Article

Data-Driven Framework for The Flow Characteristics of Self-Compacting Concrete Using Machine Learning

Nestor Ulloa [1,2,*], Kerly Vaca [3], Pablo Mancheno [4], María Albuja [1], Fredy Romero [5]

[1] Facultad de Mecánica, Escuela Superior Politécnica de Chimborazo (ESPOCH), Panamericana Sur Km 1 ½, Riobamba 060155, Ecuador; maria.albuja@espoch.edu.ec (MA)

[2] Grupo de Investigación y Desarrollo de Nanotecnología, Materiales y Manufactura (GIDENM), Escuela Superior Politécnica de Chimborazo (ESPOCH), Panamericana Sur Km 1 ½, Riobamba 060155, Ecuador

[3] Dipartimento di Ingegneria Informatica, Modellistica, Elettronica e Sistemistica-DIMES, University of Calabria, Via Pietro Bucci, Rende 87036, Italy; vcvkly99t68z605a@unical.it (KV)

[4] Facultad de Ciencias Pecuarias, Escuela Superior Politécnica de Chimborazo (ESPOCH), Panamericana Sur km. 1 ½, Riobamba 060155, Ecuador; pablo.mancheno@espoch.edu.ec (PM)

[5] Facultad de Ciencias, Escuela Superior Politécnica de Chimborazo (ESPOCH), Panamericana Sur km. 1 ½, Riobamba 060155, Ecuador; fredyd.romero@espoch.edu.ec (FR)

* Correspondence: Nestor Ulloa, Email: nestor.ulloa@espoch.edu.ec.

ABSTRACT

This study used hybrid metaheuristic machine learning techniques to predict the workability characteristics of self-compacting concrete (SCC) mixes, focusing on the influence of coarse aggregate size. Three advanced techniques were applied: genetic programming via evolutionary approach optimization (GP-EAO), evolutionary polynomial regression via genetic algorithm optimization (EPR-GAO), and artificial neural networks via modified biological neuron optimization (ANN-MBNO). Because conventional concrete constituents and the maximum nominal size (MNS) of coarse aggregates critically influence both the fresh and hardened properties of concrete, they were integrated into the SCC database. The results demonstrated that MNS significantly impacts SCC workability, with smaller sizes (10–18.5 mm) enhancing metrics such as slump flow time, height, and diameter. Among the models, ANN-MBNO prediction demonstrated the highest accuracy, with R^2, mean absolute error, and root mean squared error values of 0.964, 7 mm, and 8 mm, respectively, significantly outperforming GP-EAO and EPR-GAO. Furthermore, the ANN-MBNO model effectively captured the relative importance of all input parameters, emphasizing the critical role of MNS. This study provides a robust framework for predicting SCC workability, offering practical tools for sustainable production and optimized resource utilization.

Received: 09 Apr 2026
Accepted: 30 Jun 2026
Published: 07 Jul 2026

KEYWORDS: self-compacting concrete (SCC); workability/slump; coarse aggregate size; ANN-MBNO; GP-EAO; EPR-GAO

ABBREVIATIONS

SCC, self-compacting concrete; GP-EAO, genetic programming via evolutionary approach optimization; EPR-GAO, evolutionary polynomial regression via genetic algorithm optimization; ANN-MBNO, artificial neural networks via modified biological neuron optimization; MNS, maximum nominal size; MAE, mean absolute error; MSE, mean squared error, RMSE, root mean squared error; MAPE, mean absolute percentage error; SSE, sum of squared errors; ITZ, interfacial transition zone

INTRODUCTION

Workability is one of the most important properties of self-compacting concrete (SCC) because it determines the ability of concrete to flow, spread, and consolidate under its own weight without external vibration [1,2]. SCC is designed to possess high flowability, adequate viscosity, and segregation resistance, enabling congested reinforcement zones and complex formworks to be filled while maintaining homogeneity [3–5]. The workability characteristics of SCC are typically evaluated via experimental procedures such as slump flow, V-funnel, and flow-time tests [3–6]. These tests assess the filling ability, passing ability, and viscosity of fresh concrete mixtures, thereby providing insight into the rheological behavior of SCC under practical construction conditions [7–16].

The rheological performance of SCC is strongly influenced by mixture composition and aggregate characteristics. Factors such as water-to-powder ratio, powder content, superplasticizer dosage, viscosity modifiers, and aggregate grading significantly affect the fluidity and stability of SCC mixtures [17–19]. Proper particle packing and optimized proportions of fine and coarse aggregates contribute to improved flowability and reduced segregation [18–20]. In particular, the size, shape, texture, and amount of coarse aggregates directly influence internal friction, packing density, and concrete mobility through confined spaces [21,22]. Smaller aggregate sizes generally improve flowability and passing ability, whereas excessively large aggregates may increase blockage tendencies and prolong flow time [23–25]. Aggregate morphology also affects the interfacial transition zone (ITZ), stress transfer mechanisms, and overall structural performance of SCC.

In addition to workability, the durability and sustainability of SCC are influenced by aggregate characteristics. Proper aggregate gradation minimizes void content, improves material efficiency, and reduces segregation risk and bleeding during casting [26]. Using SCC contributes to sustainable construction because it eliminates vibration requirements, reduces labor and energy consumption, minimizes noise pollution, and improves construction efficiency [26]. Furthermore, SCC often

incorporates supplementary cementitious materials such as fly ash and ground granulated blast furnace slag, thereby reducing the environmental impact associated with conventional cement production [27]. Enhanced compaction and durability also improve service life and reduce maintenance requirements [28–34].

Because SCC behavior is dependent on complex interactions between multiple material and rheological parameters, advanced computational and machine learning (ML) techniques have been increasingly applied in predictive modeling for concrete engineering. Several studies have successfully utilized ensemble learning, artificial neural networks (ANN), and hybrid soft computing techniques to predict concrete properties such as compressive strength, bond strength, durability, and surface characteristics [35–39]. Hybrid ML frameworks combined with optimization algorithms have demonstrated superior capability in handling nonlinear relationships and improving prediction accuracy in concrete applications [38,39].

Despite these advancements, few studies have focused on predicting SCC workability characteristics under different coarse aggregate sizes using optimized hybrid artificial intelligence (AI) approaches. The influence of aggregate size on parameters such as air entrapment, slump flow behavior, and flow time remains insufficiently explored in the field of integrated ML frameworks. Therefore, this study employs enhanced AI techniques to predict the workability characteristics of multiple SCC mixtures with different coarse aggregate sizes. The developed models were optimized via advanced metaheuristic approaches, namely, modified biological neuron optimization (MBNO) for ANN—ANN-MBNO, evolutionary approach optimization (EAO) for genetic programming (GP)—GP-EAO, and genetic algorithm optimization (GAO) for evolutionary polynomial regression (EPR)—EPR-GAO. These optimization strategies were adopted to improve model robustness, prediction accuracy, and computational efficiency.

The remainder of this paper is organized as follows: “Research Gap and Novelty Statement” presents the methodology, data acquisition procedures, and ML modeling framework. “Methodology” discusses the theoretical background of the hybrid optimization approaches—GP-EAO, EPR-GAO, and ANN-MBNO. “Theoretical Framework” presents a preliminary statistical analysis of the dataset. “Baseline Statistical Analysis of Collected Database” discusses the results, model performance, and prediction accuracy. “Results and Analysis” summarizes the major conclusions and future research directions.

RESEARCH GAP AND NOVELTY STATEMENT

Although significant progress has been made in understanding the rheological and workability behaviors of SCC, certain key limitations still exist in the predictive modeling of SCC flow characteristics under varying coarse aggregate conditions. Previous studies have primarily relied on empirical correlations, laboratory-based observations, or conventional ML approaches to evaluate the behavior of fresh concrete. Although these studies have provided valuable insight into SCC performance, they often considered workability indicators independently and did not adequately capture the highly nonlinear interactions between aggregate size, particle distribution, flowability, viscosity, and stability parameters. In addition, existing ML studies on SCC and fresh concrete properties have largely focused on compressive strength prediction, durability assessment, or generalized rheological behavior, with comparatively limited attention given to the systematic prediction of V-funnel slump flow characteristics and flow time responses under varying coarse aggregate size. Furthermore, previous ML-based studies commonly employed standalone algorithms without advanced optimization strategies, thus limiting model adaptability, convergence efficiency, and predictive robustness when handling complex SCC datasets. The integration of biologically inspired metaheuristic optimization with AI-driven workability prediction remains insufficiently explored, particularly in modeling the combined influence of coarse aggregate size, air entrapment behavior, and rheological flow response in SCC systems. This study advances previous research by developing an integrated hybrid AI framework that combines ML with advanced metaheuristic optimization techniques for the predictive modeling of SCC workability characteristics. Specifically, this study introduces ANN-MBNO, GP-EAO, and EPR-GAO. Unlike conventional ML applications, the proposed framework simultaneously improves the prediction accuracy, convergence stability, computational efficiency, and interpretability of SCC flow behavior. The novelty of this research, therefore, lies not only in the application of ML to SCC but also in the establishment of a biologically inspired hybrid optimization framework capable of accurately capturing the nonlinear interdependence between coarse aggregate size and SCC workability indicators, including slump flow behavior and flow time. In addition, this study establishes interpretable predictive relationships that support intelligent SCC mix optimization for enhanced flowability, stability, constructability, and sustainability under practical construction conditions. Consequently, this study contributes to a more robust and scientifically advanced predictive methodology that extends beyond existing ML-based SCC studies and offers new routes for data-driven fresh concrete engineering.

METHODOLOGY

The methodology utilized in this study incorporated the collection of data related to the workability of concrete mixes, with an emphasis on their behavior under the influence of different coarse aggregate sizes sampled from multiple mixes [1]. The collected data were used for ML analysis and the intelligent prediction of workability characteristics by considering different metaheuristic training optimization techniques. Three different AI techniques—GP-EAO, ANN-MBNO, and EPR-GAO—were used to predict the “slump height” (Hs), “slump diameter” (Ds) and “slump flow time” (Ts) of concrete mixes with different ratios using the collected database.

The hyperparameters used in the ML models were selected by combining grid search and cross-validation techniques to ensure optimal performance and generalization. Key hyperparameters were rigorously adjusted for each model, including the number of hidden neurons (ANN-MBNO), population size and generations (GP-EAO), and polynomial degree and number of terms (EPR-GAO). Although k-fold cross-validation was employed to assess the model performance on various data subsets, the grid search method iteratively investigated a predetermined range of hyperparameter values. The final hyperparameters were selected based on minimizing error metrics such as the mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE) on the validation set, ensuring robust and accurate predictions.

All three developed models were used to predict Hs, Ds, and Ts using the concrete component contents, maximum nominal size (MNS), and air void ratio (A). Each of the three developed models was based on a different approach (evolutionary approach for GP, mimicking biological neurons for ANN, and an optimized mathematical regression technique for EPR). However, for all developed models, the prediction accuracy was evaluated in terms of MSE, RMSE, MAE, mean absolute percentage error (MAPE), and sum of squared errors (SSE). The following section discusses the results for each model. The accuracy of each developed model was evaluated by comparing the MSE, RMSE, MAE, MAPE, and SSE metrics of the predicted and calculated Hs, Ds, and Ts values.

This approach serves as the cornerstone for creating and utilizing sophisticated ML models. The theoretical underpinnings of the hybrid metaheuristic techniques used—including their mathematical formulations and optimization strategies—are presented in the following section.

THEORETICAL FRAMEWORK

GP-EAO

GP-EAO Mathematical Formulation

GP-EAO uses a fitness function that calculates prediction accuracy to determine the appropriateness of each solution [40]. The evolutionary process improves the solutions through crossing, mutation, and selection. While crossover and mutation aim to produce new solutions by merging or changing the components of preexisting solutions, selection selects high-performing individuals based on their fitness ratings. The evolutionary process continues for several generations until either a certain number of generations are reached or the solutions attain the required level of precision. The sum of terms—each with an evolution-optimized coefficient and basis function—is the final mathematical expression produced by GP-EAO.

Mathematically, the formulation is expressed as

$$f(x) = \sum_{i=1}^{n} \alpha_i \times \phi_i(x), \quad (1)$$

where a_i denotes the GP-optimized coefficients, and $\varphi_i(x)$ denotes the nonlinear basis functions selected during evolution. The GP-EAO model was developed using an evolutionary learning framework in which the optimization process was controlled by a predefined set of hyperparameters to achieve a balance between prediction accuracy and model complexity. The model was initialized using a fixed random seed to ensure reproducibility of the results throughout successive training runs. A learning rate of 0.01 was adopted in the optimization stage to facilitate stable convergence while minimizing the risk of premature stagnation. The optimization process was executed for 1000 iterations, which was sufficient for the evolutionary search to converge toward an optimal solution. The evolutionary optimizer was configured to minimize prediction error while maintaining the generalization capability of the model through iterative refinement of candidate solutions across successive generations. To control model complexity and reduce overfitting, the GP component employed a population size of 500 individuals with a maximum tree depth of 8. For parent selection, tournament selection with a tournament size of 5 was used, and crossover and mutation probabilities were set at 0.85 and 0.10, respectively, to maintain an appropriate balance between exploration and exploitation of the search space. Elitism was incorporated by preserving the top 5% of the individuals in each generation to prevent the loss of high-quality solutions. GP-EAO further utilized a selection pressure factor of 2.0 and convergence tolerance of 1×10^{-6} to guide the search toward globally optimal solutions. These hyperparameter settings enabled the GP-EAO model to efficiently explore the complex nonlinear relationships between the SCC constituent

materials and workability characteristics while maintaining computational efficiency and prediction robustness.

ANN-MBNO

The ANN-MBNO optimization procedure mimics the features of biological neural networks. The layers that comprise the structure of this model include an input layer, one or more hidden layers, and an output layer. These levels are connected using weighted connections [41]. Based on the input signals, each neuron uses an activation function—typically a hyperbolic tangent (tanh) activation function—to generate its output. Biological neural adaptation-inspired optimization techniques such as modifying learning rates or utilizing Hebbian learning principles in which weights increase in response to linked neuron activation are used to modify the weights. Both forward and backward stages are used in the ANN-MBNO learning process. The forward pass makes predictions using the weights currently in use, and backpropagation reduces prediction error by readjusting the weights to obtain a better fit. This method continually updates connections in a biologically inspired manner, allowing ANN-MBNO to capture complicated correlations in the data.

Forward prediction is expressed as

$$y = f(\sum w_i x_i + b). \quad (2)$$

Backward, the loss function (MSE) is minimized by adjusting the weights as follows:

$$y = f(WX + b), \quad (3)$$

where W denotes the weight matrix, X denotes the input vector, and f denotes the activation function. ANN-MBNO was implemented in MATLAB and trained to capture the nonlinear relationships between the SCC mixture constituents and workability responses. The network architecture and training process were governed by a set of predefined hyperparameters to ensure reproducibility and stable convergence. A fixed random seed was adopted at the beginning of each simulation to ensure consistent initialization of the weights and repeatability of the results. The training process was performed for a maximum of 1000 iterations (epochs), which was sufficient for the model to reach convergence based on the minimization of the loss function. A learning rate of 0.01 was employed to ensure stable gradient-based updates while preventing oscillations during training. The neural network weights were optimized using a hybrid MBNO-guided search integrated with a gradient-based refinement in MATLAB. The MBNO component incorporated biologically inspired evolutionary mechanisms to enhance the exploration and exploitation capabilities of the ANN training process. The population-based search strategy used had an initial population size of 50 candidate solutions, representing different network weight configurations. Tournament selection was applied with a selection pressure factor of 2 to

identify high-performing candidate solutions for reproduction. Crossover and mutation probabilities were set at 0.80 and 0.15, respectively, to maintain diversity in the solution space while ensuring convergence toward optimal weights. Elitism was included by preserving the top 5% of the solutions across generations to maintain high-quality solutions. The optimization process was executed over 100 generations, with convergence monitored through MSE stabilization. These hyperparameter settings enabled the ANN-MBNO model to efficiently learn complex nonlinear mappings in SCC workability prediction while maintaining robustness, stability, and strong generalization performance.

EPR-GAO

EPR is a hybrid optimization technique that blends evolutionary algorithms with polynomial regression. The EPR-GAO model refines polynomial models that depict the link between the input factors and expected results using a GA. This method builds polynomial equations using terms that represent variables raised to different powers, each of which is multiplied by coefficients. EPR-GAO creates an initial population of polynomial models with powers and coefficients assigned randomly for initialization. For improved performance, these models go through crossover, selection, and mutation as part of the GA evolutionary phases, focusing on a fitness function that, similar to GP-EAO, is frequently focused on minimizing prediction error selection, thereby ranking the most accurate polynomial models first. The model is further refined using new terms and combinations produced by crossover and mutation. The complicated relationships in the data are represented by polynomial equations tuned for prediction accuracy through this iterative procedure.

Mathematically, this is expressed as

$$y = \sum_{i=1}^{m} \alpha_i \times x^{\beta_i}, \quad (4)$$

where α_i and β_i are GA-optimized coefficients and exponents, respectively. EPR-GAO was implemented in MATLAB to establish explicit polynomial relationships between the SCC input variables and workability responses. The model training process was controlled using a predefined set of hyperparameters to ensure reproducibility and stable convergence of the evolutionary search process. A fixed random seed was applied at the start of each simulation to ensure consistent initialization and replicable results. The optimization process was executed over a maximum of 1000 iterations (generations), thereby providing sufficient search depth for the GA to converge toward an optimal polynomial structure. A learning rate of 0.01 was adopted within the regression fitting component to regulate parameter updates and improve numerical stability during coefficient estimation. The objective function was defined based on the minimization of the prediction error between the experimental and predicted SCC workability parameters. The GAO stage of the EPR-GAO model employed a population-based evolutionary search strategy to identify the optimal

polynomial structure and coefficients. A population size of 60 individuals was used, with tournament selection (tournament size = 3) adopted for parent selection. The crossover and mutation probabilities were set at 0.85 and 0.10, respectively, to maintain an appropriate balance between exploration and exploitation of the search space. Elitism was implemented by preserving the best 5% of the solutions in each generation to maintain high-quality candidate models. The maximum polynomial depth was limited to 5 to avoid excessive model complexity and reduce overfitting risk. The convergence criterion was defined as either reaching the maximum number of generations or achieving a predefined fitness tolerance of 1×10^{-6}. These hyperparameter settings enabled the EPR-GAO model to effectively derive interpretable polynomial expressions while maintaining predictive accuracy and computational efficiency in SCC workability prediction tasks.

The overall framework of the model project is shown in Figure 1. The computational complexities of the predictive models—GP-EAO, EPR-GAO, and ANN-MBNO—differ because of their distinct underlying algorithms and optimization approaches. As the population size and number of generations increase, GP-EAO—which is dependent on evolutionary processes including selection, crossover, and mutation—becomes computationally costly and, thus, less effective for managing extremely complicated systems or large datasets because of its restricted scalability, despite the ability to generate interpretable symbolic expressions. In contrast, EPR-GAO optimizes polynomial terms and coefficients using a GA. Compared with the GP-EAO, the fixed polynomial structure of EPR-GAO lowers the computing overhead; nonetheless, as the search space grows, its effectiveness in high-dimensional or highly nonlinear systems decreases. ANN-MBNO exhibits the highest computational complexity because of its multi-layered architecture and iterative backpropagation algorithm—which adjusts the weights across numerous neurons and layers. Despite the cost increases with the number of layers and neurons, ANN-MBNO can handle enormous datasets and capture intricate nonlinear interactions because of hardware acceleration developments such as GPUs. ANN-MBNO's ability to produce reliable and precise forecasts outweighs its computational expense and greater requirements, especially in applications such as SCC workability modeling that demand accuracy.

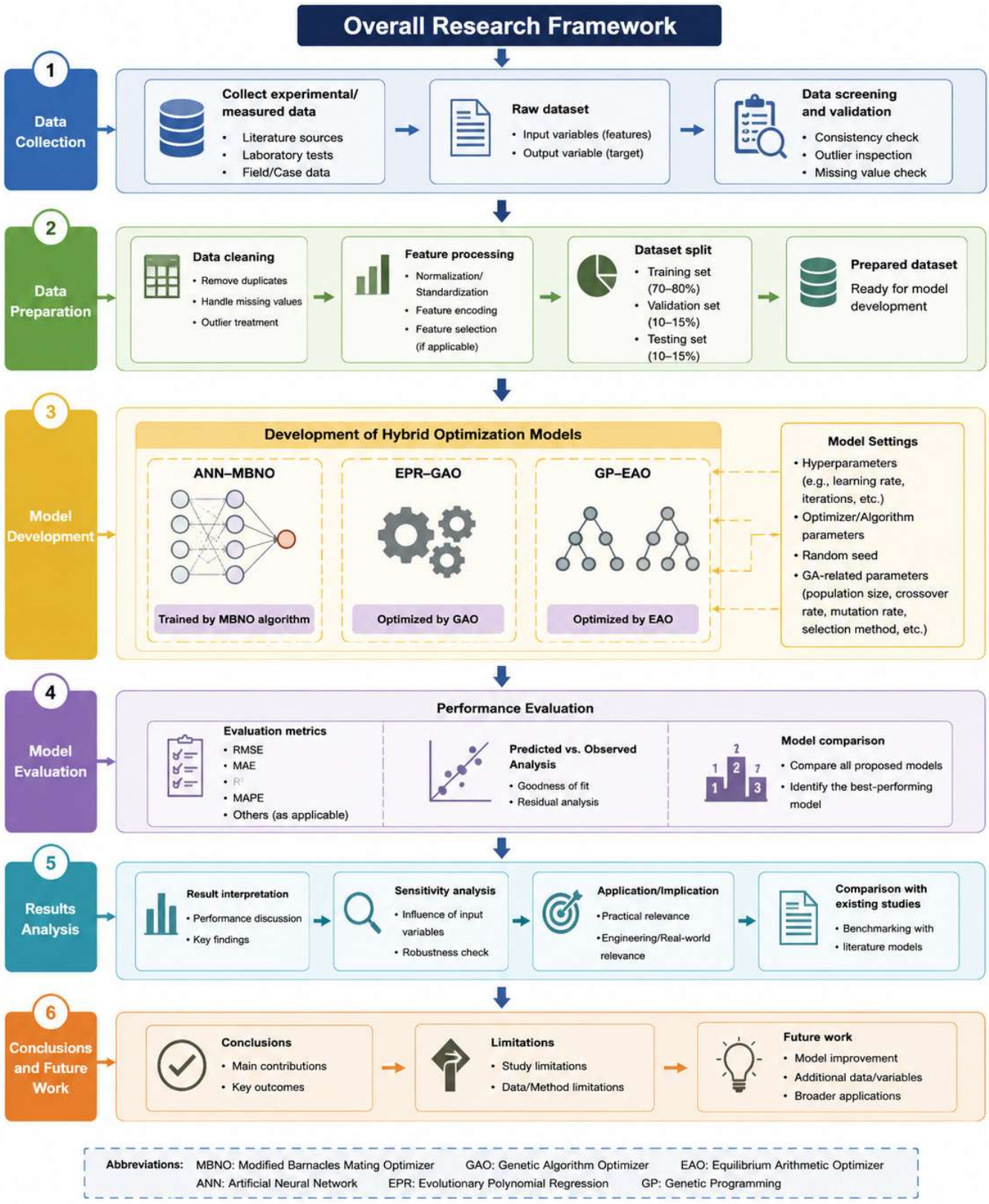


Figure 1. Overall research framework.

Prediction Accuracy Metrics

To assess the prediction performance of GP-EAO, ANN-MBNO, and EPR-GAO in line with recommendations in the literature [42,43], the key metrics used are as follows:

i. MSE

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \quad (5)$$

where y_i is the actual compressive strength, and $\hat{y}_i$ is the predicted compressive strength.

ii. RMSE

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}. \quad (6)$$

iii. MAE

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y_i}|. \quad (7)$$

iv. MAPE

$$MAPE = \frac{1}{n}\sum_{n=i}^{n}\left|\frac{y_i - \widehat{y_i}}{y_i}\right| \times 100. \quad (8)$$

v. SSE

$$SSE = \sum_{n=i}^{n}(y_i - \widehat{y_i})^2. \quad (9)$$

A quantitative measure—the a20-Index—has been used to assess the relative significance of the input variables in predictive models [44,45], especially in complex systems such as SCC workability. This measure helps identify crucial elements for SCC mix design optimization by quantifying the impact of each individual input parameter on the overall performance or output variability.

Mathematically, the a20-Index is expressed as

$$a20 = \frac{\Delta Y_{20}}{Y_{max} - Y_{min}} \times 100, \quad (10)$$

where ΔY_{20} is the change in the output parameter when the input variable is altered by ±20% of its nominal value, Y_{max} is the maximum value of the output parameter in the dataset, and Y_{min} is the minimum value of the output parameter in the dataset. The a20-Index offers a normalized sensitivity analysis that shows how each input variable proportionately affects output variability. A higher a20-Index indicates that the parameter has a greater impact on the model's predictions. The a20 index—a20 or (AUC-20) "Area Under the Curve for top 20% of ranked predictions"—is a performance metric for non-numerical or numerical discrete databases and is typically used with classifiers to reflect classification accuracy. This measure cannot be used with symbolic regression techniques such as GP, EPR, or ANN because they are not classifiers (their output being mathematical formulas rather than classes) and the databases used are

numerical-continuous [44]. The evaluation and comparison of GP-EAO, ANN-MBNO, and EPR-GAO for structural and material property predictions was facilitated by the distinct insights into model performance offered by each metric.

In this study, several steps were taken to prevent overfitting in the optimally developed model by carefully selecting hyperparameters though tuning and following the partitioning pattern for training and validation [46–48]. Through this procedure, model complexity was adjusted to prevent overfitting of the training data, which can lead to poor generalization to new data. The models' performance and resilience were further assessed using a variety of error metrics, including MSE, RMSE, MAE, MAPE, and SSE. By using a training set of 60 samples and validation set of 25 samples and employing techniques such as Pearson correlation to identify key variables, we ensured that the model was well-trained and validated for real-world applicability.

BASELINE STATISTICAL ANALYSIS OF COLLECTED DATABASE

A total of 85 records were collected literature [1] for experimentally tested samples of concrete mixes with different ratios. The Monte Carlo optimization technique was used to increase the data size. Monte Carlo optimization techniques are employed to improve the number of entries in a dataset by generating statistically representative synthetic samples through probabilistic simulation based on the underlying distribution and variability of the original data. This approach is particularly valuable in engineering and ML applications where experimental datasets are limited, expensive, or difficult to obtain. By repeatedly sampling random values within defined statistical boundaries such as mean, variance, probability density, and parameter correlations, Monte Carlo methods expand the effective dataset size while preserving the inherent stochastic characteristics and nonlinear relationships between variables. The generated samples enhance data diversity, reduce sampling bias, improve model generalization capability, and increase predictive algorithm robustness against overfitting. Furthermore, Monte Carlo simulation allows for the exploration of uncertain parameter spaces and rare combinations that may not be sufficiently represented in the original dataset, thereby improving the stability and reliability of optimization and ML models. In predictive modeling, the increased dataset density obtained through Monte Carlo optimization contributes to better convergence behavior, improved training efficiency, and enhanced statistical confidence in model predictions, particularly in complex multiparameter systems such as concrete rheology, structural mechanics, and material performance analysis. The nominal sizes of the coarse aggregates ranged between 11 and 21 mm. Note that the entrapped air behavior in concrete reflects changes in the nominal size of the coarse aggregate in the mixes. Further, the ITZ is affected by the aggregate size because the smaller the

nominal size, the greater the interfacial interaction between the aggregates and cement [1]. Each record contains the following data:

- C: cement content (kg/m^3)
- W: water content (kg/m^3)
- CAg: coarse aggregates content (kg/m^3)
- FAg: fine aggregates content (kg/m^3)
- MNS: maximum nominal size of aggregates (mm)
- A: air percent (%)
- Hs: slump height (mm)
- Ds: slump diameter (mm)
- Ts: slump flow time (s).

The collected records were divided into training (60 records) and validation (25 records) sets. Table 1 summarizes the statistical characteristics of the database, and Table 2 presents the Pearson correlation matrix. Finally, Figure 2 shows the histograms for both inputs and outputs.

Table 1. Statistical analysis of the collected database.

	C (kg/m^3)	W (kg/m^3)	Cag (kg/m^3)	FAg (kg/m^3)	MNS (mm)	A (%)	Hs (mm)	Ds (mm)	Ts (s)
Training set									
Max.	602	250	1298	1221	22	9.0	253	430	5.3
Min	237	156	689	481	11	0.00	67	198	1.5
Avg	399	200	922	783	15	4.02	176	313	2.9
SD	89	19	144	158	3	2.35	48	66	0.8
Var	0.22	0.09	0.16	0.20	0.19	0.59	0.27	0.21	0.3
Validation set									
Max.	616	241	1204	1176	22	8.8	252	438	4.0
Min	250	169	706	612	12	0.00	91	200	1.6
Avg	379	202	933	850	16	4.44	190	337	2.8
SD	94	20	109	152	3	2.39	37	59	0.5
Var	0.25	0.10	0.12	0.18	0.18	0.54	0.19	0.17	0.2

Table 2. Pearson correlation matrix.

	C	W	CAg	FAg	MNS	A	Hs	Ds	Ts
C	1.00								
W	0.54	1.00							
CAg	0.18	0.28	1.00						
FAg	−0.50	−0.05	−0.49	1.00					
MNS	0.18	0.24	0.58	−0.21	1.00				
A	−0.42	−0.47	−0.28	0.53	−0.36	1.00			
Hs	−0.46	0.29	−0.11	0.49	−0.10	−0.04	1.00		
Ds	−0.39	0.35	0.05	0.39	0.04	−0.05	0.88	1.00	
Ts	0.46	0.44	0.62	−0.50	0.40	−0.55	−0.22	−0.17	1.00

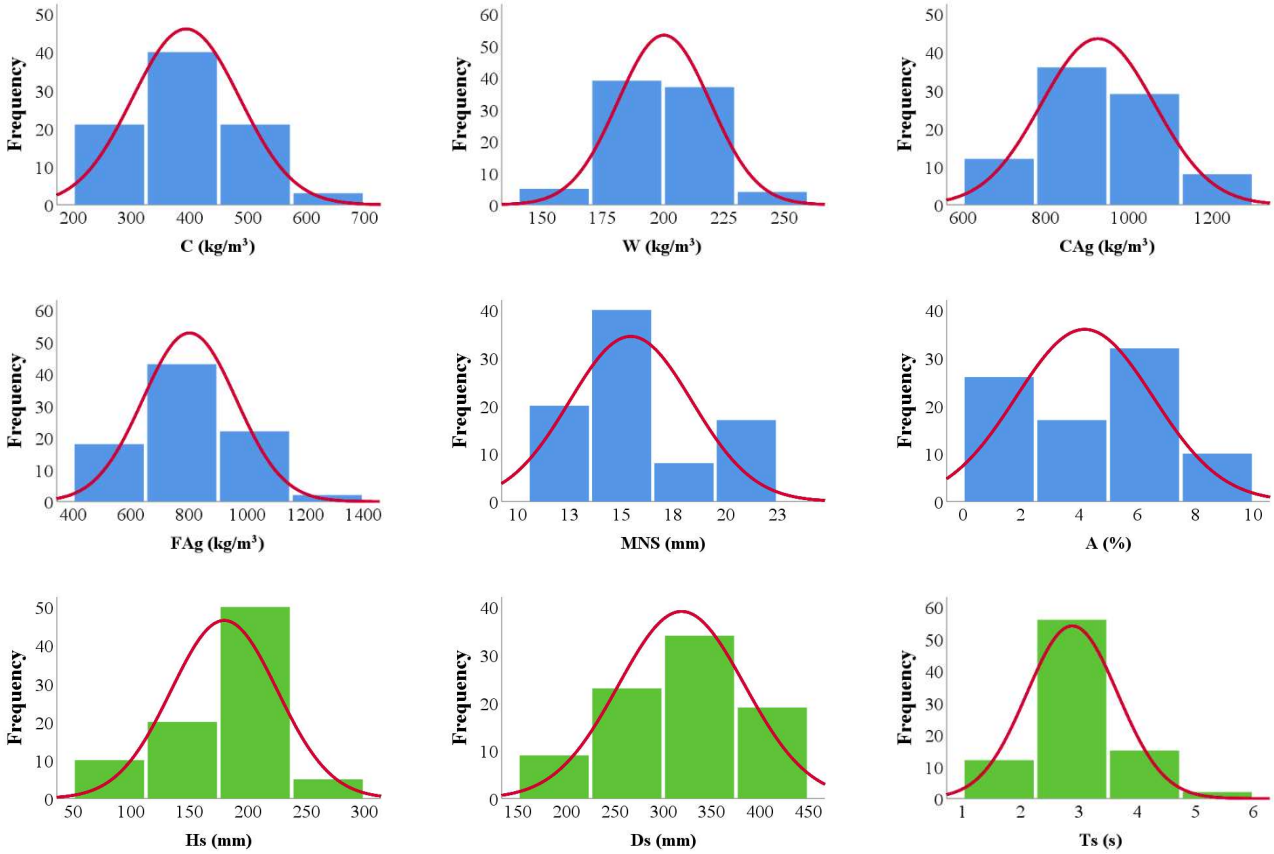


Figure 2. Distribution histograms for the inputs (in blue) and outputs (in green).

On this established theoretical basis, the results of the developed models are presented in the following section, accompanied by a comprehensive analysis of their predictive performance and accuracy.

RESULTS AND ANALYSIS

GP-EAO Model

Four GP models were developed for each output, starting with two levels of complexity and extending up to five levels of complexity. Population size, survivor size, and generation number were set at 1000, 300, and 3000, respectively. Table 3 lists the error percentage of each output for each complexity level, and Figure 3 shows the average error percentage versus the complexity level. Equations (11)–(13) present the output closed-form formulas for Hs, Ds, and Ts, respectively, for the models with five levels of complexity. Average errors (%) for the total dataset of 12.5%, 9.52% and 14.0% and R^2 values of 0.634, 0.713 and 0.679 for the Hs, Ds, and Ts predictions, respectively, were obtained.

$$Hs = \frac{(C + A)^{0.9}}{Ln^2(W) - (C + A)^{0.9} + W} - (C + A)^{0.9} + 2\,W - MNS - A, \quad (11)$$

$$Ds = \frac{C - 2.5\,W}{MNS - Ln(CAg)} - \frac{CAg + MNS^2}{C - 2.5\,W} + \frac{W + C}{2.5} + 2.5\,W - A - C + 3.0, \quad (12)$$

$$Ts = \frac{CAg}{420} + \left(\frac{C}{A + FAg}\right) + \left(\frac{C}{A + FAg}\right)^{A^A} - \frac{CAg}{420\left(\sqrt{CAg} - \frac{A^A}{FAg}\right)}. \quad (13)$$

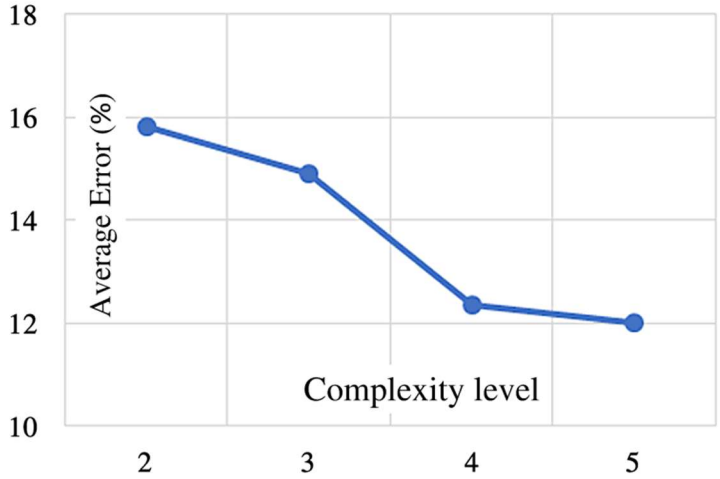


Figure 3. Average error (%) versus complexity level.

Table 3. Error (%) of each output versus complexity level.

Complexity level	Error (%)			
	Hs	Ds	Ts	Avg
2	15.8	13.1	18.5	15.8
3	15.4	11.9	17.4	14.9
4	12.7	10.3	14.1	12.4
5	12.5	9.5	14.0	12.0

EPR-GAO Model

Similarly to GP, five EPR models were developed, with the number of terms increased from 5 to 25. All EPR models were limited to a 6th level polynomial, and for 6 inputs, the 462 possible terms (252 + 126 + 56 + 21 + 6 + 1 = 462) were as follows:

$$\sum_{n=1}^{n=6}\sum_{m=1}^{m=6}\sum_{l=1}^{l=6}\sum_{k=1}^{k=6}\sum_{j=1}^{j=6}\sum_{i=1}^{i=6} X_n.X_m.X_l.X_k.X_j.X_i + \sum_{m=1}^{m=6}\sum_{l=1}^{l=6}\sum_{k=1}^{k=6}\sum_{j=1}^{j=6}\sum_{i=1}^{i=6} X_m.X_l.X_k.X_j.X_i + \sum_{l=1}^{l=6}\sum_{k=1}^{k=6}\sum_{j=1}^{j=6}\sum_{i=1}^{i=6} X_l.X_k.X_j.X_i + \sum_{k=1}^{k=6}\sum_{j=1}^{j=6}\sum_{i=1}^{i=6} X_k.X_j.X_i + \sum_{j=1}^{j=6}\sum_{i=1}^{i=6} X_j.X_i + \sum_{i=1}^{i=6} X_i + C. \quad (14)$$

Table 4 lists the error percentage of each output for each model, and Figure 4 shows the average error percentage versus the number of terms. Using the five terms models, Equations (15)–(17) present the output formulas for Hs, Ds, and Ts, respectively. Average error (%) and R^2 values of 11%, 11%, and 13% and −0.760, 0.633, and 0.68 were obtained for Hs, Ds, and Ts, respectively.

$$Hs = 253 + \frac{C{\cdot}W{\cdot}CAg{\cdot}A^2}{26\,256\,000} - \frac{C^3{\cdot}A}{81\,W^2} - \frac{FAg^2{\cdot}A}{2.5\,C^2} - \frac{6\,520\,000\,A^3}{C^2{\cdot}FAg} - \frac{11258\,CAg^2}{FAg^3}, \quad (15)$$

$$Ds = 673 - \frac{80417\,C}{W{\cdot}FAg} - \frac{C^2{\cdot}A}{13.8\,CAg} - \frac{2.3\times10^{18}}{C{\cdot}W{\cdot}CAg^3{\cdot}FAg} + \frac{5.1\times10^{14}\,A}{C{\cdot}W{\cdot}CAg^3} - \frac{3030\,FAg^2}{C^2{\cdot}W}, \quad (16)$$

$$Ts = 2.25 + \frac{4.15\,C{\cdot}FAg{\cdot}A}{W^3} + \frac{2.92\,W^2}{FAg^2} - \frac{3.69\,W^2{\cdot}A}{FAg^2} + \frac{W^3{\cdot}CAg}{7\,500\,000\,FAg} - \frac{FAg^2}{5\,W{\cdot}CAg}. \quad (17)$$

Table 4. Error (%) of each output versus number of terms.

No. of terms	Error (%)			
	Hs	Ds	Ts	Avg
5	11.1	10.5	13.4	11.7
10	11.1	10.0	12.7	11.3
15	11.9	9.3	13.3	11.5
20	10.4	10.3	13.3	11.3
25	10.4	10.3	13.3	11.3

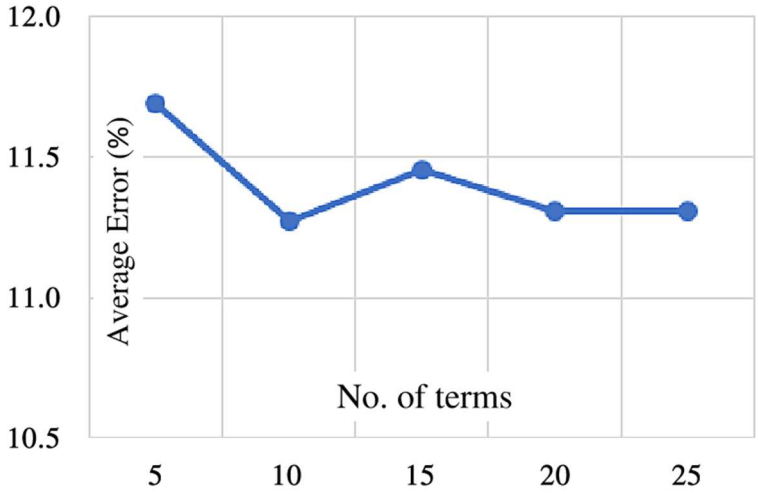


Figure 4. Average error (%) versus number of terms.

ANN-MBNO Model

Five models were developed using the ANN-MBNO technique to predict Hs, Ds, and Ts values. The biological neurons mimicking technique applied the normalization method (−1.0 to 1.0), activation function (Hyper-Tan), and back propagation training algorithm. The number of neurons in the hidden layer started with 2 and increased up to 10. Table 5 lists the error percentage of each output for each model, and Figure 5 shows the average error percentage versus the network layout. The final network layout is shown in Figure 6, and the weight matrix of the model is presented in Table 6. Average errors (%) for the total dataset of 5%, 3%, and 6% and R^2 values of 0.964, 0.980, and 0.955 were obtained for Hs, Ds, and Ts, respectively.

Table 5. Error (%) of each output versus ANN-MBNO layout.

ANN Layout	Error (%)			
	Hs	Ds	Ts	Avg
6-2-3	17.5	21.8	46.6	28.6
6-4-3	10.9	12.9	23.3	15.7
6-6-3	6.0	8.0	17.8	10.6
6-8-3	5.5	3.3	17.2	8.7
6-10-3	5.0	3.0	6.0	4.7

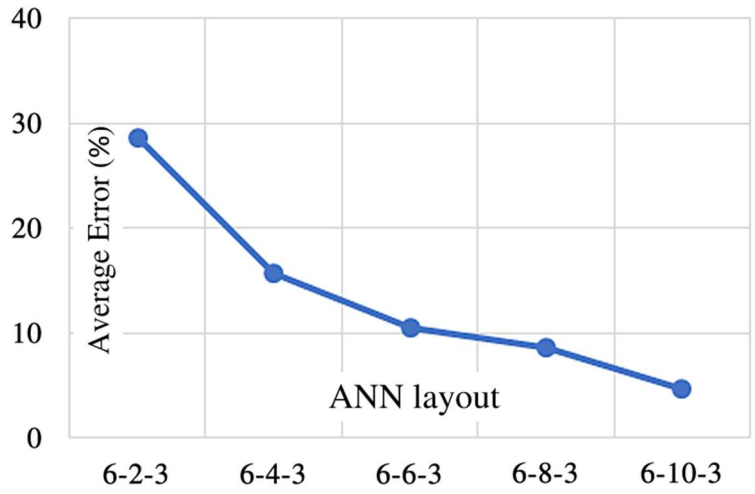


Figure 5. Average error (%) versus ANN-MBNO layout.

Table 6. Weight matrix for the developed ANN-MBNO models.

	H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	H(1:6)	H(1:7)	H(1:8)	H(1:9)	H(1:10)	
(Bias)	−0.95	7.45	−0.56	2.47	0.72	−0.38	−1.42	5.23	−25.42	−2.25	
C	−0.28	−20.53	1.03	6.53	0.68	−0.02	0.03	−9.46	9.92	−2.15	
W	−0.19	−0.35	1.10	4.06	−0.19	1.59	1.45	−9.87	−26.21	−0.47	
CAg	0.00	16.09	0.92	0.36	−0.50	0.63	−4.61	15.06	−18.95	−0.72	
FAg	−0.02	22.36	−1.52	−5.00	0.47	−1.54	3.70	−3.54	43.85	1.16	
MNS	0.17	−11.40	0.01	−5.46	−0.67	0.49	2.22	−0.72	−4.18	0.96	
A	−0.07	−11.71	−1.88	0.86	2.14	−2.30	−2.64	−11.54	7.68	0.39	
	H(1:1)	**H(1:2)**	**H(1:3)**	**H(1:4)**	**H(1:5)**	**H(1:6)**	**H(1:7)**	**H(1:8)**	**H(1:9)**	**H(1:10)**	**(Bias)**
Hs	−6.24	0.74	0.72	−0.34	−0.56	−1.01	0.33	0.19	−0.12	−0.06	−4.28
Ds	−6.40	0.90	1.88	−1.00	−0.61	−2.13	0.37	0.06	−0.06	−1.11	−4.97
Ts	−4.46	−0.12	−0.19	0.01	−1.10	−0.25	−0.46	−0.27	−0.51	0.87	−2.90

The insights derived from the results are contextualized in the Discussion section, where the implications of the findings and their alignment with prior research are explored in detail.

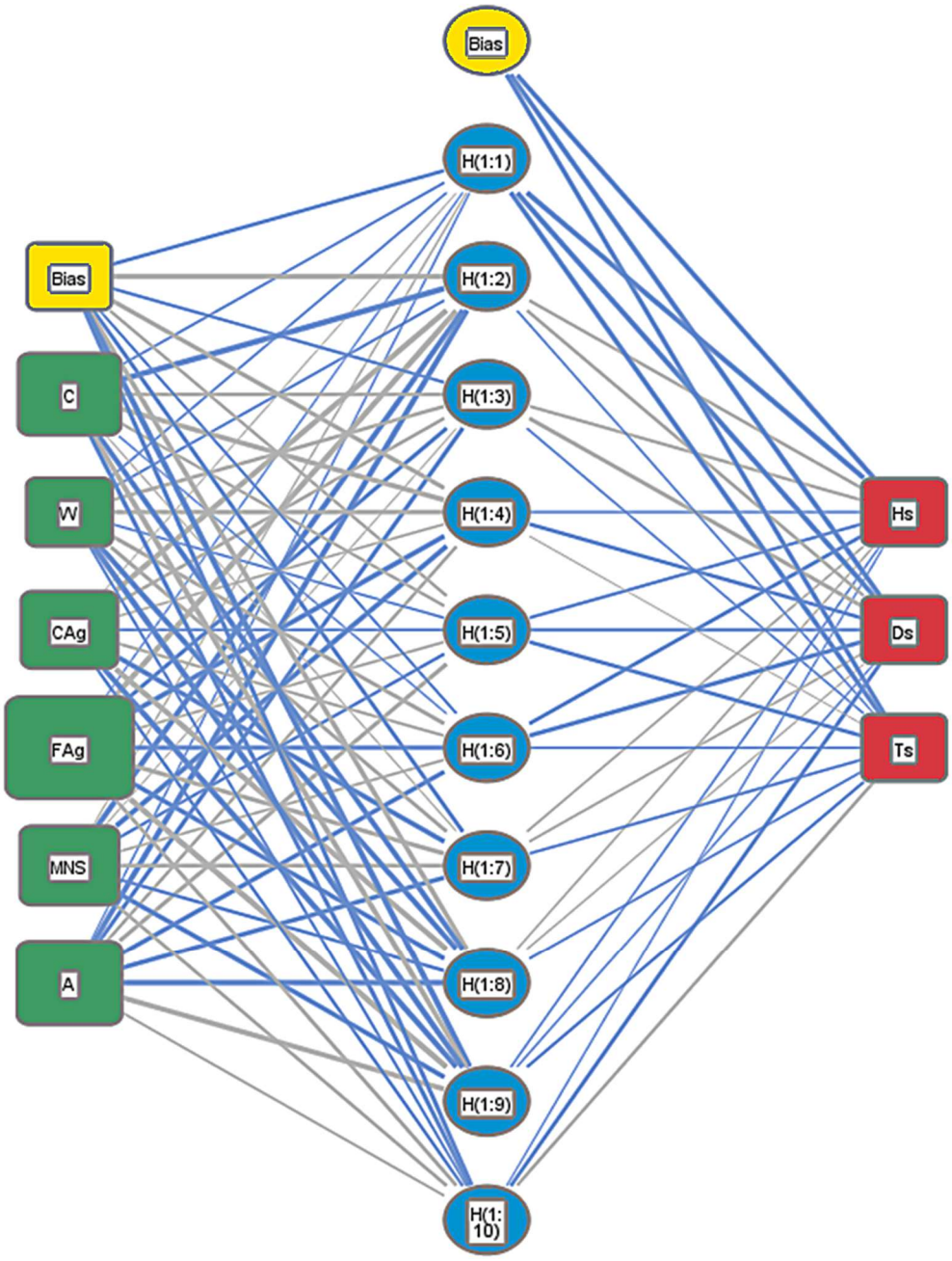


Figure 6. Layout architecture for the developed ANN-MBNO models.

General Discussion and Analysis

The relative importance values for each input parameter shown in Figure 7 indicate that all factors have almost the same importance. The sensitivity analysis of the SCC mix workability properties, with a focus on the influence of coarse aggregate sizes, provides valuable insight into the relative impact of various mix components in achieving optimal workability. Workability—a critical property of SCC—determines the ease with which concrete flows and consolidates without segregation or external compaction. The analysis revealed that the fine aggregates (FAg) exert the greatest influence on SCC workability, contributing 23% to the

variation in this property. The flowability and viscosity of concrete are controlled by the gradation and proportion of the fine aggregates, which therefore play a key role in the creation of a cohesive mix. The prominence of fine aggregates emphasizes the importance of optimizing their content to achieve the desired workability, while preventing segregation or excessive bleeding. MNS has a notable impact of 16%, reflecting the significance of aggregate gradation and size in influencing SCC flow. During flow, larger aggregates increase the risk of blocking, particularly in congested reinforcement areas, whereas smaller aggregates contribute to smoother flowability. This finding underscores the need for a carefully balanced aggregate size to achieve optimal workability without compromising the structural integrity of the mix. Coarse aggregate (CAg) contributes 15% to workability, further highlighting the importance of the aggregate properties in SCC design. Specifically, coarse aggregates affect the stability and flow characteristics of concrete. The proper selection of shape, size distribution, and content is crucial in ensuring the effective self-consolidation of the mix. Air content (A) and cement (C) both contribute 17% to workability variation, indicating their equally important roles in influencing SCC behavior. Air content affects the viscosity and stability of the mix, with excessive air producing reduced flowability and increased segregation risk. Cement content is a key factor in achieving the desired binder-to-aggregate ratio, which governs the cohesiveness and lubrication properties of the mix. Water content (W) contributes 12% to workability, reflecting its essential but less dominant role. Although water improves flowability, an excessive water content can lead to segregation and reduced durability, thus maintaining an optimal water-to-cement ratio is crucial for achieving balanced workability and strength. The findings of the sensitivity analysis are critical for optimizing SCC mix designs, particularly in applications where workability is paramount—such as heavily reinforced structures or complex formwork. By understanding the relative impact of each component, adjustments to fine aggregates, aggregate sizes, and binder content can be prioritized to achieve the desired workability without compromising other performance characteristics. Considering sustainable concrete construction, insights from this analysis will enable more efficient material usage through the identification of the critical parameters influencing workability. This will in turn contribute to the development of SCC mixes that require fewer adjustments during placement as well as reducing labor costs and minimizing resource waste. Moreover, the focus on aggregate properties aligns with sustainability goals by promoting the use of locally available or recycled aggregates, thereby further reducing the environmental footprint of construction projects. In conclusion, the sensitivity analysis underscores the complex interplay of mix components in determining SCC workability. The dominant influence of fine aggregates, along with significant contributions from aggregate size, air content, cement, and water, highlights the need for a balanced approach to mix design. The

findings herein provide a robust framework for creating SCC mixes that satisfy workability requirements while supporting sustainable construction practices. Generally, MNS—with a degree of importance of 16%—was found to have a greater influence on concrete workability than coarse aggregate density [1]. This implies that coarse aggregate sampling—which determines the nominal sizes in a concrete mix—is more important than aggregate density [25]. Tumidajski and Gong [25] reported that concrete mixes using well-graded aggregates with smaller nominal sizes between 10 and 18.5 mm produce the best concrete in terms of fresh and hardened properties. This report supports the role of the intelligent model herein—which is based on the varying nominal sizes of aggregates—as a decisive model for the design and production of SCC in sustainable concrete construction. The overall behavior of the multiple mixes with respect to workability characteristics and the nominal coarse aggregate size indicates that mixes with smaller aggregate sizes produce improved workability, especially for Ts, which decreases with reduced MNS values. This behavior agrees with the results of previous studies considering the influence of aggregate sampling on concrete workability [1,25,37,38]. The relationships between the calculated and predicted values are shown in Figure 7–9. The accuracy of each model was assessed by examining the relationship between the computed and predicted Hs values. ANN-MBNO yields a more dependable fit with an error margin of ±10%, as shown in Figure 8, whereas GP-EAO and EPR-GAO exhibit larger deviations. The efficiency of ANN-MBNO in capturing the intricate interactions between the input and output variables was thus demonstrated. Figure 8 shows that the Hs prediction produced a model best fit with relatively small outliers from the ±25%, ±20%, and ±10% best-fit envelopes for GP, EPR, and ANN, respectively. These best-fit envelopes also produced parametric lines of fit of 0.977×, 0.988×, and 0.996× for GP-EAO, EPR-GAO, and ANN-MBNO, respectively. These models further produced evaluation of performance metrics—R^2 of 0.634, 0.760, and 0.964, MAE of 18, 16, and 7.0 mm, and RMSE of 22, 20, and 8.0 mm for GP-EAO, EPR-GAO, and ANN-MBNO, respectively. Similarly, a com parison of the computed and projected Ds values are shown in Figure 9. The accuracy of ANN-MBNO is confirmed by its decreased error margin of ±6%, as opposed to ±15% and ±20% for GP-EAO and EPR-GAO, respectively. This outcome highlights how well ANN-MBNO models droop diameter variations induced by adjustments to the input parameters.

The fit envelopes were used to produce parametric lines of 0.981×, 0.989×, and 0.999×, respectively, and performance evaluation metrics—R^2 of 0.713, 0.633, and 0.980, MAE of 25, 27, and 7.0 mm, and RMSE of 30, 34, and 9.0 mm for GP-EAO, EPR-GAO, and ANN-MBNO, respectively. Finally, Figure 10 shows the accuracy of the Ts predictions. ANN-MBNO performs the best, with an error envelope of ±12%, while GP-EAO and EPR-GAO have error envelopes of ±30% and ±25%, respectively. This indicates that ANN-

MBNO can predict the SCC time-sensitive workability characteristics more reliably.

These fitting envelopes were used to further produce parametric lines of 0.976×, 0.983×, and 0.996×, respectively, and model performance metrics—R^2 of 0.679, 0.680, and 0.955, MAE of 0.29, 0.30, and 0.13 s, and RMSE of 0.40, 0.38, and 0.16 s overall for GP-EAO, EPR-GAO, and ANN-MBNO, respectively. Generally, ANN-MBNO is observed to outperformed GP-EAO and EPR-GAO in a wide range of performance metrics and, hence, became the decisive model in the prediction of the workability properties of the concrete mixes considered [20–49]. Note that the introduction of the metaheuristic training algorithms in the ANN, GP, and EPR prediction interfaces increased their robustness, handling efficiency, and speed as the models were executed within a 30 s runtime compared to hours and days of runtime without the introduction of the hybridized execution interface. Using Taylor diagrams, Figure 11 shows a comparison of the performances of the models and their outcomes, which are in agreement with the line of fit presented earlier. Finally, the results of all developed models are summarized in Table 7. Table 8 provides a comparison of the performances of the three models used in this study.

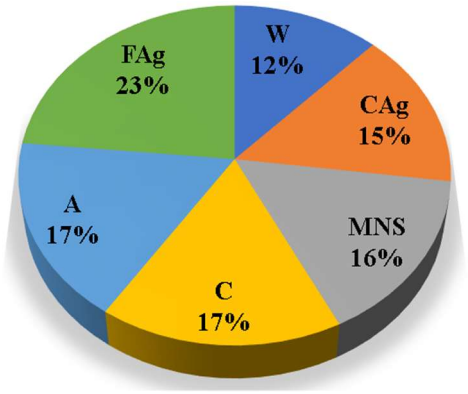


Figure 7. Individual influence of input parameters in the model relative to each other.

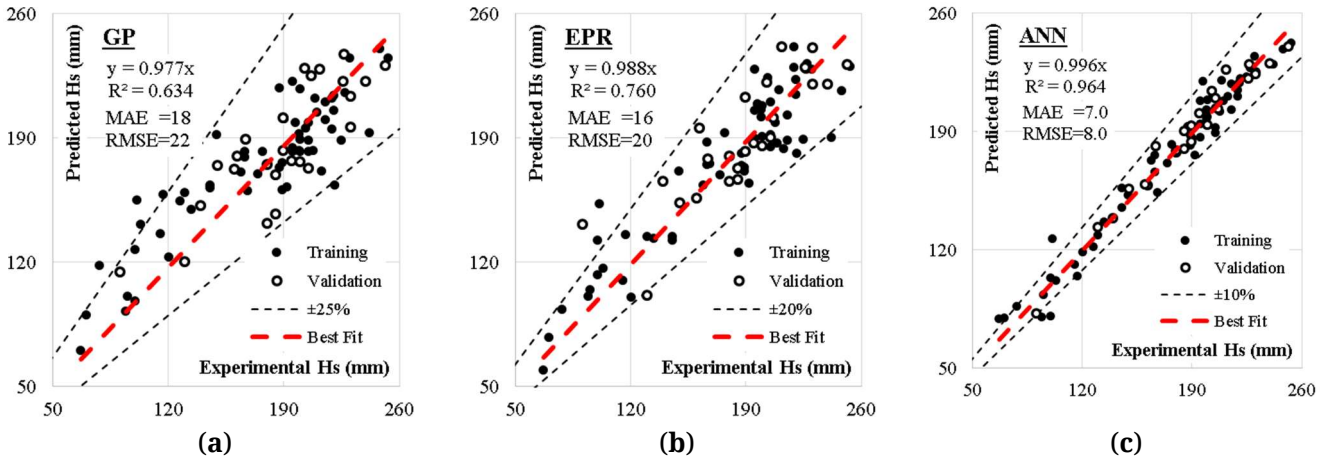


Figure 8. Relationship between the predicted and calculated Hs values based on the developed models: (**a**) GP-EAO model, (**b**) EPR-GAO model, and (**c**) ANN-MBNO model.

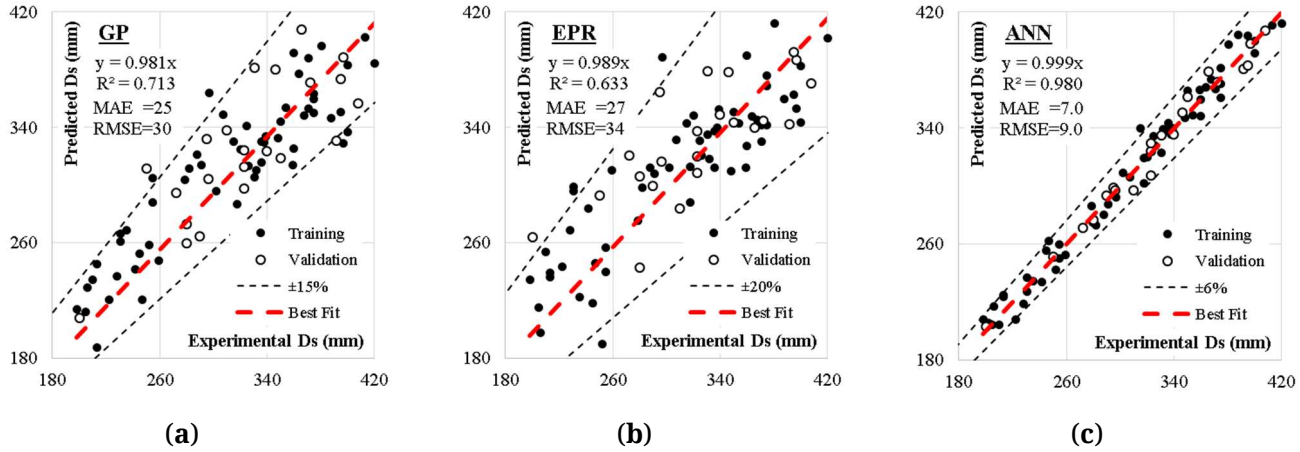


Figure 9. Relationship between predicted and calculated Ds values based on the developed models: (**a**) GP-EAO model, (**b**) EPR-GAO model, and (**c**) ANN-MBNO model.

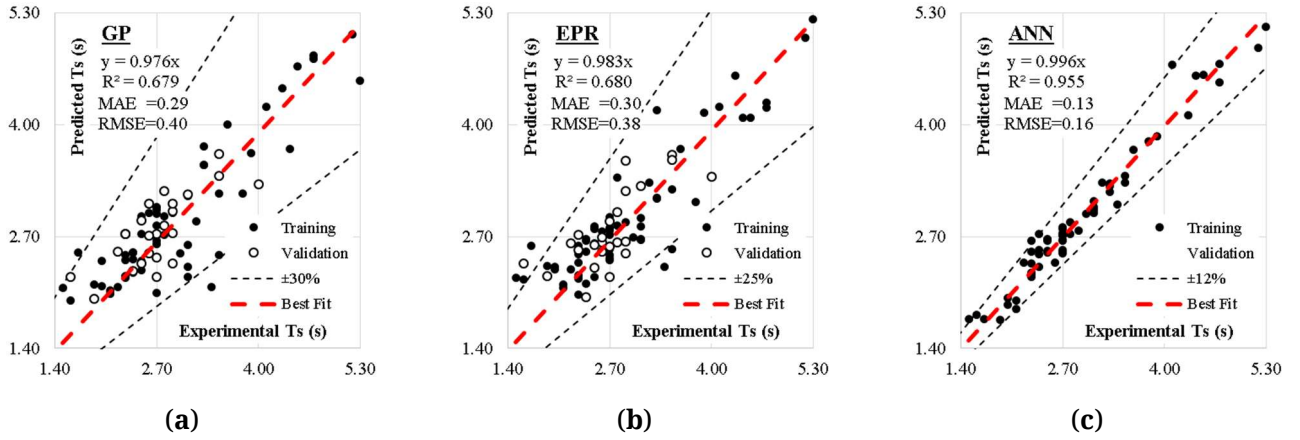
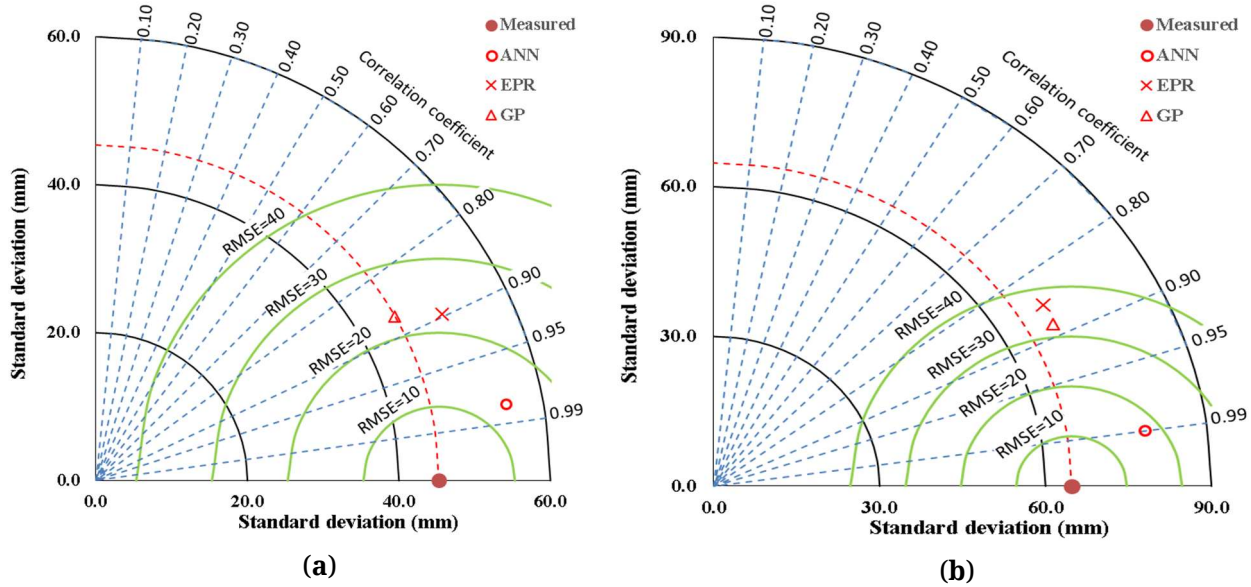
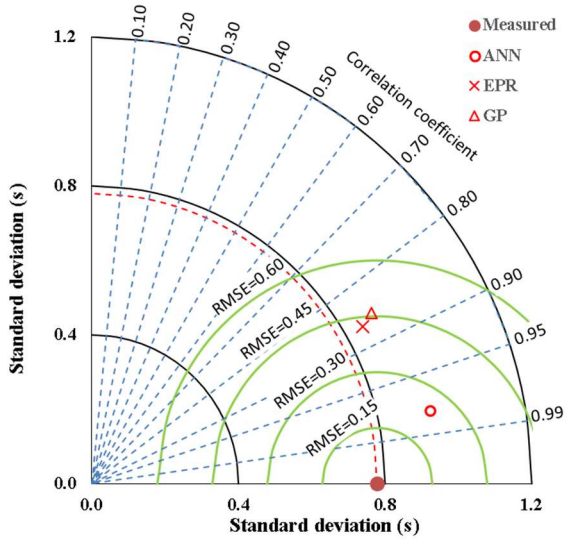


Figure 10. Relataionship between predicted and calculated Ts values based on the developed models: (**a**) GP-EAO model, (**b**) EPR-GAO model, and (**c**) ANN-MBNO model.





(c)

Figure 11. Comparison of the accuracies of the developed models using Taylor charts for (a) Hs, (b) Ds, and (c) Ts.

Table 7. Comparison of the accuracies of the developed models.

Item	Technique	Model	SSE	MAE	MSE	RMSE	Avg. Error (%)	R^2
Hs (mm)	GP-EAO	Equation (1)	42 777	18	503	22	12	0.634
	EPR-GAO	Figure 5, Table 6	34 093	16	401	20	11	0.760
	ANN-MBNO	Equation (4)	6 087	7	72	8	5	0.964
Ds (mm)	GP-EAO	Equation (2)	78 568	25	924	30	10	0.713
	EPR-GAO	Figure 5, Table 6	96 256	27	1132	34	11	0.633
	ANN-MBNO	Equation (5)	7 026	7	83	9	3	0.980
Ts (s)	GP-EAO	Equation (3)	13.84	0.29	0.16	0.40	14	0.679
	EPR-GAO	Figure 5, Table 6	12.66	0.30	0.15	0.38	13	0.680
	ANN-MBNO	Equation (6)	2.22	0.13	0.03	0.16	6	0.955

Table 8. Comparison of the GP-EAO, ANN-MBNO, and EPR-GAO models.

Feature	GP-EAO	ANN-MBNO	EPR-GAO
Model Basis	Symbolic regression using evolutionary operations (crossover, mutation).	Layered neural network inspired by biological neurons with adaptive weight optimization.	Polynomial regression optimized using GA for accuracy.
Optimization Technique	Evolutionary approach to refine symbolic expressions.	Mimicking biological learning mechanisms (e.g., Hebbian principles and adaptive learning rates).	GA for optimizing polynomial terms and coefficients.
Mathematical Complexity	Produces interpretable closed-form expressions.	High complexity; captures nonlinear relationships via weighted neuron connections.	Produces polynomial models with variable exponents and coefficients.
Performance Accuracy	Moderate; R^2 = 0.634–0.713	High; R^2 = 0.955–0.980	Moderate; R^2 = 0.633–0.760
Strengths	Provides interpretable formulas; useful for identifying broad trends.	Highly accurate; captures complex nonlinear relationships; robust to data variability.	Balances interpretability and accuracy; suitable for moderately complex systems.
Weaknesses	Limited accuracy for highly nonlinear systems.	Less interpretable; requires extensive tuning.	Struggles with extreme nonlinearity; limited flexibility compared to ANN-MBNO.
Runtime Efficiency	Fast but scales with symbolic complexity.	Slightly slower owing to backpropagation and weight adjustments.	Moderate; depends on polynomial complexity.
Application Suitability	Useful for exploratory analysis and identifying dominant variables.	Best for highly complex systems requiring high prediction accuracy.	Effective for systems with moderate nonlinearity and parameter interactions.

Overall, the ANN-MBNO model outperforms the GP-EAO and EPR-GAO models because of its robust optimization methods, ability to handle highly nonlinear relationships, ability to integrate features holistically, adaptability to data complexity, and efficient error minimization. Although the numerical findings quantify these benefits, the underlying mechanisms and design concepts offer a solid theoretical basis for the superior performance of ANN-MBNO in predicting the SCC workability. R^2 is highlighted because it provides an easy-to-understand metric for evaluating performance and concisely expresses the capacity of the model to account for data variability. Although less appropriate for assessing the overall fit and explanatory power of the model, RMSE, MAE, and MAPE supplement R^2 by providing additional information regarding the size and distribution of errors. This study strikes a balance between interpretability and precision by combining various measurements. However, R^2 is still the primary parameter used to compare model performance. The usefulness of the ANN-MBNO model in revolutionizing SCC design and production is demonstrated by the claimed accuracy. The quality and performance of SCC in a variety of building applications are ensured, while promoting material efficiency, cost savings, and sustainability through accurate predictions. The discussion highlights the broader significance of the findings and sets the stage for the conclusions. The final section will summarize the key outcomes and proposed directions for future research to advance SCC workability prediction.

The findings of this study align well with the recent advances in explainable and knowledge-guided ML for concrete materials, particularly those reported by Tan et al. (2026) [50] and Guo et al. (2025) [51]. Tan et al. (2026) [50] demonstrated that the integration of explainable AI techniques, particularly SHAP-based interpretation, provides valuable insights into the relative contribution of input variables toward predicting concrete performance. Although their work focused on the compressive strength of rubberized concrete, the underlying principle is directly applicable to the current SCC workability investigation. Similar to the SHAP interpretation adopted in their study, the sensitivity analysis conducted herein quantified the influence of SCC constituent materials on workability indices, revealing that fine aggregates (23%), air content and cement (17% each), maximum nominal aggregate size (16%), coarse aggregates (15%), and water content (12%) collectively govern the flow behavior of SCC. The prominence of the MNS identified in this study further validates the importance of explainable ML frameworks in uncovering physically meaningful relationships between material characteristics and concrete performance. Consequently, the sensitivity analysis serves not only as a model validation tool, but also as an engineering decision-support mechanism for SCC mix optimization. Furthermore, the work of Guo et al. (2025) [51] on multi-agent collaboration for the knowledge-guided data-driven design of ultra-high-performance concrete (UHPC) highlights the growing transition from purely predictive ML models toward intelligent

design frameworks capable of integrating domain knowledge, material science principles, and optimization algorithms. This concept is particularly relevant to the present research, where hybrid metaheuristic learning techniques (GP-EAO, EPR-GAO, and ANN-MBNO) were employed not only for prediction but also for identifying the relative significance of the SCC mixture parameters affecting workability. The superior performance of the ANN-MBNO model, combined with the sensitivity analysis results, demonstrates the potential of ML systems to support data-driven SCC mix design by identifying the critical parameters requiring optimization. Similar to the knowledge-guided design strategy proposed by Guo et al. (2025) [51], the present framework enables engineers to prioritize aggregate grading, nominal aggregate size selection, and fine aggregate proportions during SCC production, thereby reducing trial-and-error experimentation and improving resource efficiency. From a practical perspective, both studies reinforce the importance of interpretable ML in sustainable concrete technology. The sensitivity analysis presented in this work provides actionable engineering knowledge by demonstrating that workability can be significantly enhanced through careful control of aggregate characteristics, particularly the fine aggregate content and nominal aggregate size. This finding supports the observations of Tumidajski and Gong (2006) [25] regarding the superior fresh concrete performance associated with lower aggregate nominal sizes. Moreover, the ability of the ANN-MBNO model to accurately predict Hs, Ds, and Ts suggests that such intelligent frameworks can be integrated into future digital concrete design platforms. In this regard, the present study represents an important step toward the development of explainable, knowledge-guided, and sustainability-oriented ML systems for SCC design, consistent with the emerging research direction established by Tan et al. (2026) [50] and Guo et al. (2025) [51]. Xing et al. (2026) [52] demonstrated the effectiveness of interpretable ML techniques, particularly SHAP (SHapley Additive exPlanations), in revealing the internal decision-making behavior of predictive models used for estimating the splitting strength of asphalt concrete. Their study highlighted the importance of moving beyond black-box prediction accuracy to include feature-level interpretability, thereby improving trust, transparency, and engineering insight. This contribution is directly relevant to the present study because it supports the integration of advanced ML models with interpretability considerations in predicting the mechanical properties of construction materials. Similarly, the present research emphasizes not only predictive performance but also the underlying influences of input variables, which are essential for engineering decision-making and model validation. Similarly, Tan et al. (2026) [53] extended the application of explainable ML by integrating predictive modeling with life cycle assessment for the sustainable design of fiber-reinforced asphalt concrete. Their study provided a comprehensive framework that linked ML predictions with

environmental performance indicators, thereby enabling a more holistic evaluation of material sustainability. This contribution is particularly significant to the present study because it demonstrates how hybrid and data-driven approaches can be effectively used in civil engineering materials research to support both performance optimization and sustainability considerations. Furthermore, their emphasis on explainability strengthens the methodological foundation of the present study, wherein hybrid optimization models are developed not only for accuracy improvement but also for enhanced interpretability and engineering relevance.

CONCLUSIONS

This study developed and assessed three hybrid ML frameworks—GP-EAO, EPR-GAO, and ANN-MBNO—for predicting the workability characteristics of SCC incorporating varying MNS of coarse aggregate. Based on the results obtained, the following conclusions can be drawn.

1. The developed models successfully captured the complex nonlinear relationships between the SCC constituent materials and workability indices, including Hs, Ds, and Ts. The inclusion of maximum nominal aggregate size as an input parameter significantly enhanced the predictive capability of the models, confirming its importance in SCC workability assessment.
2. The sensitivity analysis revealed that fine aggregate content was the most influential parameter affecting SCC workability, contributing 23% to the overall variation, followed by cement content and air content, each contributing 17%, maximum nominal aggregate size contributing 16%, coarse aggregate content contributing 15%, and water content contributing 12%. These findings demonstrate that aggregate characteristics, particularly grading and nominal size, play a dominant role in controlling SCC flow behavior.
3. The results confirmed that SCC mixtures with a smaller nominal aggregate size generally exhibit superior workability characteristics. Aggregates within the lower nominal size range (10–18.5 mm) promoted improved flowability, increased Hs and Ds, and reduced Ts compared with mixtures with a larger aggregate size.
4. Among the three predictive models, ANN-MBNO consistently achieved the highest prediction accuracy. For Hs prediction, ANN-MBNO attained an R^2 value of 0.964, MAE of 7 mm, and RMSE of 8 mm compared with GP-EAO (R^2 = 0.634, MAE = 18 mm, RMSE = 22 mm) and EPR-GAO (R^2 = 0.760, MAE = 16 mm, RMSE = 20 mm). For Ds prediction, ANN-MBNO achieved an R^2 of 0.980, MAE of 7 mm, and RMSE of 9 mm, outperforming GP-EAO (R^2 = 0.713, MAE = 25 mm, RMSE = 30 mm) and EPR-GAO (R^2 = 0.633, MAE = 27 mm, RMSE = 34 mm). Similarly, for Ts prediction, ANN-MBNO produced an R^2 of 0.955, MAE of 0.13 s, and RMSE of 0.16 s, significantly outperforming GP-EAO (R^2 = 0.679, MAE =

0.29 s, RMSE = 0.40 s) and EPR-GAO (R^2 = 0.680, MAE = 0.30 s, RMSE = 0.38 s).

5. The ANN-MBNO model further demonstrated superior fitting characteristics, producing best-fit parametric slopes of 0.996, 0.999, and 0.996 for Hs, Ds, and Ts, respectively, with the narrowest prediction error envelopes of ±10%, ±6%, and ±12%. These results confirmed the robustness and reliability of this method for practical SCC workability prediction.
6. The findings highlighted the importance of aggregate size selection in SCC mix design and demonstrated the capability of hybrid metaheuristic ML approaches in reducing experimental effort, accelerating mixture optimization, and supporting sustainable concrete production through efficient material utilization.
7. Although the developed models exhibited excellent predictive performance, their applicability remained dependent on the quality and range of the available database. Future studies should expand the dataset to include wider SCC compositional ranges, supplementary cementitious materials, recycled aggregates, and diverse environmental conditions. The integration of explainable AI techniques such as SHAP analysis and knowledge-guided learning frameworks as well as the development of user-friendly design interfaces would further enhance model transparency and facilitate the industrial implementation of intelligent SCC mix design systems.

Practical Application of Research

The practical application of this research lies in providing engineers, researchers, and construction practitioners with an intelligent predictive framework for evaluating the workability behavior of SCC mixtures containing different coarse aggregate sizes without relying solely on extensive laboratory experimentation. The developed hybrid ML models, particularly the ANN-MBNO framework, can support the rapid prediction of Hs, Ds, and flow time during SCC mix design and quality control processes. This enables more efficient optimization of aggregate selection, material proportioning, and rheological performance prior to field implementation.

The proposed predictive models may also assist concrete producers and construction engineers in minimizing the trial-and-error procedures commonly associated with SCC production, thereby reducing material wastage, experimental costs, and construction delays. By accurately estimating the workability characteristics under varying coarse aggregate conditions, the framework can contribute to improved constructability in projects involving congested reinforcement, complex formworks, and difficult placement conditions, where SCC performance is critical.

Furthermore, the integration of biologically inspired optimization techniques with ML provides a practical pathway for implementing data-driven decision support systems in modern concrete technology. The

developed models can potentially be incorporated into automated batching systems, intelligent construction monitoring platforms, and mix proportioning software for preliminary SCC performance assessment.

From a sustainability perspective, the framework supports the more efficient utilization of construction materials by facilitating optimized aggregate selection and improved workability management, which may contribute to reduced material waste and enhanced construction efficiency. Nevertheless, further validation using larger and more diverse industrial datasets is recommended prior to large-scale field deployment of the developed predictive models.

Limitations and Future Work

Despite the encouraging results of the proposed soft computing models, a few drawbacks should be noted. The models' ability to generalize to real-world situations may be limited by their training requirements, which may require large datasets that are not always easily accessible. Additional verification using experimental data from various settings may improve their dependability. Future research should concentrate on enhancing model resilience, adding increasingly intricate datasets, and investigating hybrid strategies that combine soft computing and traditional engineering methodologies. Furthermore, there is a need to investigate how well these models adapt to various concrete types and situations.

DATA AVAILABILITY

The data used in this study will be made available upon reasonable request to the corresponding author.

AUTHOR CONTRIBUTIONS

Conceptualization, NU and KV; methodology, NU, KV and PM; software, KV and MA; validation, NU, KV, MA and FR; formal analysis, NU and KV; investigation, NU, PM and MA; resources, PM and MA; data curation, NU and FR; writing—original draft preparation, NU; writing—review and editing, KV, PM, MA and FR; visualization, NU and KV; supervision, KV; project administration, NU. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

1. Mahmoodzadeh F, Chidiac SE. Rheological models for predicting plastic viscosity and yield stress of fresh concrete. Cem Concr Res. 2013;49:1-9. doi: 10.1016/j.cemconres.2013.03.004
2. EFNARC. Specification and guidelines for self-compacting concrete. European Federation of National Associations Representing Producers and Applicators of Specialist Building Products for Concrete. 2002. Available from: https://wwwp.feb.unesp.br/pbastos/c.especiais/Efnarc.pdf. Accessed on 2026 Feb 20.
3. Ahmed S, Yehia S. Evaluation of workability and structuration rate of locally developed 3D printing concrete using conventional methods. Materials. 2022;15(3):1243. doi: 10.3390/ma15031243
4. Alghamdi SJ. Determining the mix design method for normal strength concrete using machine learning. J Umm Al-Qura Univ Eng Archit. 2023;14:95-104. doi: 10.1007/s43995-023-00022-4
5. Anand-Kumar BG. Assessment of concrete workability using conical funnel. In: Nandagiri L, Narasimhan MC, Marathe S, editors. Recent advances in civil engineering. CTCS 2021. Lecture notes in civil engineering, vol. 256. Singapore (Singapore): Springer; 2023. p. 23-31. doi: 10.1007/978-981-19-1862-9_3
6. Chamroeun C, Young KK, Seong JH, Seung WL. Evaluation on compactibility and workability of roller-compacted concrete for pavement. Int J Pavement Eng. 2019;20(8):905-10. doi: 10.1080/10298436.2017.1366762
7. Chen G, Suhail SA, Bahrami A, Sufian M, Azab M. Machine learning-based evaluation of parameters of high-strength concrete and raw material interaction at elevated temperatures. Front Mater. 2023;10:1187094. doi: 10.3389/fmats.2023.1187094
8. Dinesh A, Anithaselvasofia SD, Datcheen KS, Rakheshvarshan D. Machine learning for strength evaluation of concrete structures: a critical review. Mater Today Proc. 2023. doi: 10.1016/j.matpr.2023.04.090
9. Gamil Y. Machine learning in concrete technology: a review of current researches, trends, and applications. Front Built Environ. 2023;9:1145591. doi: 10.3389/fbuil.2023.1145591
10. Han ID, Han D. Introducing a new quantitative evaluation method for segregation of normal concrete. Int J Concr Struct Mater. 2021;15:21. doi: 10.1186/s40069-021-00458-9
11. Kaveh A, Khavaninzadeh N. Efficient training of two ANNs using four meta-heuristic algorithms for predicting the FRP strength. Structures. 2023;52:256-72. doi: 10.1016/j.istruc.2023.03.178
12. Kazemi R. Artificial intelligence techniques in advanced concrete technology: a comprehensive survey on 10 years research trend. Eng Rep. 2023;e12676. doi: 10.1002/eng2.12676
13. Mohamed YM, Fathelrahman MA. Evaluation of the concrete works produced locally in River Nile State. J Civ Eng Res. 2015;5(3):67-73. doi: 10.5923/j.jce.20150503.03

14. Nhat-Duc H, Anh-Duc P. Estimating concrete workability based on slump test with least squares support vector regression. J Const Eng. 2016;2016:5089683. doi: 10.1155/2016/5089683
15. Neeraja D, Arshad SM, Nadaf AKN, Reddy MK. Evaluation of workability and strength of green concrete using waste steel scrap. IOP Conf Ser Mater Sci Eng. 2017;263:032013. doi: 10.1088/1757-899X/263/3/032013
16. Onyelowe KC, Ebid AM, Riofrio A, Baykara H, Soleymani A, Mahdi A, et al. Multi-objective prediction of the mechanical properties and environmental impact appraisals of self-healing concrete for sustainable structures. Sustainability. 2022;14(15):9573. doi: 10.3390/su14159573
17. Onyelowe KC, Kontoni DP, Ebid AM. Flow simulation of self-consolidating concrete through V-funnel for sustainable buildings. IOP Conf Ser Earth Environ Sci. 2022;1123:012044. doi: 10.1088/1755-1315/1123/1/012044
18. Onyelowe KC, Kontoni DP, Ebid AM. Simulation of self-compacting concrete (SCC) passing ability using the L-box model for sustainable buildings. IOP Conf Ser Earth Environ Sci. 2022;1123:012065. doi: 10.1088/1755-1315/1123/1/012065
19. Qian Y, Sufian M, Accouche O, Azab M. Advanced machine learning algorithms to evaluate the effects of the raw ingredients on flowability and compressive strength of ultra-high-performance concrete. PLoS One. 2022;17(12):e0278161. doi: 10.1371/journal.pone.0278161
20. Rathakrishnan V, Beddu SB, Ahmed AN. Predicting compressive strength of high-performance concrete with high volume ground granulated blast-furnace slag replacement using boosting machine learning algorithms. Sci Rep. 2022;12:9539. doi: 10.1038/s41598-022-12890-2
21. SCC EPG. The European guidelines for self-compacting concrete—specification, production and use. Self-Compacting Concrete European Project Group (SCC 028). Available from: https://www.theconcreteinitiative.eu/images/ECP_Documents/EuropeanGuidelinesSelfCompactingConcrete.pdf. Accessed on 2026 Jan 12.
22. Sharma N, Upadhya A, Thakur MS, Sihag P. Comparison of machine learning algorithms to evaluate strength of concrete with marble powder. Adv Mater Res. 2022;11(1):75-90. doi: 10.12989/amr.2022.11.1.075
23. Chou JS, Pham AD. Enhanced artificial intelligence for ensemble approach to predicting high-performance concrete compressive strength. Constr Build Mater. 2013;49:554–63. doi:10.1016/j.conbuildmat.2013.08.078.
24. Wang LJ, Gu X, Hu M. Workability evaluation of self compacting concrete. In: Applied mechanics and materials. Vol. 44-47. Stafa (Switzerland): Trans Tech Publications; 2010. p. 2359-63. doi: 10.4028/www.scientific.net/AMM.44-47.2359
25. Tumidajski PJ, Gong B. Effect of coarse aggregate size on strength and workability of concrete. Can J Civ Eng. 2006;33(2):206-13. doi: 10.1139/l05-090
26. Sonebi M, Grünewald C, Cevik A, Walraven J. Neural network technique: modelling fresh properties of self-compacting concrete. Comput Concr. 2016;18(6):903-20. doi: 10.12989/cac.2016.18.6.903

27. Sonebi M, Grünewald C, Cevik W. Neural network technique: modelling fresh properties of self-compacting concrete. Comput Concr. 2016;18(6):903-20. doi: 10.12989/cac.2016.18.6.903
28. Siddique R, Aggarwal P, Aggarwal Y, Gupta SM. Modeling properties of self-compacting concrete: support vector machines approach. Comput Concr. 2008;5(5):461-73. doi: 10.12989/cac.2008.5.5.461
29. Yaman MA, Elaty MA, Taman M. Predicting the ingredients of self-compacting concrete using artificial neural network. Alex Eng J. 2017;16(4):523-32. doi: 10.1016/j.aej.2017.04.007
30. Ismael Jaf DK. Soft computing and machine learning-based models to predict the slump and compressive strength of self-compacted concrete modified with fly ash. Sustainability. 2023;15(15):11554. doi: 10.3390/su151511554
31. Mai HVT, Nguyen MH, Trinh SH, Ly HB. Optimization of machine learning models for predicting the compressive strength of fiber-reinforced self-compacting concrete. Front Struct Civ Eng. 2023;17:284-305. doi: 10.1007/s11709-022-0901-6
32. Ghafor K. Multifunctional models, including an artificial neural network, to predict the compressive strength of self-compacting concrete. Appl Sci. 2022;12(16):8161. doi: 10.3390/app12168161
33. Asri YE, Aicha MB, Zaher M, Alaoui AH. Modelization of the rheological behavior of self-compacting concrete using artificial neural networks. Mater Today Proc. 2022;58(Pt 4):1114-21. doi: 10.1016/j.matpr.2022.01.257
34. Onyelowe KC, Kontoni DP. The net-zero and sustainability potential of SCC development, production and flowability in structures design. Int J Low Carbon Technol. 2023;18:530-41. doi: 10.1093/ijlct/ctad033
35. Onyelowe KC, Kontoni DP. A critical review of rheological models in self-compacting concrete for sustainable structures. Sci Rep. 2023;13:21296. doi: 10.1038/s41598-023-48673-6
36. Cavaleri L, Barkhordari MS, Repapis CC, Armaghani DJ, Ulrikh DV, Asteris PG. Convolution-based ensemble learning algorithms to estimate the bond strength of the corroded reinforced concrete. Constr Build Mater. 2022;359:129504. doi: 10.1016/j.conbuildmat.2022.129504
37. Nguyen NH, Vo TP, Lee S, Asteris PG. Heuristic algorithm-based semi-empirical formulas for estimating the compressive strength of the normal and high performance concrete. Constr Build Mater. 2021;304:124467. doi: 10.1016/j.conbuildmat.2021.124467
38. Cavaleri L, Chatzarakis GE, Di Trapani F, Douvika MG, Roinos K, Vaxevanidis NM, et al. Modeling of surface roughness in electro-discharge machining using artificial neural networks. Adv Mater Res. 2017;6(2):169-84. doi: 10.12989/amr.2017.6.2.169
39. Ashrafian A, Panahi E, Salehi S, Karoglou M, Asteris PG. Mapping the strength of agro-ecological lightweight concrete containing oil palm by-product using artificial intelligence techniques. Structures. 2023;48:1209-29. doi: 10.1016/j.istruc.2022.12.108

40. Alkayem NF, Shen L, Mayya A, Asteris PG, Fu R, Di Luzio G, et al. Prediction of concrete and FRC properties at high temperature using machine and deep learning: a review of recent advances and future perspectives. J Build Eng. 2024;83:108369. doi: 10.1016/j.jobe.2023.108369
41. Koza JR. Genetic programming as a means for programming computers by natural selection. Stat Comput. 1994;4(2):87-112. doi: 10.1007/BF00175355
42. Goodfellow I, Bengio Y, Courville A. Deep learning. Cambridge (MA, US): MIT Press; 2016.
43. Poli R, Langdon WB, McPhee NF. A field guide to genetic programming. Morrisville (NC, US): Lulu Enterprises; 2008.
44. Chai T, Draxler RR. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. Geosci Model Dev. 2014;7:1247-50. doi: 10.5194/gmd-7-1247-2014
45. Asteris PG, Koopialipoor M, Armaghani DJ, Kotsonis EA, Lourenço PB. Prediction of cement-based mortars compressive strength using machine learning techniques. Neural Comput Appl. 2021;33(19):13089-121. doi: 10.1007/s00521-021-06004-8
46. Asteris G, Karoglou M, Skentou AD, Vasconcelos G, He M, Bakolas A, et al. Predicting uniaxial compressive strength of rocks using ANN models: incorporating porosity, compressional wave velocity, and Schmidt hammer data. Ultrasonics. 2024;141:107347. doi: 10.1016/j.ultras.2024.107347
47. Asteris PG, Skentou AD, Bardhan A, Samui P, Pilakoutas K. Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models. Cem Concr Res. 2021;145:106449. doi: 10.1016/j.cemconres.2021.106449
48. Benzaamia A, Ghrici M, Rebouh R, Zygouris N, Asteris PG. Predicting the shear strength of rectangular RC beams strengthened with externally-bonded FRP composites using constrained monotonic neural networks. Eng Struct. 2024;313:118192. doi: 10.1016/j.engstruct.2024.118192
49. Ahmad J, Zhou Z, Deifalla AF. Steel fiber reinforced self-compacting concrete: a comprehensive review. Int J Concr Struct Mater. 2023;17(1):51. doi: 10.1186/s40069-023-00602-7
50. Tan X, Xing J, Wang Y, Qiu H, Mahjoubi S, Guo P. Explainable machine learning for predicting compressive strength of rubberized concrete: SHAP interpretation, lifecycle assessment, and design recommendations. J Clean Prod. 2026;538:147338. doi: 10.1016/j.jclepro.2025.147338
51. Guo P, Jiang Z, Meng W, Bao Y. Multi-agent collaboration for knowledge-guided data-driven design of ultra-high-performance concrete (UHPC) incorporating solid wastes. Cem Concr Compos. 2025;164:106230. doi: 10.1016/j.cemconcomp.2025.106230
52. Xing J, Tan X, Li Y, Jin D, Guo P, Wang Y, et al. Interpretable machine learning for predicting splitting strength of asphalt concrete: insights from SHAP analysis. Materials. 2026;19(8):1636. doi: 10.3390/ma19081636

53. Tan X, Xing J, Mahjoubi S, Guo P, Wei Z, Wang Y, et al. Explainable machine learning and life cycle assessment for sustainable design of fiber-reinforced asphalt concrete. J Clean Prod. 2026;547:147759. doi: 10.1016/j.jclepro.2026.147759

How to cite this article:
Ulloa N, Vaca K, Mancheno P, Albuja M, Romero F. Data-Driven Framework for The Flow Characteristics of Self-Compacting Concrete Using Machine Learning. J Sustain Res. 2026;8(3):e260061. https://doi.org/10.20900/jsr20260061.