > 0.95$). The very high training accuracy is characteristic of high-capacity tree-based learners, which fit the training data closely; the testing performance and its consistency across configurations are therefore the more informative indicators, confirming that integrating multiple learners improves robustness and accuracy relative to the individual models. The gap between the training and testing metrics is consistent across all four configurations and reflects the inherent tendency of tree-based ensembles to fit training data closely rather than genuine overfitting.

Table 5: Performance of stacked ML models used in this study.

Models	Notations	R^2		PCC		MAE		RMSE	
		Train	Test	Train	Test	Train	Test	Train	Test
DT-RF	**DR**	0.999	0.961	0.999	0.982	0.294	2.495	0.368	3.314
DT-XGB	**DX**	**0.999**	**0.965**	**0.999**	**0.982**	**0.304**	**2.362**	**0.441**	**3.168**
RF-XGB	**RX**	0.999	0.953	0.999	0.981	0.278	2.689	0.354	3.657
DT-RF-XGB	**DRX**	0.999	0.961	0.999	0.983	0.189	2.425	0.241	3.357

Note: Bold values indicate the best-performing model/metric among the evaluated stacked ensemble models.

Among them, DX performed best overall, giving the highest R^2 (0.965) and the lowest errors (MAE = 2.362, RMSE = 3.168), indicating the most effective complementarity in reducing prediction deviations. DR and DRX followed closely (R^2 = 0.961), with DRX showing the highest PCC (0.983) but without further reduction in error. In contrast, RX yielded the weakest performance (R^2 = 0.953 and the highest error values), suggesting limited added benefit when combining RF and XGB in this dataset. Therefore, based on its consistently high accuracy and error minimization, DX is selected as the final best stacking model for subsequent analyses.

Fig. 6 compares the predicted and measured compressive strength (CS) values for both training and testing datasets. The four ensemble stacked models developed in this study, DR, DX, RX, and DRX, exhibit excellent predictive accuracy and strong generalization capability.

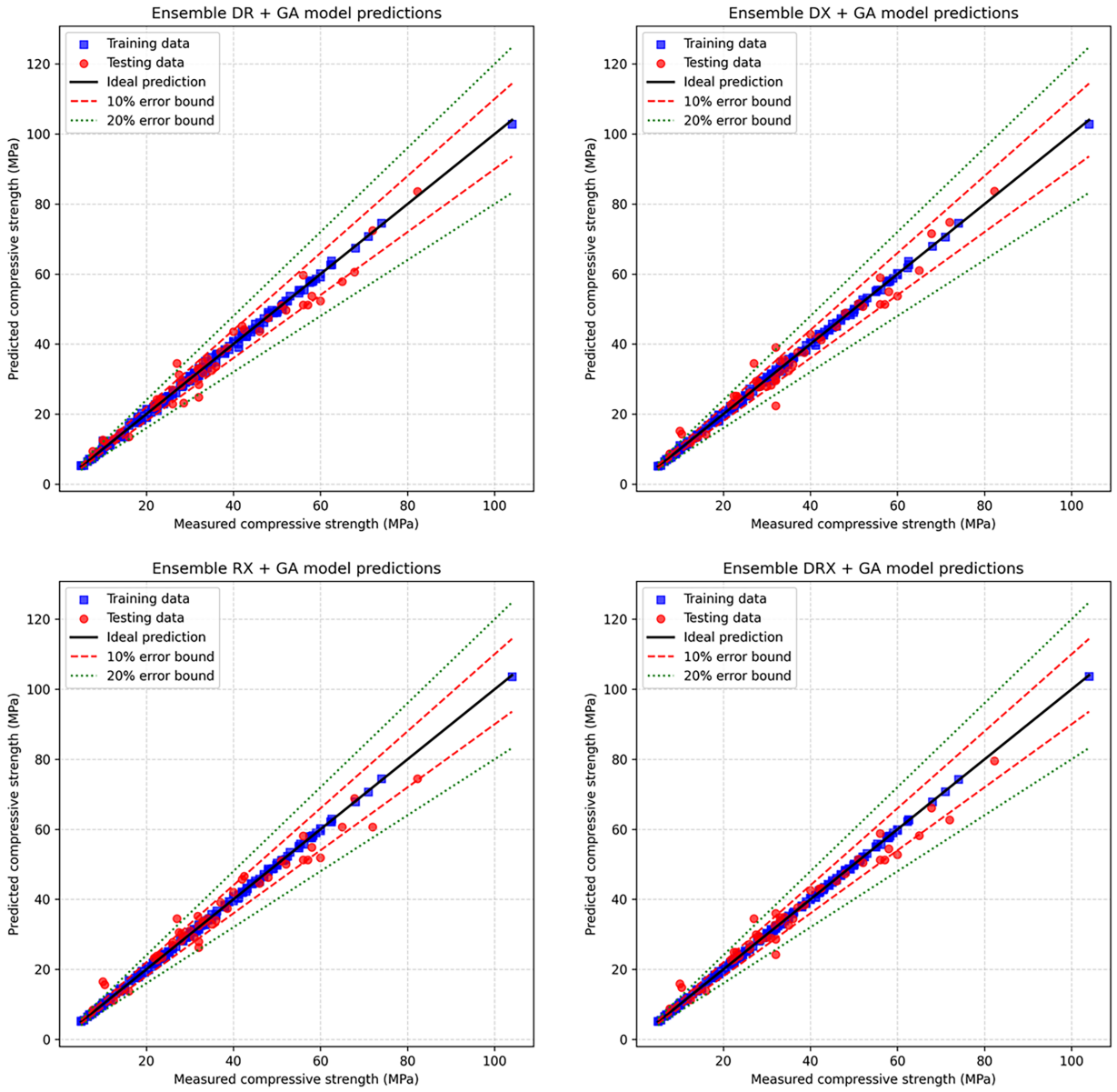

Figure 6: Performance comparison of stacked ensemble models.

For all models, the majority of data points from the training set (blue squares) and testing set (red circles) are tightly clustered around the ideal prediction line, with most predictions falling within the ±20% error bounds. Moreover, a large proportion of both training and testing data lie within the ±10% error range, particularly in the mid-range of compressive strength values, indicating a high level of agreement between predicted and experimental results. Only a small number of outliers are observed, primarily in the testing dataset. The similar distribution of training and testing points in Fig. 6 indicates in-distribution agreement under a random split and is not interpreted as evidence of generalization to unseen source clinkers, which is assessed separately in Table 6.

Table 6: Cross-validation performance of the DX ensemble under random, grouped, and leave-one-group-out schemes; grouping is by source study proxied by clinker mineralogy signature.

Validation Strategy	R^2	MAE	RMSE
Random five-fold	0.93	2.84	4.35
Grouped five-fold	0.40	10.44	12.91
Leave-one-group-out	0.41	10.10	12.77

These visual observations are fully consistent with the quantitative performance indicators reported in the results tables, including high coefficients of determination ($R^2 \geq 0.95$) and low RMSE values. Overall, the integration of ensemble stacked learning with genetic algorithm optimization results in robust and reliable models capable of accurately predicting compressive strength over a wide range of values.

A random division of the records into training and testing subsets can place replicate specimens and identical mix designs from a single source study on both sides of the split, so that the reported accuracy reflects interpolation within source clinkers rather than prediction of unseen ones. To separate these two situations, the DX ensemble is evaluated under three schemes. In the random five-fold scheme, the records are shuffled without regard to their origin. In the grouped five-fold and leave-one-group-out schemes, the records are partitioned by source study, proxied by the unique clinker mineralogy signature defined by the paired ye'elimite and belite contents, so that all records sharing a clinker signature are confined to a single fold, and every prediction concerns a clinker that is absent from the corresponding training set. Table 6 reports the three evaluations. The random scheme attains a coefficient of determination of 0.93 with a root-mean-square error of 4.4 MPa, while the grouped and leave-one-group-out schemes attain coefficients of determination of 0.40 and 0.41 with root-mean-square errors of 12.9 and 12.8 MPa. The contrast quantifies the optimism of the random split and identifies the grouped estimate as the value that describes generalization to clinker systems not represented in training. Fig. 7 shows the grouped-scheme parity, in which the larger scatter at the extremes of the strength range reflects the difficulty of extrapolating to unseen clinker chemistries. The higher grouped error is consistent with the modest corpus size and the small number of distinct source clinkers, and it defines the operating regime within which the optimized formulations are to be interpreted.

4.3 Residual Diagnostics and Heteroscedasticity Assessment

Residual behavior is assessed through the diagnostics in Fig. 8, which comprises a residuals-vs.-predicted plot, a residual distribution with a fitted normal density, and a normal quantile-quantile plot. The residuals scatter about zero with a mean of −0.2 MPa and a standard deviation of 4.4 MPa, and the reference band at plus and minus one point nine six standard deviations contains the large majority of cases. The distribution is approximately symmetric but carries a right tail, and a Shapiro-Wilk test returns a statistic of 0.89 with a p-value below 0.001, so the normality of the residuals is rejected and a small number of high-strength mixtures are under-predicted. A Spearman association between the absolute residuals and the predicted values gives a coefficient of 0.21 with a p-value of 0.001, which indicates mild heteroscedasticity in the form of slightly greater dispersion at higher predicted strength. The pattern is consistent with the sparser coverage of high-strength, low water-to-cement mixtures in the corpus, and it bounds the confidence that attaches to predictions in that region.

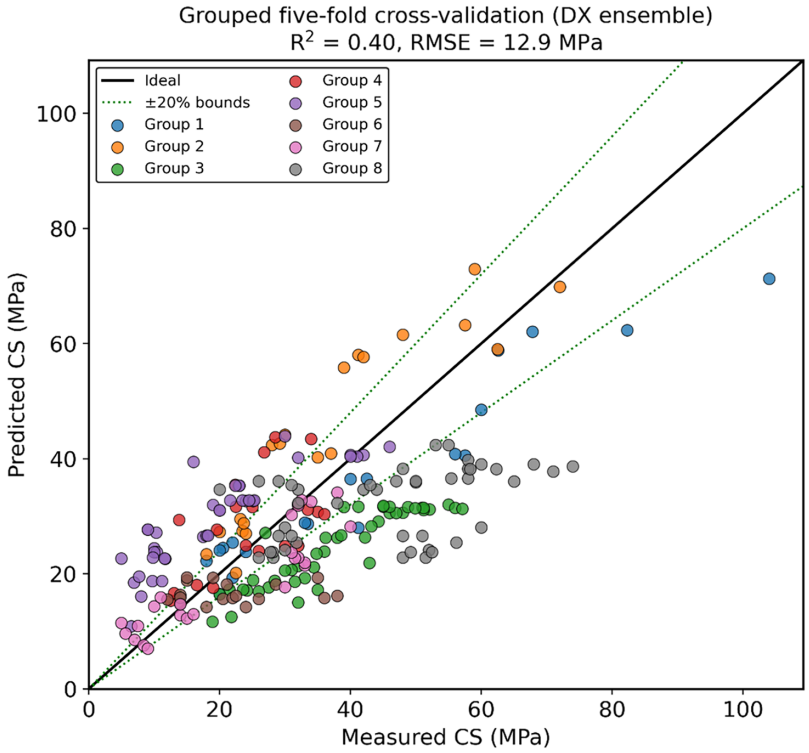

Figure 7: Parity plot of the DX ensemble under grouped five-fold cross-validation, with records colored by source-study group.

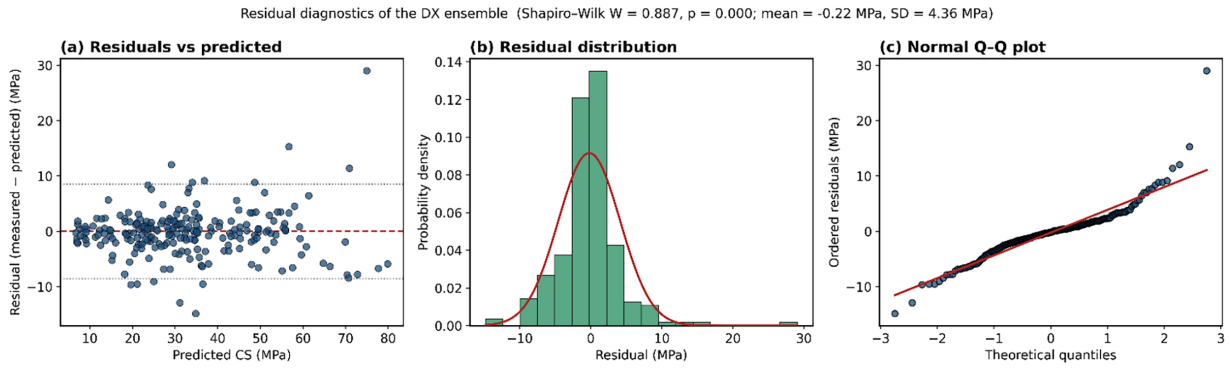

Figure 8: Residual diagnostics of the DX ensemble (residuals vs. predicted, residual distribution with fitted normal density, and normal Q–Q plot).

4.4 Sensitivity Analysis Using SHAP

Fig. 9 provides a global interpretation of the DX ensemble model by showing the overall importance and directional influence of each input variable across the entire dataset. Fig. 9 presents the SHAP feature importance and SHAP summary analysis of the DX ensemble model, which was selected for interpretation due to its superior predictive performance among all developed models.

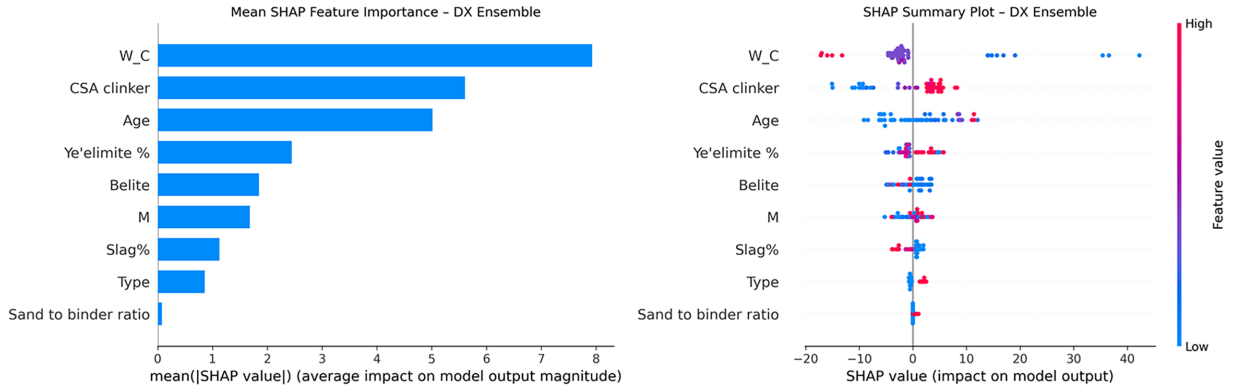

Figure 9: SHAP summary plots of ensemble DX ML model.

The SHAP feature importance analysis reveals a clear hierarchy of controlling variables governing compressive strength prediction. Among all input parameters, the water-to-cement ratio (W/C) is identified as the most influential feature, with the highest mean SHAP value of approximately 8. This indicates that water availability plays the dominant role in controlling hydration kinetics, pore structure development, and strength evolution in CSA cement systems. The CSA clinker content is the second most influential parameter, with a mean SHAP value of approximately 5.7, demonstrating the strong contribution of reactive clinker phases toward strength development. Curing age also exhibits significant influence, with a SHAP importance close to 5, confirming the continuous hydration and microstructural densification occurring over time. Secondary variables, including ye'elimite content, belite content, M value, and slag percentage, show comparatively lower SHAP contributions, generally below 3. Among these, ye'elimite content exhibits greater importance than slag and Belite, indicating its critical role in ettringite formation and early hydration reactions. In contrast, the sand-to-binder ratio and sample type contribute minimally to the prediction output, suggesting limited influence on compressive strength compared with binder chemistry and water availability.

The SHAP summary plot further illustrates both the magnitude and directional influence of each variable on model predictions. For W/C ratio, higher feature values (red points) are predominantly associated with negative SHAP values, indicating that increasing W/C reduces predicted compressive strength due to increased porosity and reduced matrix densification. Conversely, lower W/C values (blue points) shift the SHAP values toward positive regions, confirming their beneficial effect on strength enhancement. For CSA clinker content and curing age, higher values generally correspond to positive SHAP contributions, demonstrating that increased clinker dosage and longer curing duration promote hydration product formation and strength development. Similarly, higher ye'elimite content tends to positively influence compressive strength through accelerated ettringite generation. The influence of slag, Belite, and M value appears more distributed and nonlinear, reflecting their secondary but still important contributions to hydration and microstructural evolution.

On the whole, the SHAP analysis confirms that the DX model predictions are primarily governed by W/C ratio, CSA clinker content, and curing age, while other compositional variables contribute comparatively smaller effects. The consistency between SHAP importance ranking and known hydration mechanisms further validates the reliability and physical interpretability of the DX ensemble model for predicting the compressive strength behavior of CSA cement systems.

To further investigate local prediction behavior, Fig. 10 presents the SHAP waterfall plot for a representative prediction instance of the DX ensemble model. Unlike the global SHAP summary plot in Fig. 8,

the waterfall plot provides a local interpretation by explaining how each input feature contributes to one specific prediction. The predicted compressive strength for this sample was 30.43 MPa, compared with the dataset baseline prediction of approximately 34.03 MPa. The waterfall plot demonstrates how individual features contribute positively or negatively to the final prediction. Among all variables, Belite content (18%) exhibited the strongest negative contribution (−3.55 MPa), followed by Ye'elimite content (−2.29 MPa) and W/C ratio (−2.16 MPa). These variables collectively reduced the predicted compressive strength. In contrast, CSA clinker content (90%) showed the largest positive contribution (+3.43 MPa), confirming the beneficial effect of reactive clinker phases on strength development. Slag content (+0.72 MPa), curing age (+0.74 MPa), and M value (+0.20 MPa) also contributed positively, although to a smaller extent. Overall, the lower predicted strength resulted from the combined negative effects of Belite content, ye'elimite proportion, and W/C ratio outweighing the positive contributions from CSA clinker and slag.

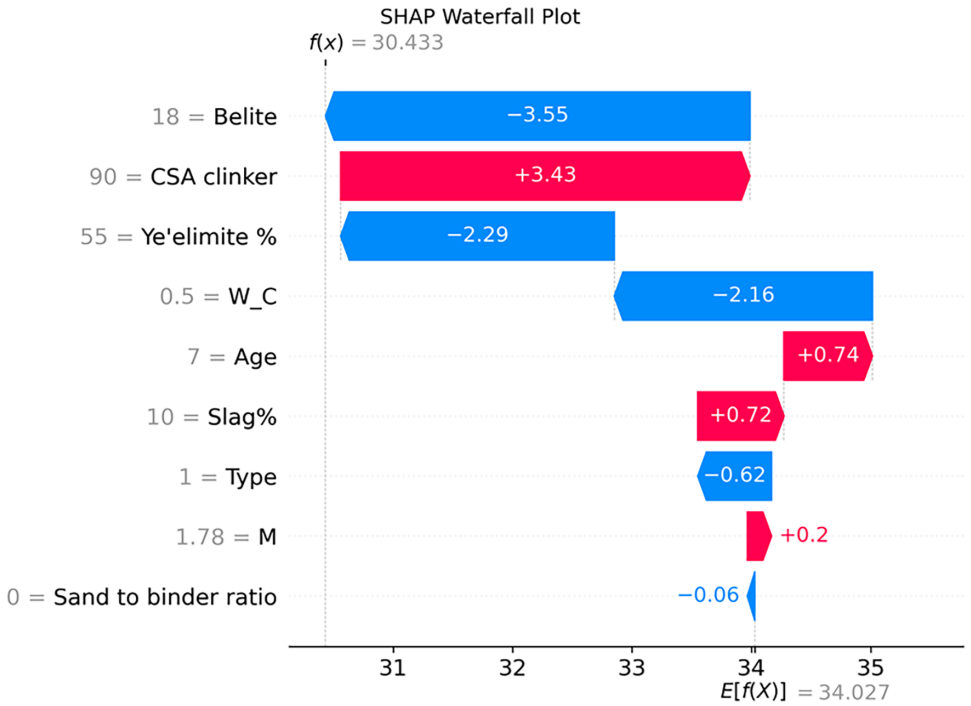

Figure 10: SHAP waterfall plot.

Fig. 11 illustrates the distribution of SHAP values for the molar ratio M, highlighting its interaction effects through the vertical spread of SHAP values color-coded by a second interacting variable (CSA clinker content in Fig. 11a and slag content in Fig. 11b). A vertical color gradient at a given M value thus reflects how the influence of M on the predicted compressive strength varies with the interacting parameter. In Fig. 11a (M vs. CSA clinker), the SHAP values for M vary significantly with changes in CSA clinker content. At lower M values (approximately 0.7–1.0), most SHAP values remain negative, particularly for mixtures with lower CSA clinker content, indicating that insufficient M dosage adversely affects compressive strength. However, as the M value increases to approximately 1.7–2.4, the SHAP values become predominantly positive, especially for mixtures containing higher CSA clinker contents (70%–100%). This behavior suggests a synergistic interaction between M and CSA clinker, where higher clinker content enhances the beneficial contribution of M toward strength development. The positive SHAP values at higher M levels indicate improved hydration and microstructural densification within the optimized compositional range.

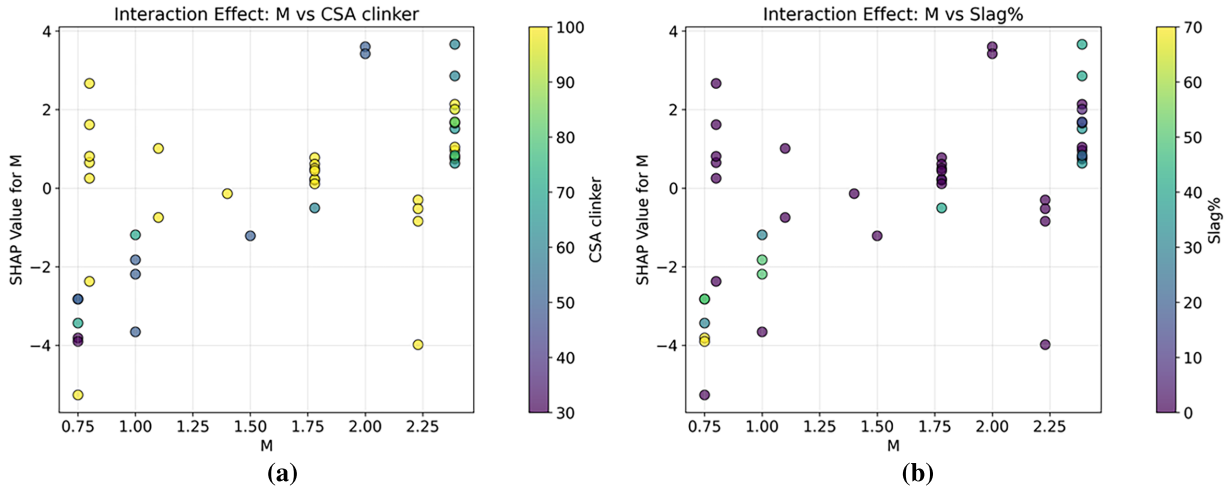

Figure 11: SHAP dependency plot for (**a**) interaction effect of M and CSA clinker, and (**b**) interaction effect of M and Slag on CS.

Fig. 11b shows the interaction between M and slag content. At low M values, mixtures with higher slag contents generally exhibit strongly negative SHAP values, reaching approximately −5, indicating that excessive slag combined with insufficient M adversely influences strength development. In contrast, increasing the M value shifts the SHAP contributions toward positive values, particularly at moderate slag contents. This suggests that adequate M dosage compensates for the slower reactivity of slag and contributes to improved hydration synergy and matrix development. Nevertheless, the interaction appears more scattered than the CSA clinker relationship, indicating a comparatively weaker but still important interaction effect between M and slag.

Collectively, the interaction analysis demonstrates that the effect of M on compressive strength is not independent but strongly influenced by binder composition. Higher M values generally contribute positively to strength prediction when combined with sufficient CSA clinker and moderate slag contents, highlighting the importance of compositional balance in optimizing CSA cement systems.

4.5 Multi-Objective Optimization Using NSGA II

In the previous section, reliable ML models were developed to predict the compressive strength of CSA cement systems. However, the practical design of sustainable cementitious materials requires not only high mechanical performance but also reduced material cost and lower embodied CO_2 emissions. Since CS, cost, and environmental impact are all strongly influenced by mixture composition, a multi-objective optimization framework was employed to identify optimal CSA mixtures with balanced mechanical, economic, and environmental performance.

Fig. 12 and Table 7 present the multi-objective optimization results obtained using the NSGA-II for identifying the optimal formulation of the CSA-slag binder system at different water-to-cement (W/C) ratios. The optimization simultaneously considered three conflicting objectives: maximizing CS while minimizing cost and embodied CO_2 emissions. The 3D Pareto front shown in Fig. 11 illustrates the trade-off relationship among these parameters for all optimized mixtures at different W/C ratios. Each colored marker represents a non-dominated solution generated by the NSGA-II algorithm, where improvement in one objective is generally accompanied by deterioration in another. It should be noted that the mixtures listed in Table 7 should not be interpreted as a single-variable W/C sequence. Rather, each row represents an

independent Pareto-optimal solution in which W/C ratio, CSA clinker content, slag content, and M-value vary simultaneously under the optimization constraints. Therefore, the compressive-strength variation is not expected to follow a strictly monotonic trend with W/C ratio or CSA clinker content alone. For example, mixtures with lower W/C ratios exhibit significantly higher compressive strength, driven primarily by the reduced water-to-cement ratio; the associated increase in cost and CO_2 emissions arises from the greater binder mass per unit volume at low W/C rather than from a higher CSA clinker fraction, since the strongest low-W/C mixtures in Table 7 retain a high slag replacement level. Conversely, mixtures with higher W/C ratios tend to show lower environmental impact and cost but reduced mechanical performance.

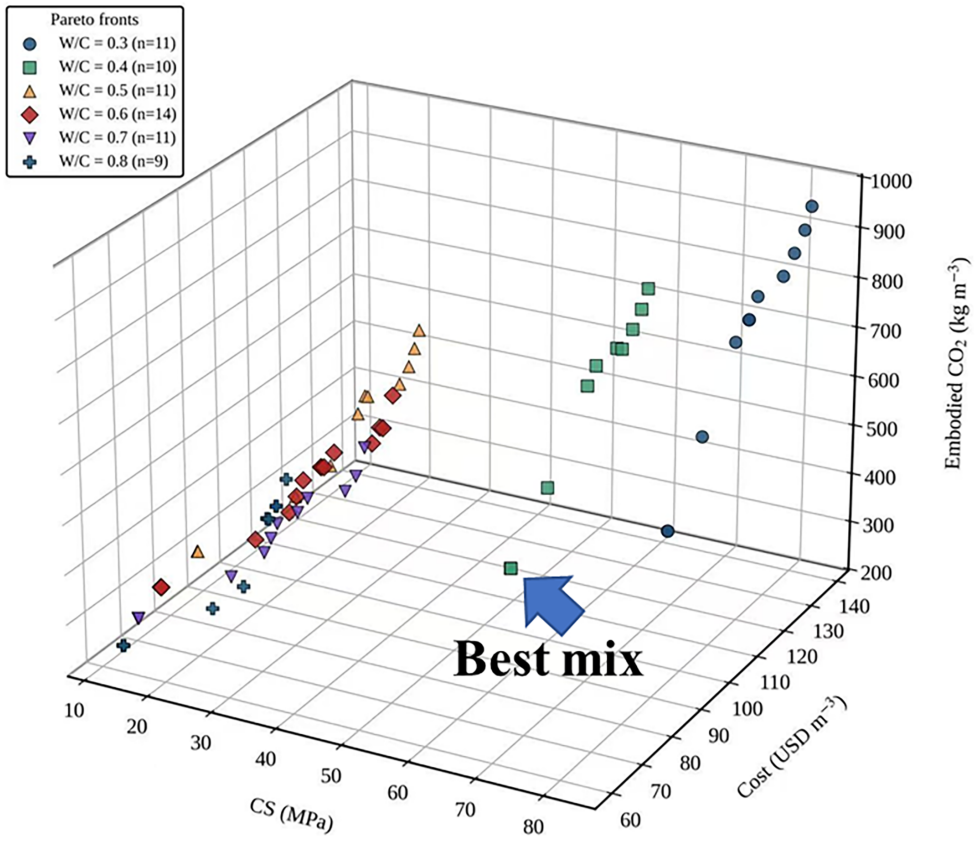

Figure 12: Pareto-optimal solution using NSGA II.

Table 7: Optimized formulation at each W/C ratio.

W/C	CSA clinker (%)	Slag (%)	M (mol/mol)	CS (MPa)	Cost (USD/m^3)	CO_2 (kg/m^3)
0.3	30	70	1.76	75.4	105.7	437
0.4	**30**	**70**	0.86	**57.9**	**91.8**	**379**
0.5	60	40	2.04	31.6	91.3	503
0.6	65	35	2.39	28.5	84	475
0.7	65	35	2.39	28.5	76.1	430
0.8	60	40	2.39	28.2	68.3	372

Note: The bold row indicates the selected optimized mixture with the best overall balance among compressive strength, material cost, and embodied CO_2 emissions.

Among all Pareto-optimal solutions, the model-predicted formulation with a W/C ratio of 0.4 was identified as the most balanced formulation, as highlighted in Fig. 11 by the "Best mix" marker. As summarized in Table 7, this optimized mixture contains 30% CSA clinker and 70% slag, achieving a compressive strength of 57.9 MPa with a relatively low cost of 91.8 USD/m^3 and embodied CO_2 emissions of 379 kg/m^3. Although the W/C = 0.3 mixture achieved the highest compressive strength (75.4 MPa), it also exhibited substantially higher cost and CO_2 emissions. In contrast, mixtures with W/C ratios of 0.6–0.8 showed lower cost and environmental impact but insufficient mechanical performance. Therefore, the W/C = 0.4 mixture represents the optimum balance between mechanical, economic, and environmental performance, demonstrating the effectiveness of the NSGA-II approach for sustainable mix design optimization of low-carbon CSA-slag binders.

5 Conclusion

- In this study, a comprehensive machine learning-assisted framework was developed for the prediction and multi-objective optimization of the compressive strength, cost, and embodied CO_2 emissions of sustainable CSA cement systems. Four stacked ensemble learning models (DR, DX, RX, and DRX) were established and systematically evaluated for compressive strength prediction. Among them, the DX model demonstrated the best predictive performance, achieving the highest accuracy (R^2 = 0.965) together with the lowest prediction errors (MAE = 2.362 MPa and RMSE = 3.168 MPa).
- SHAP-based interpretability analysis revealed that the water-to-cement ratio, CSA clinker content, and curing age were the dominant variables controlling compressive strength prediction, whereas slag content, M value, and clinker phase composition showed secondary but important contributions. SHAP interaction analysis further demonstrated strong synergistic relationships between M value, CSA clinker content, and slag dosage, confirming that the influence of individual parameters strongly depends on overall binder composition.
- To further balance engineering performance with sustainability objectives, a multi-objective optimization framework based on the NSGA-II algorithm was employed to simultaneously optimize compressive strength, cost, and embodied CO_2 emissions. The generated Pareto-optimal solutions clearly demonstrated the trade-offs between mechanical performance, environmental impact, and economic feasibility. Among all optimized mixtures, the formulation with a W/C ratio of 0.4 containing 30% CSA clinker and 70% slag exhibited the best overall balance, achieving a compressive strength of 57.9 MPa with relatively low embodied CO_2 emissions (379 kg/m^3) and cost (91.8 USD/m^3).

All in all, the proposed integrated framework provides an efficient and reliable strategy for the intelligent design and optimization of low-carbon CSA cement systems with enhanced mechanical, economic, and environmental performance.

Although the proposed framework provides an efficient data-driven strategy for identifying promising low-carbon CSA-slag cement formulations, the optimized mixture was not experimentally validated in this study. Therefore, future experimental work is required to verify the predicted optimum formulation and assess its practical performance and long-term reliability. The present optimization framework considered compressive strength as the only mechanical performance indicator. Future studies should include shrinkage, setting time, carbonation resistance, and durability-related properties when sufficient experimental data are available.

Acknowledgement: The authors would like to acknowledge the Research Center of Basic Sciences, Engineering and High Altitude, Taif University for funding this work.

Funding Statement: The authors appreciate the financial support from Research Center of Basic Sciences, Engineering and High Altitude, Taif University for funding this work.

Author Contributions: Ahmad Nawaz: conceptualization, methodology, formal analysis, software, investigation, writing—original draft; Hilal Khan: methodology, investigation, software, writing—review & editing; Salamat Ullah: methodology, investigation, writing—review & editing; Hamad Almujibah: writing—review & editing, supervision, funding acquisition; Ali E. A. Elshekh: writing—review & editing, supervision, funding acquisition; Maaz Osman Bashir: writing—review & editing, supervision, funding acquisition. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data used in this study were collected from published literature sources cited in the manuscript. The processed dataset can be made available from the corresponding author upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ren M, Ma T, Fang C, Liu X, Guo C, Zhang S, et al. Negative emission technology is key to decarbonizing China's cement industry. Appl Energy. 2023;329:120254. doi:10.1016/j.apenergy.2022.120254.
2. Dinga CD, Wen Z. China's green deal: can China's cement industry achieve carbon neutral emissions by 2060? Renew Sustain Energy Rev. 2022;155:111931. doi:10.1016/j.rser.2021.111931.
3. Gartner E. Industrially interesting approaches to low-CO_2 cements. Cem Concr Res. 2004;34(9):1489–98. doi:10.1016/j.cemconres.2004.01.021.
4. Gartner E, Hirao H. A review of alternative approaches to the reduction of CO_2 emissions associated with the manufacture of the binder phase in concrete. Cem Concr Res. 2015;78(7):126–42. doi:10.1016/j.cemconres.2015.04.012.
5. Kurtis KE, Alapati P, Burris LE, Ley MT, Peery J, Hajibabaee A, et al. Novel alternative cementitious materials for development of the next generation of sustainable transportation infrastructure [TechNote]. Washington, DC, USA: U.S. Department of Transportation Federal Highway Administration; 2014. doi:10.21949/1521792.
6. Deng C, Li L, Wang X, Tan Y, Zhou Y, Xu Z, et al. Fluidity and strength characteristics of grouts utilizing waste glass powder as cementitious substitute: effects of content and alkaline activator. Waste Biomass Valorization. 2026;13(3):2071. doi:10.1007/s12649-026-03501-6.
7. Haq IU, Bibi T, Nawaz A, Khalid HR, Elahi A. Calcined clay—substituted sustainable cement binders: a holistic review. Environ Sci Pollut Res. 2025;32(18):11317–49. doi:10.1007/s11356-025-36347-x.
8. Gong W, Wang N, Zhang N. Effect of metakaolin on water resistance of magnesium oxychloride cement. ACI Mater J. 2022;119(1):47–57. doi:10.14359/51734254.
9. Nawaz A, Kim N, Seo J, Farooq SZ, Lee HK. Hydration and phase conversion of MgO-modified calcium aluminate cement. Constr Build Mater. 2023;369(8):130425. doi:10.1016/j.conbuildmat.2023.130425.
10. Juenger MCG, Winnefeld F, Provis JL, Ideker JH. Advances in alternative cementitious binders. Cem Concr Res. 2011;41(12):1232–43. doi:10.1016/j.cemconres.2010.11.012.
11. Walling SA, Provis JL. Magnesia-based cements: a journey of 150 years, and cements for the future? Chem Rev. 2016;116(7):4170–204. doi:10.1021/acs.chemrev.5b00463.
12. Seo J, Nawaz A, Jang JG, Lee HK. Modifications in hydration kinetics and characteristics of calcium aluminate cement upon blending with calcium sulfoaluminate cement. Constr Build Mater. 2022;342(3):127958. doi:10.1016/j.conbuildmat.2022.127958.
13. Péra J, Ambroise J. New applications of calcium sulfoaluminate cement. Cem Concr Res. 2004;34(4):671–6. doi:10.1016/j.cemconres.2003.10.019.
14. Martin LHJ, Winnefeld F, Tschopp E, Müller CJ, Lothenbach B. Influence of fly ash on the hydration of calcium sulfoaluminate cement. Cem Concr Res. 2017;95:152–63. doi:10.1016/j.cemconres.2017.02.030.

15. Ke G, Zhang J, Xie S, Pei T. Rheological behavior of calcium sulfoaluminate cement paste with supplementary cementitious materials. Constr Build Mater. 2020;243(4):118234. doi:10.1016/j.conbuildmat.2020.118234.
16. Juenger MCG, Snellings R, Bernal SA. Supplementary cementitious materials: new sources, characterization, and performance insights. Cem Concr Res. 2019;122:257–73. doi:10.1016/j.cemconres.2019.05.008.
17. Guo Y, Chen S, Lakhiar MT, Zhuang S, Lakhiar IW. Ternary binders and recycled turbine blade fibres in mortar: reducing embodied carbon in coastal construction. Constr Build Mater. 2025;501(12):144277. doi:10.1016/j.conbuildmat.2025.144277.
18. Gao D, Meng Y, Yang L, Tang J, Lv M. Effect of ground granulated blast furnace slag on the properties of calcium sulfoaluminate cement. Constr Build Mater. 2019;227(2):116665. doi:10.1016/j.conbuildmat.2019.08.046.
19. Ali MK, Qader DN, Ali OAA, Raheem AM, Naser IJ, Al-Luhybi AS. Mechanical and durability performance of sustainable cement mortars: experimental optimization and multi-regression modelling. Next Mater. 2026;12(5):102418. doi:10.1016/j.nxmate.2026.102418.
20. García-Maté M, De la Torre AG, León-Reina L, Aranda MAG, Santacruz I. Hydration studies of calcium sulfoaluminate cements blended with fly ash. Cem Concr Res. 2013;54:12–20. doi:10.1016/j.cemconres.2013.07.010.
21. Seo J, Kim S, Yoon HN, Lee HK. Effect of the molar ratio of calcium sulfate over ye'elimite on the reaction of CSA cement/slag blends under an accelerated carbonation condition. J Build Eng. 2022;46:103785. doi:10.1016/j.jobe.2021.103785.
22. Yoon HN, Seo J, Kim S, Lee HK, Park S. Hydration of calcium sulfoaluminate cement blended with blast-furnace slag. Constr Build Mater. 2021;268:121214. doi:10.1016/j.conbuildmat.2020.121214.
23. Bertola F, Gastaldi D, Canonico F, Paul G. CSA and slag: towards CSA composite binders. Adv Cem Res. 2019;31(3):147–58. doi:10.1680/jadcr.18.00105.
24. Luo X, Li Y, Wang Q, Mu J, Liu Y. Machine learning based modeling for predicting the compressive strength of solid waste material-incorporated Magnesium Phosphate cement. J Clean Prod. 2024;442(3):141172. doi:10.1016/j.jclepro.2024.141172.
25. Balasooriya Arachchilage C, Huang G, Fan C, Liu WV. Forecasting unconfined compressive strength of calcium sulfoaluminate cement mixtures using ensemble machine learning techniques integrated with shapely-additive explanations. Constr Build Mater. 2023;409(7):134083. doi:10.1016/j.conbuildmat.2023.134083.
26. Jing M, Jia H, Liu Q, Zhang K, Xu S, Zheng X, et al. Multi-objective optimization design of cement-based materials for low-carbon goals. Mater Today Commun. 2025;44:112135. doi:10.1016/j.mtcomm.2025.112135.
27. Chen X, Chen W, Li Z, Zhang P. A hybrid prediction and multi-objective optimization framework for limestone calcined clay cement concrete mixture design. Sci Rep. 2025;15(1):22120. doi:10.1038/s41598-025-05288-3.
28. Nafiuzzaman M, Jakir TI, Aditi IJ, Kabir A, Ahsan KA. Different machine learning approaches to predict the compressive strength of composite cement concrete. J Build Pathol Rehabil. 2025;10(2):88. doi:10.1007/s41024-025-00598-5.
29. Jueyendah S, Lezgy-Nazargah M, Eskandari-Naddaf H, Emamian SA. Predicting the mechanical properties of cement mortar using the support vector machine approach. Constr Build Mater. 2021;291(2):123396. doi:10.1016/j.conbuildmat.2021.123396.
30. Nawaz A, Nawaz MZ, Lv L, Zhou Y. Machine learning based predictive insights into compressive strength and strength retention of magnesium oxychloride cement. Case Stud Constr Mater. 2025;23(5):e05135. doi:10.1016/j.cscm.2025.e05135.
31. Li Y, Li H, Jin C, Shen J. The study of effect of carbon nanotubes on the compressive strength of cement-based materials based on machine learning. Constr Build Mater. 2022;358(7):129435. doi:10.1016/j.conbuildmat.2022.129435.
32. Ali Shah SF, Chen B, Zahid M, Ahmad MR. Compressive strength prediction of one-part alkali activated material enabled by interpretable machine learning. Constr Build Mater. 2022;360(5):129534. doi:10.1016/j.conbuildmat.2022.129534.
33. Verma S, Pant M, Snasel V. A comprehensive review on NSGA-II for multi-objective combinatorial optimization problems. IEEE Access. 2021;9:57757–91. doi:10.1109/ACCESS.2021.3070634.

34. Zhang P, Qian Y, Qian Q. Multi-objective optimization for materials design with improved NSGA-II. Mater Today Commun. 2021;28:102709. doi:10.1016/j.mtcomm.2021.102709.
35. Seo J, Kim S, Park S, Yoon HN, Lee HK. Carbonation of calcium sulfoaluminate cement blended with blast furnace slag. Cem Concr Compos. 2021;118:103918. doi:10.1016/j.cemconcomp.2020.103918.
36. Jeong Y, Hargis CW, Chun SC, Moon J. The effect of water and gypsum content on strätlingite formation in calcium sulfoaluminate-belite cement pastes. Constr Build Mater. 2018;166(9):712–22. doi:10.1016/j.conbuildmat.2018.01.153.
37. Yoon HN, Seo J, Park S. Effects of M-value on long-term microstructural evolution of CSA cement blended with slag. J Build Eng. 2024;82:108241. doi:10.1016/j.jobe.2023.108241.
38. Burris LE, Kurtis KE. Water-to-cement ratio of calcium sulfoaluminate belite cements: hydration, setting time, and strength development. Cement. 2022;8:100032. doi:10.1016/j.cement.2022.100032.
39. Chen IA, Juenger MCG. Synthesis and hydration of calcium sulfoaluminate-belite cements with varied phase compositions. J Mater Sci. 2011;46(8):2568–77. doi:10.1007/s10853-010-5109-9.
40. Gálvez-Martos JL, Chaliulina R, Elhoweris A, Mwanda J, Hakki A, Al-horr Y. Techno-economic assessment of calcium sulfoaluminate clinker production using elemental sulfur as raw material. J Clean Prod. 2021;301:126888. doi:10.1016/j.jclepro.2021.126888.
41. Huang Y, Huo Z, Ma G, Zhang L, Wang F, Zhang J. Multi-objective optimization of fly ash-slag based geopolymer considering strength, cost and CO_2 emission: a new framework based on tree-based ensemble models and NSGA-II. J Build Eng. 2023;68(7):106070. doi:10.1016/j.jobe.2023.106070.
42. Luo X, Li Y, Lin H, Li H, Shen J, Pan B, et al. Research on predicting compressive strength of magnesium silicate hydrate cement based on machine learning. Constr Build Mater. 2023;406(1):133412. doi:10.1016/j.conbuildmat.2023.133412.
43. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32. doi:10.1023/a:1010933404324.
44. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016 Aug 13–17; San Francisco, CA, USA. p. 785–94. doi:10.1145/2939672.2939785.
45. Holland JH. Genetic algorithms. Sci Am. 1992;267(1):66–72. doi:10.1038/scientificamerican0792-66.
46. Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput. 2002;6(2):182–97. doi:10.1109/4235.996017.