scientific reports

Check for updates

OPEN

ANN trained by BBO for modeling of fly ash cementitious systems with high range water reducing admixtures

Naz Mardani[1], Ramin Kazemi[2], Murteda Unverdi[3], Ali Mardani[3] & Seyedali Mirjalili[4,5,6]✉

This study aims to develop artificial intelligence (AI) models for predicting the compressive strength and flow value of cementitious systems containing fly ash, influenced by various high-range water-reducing admixtures (HRWRAs) that differ in molecular weight and chain length. A database comprising 180 mixes was created, encompassing cement and fly ash dosages, HRWRA characteristics (including molecular weight, main and side chain lengths) curing period, and flow time. Two AI-based modelling approaches were employed: a classical artificial neural network (ANN) and a new hybrid model that integrates ANN with biogeography-based optimisation (ANN–BBO). The modeling results showed that the hybrid model achieved a compressive strength performance with an R^2 of approximately 0.99 and an RMSE of around 1.37 MPa, while the single ANN model attained an R^2 of about 0.91 and an RMSE of 4.40 MPa. For flow value prediction, the ANN–BBO model also demonstrated high accuracy ($R^2 \approx 0.98$; RMSE ≈ 0.32 cm). Furthermore, the ANN–BBO model reduced the prediction error by approximately 60% across the evaluation criteria compared to the single ANN model, highlighting its enhanced performance. The importance of the input variables indicated that curing time and cement content have the greatest impact on compressive strength, while flow time and the molecular weight of the HRWRA significantly influence the flow value. Since AI models rely solely on virtual trials, they significantly reduce laboratory time and material usage while aiding in the design of mixes with lower water-to-binder ratios and higher fly ash content, which ultimately helps to reduce the CO_2 footprint. The proposed models provide a practical route to low-clinker, FA-rich mix designs that satisfy strength/workability targets with less cement, supporting embodied-carbon reductions and straightforward integration into ready-mix/precast quality-control workflows.

Keywords Cementitious systems, Fly ash, High-range water-reducing admixture, Artificial intelligence, Artificial neural network, Biogeography-based optimisation

Abbreviations

WRA	Water-reducing admixture
HWRA	High-range water-reducing admixture
CS	Compressive strength
AI	Artificial intelligence
ANN	Artificial neural network
LR	Linear regression
NLR	Nonlinear regression
RFR	Random forest regression
ETR	Extra trees regression
GBR	Gradient boosting regression
XGB	EXtreme gradient boosting
XGBR	EXtreme gradient boosting regression

[1]Mathematics Education Department, Bursa Uludag University, Bursa, Turkey. [2]Independent Researcher, Sabzevar, Iran. [3]Civil Engineering Department, Bursa Uludag University, Bursa, Turkey. [4]Centre for Artificial Intelligence Research and Optimisation, Torrens University Australia, Brisbane, Australia. [5]Faculty of Electrical Engineering and Computer Science, VSB – Technical University of Ostrava, Ostrava, Czech Republic. [6]University Research and Innovation Center, Obuda University, Budapest, Hungary. ✉email: ali.mirjalili@torrens.edu.au

LSSVM	Least square support vector machine
CSA	Coupled simulated annealing
MPMR	Minimax probability machine regression
RVM	Relevance vector machine
ENN	Emotional neural network
GP	Genetic programming
MLR	Multi-linear regression
PSO	Particle swarm optimization
ANFIS	Adaptive neuro-fuzzy inference system
GA	Genetic algorithm
DT	Decision tree
SVM	Support vector machines
MLP	Multi-layer perceptron
BP	Backpropagation
FQ	Full-quadratic
IN	Interaction
KNN	K-nearest neighbors
BBO	Biogeography-based optimization
CNN	Convolutional neural networks
ELM	Extreme learning machine
GEP	Gene expression programming
M5p	M5prime
MAE	Mean absolute error
NSE	Nash–Sutcliffe efficiency
RF	Random forest
RMSE	Root mean squared error
SIVs	Suitability index variables
SVR	Support vector machines
VAF	Variance accounted for

Water-reducing admixtures (WRAs) have become indispensable in contemporary concrete technology, as they enhance workability and durability while reducing the water content in cementitious mixtures. These admixtures contribute to improved concrete performance by enabling lower water-to-cement (w/c) ratios without adversely affecting workability. The optimization of w/c ratios is essential for attaining the desired properties of concrete, given its direct influence on the strength and durability of the final product. In recent developments, the design of advanced WRAs has increasingly focused on modifying their chemical architecture to enhance performance across diverse cementitious systems.

WRAs have undergone substantial advancements, with scholarly investigations demonstrating that the precise chemical formulation and architecture of these additives are critical in augmenting the characteristics of cementitious systems[1]. This persistent innovation seeks to optimize the functionality of WRAs, particularly by resolving compatibility challenges and enhancing their impact on both the properties of fresh and hardened concrete. The ongoing enhancement of these additives is imperative for satisfying the escalating requirements of contemporary construction methodologies[2].

Cementitious systems play a pivotal role in construction, delivering the structural strength and stability (flow retention over time) required for various applications. The selection of constituent materials and their proportions significantly impacts both the mechanical properties and durability of these systems[3]. Beyond the mechanical aspects of concrete, carbon emissions have emerged as a major environmental concern[4,5]. In particular, in alignment with the European Green Deal, the construction sector is placing greater emphasis on sustainable approaches, such as incorporating supplementary materials like fly ash to improve concrete performance while simultaneously reducing its ecological footprint[6,7].

Fly ash plays a crucial role in enhancing the sustainability of cementitious materials by lowering carbon emissions and improving mechanical performance, thereby promoting environmentally friendly construction practices[8]. This strategy aligns with the increasing focus on sustainable building materials that minimize the environmental impact of concrete production, as demonstrated by the notable advantages linked to the use of fly ash[9–11]. In concretes incorporating fly ash, considerations regarding workability and strength are of critical importance, since an appropriate proportion of fly ash can enhance performance without compromising durability[12]. Moreover, the integration of advanced WRAs with fly ash has the potential to produce optimized concrete mixtures that satisfy both structural performance and environmental sustainability criteria[13]. Current research continues to explore how various WRA chemical structures influence their interaction with fly ash, with the goal of further improving both the mechanical performance and environmental sustainability of concrete[14].

The application of high-range water-reducing admixtures (HRWRAs) can markedly improve the mechanical performance of concrete, especially when used in conjunction with supplementary cementitious materials such as fly ash. This combination not only enhances workability but also significantly contributes to the long-term durability of concrete structures, increasing their resilience to environmental stressors[15]. The synergistic effect (combined effect greater than the sum of individual effects; positive interaction) of HRWRAs and fly ash not only optimizes concrete performance but also plays a vital role in decreasing the carbon emissions linked to cement production[16,17]. This reduction is particularly important considering the substantial share of the construction industry in global greenhouse gas emissions[18–21].

A laboratory-based approach offers valuable insights into the compressive strength (CS) of concrete and the interaction mechanisms between HRWRAs and fly ash within cementitious matrices, thereby supporting the development of more environmentally friendly concrete alternatives[22]. However, the success of these formulations is highly dependent on the specific chemical structures of the HRWRAs employed, which can exhibit varying degrees of compatibility with fly ash. Consequently, laboratory experimentation may encounter several drawbacks, including time constraints, consumption of raw materials, variability in material properties, limited formulation scope, and the need for extensive data to comprehensively assess performance outcomes[23].

To overcome these limitations, machine learning techniques are increasingly utilized to optimize the formulation of cementitious systems. These approaches enhance predictive capabilities and material efficiency, thereby contributing to more sustainable construction practices[24]. The integration of advanced artificial intelligence (AI) technologies into optimization processes allows for more accurate performance predictions and more efficient use of materials in concrete production[25]. These technologies not only improve prediction accuracy but also promote sustainability by minimizing waste and resource consumption in construction. Within these frameworks, algorithms evaluate the influence of various input parameters on construction outcomes, thus enhancing decision-making and operational efficiency. Through the application of machine learning, construction professionals can reduce project timelines and costs while maintaining specification-compliant performance in materials and design[26].

Machine learning approaches, particularly artificial neural networks (ANN), have demonstrated substantial potential in forecasting the performance of cementitious systems by capturing complex interactions between input parameters and resulting properties. This predictive capability facilitates more informed decisions in concrete mixture design, ultimately contributing to improved sustainability and efficiency in construction practices[27,28]. Artificial intelligence technologies enhance design efficiency and broaden the scope for material exploration, surpassing the limitations of conventional methods. This advancement supports sustainable construction by optimizing resource use and minimizing material waste. The evolution of AI tools continues to boost productivity, reduce waste generation, and enable the development of tailored concrete formulations that meet specific project requirements[29].

Recent studies underscore a growing focus on modeling cementitious systems, especially those incorporating fly ash. These studies highlight the vital importance of advanced modeling techniques in precisely assessing the mechanical behavior of such systems. Table 1 presents a summary of key AI-driven modeling research relevant to cementitious materials.

Previous research has confirmed the effectiveness of various AI methods, particularly ANN, in predicting the CS of concrete mixtures containing fly ash and other supplementary materials, thereby demonstrating their potential to enhance performance and promote sustainability in construction practices[24,34]. Moreover, these techniques continue to advance, offering valuable insights into the optimization of cementitious systems to achieve improved environmental outcomes, ultimately contributing to the reduction of greenhouse gas emissions associated with concrete production and supporting sustainable construction initiatives[44,45]. The capacity of ANNs to identify and learn from complex data patterns enables more precise predictions of concrete behavior, thus fostering the development of innovative and environmentally friendly solutions within civil engineering applications[38,46]. In addition, ANNs facilitate the exploration of new concrete formulations aligned with sustainability objectives, thereby improving both the efficiency and ecological performance of construction projects[47,48].

The limitations of current AI methodologies are often linked to their dependence on large datasets and the inherent difficulties in obtaining high-quality, representative data for training. Moreover, insufficient data availability can impair the predictive capabilities of machine learning models in evaluating concrete

References	Product or waste material used in mixtures	Dataset	AI method	The best model for CS (R^2 value)
[30]	HRWR	54	ANN-1, ANN-2	*ANN-1 (0.9431)
[31]	High volume fly ash	450	LR, NLR, ANN, M5P-tree,	*NLR (0.9969)
[32]	Phase change materials	154	RFR, ETR, GBR, XGBR,	*GBR (09,884)
[33]	Blast furnace slag, superplasticizer	1030	LSSVM-CSA, GP	*LSSVM-CSA (0.9767)
[34]	Fly ash	112	MPMR, RVM, ENN, GP, ELM	*MPMR (0.9920)
[35]	Fly ash	196	MLR, PSO-ANFIS, ANN, GA-ANFIS	*GA-ANFIS (0.9726)
[26]	Fly ash, blast furnace slag	1030	Bagging, GEP, AdaBoost, DT	*Bagging (0.9200)
[36]	Metakaolin	276	ANN, ANFIS	*ANFIS (0.9909)
[37]	Calcined sludge	2160	LR, SVM, RF, Ensemble, CNN, MLP-ANN	*CNN (0.9949)
[38]	Rice husk ash, superplasticizer, fly ash	93	GEP, ANN	*ANN (0.9695)
[39]	Fly ash, superplasticizer	100	GEP, ANN, DT, Bagging regressor	*Bagging regressor (0.9849)
[40]	RCA, fly ash, silica fume	380	SVM, ANN, AdaBoost, CNN	*CNN (0.9670)
[41]	Fly ash, superplasticizer	190	BP-ANN, RF, XGBoost, BP-GA, BP-PSO	*BP-GA (0.9731)
[42]	Fly ash, superplasticizer	130	FQ, IN, M5P-tree	*FQ (0.9849)
[43]	Fly ash	315	ANN, KNN, SVR, XGB	*ANN (0.9758)

Table 1. AI-based modeling studies selected from the literature in civil engineering. *Represents the best model for CS.

performance. Many AI techniques, particularly deep learning neural networks, face significant challenges when confronted with limited data, emphasizing the need to explore alternative strategies that maintain predictive reliability. Ongoing research seeks to enhance the robustness of AI models by incorporating transfer learning and data augmentation techniques, which can improve prediction accuracy and model generalization under data-constrained conditions[11]. Effective AI integration in engineering requires expertise across both AI and civil engineering domains, along with a collaborative approach to overcome the complex challenges faced in the construction industry. Highlighting the importance of interdisciplinary knowledge is essential for the successful application of AI technologies in sustainable construction practices[24,49,50].

Researchers are actively exploring innovative techniques, such as synthetic data generation and transfer learning, to enhance data acquisition and model training in light of these existing limitations. Additionally, the use of binary AI models improves data interpretability for engineers, thereby strengthening decision-making and model resilience in complex real-world conditions and expanding the scope of AI applications in civil engineering projects[19,41].

In the modeling of cementitious systems, most proposed models function in isolation, and hybrid approaches integrating meta-heuristic optimization techniques have yet to be comprehensively explored. Furthermore, a review of current modeling studies reveals a noticeable gap in research focusing on the flow value (FV) of cement mortars, a critical parameter for evaluating the workability and overall performance of cementitious mixtures. Considering these identified gaps, there is a clear need for comprehensive optimization strategies and innovative modeling approaches that can effectively address the FV of cement mortars and improve the predictive performance of ANN techniques within cementitious systems. The integration of ANN with biogeography-based optimization (BBO) presents a novel methodology aimed at enhancing predictive accuracy by optimizing neural network parameters. This combined approach not only increases the precision of performance predictions but also contributes to the formulation of more efficient and sustainable concrete mixtures for civil engineering applications[29]. The current absence of the ANN–BBO model in studies concerning cementitious systems will be addressed in this research, thereby improving the analysis of FV and enriching the literature on flow value modeling. Figure 1. depicts the conceptual framework of the current study.

In this study, the use of ANN, along with its integration with the BBO algorithm, is expected to yield notable advancements in predicting both the CS and FV of cementitious systems, ultimately supporting sustainable construction efforts. The data utilized in this research were obtained from three studies examining the influence of side and main chain lengths of HRWRAs in fly ash-based cementitious systems and were used as input for the modeling process. This study aims to assess the predictive performance and engineering relevance of the proposed models. The results are anticipated to expand existing knowledge and offer practical insights into improving cement mortar properties. Ultimately, the research will highlight the benefits of integrating artificial intelligence with optimization techniques for advancing effective applications in civil engineering. An overview of the research framework proposed in this study is illustrated in Fig. 2.

Conventional prediction methods in cementitious systems -w/c-based rules, semi-empirical models, and low-order regressions- struggle to represent nonlinear, time-dependent, and structure–property effects arising from HRWRA molecular architecture and FA replacement. Our AI approach (ANN and ANN–BBO) learns these interactions directly from data, improving accuracy and enabling fast "what-if" optimization of strength–flow–cement content, which traditional tools cannot reliably deliver.

Experimental dataset, variables description and mix design

This study explores the predictive modeling of CS and FV in fly ash-blended cementitious mortars incorporating HRWRAs, utilizing a dataset comprising 180 distinct mortar mixtures. The dataset was derived from a series of comprehensive and systematically executed experimental investigations[13,14,51], which examined the molecular characteristics of HRWRAs, specifically the influence of main chain length, side chain length, and molecular weight, on the fresh and hardened properties of cementitious systems.

Materials and admixture characteristics

The experimental program employed CEM I 42.5R Portland cement, Class F fly ash, standardized quartz sand as the fine aggregate, and three different polycarboxylate ether (PCE)-based HRWRAs. The physical, mechanical, and chemical properties of the cement and fly ash used in the experimental and modeling studies are presented in Table 2. [13,14,51]. The distinguishing parameters among the HRWRAs were their molecular weights (kg/mol), main chain lengths, and side chain lengths (g/mol). This controlled variation enabled a focused examination of the influence of molecular architecture on the performance of the cementitious systems.

Mortar mix design and proportions

A range of cement mortars were prepared, incorporating different admixture molecular structures and fly ash replacement rates of 0%, 15%, 30%, and 45% by total cement weight. All mixtures were produced with a constant w/c ratio of 0.485 and a sand-to-binder ratio of 2.75, following the specifications of ASTM C109.

Sample preparation and curing

The mortars were mixed and cast under controlled laboratory conditions to guarantee uniformity. Following casting, specimens were compacted using standardized procedures and demolded after 24 h. Subsequently, curing was performed in lime-saturated water at 23 ± 2 °C, adhering to conventional curing protocols to replicate typical field conditions.

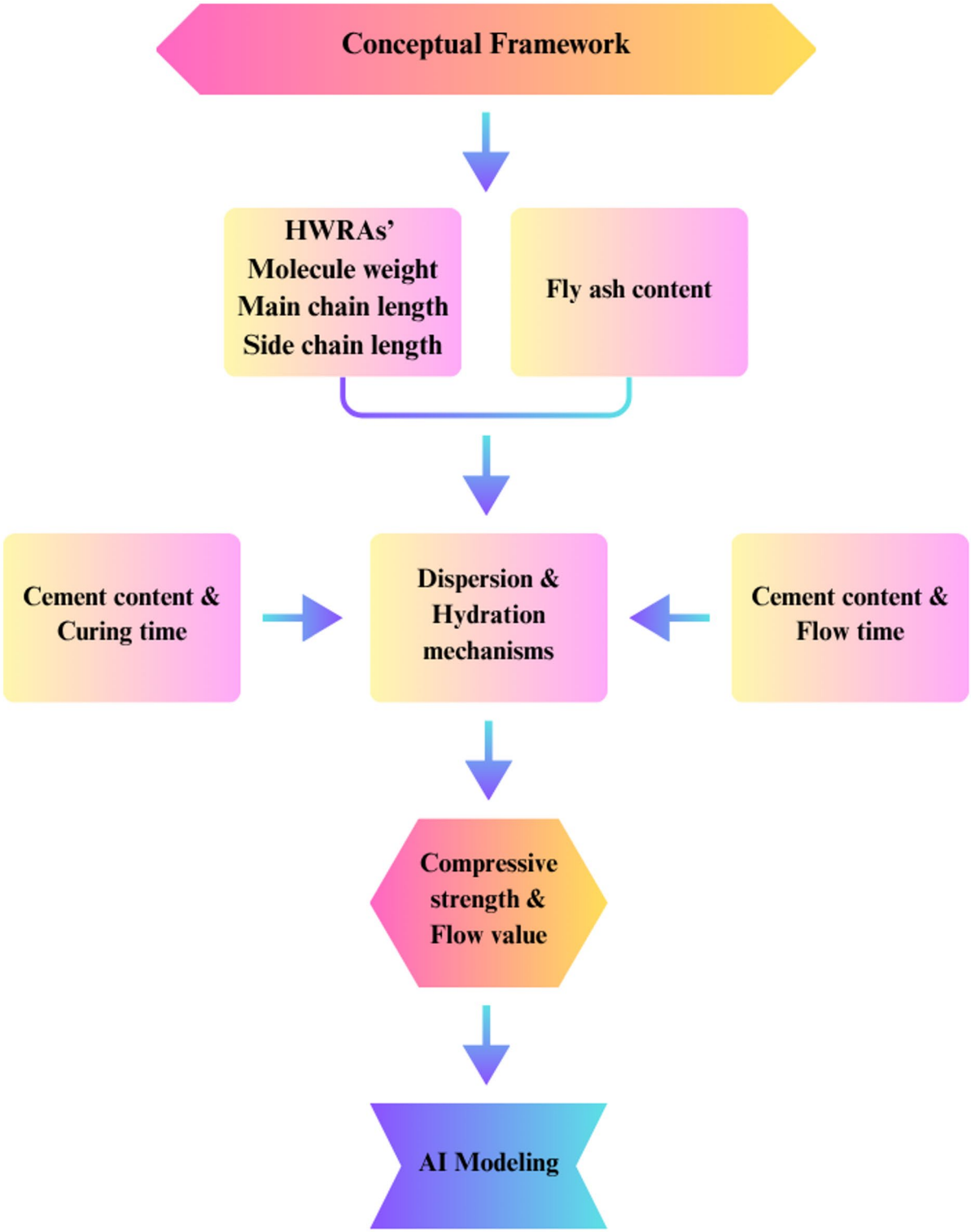


Fig. 1. Conceptual framework of the current study.

Testing of fresh and hardened properties

Flow properties were measured in accordance with ASTM C1437-20, evaluating the time-dependent spread behavior of the mortar mixtures. Compressive strength tests were conducted at various curing intervals (1, 3, 7, 28, and 90 days) following the ASTM C109 standard. Each testing protocol was replicated across different admixture formulations and fly ash replacement levels to ensure the reliability and statistical significance of the results.

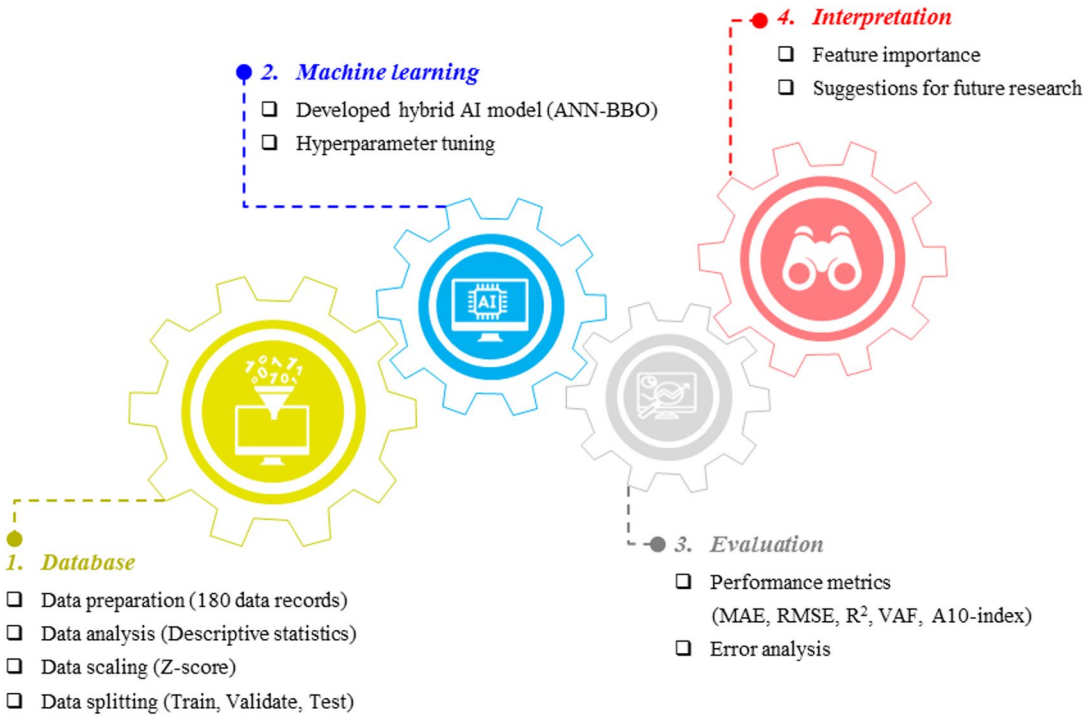


Fig. 2. Research process in this study.

Item	Amount (%)	
	Cement	Fly ash
SiO_2	18.86	59.22
Al_2O_3	5.71	22.86
Fe_2O_3	3.09	6.31
CaO	62.7	3.09
MgO	1.16	1.31
SO_3	2.39	0.17
$Na_2O + 0.658\ K_2O$	0.92	1.4
Cl^-	0.01	0.001
Insoluble residue	0.32	0.32
Loss on ignition	3.2	3.2
Free CaO	1.26	0
Physical properties		
Specific gravity	3.15	2.31
Blaine specific surface (cm^2/g)	3,530	4,300
Residual on 0.045 mm sieve (%)	7.6	10
Mechanical properties		
Compressive strength (MPa)		
1-day	14.7	-
2-day	26.8	-
7-day	49.8	85.9
28-day	58.5	100.7
90-day	-	110.2

Table 2. Chemical composition, physical and mechanical properties of cement and fly ash.

Definition of input and output variables

The experimental results were carefully recorded. The key parameters considered in this study included the amount of cement (X_1), fly ash (X_2), molecular weight of the HRWRA (X_3), main chain length of the HRWRA (X_4), side chain length of the HRWRA (X_5), curing time (X_6), and flow time (X_7), all of which were rigorously

analyzed and treated as input variables influencing the overall performance of the mortar mixture. Conversely, the variables related to performance outcomes, specifically CS (Y_1) and FV (Y_2), were systematically classified as dependent output variables, expected to quantitatively reflect the effects of the aforementioned input parameters. This organized categorization of variables enabled precise modeling of the relationships among admixture properties, mix design parameters, and mortar performance indicators.

Data normalization and statistical processing

For machine learning applications, all variables were normalized using the Z-score standardization method. This procedure eliminated dimensional biases and facilitated the convergence of predictive algorithms. Figure 3 depicts the histograms of the frequency distribution of the variables. Furthermore, key descriptive statistics (including the mean, standard deviation, kurtosis, and skewness) are presented in Table 3 to assess the distribution and variability of each variable in the dataset.

Outlier detection and treatment were carefully conducted prior to model development. Initially, the dataset was screened using both statistical (z-score analyse) and visual methods (Histogram plots) to identify abnormal values. Observations with unrealistic or inconsistent measurements were cross-checked against the original laboratory records, and only those confirmed as experimental errors were removed. All valid data points representing natural variability in material behavior were retained to preserve the representativeness of the dataset.

This systematic and scientifically rigorous data preparation established a robust foundation for the subsequent implementation of both ANN and ANN–BBO hybrid models. The insights derived from this process contribute to the advancement of high-performance, sustainable cementitious systems optimized through AI.

Selection of influencing variables

The input set was chosen to (i) be available at mix-design time, (ii) maintain a clear mechanistic link to strength and workability in FA–HRWRA systems, and (iii) avoid confounding with controlled factors. Variables were standardized prior to modeling.

- X_1: Cement—Principal contributor of clinker and early hydration products; governs paste reactivity and is a primary driver of compressive strength and water demand.
- X_2: Fly ash (FA)—Modulates the dilution–pozzolanic balance and particle packing; at fixed w/c, higher FA typically lowers early strength but can improve later-age strength and influences flow via particle morphology and glass content. Modeling Cement and FA separately (rather than as a single "binder%") preserves their counter-varying effects.
- X_3: Molecular weight (HRWRA)—Encodes hydrodynamic size/distribution of the PCE and relates to adsorption layer thickness and dispersion efficiency, affecting flow and early paste structure.
- X_4: Main chain length—Proxies backbone anchoring/adsorption on cement surfaces and competitive adsorption under changing ionic strength; impacts steric stabilization and dispersion robustness.
- X_5: Side chain length—Governs steric hindrance and flow retention; longer side chains generally mitigate re-flocculation and improve time-dependent workability.
- X_6: Curing time—Captures hydration progress and secondary reactions (including FA contribution), the dominant driver of compressive strength evolution.
- X_7: Flow time—Captures thixotropy/structuration and HRWRA retention effects, the dominant driver of flow evolution.

The water-to-cement ratio (w/c), sand ratio, temperature, and mixing protocol were held constant in the experimental program and therefore not modeled to avoid confounding. HRWRA dosage was fixed within series; consequently, the analysis emphasizes architecture descriptors (X_3–X_5) as the differentiating admixture factors.

Because Cement and FA sum to the binder and can be collinear, both were retained for interpretability and monitored via diagnostic checks (e.g., feature-importance stability and sensitivity analyses). Resulting importances were consistent with mechanism: Curing time and Cement/FA dominate compressive strength, whereas Flow time and Side chain length most strongly influence flow, with Molecular weight and Main chain length providing secondary but meaningful contributions.

These choices reflect practical levers available to mix-design engineers and align with the controlled domain of this study (Class F FA; PCE-type HRWRA). Extensions to variable w/c, broader admixture families, and multi-lab conditions are reserved for future work. The complete mix desing is provided in Supplementary Table S1.

Overview of the AI methods employed

Artificial neural network (ANN)

The first attempt to develop a mathematical model of an ANN was undertaken by McCulloch and Pitts[52]. The concept of ANNs was inspired by the structure and function of biological neural networks found in the human brain. The computational structure of ANNs are modeled on the signal transmission mechanisms observed in nerve cells (neurons) in the human brain. They simulate the process of receiving input signals and producing corresponding outputs through interconnected layers of artificial neurons. An ANN is generally composed of three primary layers:

- An input layer (*IL*), which receives data and signals from outside sources, with a neuron count equal to the number of input variables;

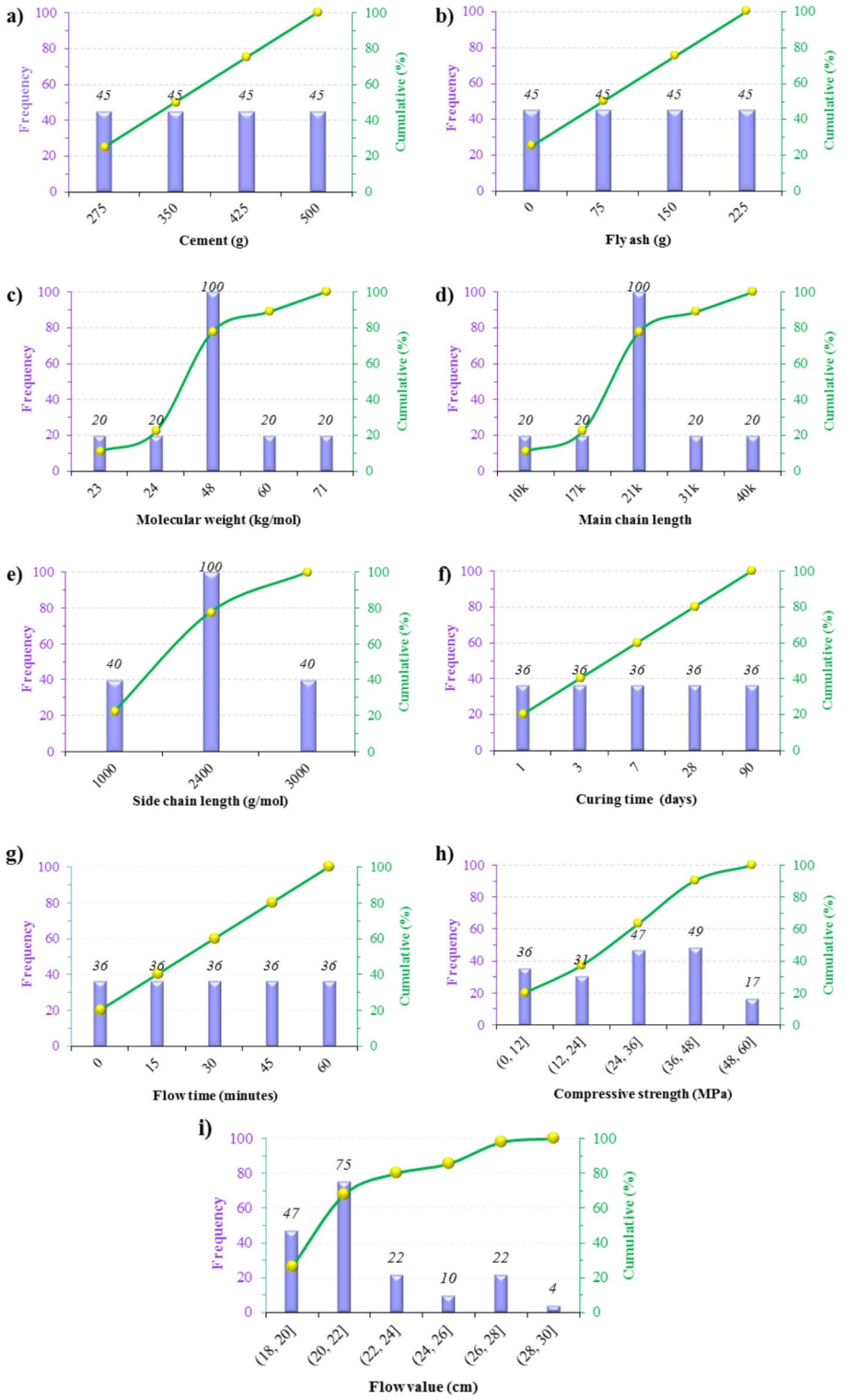


Fig. 3. Histograms of the frequency of variables: inputs (a-g) and outputs (h & i).

- An output layer (*OL*), which delivers the final predictions or outcomes of the model, with its neurons representing the number of output variables of the modeled problem; and
- Hidden layer/s (*HL/s*) positioned between layers of *IL* and *OL*, handle the internal processing of information within the network. The number of neurons in these *HL/s* can vary, and determining an optimal number is crucial for ensuring high model accuracy[53,54]. At a glance, Fig. 4 provides an overview of the ANN model inspired by the human nervous system and its structure and function.

ID No.: variables	Category	Unit	Range	Mean	Standard deviation	Kurtosis	Skewness
X_1: Cement	Input for Y_1 & Y_2	g	275–500	387.50	83.85	-1.36	0.00
X_2: Fly ash	Input for Y_1 & Y_2	g	0–225	112.50	83.85	-1.36	0.00
X_3: Molecular weight	Input for Y_1 & Y_2	kg/mol	23–71	46.44	14.35	-0.48	-0.26
X_4: Main chain length	Input for Y_1 & Y_2	-	10–40	22.56	8.00	0.39	0.82
X_5: Side chain length	Input for Y_1 & Y_2	g/mol	1000–3000	2222.22	695.67	-0.51	-0.89
X_6: Curing time	Input for Y_1	day	1–90	25.80	33.51	-0.15	1.24
X_7: Flow time	Input for Y_2	minute	0–60	30.00	21.21	-1.303	0.00
Y_1: Compressive strength	Output of Y_1	MPa	1.8–57.7	21.967	2.463	0.280	1.133
Y_2: Flow value	Output of Y_2	cm	18.3–28.5	21.968	2.607	-0.089	1.007

Table 3. Descriptive statistics of the variables.

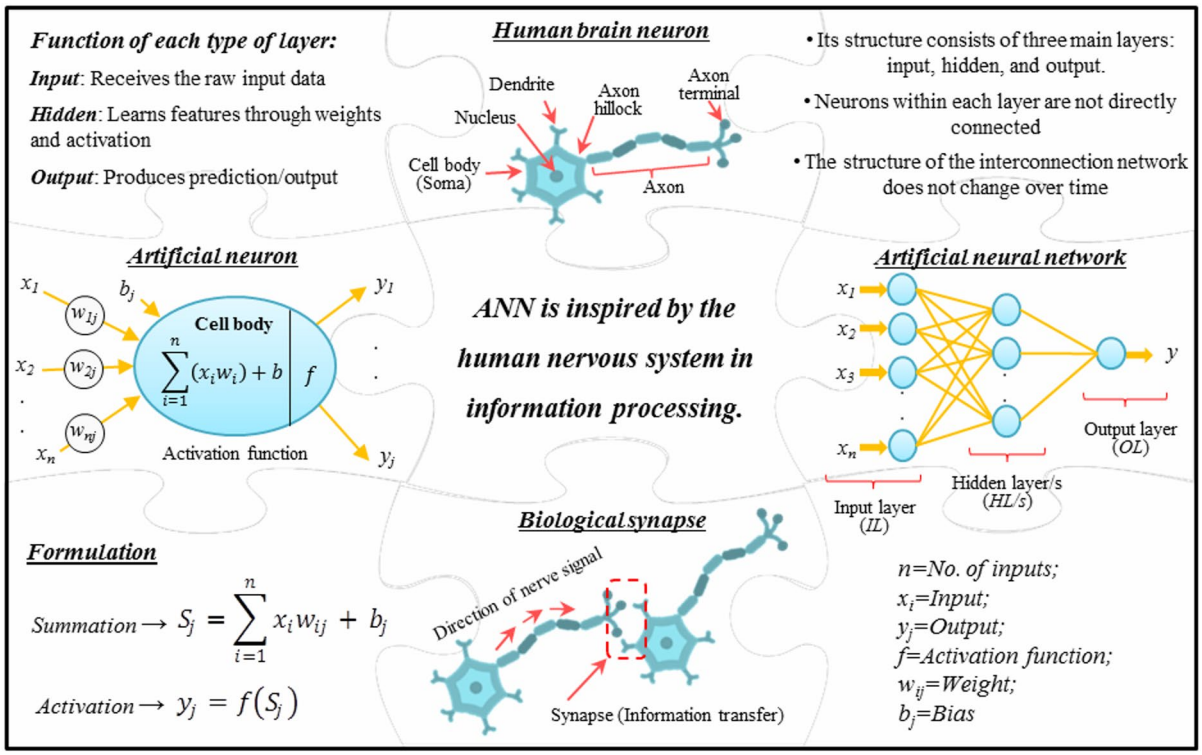


Fig. 4. A schematic illustrating the introduction and performance of the ANN model.

Biogeography-based optimization (BBO)

In 2008, BBO was developed by Simon[55] as an evolutionary algorithm inspired by the concept of biogeography. Biogeography deals with how species evolve from their distribution in ecosystems to their displacement and extinction, where habitats are small ecosystems composed of diverse populations of creatures that are typically segregated from other habitats. Regions that can support a diverse range of creatures are evaluated using a habitat suitability index (HSI). The HSI of a given habitat is influenced by characteristics such as climate, vegetation, etc. These characteristics, known as suitability index variables (SIVs), play a key role in determining the HSI. A habitat with a high HSI suggests it offers suitable conditions for supporting larger populations, whereas a low HSI reflects poor conditions that often lead to increased migration and population decrease. Consequently, regions with high HSI typically exchange SIV with those having low HSI. The two main operators considered in this optimization method are:

- Migration: This process facilitates the finding of new regions in the search space by switching solutions between habitats through emigration and immigration rates to enable global search capabilities.
- Mutation: By creating random conditions in the solution search, it not only increases the diversity in the population but also prevents stagnation in local optima.

Figure 5 provides an overview of the BBO method inspired by the principles of biogeography and its mathematical formulation.

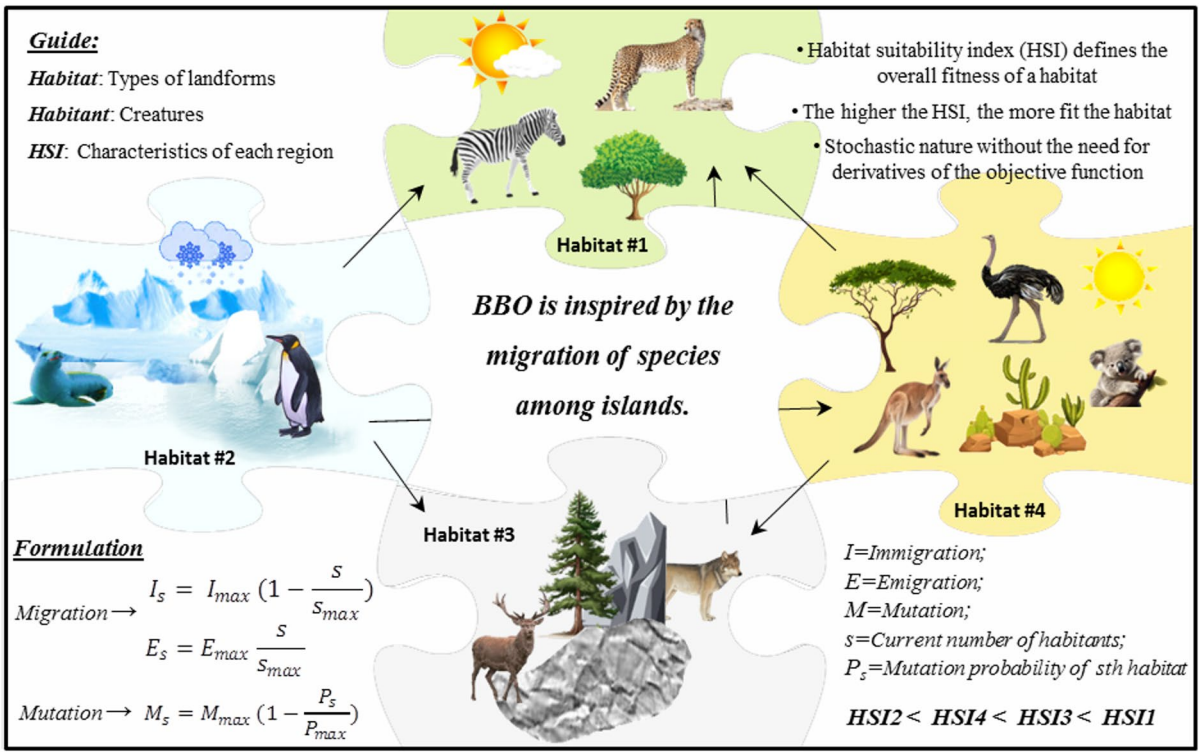


Fig. 5. A schematic illustrating the introduction and performance of the BBO model.

After gaining a comprehensive understanding of the BBO algorithm, its distinctive advantages over other optimization methods become evident. The superior performance of BBO primarily arises from its unique migration and mutation mechanisms, which enable the algorithm to effectively escape local minima and maintain population diversity[56]. These operators introduce abrupt yet beneficial perturbations in the candidate solutions, thereby strengthening the algorithm's exploration capability. In a comparative study by Mirjalili et al.[56] the performance of BBO was evaluated against several established algorithms, including particle swarm optimization (PSO), genetic algorithm (GA), ant colony optimization (ACO), evolutionary strategies, and probability-based incremental learning. The findings demonstrated that, unlike these methods, BBO incorporates operators that foster sudden transitions in candidate solutions, reducing the likelihood of premature convergence to local optima. Furthermore, BBO exhibits desirable characteristics such as simplicity, flexibility, and computational efficiency, and, due to its stochastic, derivative-free nature, it can handle complex optimization problems without requiring gradient information. Previous research has confirmed that BBO is a robust and competitive algorithm capable of addressing a broad range of real-world engineering and scientific optimization challenges[57].

Model development

This section aims to provide a comprehensive description of the model development carried out to propose an AI approach for predicting the CS and FV of cementitious systems incorporating fly ash. To accomplish this, two AI modeling strategies are implemented: a classical ANN and a hybrid model that integrates ANN with BBO. The comparison between these models provides important insights into the performance of the hybrid approach relative to the classical model. The metaheuristic optimization method of BBO is used to determine the optimal parameters, aiming to improve the prediction model's accuracy and reduce errors. Data preprocessing is known as a crucial early step in the modeling process, as it greatly enhances data quality by eliminating biases resulting from differences in scale or units among variables. This, in turn, enhances the performance of the developed models. Accordingly, all data were normalized utilizing the Z-score method, as shown in Eq. (1).

$$Z = \frac{X - \mu}{\sigma} \tag{1}$$

Here, X represents the value of each variable, while μ and σ denote the mean and standard deviation of the data for that variable, respectively.

When evaluating AI models, an important consideration is whether the developed model is the optimal choice within its hypothesis space, especially regarding its ability to generalize to new data. This issue is closely linked to how the data is divided, which is a vital aspect in the model development process. To ensure proper model generalization and prevent biased performance estimation, the present study employed a three-way holdout validation strategy consisting of separate training, validation, and testing datasets[58]. This approach minimizes the risk of over-optimistic results that may arise when the same dataset is repeatedly used for both

model development and evaluation. According to Mehlig[59], an effective data split involves three distinct subsets: (i) a training set to build the model and capture underlying data patterns; (ii) a validation set to fine-tune the model and control overfitting; and (iii) a testing set to independently assess generalization capability. Because model performance can be highly sensitive to the composition of training data, the available parameters were sampled across a broad range within each subset. Consequently, the complete dataset—comprising 180 records for both compressive strength (CS) and flow value (FV)—was randomly divided into training (134 records), validation (23 records), and testing (23 records) subsets to ensure an unbiased and statistically representative evaluation of the proposed models. Numerous studies evaluating various learning algorithms have reported that the Levenberg–Marquardt algorithm outperforms others[60]. Therefore, it has been selected for use in this study. Moreover, the hyperbolic tangent is employed as the transfer function, as outlined in Haykin's highly cited book[61]. The architecture of the model, particularly the number of hidden layers (HL), affects the effective convergence of the model in the training process[62]. Based on the universal approximation theorem[59], the models are designed with a single HL. Accordingly, the initial architecture of the models used for CS and FV prediction is structured as *6-*$HL_{neurons}$*-1*, as illustrated in Fig. 6. To establish the final architecture of the model (i.e., $HL_{neurons}$), a detailed analysis will be conducted at the beginning of the results section.

Evaluation criteria

To evaluate the performance of the proposed models and determine the most effective one, various statistical metrics have been utilized. Table 4 provides a concise description of the metrics along with their mathematical formula and ideal values.

Results and discussion

Establishing the final architecture of the models

As stated in the model development section, identifying the appropriate number of $HL_{neurons}$ and finalizing the model architecture involves thorough analysis. To do this, models with different architectures ranging from 10 to 20 $HL_{neurons}$ are assessed. To guaranty a fair comparison, the models are run using identical subsets. The evaluation criteria R^2 and RMSE are calculated for the models and ranked according to their performance results. The ranking method is that the top is assigned to the highest R^2 and the lowest RMSE. Ultimately, the model that achieves the top overall ranking score is regarded as the best model. Table 5 presents the results of this analysis. According to these results, the model containing 16 $HL_{neurons}$ achieves the top ranking score, indicating its superior performance over other $HL_{neurons}$ architectures. Therefore, the model with the final architecture of *6-16-1* is selected.

Table 6 outlines the parameter settings used to enhance the clarity of the presented information. All model developments were carried out using MATLAB software.

Evaluating the performance of models

Figures 7 and 8 present the scatter plots and error distribution plots of CS modeling for the training, validating, and testing sets using the two developed models: ANN and ANN–BBO, respectively. The ANN model attained R^2 values of 0.9325, 0.9048, and 0.9180 for the training, validating, and testing sets, respectively. In comparison, the ANN–BBO model achieved higher R^2 values of 0.9917, 0.9897, and 0.9903 for the same sets. The results depict a stronger correlation between the measured and predicted values in the ANN–BBO model, with its data points aligning more closely with the ideal regression line ($y = x$) compared to those of the ANN model. To further assess the models, the figures illustrate the error distribution for each model across all three performance sets. A detailed analysis of the error margins reveals that for the ANN–BBO model, 85%, 95%, and 87% of predictions in the training, validating, and testing sets, respectively, fell within an error range of [-10%, 10%].

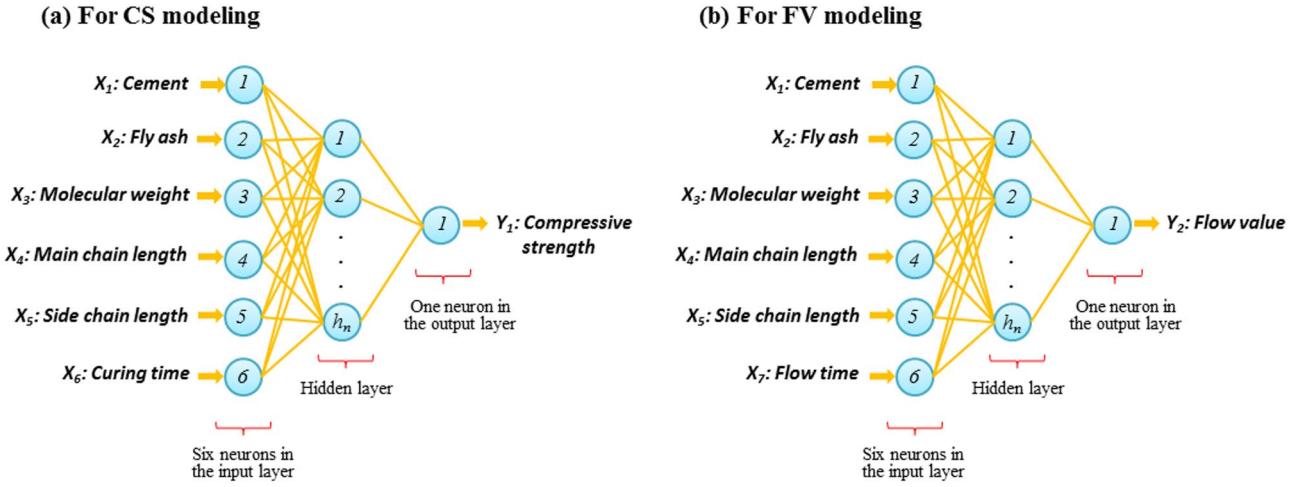


Fig. 6. Model architecture for predicting (**a**) compressive strength and (**b**) flow value of cementitious systems containing fly ash.

Metric formula	Brief description	Range ideal value
Mean absolute error (MAE) $MAE = \frac{1}{N}\sum_{i=1}^{N}\left\|A_i - P_{p_i}\right\|$	The average magnitude of the errors between the predicted and actual values, regardless of their direction	$(0, +\infty)$ 0
Root mean squared error (RMSE) $RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(A_i - P_i)^2}$	The measure of how widely errors are dispersed around the line of best fit	$(0, +\infty)$ 0
Coefficient of determination (R^2) $R^2 = \left(\frac{\sum_{i=1}^{N}\left(A_i-\overline{A}\right)\left(P_i-\overline{P}\right)}{\sqrt{\left[\sum_{i=1}^{N}\left(A_i-\overline{A}\right)^2\right]\left[\sum_{i=1}^{N}\left(P_i-\overline{P}\right)^2\right]}}\right)^2$	The metric used to evaluate how well the model explains the variability of the response variable	(0, 1) 1
Nash–Sutcliffe efficiency (NSE) $NSE = 1 - \frac{\sum_{i=1}^{N}(A_i-P_i)^2}{\sum_{i=1}^{N}\left(A_i-\overline{A}\right)^2}$	Measures how closely a model's predictions align with the actual values by comparing the relative magnitude of the prediction error variance to the variance of the actual values	$(-\infty, 1)$ 1
Variance Accounted For (VAF) $VAF = \left(1 - \frac{var(A-P)}{var(A)}\right) \times 100$	The metric to describe the variability of the dataset by the model	(0, 100) 100%
A10-index $\frac{m10}{N}$ where m10 = No. of data with $0.90 \le A/P \le 1.10$	The proportion of data records whose predicted values deviate by ± 10% from the actual values	(0, 1) 1

Table 4. Summary of statistical metrics used. The number of data related to N = total, N_{Tr} = training, N_{Va} = validating, and N_{Te} = testing sets. A_i and P_i = The actual and predicted values of the *i*th data, respectively. $\overline{A}$ and $\overline{P}$ = The average of total A_i and P_i, respectively.

	Evaluation criteria values				Ranking of criteria				
	Compressive strength		Flow value		Compressive strength		Flow value		
No. of $HL_{neurons}$	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	Overall ranking score
HL_{10}	0.953	1.485	0.931	0.412	2	3	2	1	8
HL_{11}	0.966	1.463	0.947	0.385	3	4	3	2	12
HL_{12}	0.972	1.417	0.961	0.376	6	6	5	4	21
HL_{13}	0.978	1.404	0.970	0.355	7	7	7	6	27
HL_{14}	0.980	1.396	0.975	0.344	8	8	8	8	32
HL_{15}	0.988	1.372	0.978	0.320	10	10	9	11	40
HL_{16}	0.991	1.365	0.986	0.325	11	11	11	10	43
HL_{17}	0.985	1.375	0.982	0.334	9	9	10	9	37
HL_{18}	0.970	1.422	0.968	0.349	5	5	6	7	23
HL_{19}	0.967	1.488	0.957	0.360	4	2	4	5	15
HL_{20}	0.945	1.525	0.923	0.384	1	1	1	3	6

Table 5. Model performance and rankings based on varying numbers of $HL_{neurons}$.

Model	Parameter	Setting
ANN	Dataset segmentation	
	Training set	134 data
	Validation set	23 data
	Testing set	23 data
	No. of input parameter	6
	No. of hidden layer/nodes	1/16
	Learning algorithm	Levenberg–Marquardt
	Transfer function	Hyperbolic tangent
	Max. no. of epochs	500
BBO	Population size	100
	Max emigration rate	1
	Habitat modification	1
	Mutation probability	0.005

Table 6. Settings of the parameters used.

In comparison, the ANN model shows lower percentages of 44%, 65%, and 26% for the same sets. This clearly demonstrates the enhanced accuracy of the ANN–BBO model in predicting the CS of cementitious systems incorporating fly ash within a tighter error margin.

To better understand the models' performance, Fig. 9 presents the ratios of predicted to measured CS values, offering a clear view of the model outputs for each data record in both developed models. Values on the vertical axis that are closer to 1 are considered desirable, as they reflect a strong similarity between the model predictions and the measured data. Figure 9 shows that the CS values predicted by the proposed ANN–BBO model align more closely with the measured CS values compared to those predicted by the classical ANN model. Specifically, 156 out of 180 data records (87%) fall within the 0.9–1.1 range for the hybrid ANN–BBO model, whereas only 81 out of 180 data records (45%) fall within this range for the single ANN model.

A similar analysis is performed for the FV property of cementitious systems incorporating fly ash in Figs. 10, 11, 12. The correlation values for the single ANN model's performance are 0.9219 for the training set, 0.91 for the validating set, and 0.9183 for the testing set (Fig. 10), while the corresponding values for the hybrid ANN–BBO model are 0.9845, 0.9905, and 0.9823, respectively (Fig. 11). A comparison of the two figures shows that the hybrid model yields a higher concentration of data records within the region bounded by the black dotted lines—representing a ± 10% deviation from the $y = x$ line—indicating that the ANN–BBO model outperforms the single ANN model in estimating the FV of cementitious systems containing fly ash. This is further evident in the error distribution graphs, which show that the predictions from the single model exhibit a wider spread (indicating higher error) compared to those from the hybrid model.

Figure 12 illustrates that the predicted to measured FV value ratio varies roughly between 0.85 and 1.1 for the single ANN model, while it falls within a narrower range of approximately 0.95 to 1.05 for the hybrid ANN–BBO model. The ratio values in the hybrid model are noticeably closer to the ideal line of $y = 1$, indicating strong alignment of the predicted to the measured values. The results of FV performance modeling are consistent with earlier results regarding the effectiveness of the ANN–BBO model in predicting the CS of cementitious systems incorporating fly ash.

To gain deeper insight into the performance of the proposed models, several evaluation criteria were applied, with the results summarized in Table 7. Grasping the ideal value of these criteria enables more effective evaluation. Specifically, the best performance for MAE and RMSE is indicated by values nearing zero, while ideal values for NSE are close to 1, and for VAF and the A10-index, they should approach 100%. By comparing the evaluation criteria across all sets, it is evident that the hybrid proposed model has considerably reduced the prediction errors and consequently achieved higher accuracy than the classical ANN model. For example, the results emphasize that the ANN–BBO model outperformed the standalone ANN model, demonstrating approximately a 67% reduction in both MAE (dropping from 0.1486 in ANN to 0.0490 in ANN–BBO) and RMSE (from 4.40 in ANN to 1.36 in ANN–BBO) across all CS data. Likewise, the FV modeling results exhibit a decrease of around 50%. In addition, as presented in Table 7, the hybrid ANN–BBO model consistently outperformed the standalone ANN across all datasets for both output variables. For CS prediction, the ANN–BBO achieved remarkably higher R^2 (0.992 [95% CI 0.988–0.995] for all data) and substantially lower RMSE values (1.36 [95% CI 1.11–1.66]), indicating superior accuracy and generalization capability. Similarly, for FV prediction, the ANN–BBO model maintained excellent agreement with experimental data ($R^2 = 0.985$ [95% CI 0.981–0.989]; RMSE = 0.32 [95% CI 0.28–0.36]), surpassing the standalone ANN by a significant margin. The narrow 95% confidence intervals confirm the statistical robustness, reliability, and stability of the proposed ANN–BBO model. Therefore, the evaluation metrics provide additional support for the earlier findings and reinforce the effectiveness of the proposed hybrid model in predicting the properties of cementitious systems containing fly ash.

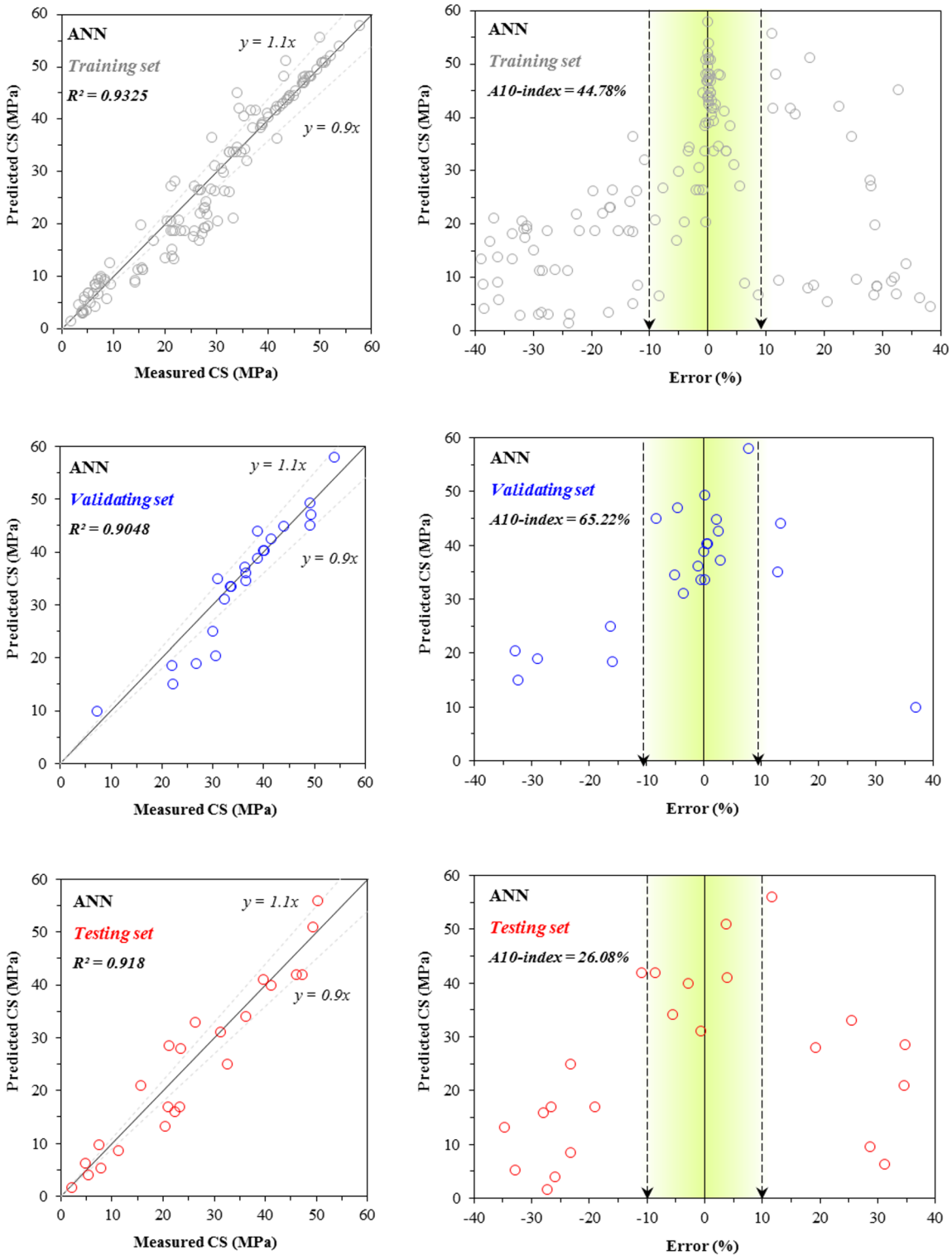


Fig. 7. Comparison between the measured and predicted ANN values for CS, including the error distribution.

Contribution of input variables in modeling the CS and FV of cementitious systems containing fly ash

To evaluate the contribution of input variables to the performance of models predicting the CS and FV of cementitious systems containing fly ash, a comparative analysis was performed. Accordingly, the impact of each input on the best model's performance is assessed by evaluating how the model behaves when that specific input is excluded from the modeling process. This method helps identify variables whose removal significantly reduces the model's performance, indicating their important role in the prediction process. Figure 13 presents

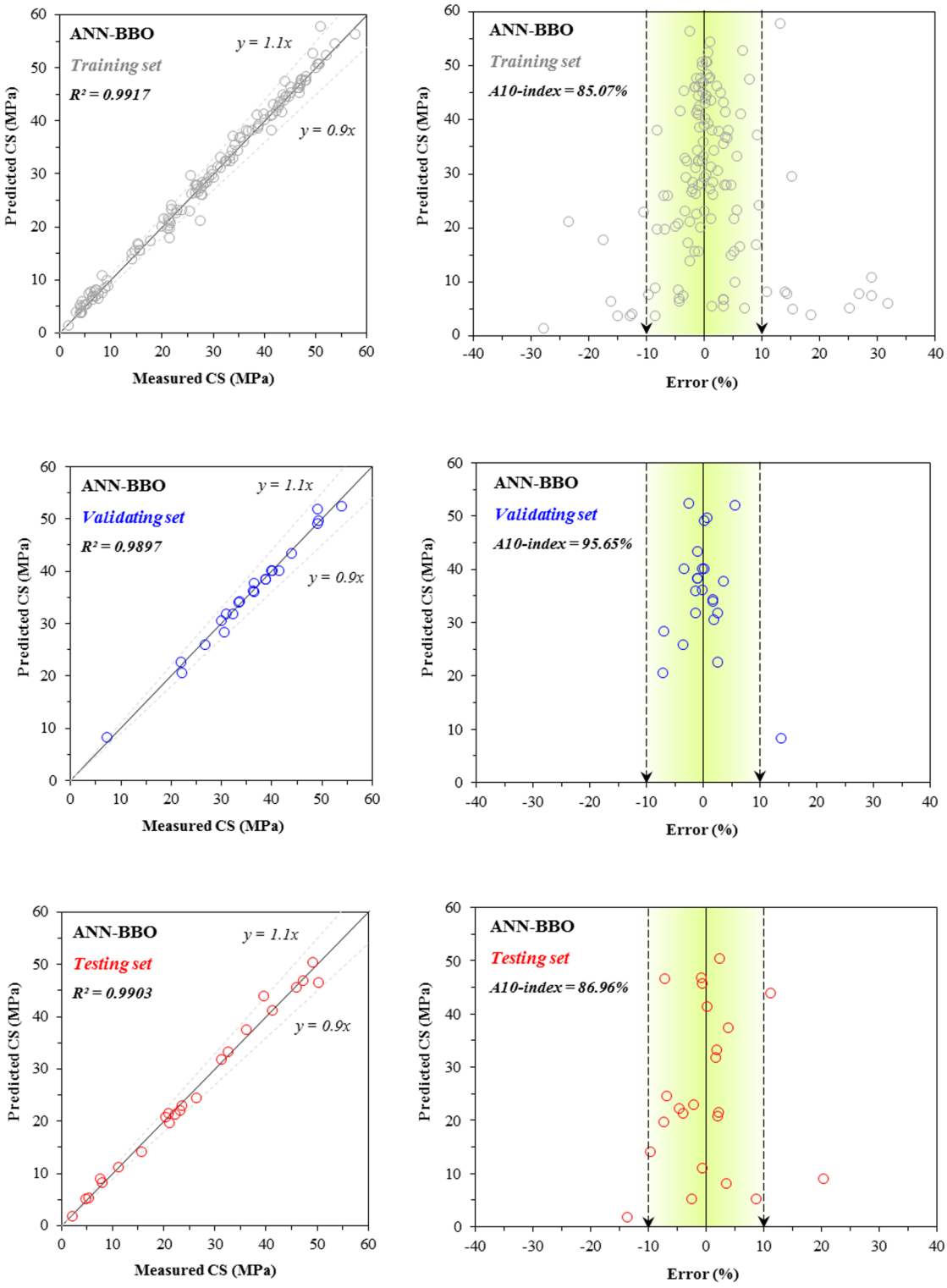


Fig. 8. Comparison between the measured and predicted ANN–BBO values for CS, including the error distribution.

the importance of the variables considered in the most effective model (ANN–BBO). While all input variables play a role in the model's performance, their levels of influence differ. As shown in Fig. 13, curing time (X_6) and flow time (X_7) are the most influential variables affecting CS and FV, contributing 27% and 30%, respectively. Next in importance for predicting CS are cement ($X_1 = 20\%$) and fly ash ($X_2 = 16\%$) as cementitious materials. Similarly, molecular weight ($X_3 = 20\%$) and main chain length ($X_4 = 15\%$) are key characteristics of HRWRAs influencing the prediction of FV. Altogether, the three input variables—curing time, cement, and fly ash—for CS,

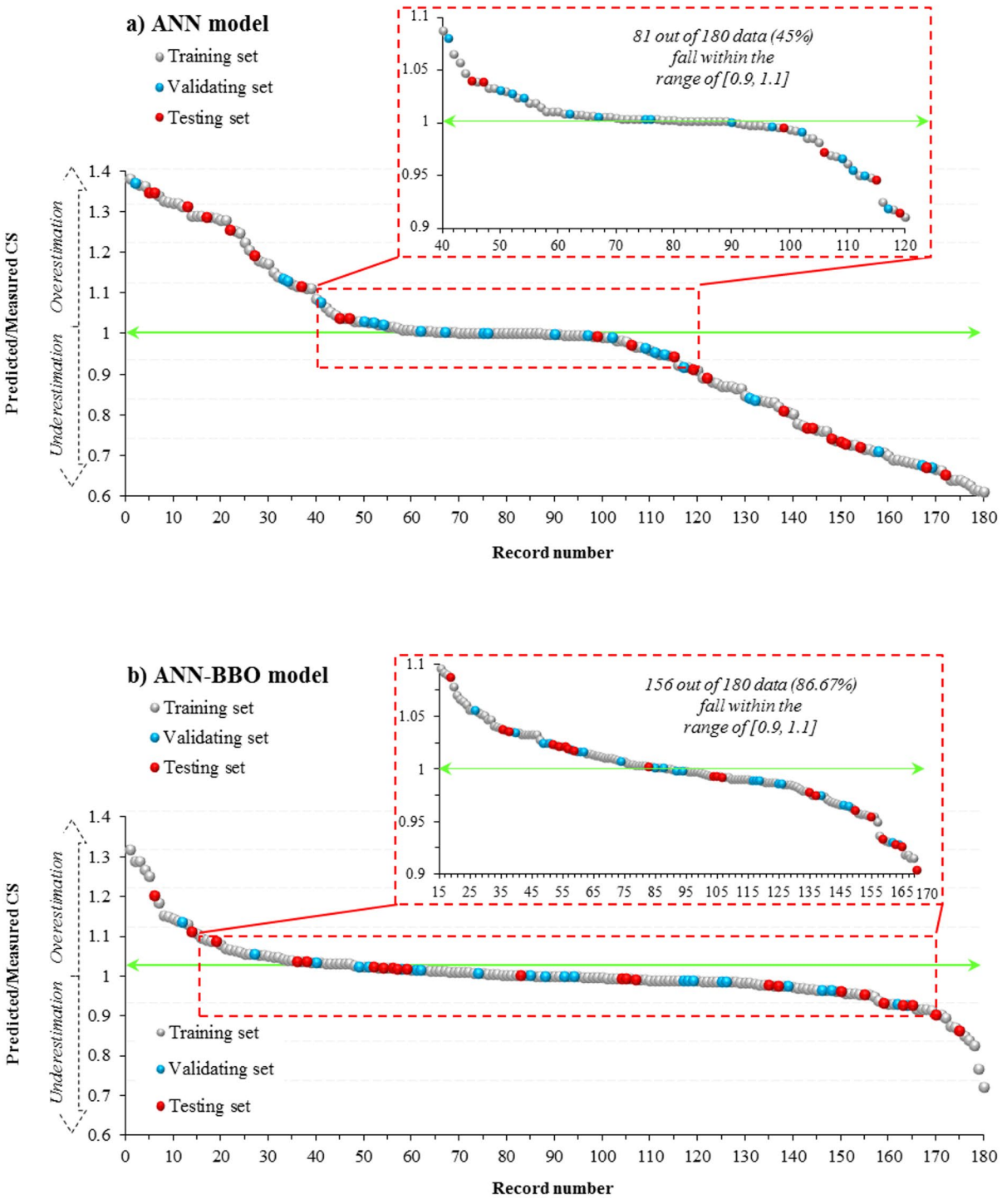


Fig. 9. Ratios of predicted/measured CS value for the models (**a**) ANN and (**b**) ANN–BBO.

and flow time, molecular weight, and main chain length for FV, contribute approximately 65% to the predictive accuracy for estimating CS and FV in cementitious systems containing fly ash. The remaining variables have a comparatively lower impact on the predictions.

To gain deeper insights into the contribution of individual input parameters, the SHapley Additive exPlanations (SHAP) technique—derived from cooperative game theory—was applied to quantify and interpret each variable's influence on the model output[63]. Figure 14 visualizes the relative importance of all features by depicting their respective impacts on the prediction outcomes. The color scale represents the magnitude of each feature, transitioning from low (yellow) to high (dark purple) values. The horizontal axis indicates the directional effect of each variable on the model prediction, where positive and negative contributions correspond to increases

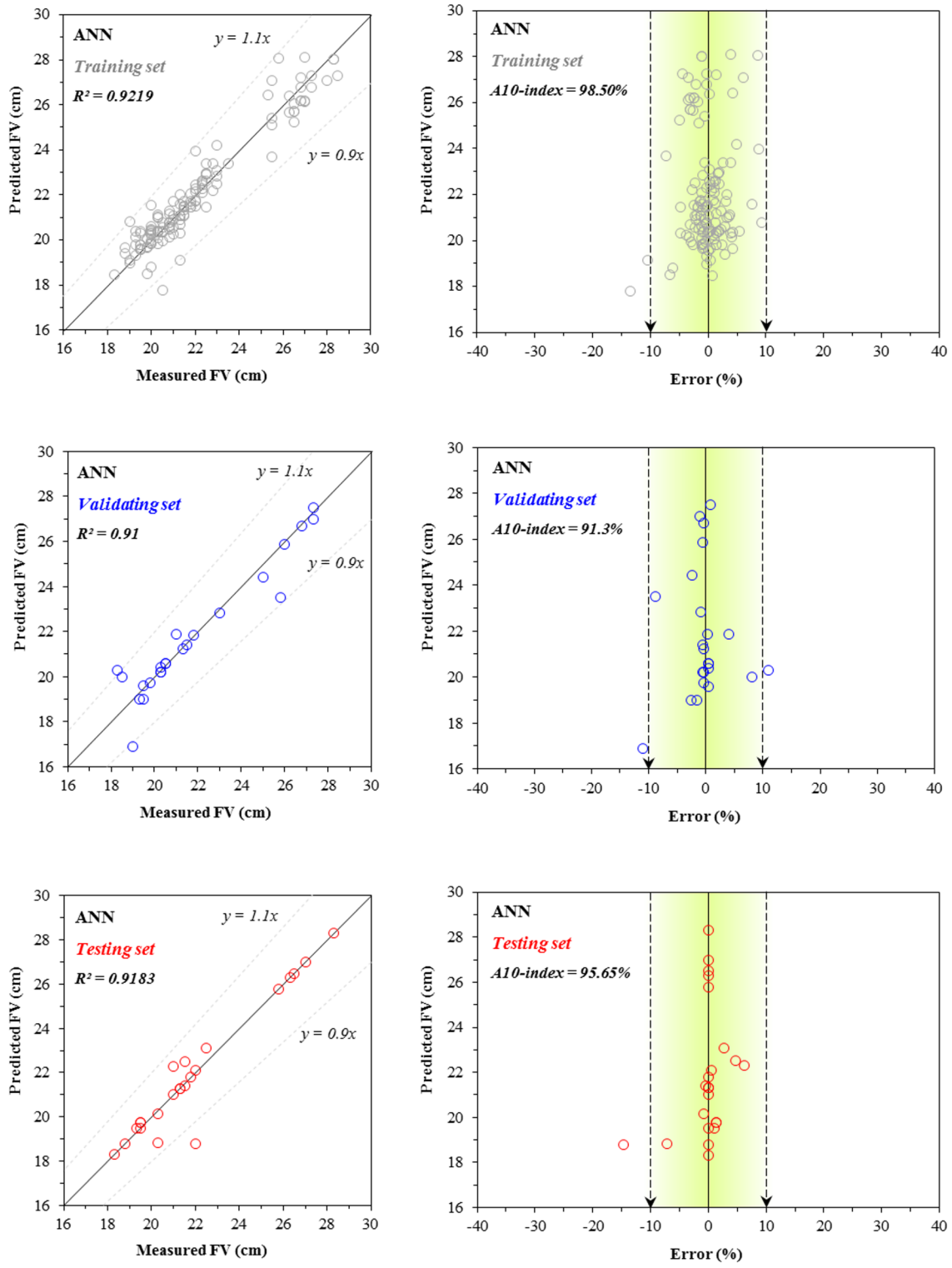


Fig. 10. Comparison between the measured and predicted ANN values for FV, including the error distribution.

or decreases in the predicted response, respectively. For CS, the curing time (X_6) emerged as the most influential parameter, with the dominance of dark-purple points on the positive side confirming its enhancing effect on strength development. The subsequent key contributors were cement content (X_1) and fly ash (X_2), both of which are primary constituents governing the mechanical performance of cementitious systems. Conversely, for FV, the most critical variable was the flow time (X_7), exhibiting a predominant negative influence. The following influential parameters for FV were the HRWRA-related features—molecular weight (X_3) and main chain length

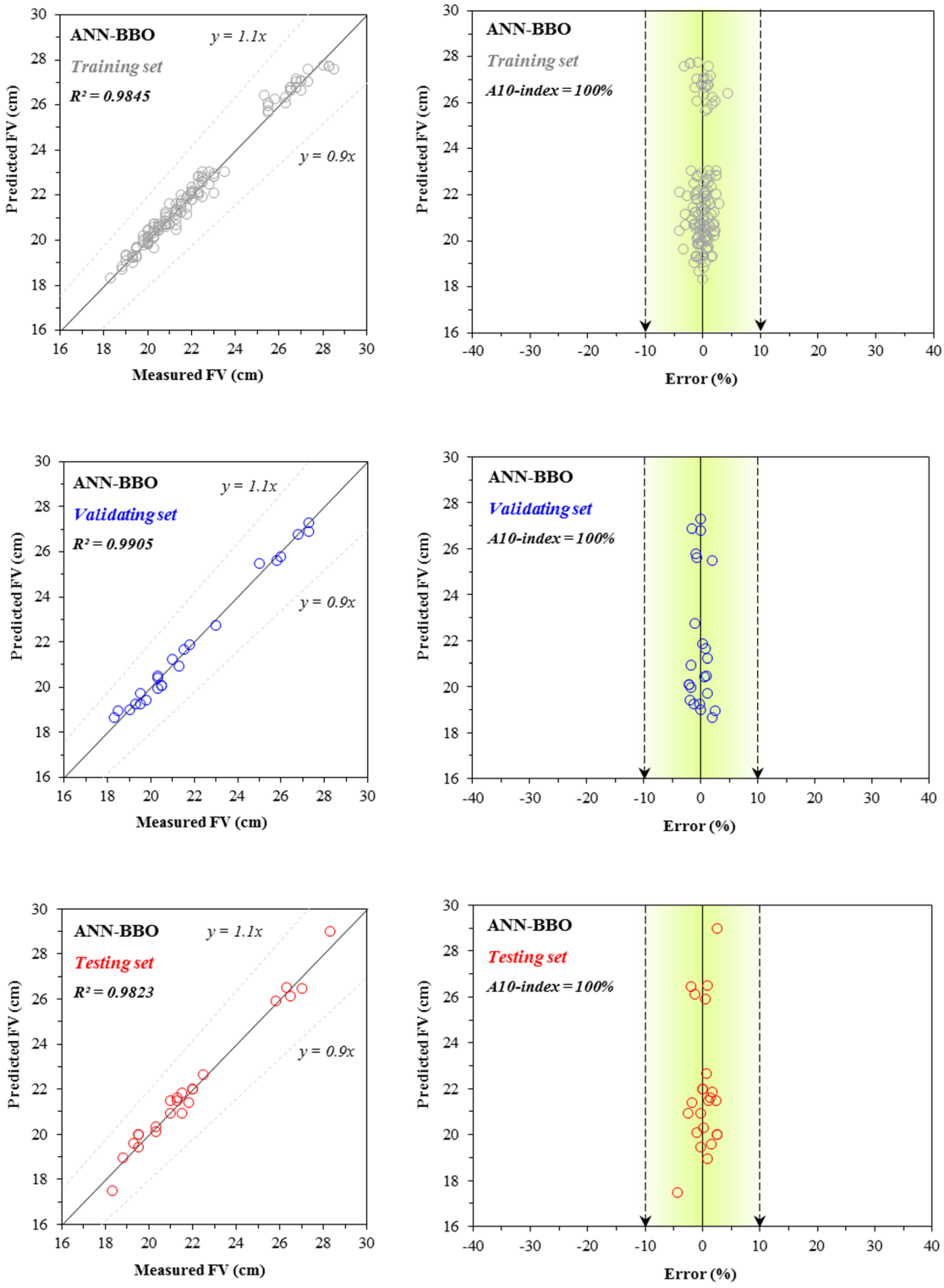


Fig. 11. Comparison between the measured and predicted ANN–BBO values for FV, including the error distribution.

(X_4)—which strongly affect the rheological behavior of the mixtures. These findings collectively demonstrate that the SHAP analysis provides a consistent and interpretable framework for identifying the dominant factors governing both mechanical and rheological properties in fly ash–based cementitious composites.

In this regard, Fig. 15 enables a more expressive analysis of the effect of the variables on the CS and FV outputs. As can be seen in Fig. 15a,b, CS exhibits the highest sensitivity to changes in curing time, while FV is most sensitive to variations in flow time. The main difference lies in the fact that the CS of cementitious systems

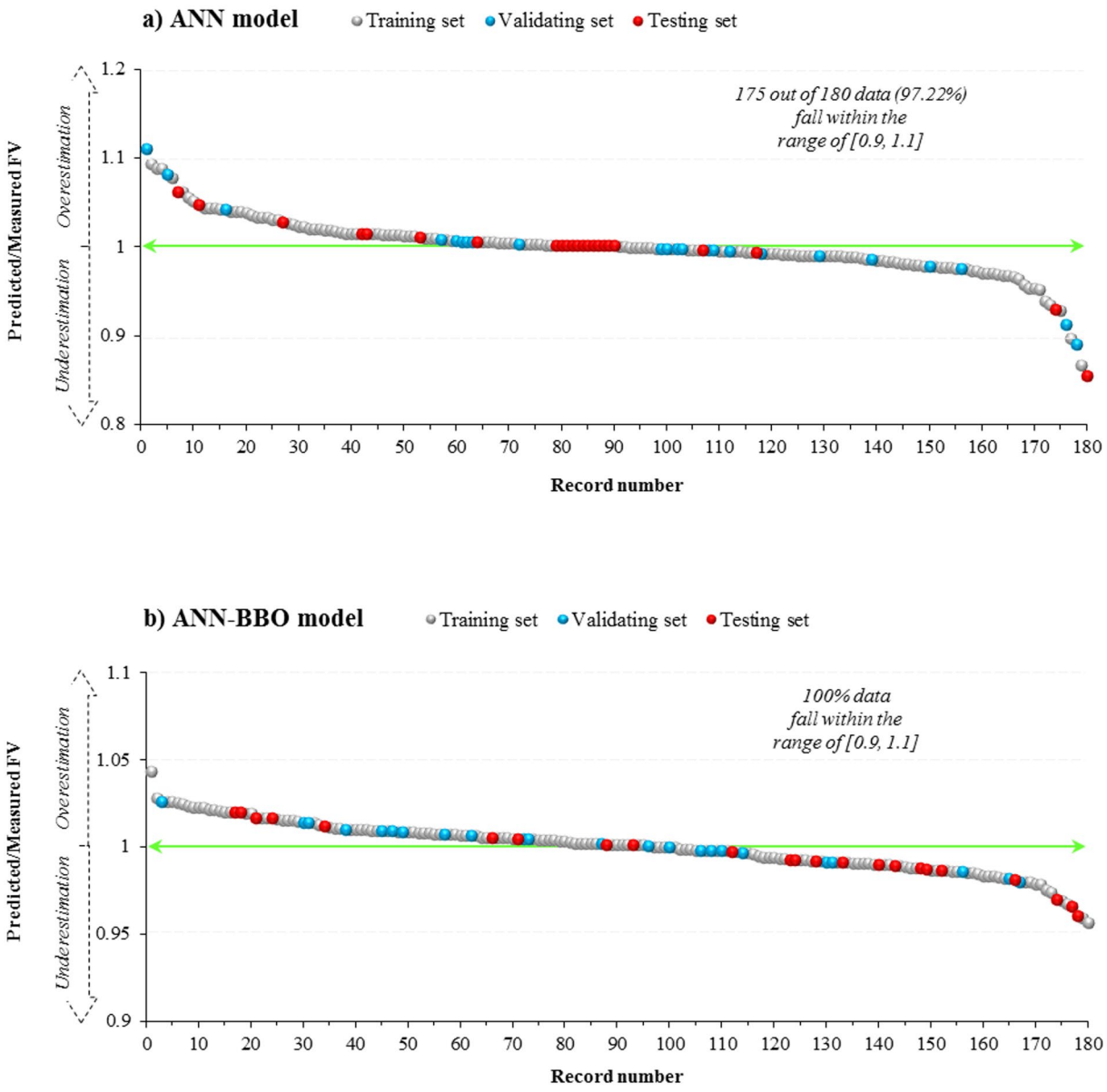


Fig. 12. Ratios of predicted/measured FV value for the models (**a**) ANN and (**b**) ANN–BBO.

increases with longer curing time. On the other hand, flow time is inversely related to FV, meaning that greater flow times reflect lower flowability. For example, the most notable changes in CS take place in the early stages of curing, especially from day 1 to day 7, whereas the most pronounced variations in FV are observed within the first 30 min. Mixtures containing 425 g of cement and 150 g of fly ash demonstrated the highest CS. Following the flow time variable, the FV is influenced by two factors of molecular weight and main chain length. The analyses presented in this section can be valuable because they clarify the impact of each variable used on the model's output, providing clearer insight into the role of these variables and how they affect the model's predictions.

Conclusions

The compressive strength and flow value of cementitious systems were statistically modeled using data collected from the experimental investigation of the effects of fly ash substitution and the chemical structure of HRWRAs.

In conclusion, the ANN model achieved R^2 values of 0.9325, 0.9048, and 0.9180 for CS during training, validation, and testing phases, respectively. For FV, the ANN model predicted R^2 values of 0.9219, 0.9100, and 0.9183 at training, validation, and testing stages, respectively. In contrast, the ANN–BBO hybrid model substantially outperformed the ANN alone, attaining R^2 values of 0.9917, 0.9897, and 0.9903 for CS across training, validation, and testing sets, respectively. For FV, the ANN–BBO model yielded R^2 values of 0.9845, 0.9905, and 0.9823 at the corresponding phases, indicating an almost perfect correlation with the ideal regression line ($y = x$). Error distribution analysis further confirmed that over 85% of ANN–BBO predictions

Output	Proposed model	Data set	Metrics				95% confidence intervals [95% CI]	
			MAE*	NSE	VAF (%)	A10-index (%)	R^2 [95% CI]	RMSE [95% CI]
Y_1: Compressive strength	ANN	Train	0.1473	0.9171	91.9012	44.78	0.933[0.909, 0.955]	4.36[3.58, 5.08]
		Validate	0.1000	0.8547	86.4957	65.22	0.905[0.829, 0.959]	3.89[2.64, 5.06]
		Test	0.1850	0.9010	91.1665	26.08	0.918[0.844, 0.960]	4.47[3.51, 5.32]
		All	0.1486	0.9123	91.6403	44.44	0.930[0.909, 0.950]	4.40[3.74, 5.00]
	ANN–BBO	Train	0.0522	0.9915	99.1561	85.07	0.992[0.987, 0.995]	1.38[1.06, 1.73]
		Validate	0.0277	0.9895	98.9570	95.65	0.990[0.979, 0.996]	1.02[0.68, 1.35]
		Test	0.0501	0.9902	99.0193	86.96	0.991[0.981, 0.997]	1.44[0.81, 2.04]
		All	0.0490	0.9915	99.1548	86.67	0.992[0.988, 0.995]	1.36[1.11, 1.66]
Y_2: Flow value	ANN	Train	0.0226	0.9214	92.1574	98.50	0.921[0.888, 0.949]	0.71[0.59, 0.83]
		Validate	0.0249	0.9092	91.0000	91.30	0.911[0.794, 0.975]	0.85[0.43, 1.21]
		Test	0.0151	0.9113	91.1644	95.65	0.918[0.741, 0.991]	0.78[0.32, 1.32]
		All	0.0223	0.9180	91.8287	97.22	0.919[0.888, 0.919]	0.74[0.62, 0.86]
	ANN–BBO	Train	0.0107	0.9844	98.4403	100.00	0.985[0.981, 0.990]	0.31[0.26, 0.35]
		Validate	0.0115	0.9902	99.0445	100.00	0.991[0.982, 0.995]	0.29[0.23, 0.34]
		Test	0.0110	0.9817	98.1931	100.00	0.982[0.963, 0.992]	0.37[0.28, 0.47]
		All	0.0112	0.9849	98.4930	100.00	0.985[0.981, 0.989]	0.32[0.28, 0.36]

Table 7. Comparing the evaluation criteria. *The unit of MAE and RMSE metrics is megapascal (MPa) for compressive strength and centimeter (cm) for flow value.

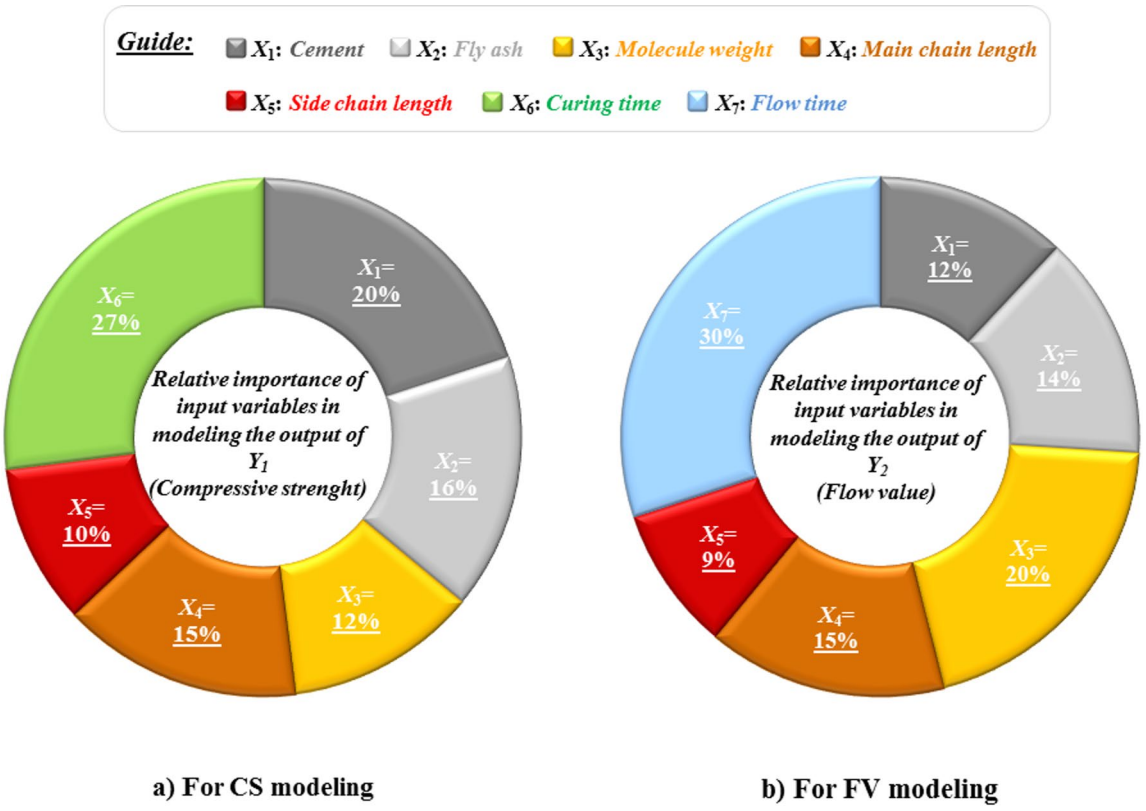


Fig. 13. Importance of input variables in modeling the (**a**) CS and (**b**) FV.

for CS and 100% for FV fell within narrow deviation intervals of [-10%, 10%], demonstrating the hybrid model's performance accuracy and reliability.

Analysis of performance metrics across all data subsets revealed significant enhancements in prediction accuracy and consistency when using the ANN–BBO approach. Specifically, metrics including mean absolute error (MAE), root mean square error (RMSE), Nash–Sutcliffe efficiency (NSE), variance accounted for (VAF), and A10-index confirmed reductions in error magnitude and improvements in model robustness, with ANN–

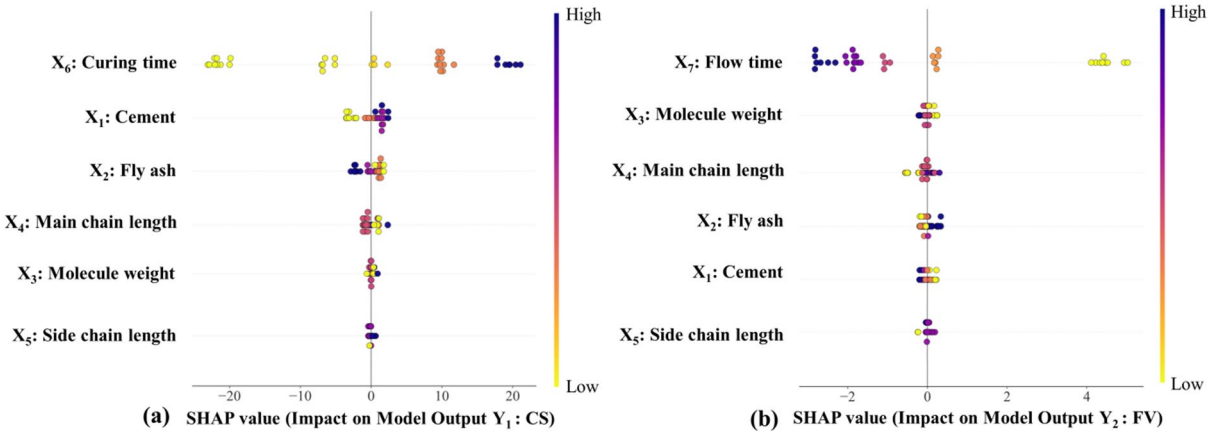


Fig. 14. SHAP summary plot for (**a**) CS and (**b**) FV.

BBO reducing prediction errors by more than 60% in certain metrics compared to the standalone ANN. These results validate the efficacy of integrating metaheuristic optimization with neural network models, particularly for complex cementitious systems involving multiple interacting variables.

The ANN–BBO hybrid methodology offers a more precise, stable, and generalizable predictive framework for modeling the performance of fly ash-based cementitious mortars containing HRWRAs with varied molecular structures. This modeling technique facilitates the development of sustainable, high-performance construction materials and serves as a powerful tool for optimizing admixture design within modern concrete technology.

This study demonstrates that ANN–BBO surrogates can reliably predict compressive strength and flow of FA-incorporated mortars while explicitly accounting for the molecular architecture of PCE-type HRWRAs. In practice, these models enable producers to meet target performance with lower binder demand and/or higher FA replacement, thereby reducing clinker intensity and embodied CO_2. Because RMSE directly translates into reliability-based safety margins, the observed error reductions versus a baseline ANN indicate a smaller overdesign requirement, which helps curb unnecessary cement additions during mix qualification and production. The workflow is straightforward to integrate into ready-mix and precast QC, supporting data-driven proportioning, fewer re-tests/re-tempering events, and more consistent field performance—all aligned with sustainable concrete production.

Future works

Our ANN–BBO surrogates accurately predict compressive strength and flow of FA-incorporated mortars while encoding HRWRA molecular architecture, enabling target performance with less cement or higher FA, thereby reducing clinker intensity and embodied CO_2. Because RMSE reductions map directly to smaller reliability-based safety margins, the hybrid model curtails overdesign and unnecessary binder additions in production. Going forward, we will broaden the dataset across cements (CEM I/II, LC[3]), wider FA ranges, additional HRWRA families, and multi-lab sources to assess domain shift. We will benchmark alternative learners (gradient-boosted trees, GPR, transformer-style models) with Bayesian/evolutionary hyperparameter search, ensembles, and multi-objective optimization of strength–flow–CO_2–cost. Finally, we will quantify uncertainty and calibration (e.g., conformal intervals), couple the surrogate with LCA/LCCA, and conduct field trials with active-learning–guided experiments to accelerate deployment.

While the present dataset centers on Class F FA, the proposed ANN/ANN–BBO framework is SCM-agnostic provided the inputs include descriptors that capture the SCM's physicochemical action. As FA supply tightens, the model can be extended to blends and FA-lean/FA-free systems (e.g., slag, metakaolin/calcined clay, limestone, ground glass pozzolan) by replacing X_2 (Fly ash) with a feature vector of SCM properties (e.g., oxide composition, glass content, fineness/PSD, LOI, specific surface, morphology). We outline a practical transfer-learning route: (i) pretrain on the FA corpus, (ii) add new SCM data, (iii) fine-tune with domain-adaptation regularization and monitor drift. For deployment, we recommend domain-shift checks (error/feature monitors) and light local recalibration when the SCM set changes. The surrogate also remains useful for legacy FA concretes and reclaimed/beneficiated ashes where they are available. In future work we will build a multi-SCM surrogate with SCM embeddings and active-learning to target under-sampled regions of the mix space, enabling performance–carbon optimization regardless of the specific SCM portfolio.

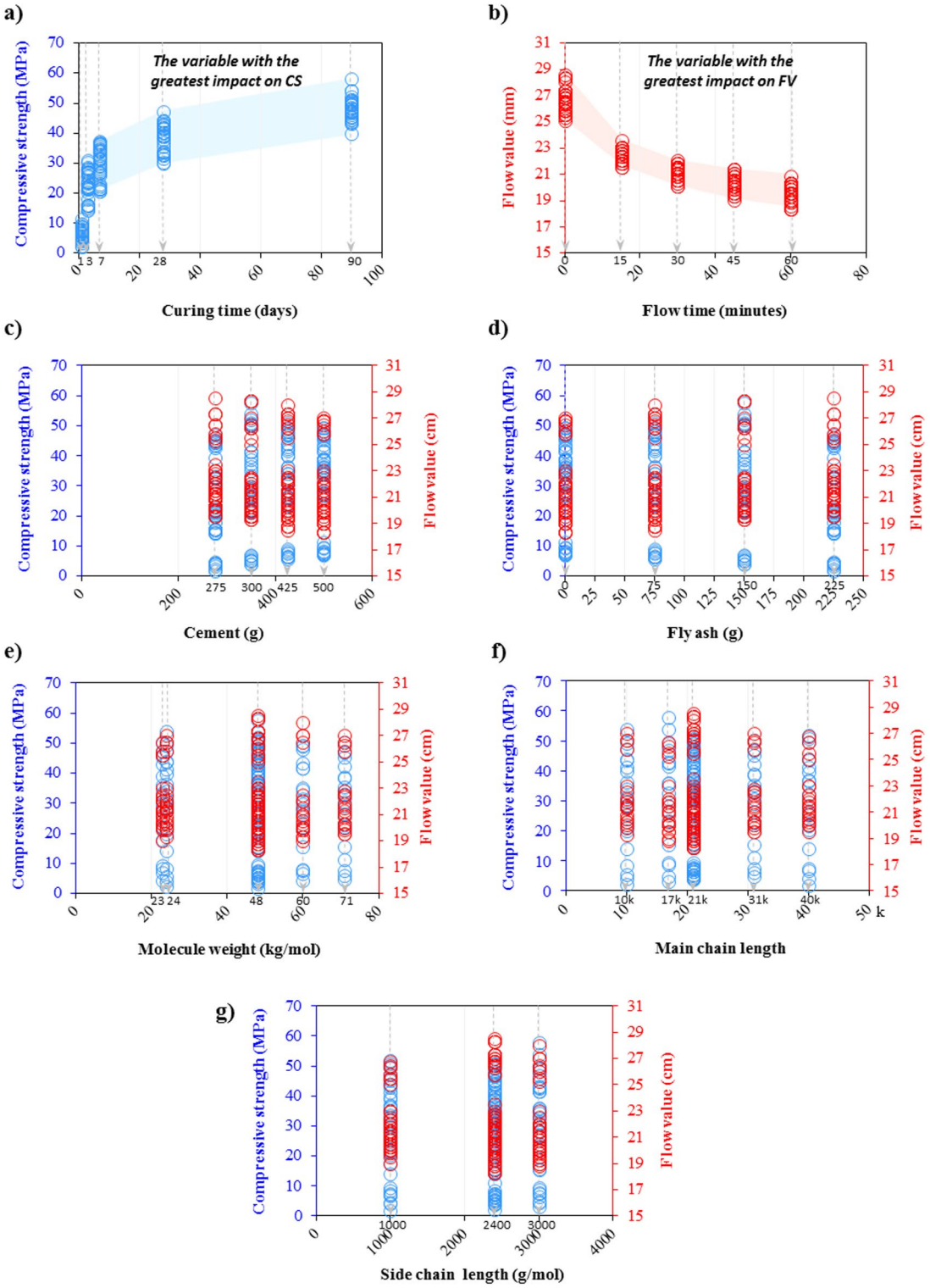


Fig. 15. Effect of variables used in modeling on CS and FV outputs.

Data availability

The database used in this study is available in the supplementary file.

Received: 13 August 2025; Accepted: 15 December 2025
Published online: 01 February 2026

References

1. Mansor, A. M., Borg, R. P., M Hamed, A. M., Gadeem, M. M. & Saeed, M. M. The effects of water-cement ratio and chemical admixtures on the workability of concrete. *IOP Conf. Ser. Mater. Sci. Eng.* **442** (2018).
2. Cavusoglu, I., Yilmaz, E. & Yilmaz, A. O. Additivity effect on properties of cemented coal fly ash backfill containing water-reducing admixtures. *Constr. Build. Mater.* **267**, 121021 (2021).
3. Wang, X. Y. Effect of fly ash on properties evolution of cement based materials. *Constr. Build. Mater.* **69**, 32–40 (2014).
4. Ahmed, O., Ahmad, S. & Adekunle, S. K. Carbon dioxide sequestration in cementitious materials: A review of techniques, material performance, and environmental impact. *J. CO2 Util.* **83**, 102812 (2024).
5. Yang, K. H., Jung, Y. B., Cho, M. S. & Tae, S. H. Effect of supplementary cementitious materials on reduction of CO2 emissions from concrete. *J. Clean. Prod.* **103**, 774–783 (2015).
6. Roshani, M. M., Kargar, S. H., Farhangi, V. & Karakouzian, M. Predicting the effect of fly ash on concrete's mechanical properties by ann. *Sustain.* **13**, 1–16 (2021).
7. Liu, X., Wu, P., Liu, X., Zhang, Z. & Ai, X. The utilization of carbonated steel slag as a supplementary cementitious material in cement. *Materials (Basel).* **17** (2024).
8. Snellings, R., Suraneni, P. & Skibsted, J. Future and emerging supplementary cementitious materials. *Cem. Concr. Res.* **171** (2023).
9. Danish, A., Torres, A. S., Moro, C. & Salim, M. U. Hope or hype? Evaluating the environmental footprint of reclaimed fly ash in geopolymer production. *Resour. Conserv. Recycl.* **205**, 107564 (2024).
10. O'Brien, K. R., Ménaché, J. & O'Moore, L. M. Impact of fly ash content and fly ash transportation distance on embodied greenhouse gas emissions and water consumption in concrete. *Int. J. Life Cycle Assess.* **14**, 621–629 (2009).
11. Bhagath Singh, G. V. P. & Durga Prasad, V. Environmental impact of concrete containing high volume fly ash and ground granulated blast furnace slag. *J. Clean. Prod.* **448**, 141729 (2024).
12. Orozco, C., Babel, S., Tangtermsirikul, S. & Sugiyama, T. Comparison of environmental impacts of fly ash and slag as cement replacement materials for mass concrete and the impact of transportation. *Sustain. Mater. Technol.* **39**, e00796 (2024).
13. Özen, S., Altun, M. G., Mardani-Aghabaglou, A. & Ramyar, K. Multi-effect of superplasticisers main and side-chain length on cementitious systems with fly ash. *Mag. Concr. Res.* **74**, 727–739 (2022).
14. Altun, M. G., Özen, S., Mardani, A. & Ramyar, K. Flowability, strength, and water absorption of mortars containing fly ash and WRA having varying main chain lengths. *J. Mater. Civ. Eng.* **35** (2023).
15. Huang, F. et al. Influence of fly ash on the long-term performance of low water-to-binder ratio paste containing silica fume in water curing and drying curing regimes. *J. Build. Eng.* **104**, 112255 (2025).
16. Nath, P., Sarker, P. K. & Biswas, W. K. Effect of fly ash on the service life, carbon footprint and embodied energy of high strength concrete in the marine environment. *Energy Build.* **158**, 1694–1702 (2018).
17. Tosti, L., van Zomeren, A., Pels, J. R. & Comans, R. N. J. Technical and environmental performance of lower carbon footprint cement mortars containing biomass fly ash as a secondary cementitious material. *Resour. Conserv. Recycl.* **134**, 25–33 (2018).
18. Laxman Kudva, P., Nayak, G., Shetty, K. K. & Sugandhini, H. K. A sustainable approach to designing high volume fly ash concretes. *Mater. Today Proc.* **65**, 1138–1145 (2022).
19. Onyelowe, K. C. et al. Global warming potential-based life cycle assessment and optimization of the compressive strength of fly ash-silica fume concrete; Environmental impact consideration. *Front. Built Environ.* **8**, 1–15 (2022).
20. Amran, M. et al. Fly ash-based eco-efficient concretes: A comprehensive review of the short-term properties. *Materials (Basel).* **14**, 1–41 (2021).
21. Zhao, H., Sun, W., Wu, X. & Gao, B. Sustainable self-compacting concrete containing high-amount industrial by-product fly ash as supplementary cementitious materials. *Environ. Sci. Pollut. Res.* **29**, 3616–3628 (2022).
22. Keulen, A., Yu, Q. L., Zhang, S. & Grünewald, S. Effect of admixture on the pore structure refinement and enhanced performance of alkali-activated fly ash-slag concrete. *Constr. Build. Mater.* **162**, 27–36 (2018).
23. Huang, J. *et al.* Predicting the compressive strength of the cement-fly ash–slag ternary concrete using the firefly algorithm (FA) (2022).
24. Kazemi, R. & Mirjalili, S. An AI-driven approach for modeling the compressive strength of sustainable concrete incorporating waste marble as an industrial by-product. *Sci. Rep.* **14**, 1–29 (2024).
25. Ahmed, M. et al. Artificial neural networks for sustainable development of the construction industry. *Sustainability* **14**, 1–21 (2022).
26. Ahmad, W. *et al.* Application of advanced machine learning approaches to predict the compressive strength of concrete containing supplementary cementitious materials. *Materials (Basel).* **14** (2021).
27. Stegemann, J. A. & Buenfeld, N. R. Prediction of unconfined compressive strength of cement paste containing industrial wastes. *Waste Manag.* **23**, 321–332 (2003).
28. Topçu, I. B. & Saridemir, M. Prediction of compressive strength of concrete containing fly ash using artificial neural networks and fuzzy logic. *Comput. Mater. Sci.* **41**, 305–311 (2008).
29. Bansal, A. K., Sangtani, V. S., Dadheech, P., Aneja, N. & Yahya, U. Biogeography-based optimization of artificial neural network (BBO-ANN) for solar radiation forecasting. *Appl. Artif. Intell.* **37** (2023).
30. Eskandari-Naddaf, H. & Kazemi, R. ANN prediction of cement mortar compressive strength, influence of cement strength class. *Constr. Build. Mater.* **138**, 1–11 (2017).
31. Mohammed, A. et al. ANN, M5P-tree and nonlinear regression approaches with statistical evaluations to predict the compressive strength of cement-based mortar modified with fly ash. *J. Mater. Res. Technol.* **9**, 12416–12427 (2020).
32. Marani, A. & Nehdi, M. L. Machine learning prediction of compressive strength for phase change materials integrated cementitious composites. *Constr. Build. Mater.* **265**, 120286 (2020).
33. Salami, B. A., Olayiwola, T., Oyehan, T. A. & Raji, I. A. Data-driven model for ternary-blend concrete compressive strength prediction using machine learning approach. *Constr. Build. Mater.* **301**, 124152 (2021).
34. Khursheed, S., Jagan, J., Samui, P. & Kumar, S. Compressive strength prediction of fly ash concrete by using machine learning techniques. *Innov. Infrastruct. Solut.* **6**, 1–21 (2021).
35. Sevim, U. K., Bilgic, H. H., Cansiz, O. F., Ozturk, M. & Atis, C. D. Compressive strength prediction models for cementitious composites with fly ash using machine learning techniques. *Constr. Build. Mater.* **271**, 121584 (2021).
36. Armaghani, D. J. & Asteris, P. G. *A Comparative Study of ANN and ANFIS Models for the Prediction of Cement-Based Mortar Materials Compressive Strength*. *Neural Computing and Applications* vol. 33 (Springer London, 2021).
37. Zhang, J., Niu, W., Yang, Y., Hou, D. & Dong, B. Machine learning prediction models for compressive strength of calcined sludge-cement composites. *Constr. Build. Mater.* **346**, 128442 (2022).
38. Al-hashem, M. N. *et al.* Predicting the compressive strength of concrete containing. *Materials (Basel).* (2022).
39. Mahajan, L. & Bhagat, S. Machine learning approaches for predicting compressive strength of concrete with fly ash admixture. *Res. Eng. Struct. Mater.* **9**, 431–456 (2023).
40. Zeng, Z. et al. Accurate prediction of concrete compressive strength based on explainable features using deep learning. *Constr. Build. Mater.* **329**, 127082 (2022).
41. Liang, W. et al. Mixed artificial intelligence models for compressive strength prediction and analysis of fly ash concrete. *Adv. Eng. Softw.* **185**, 103532 (2023).
42. Omer, B. et al. Exploring the potential of soft computing for predicting compressive strength and slump flow diameter in fly ash-modified self-compacting concrete. *Arch. Civ. Mech. Eng.* **24**, 1–22 (2024).

43. Sathiparan, N., Jeyananthan, P. & Subramaniam, D. N. Predicting compressive strength of pervious concrete with fly ash: a machine learning approach and analysis of fly ash compositional influence. *Multiscale Multidiscip. Model. Exp. Des.* **7**, 5651–5671 (2024).
44. Bakhoum, E. S., Amir, A., Osama, F. & Adel, M. Prediction model for the compressive strength of green concrete using cement kiln dust and fly ash. *Sci. Rep.* **13**, 1–14 (2023).
45. Asteris, P. G., Koopialipoor, M., Armaghani, D. J., Kotsonis, E. A. & Lourenço, P. B. *Prediction of Cement-Based Mortars Compressive Strength Using Machine Learning Techniques. Neural Computing and Applications* vol. 33 (Springer London, 2021).
46. Rathnayaka, M. et al. Machine learning approaches to predict compressive strength of fly ash-based geopolymer concrete: A comprehensive review. *Constr. Build. Mater.* **419**, 135519 (2024).
47. Ghafor, K. *et al.* Computing models to predict the compressive strength of engineered cementitious composites (ECC) at various mix proportions. *Sustainability* **14** (2022).
48. Hameed, M. M., Abed, M. A., Al-Ansari, N. & Alomar, M. K. Predicting compressive strength of concrete containing industrial waste materials: Novel and hybrid machine learning model. *Adv. Civ. Eng.* **2022** (2022).
49. Ahmad, A. et al. Prediction of compressive strength of fly ash based concrete using individual and ensemble algorithm. *Materials (Basel).* **14**, 1–21 (2021).
50. Shanmugasundaram, N., Praveenkumar, S., Gayathiri, K. & Divya, S. Prediction on compressive strength of Engineered Cementitious composites using Machine learning approach. *Constr. Build. Mater.* **342**, 127933 (2022).
51. Altun, M. G., Özen, S. & Mardani-Aghabaglou, A. Effect of side chain length change of polycarboxylate-ether–based high-range water-reducing admixture on properties of cementitious systems containing fly ash. *J. Mater. Civ. Eng.* **33**, 1–11 (2021).
52. McCulloch, W. S. & Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **5**, 115–133 (1943).
53. Panchal, G., Ganatra, A., Kosta, Y. P. & Panchal, D. Behaviour analysis of multilayer perceptronswith multiple hidden neurons and hidden layers. *Int. J. Comput. Theory Eng.* 332–337 (2011). https://doi.org/10.7763/ijcte.2011.v3.328.
54. Uzair, M. & Jamil, N. Effects of hidden layers on the efficiency of neural networks. in *Proceedings—2020 23rd IEEE International Multi-Topic Conference, INMIC 2020* 1–6 (2020). https://doi.org/10.1109/INMIC50486.2020.9318195.
55. Simon, D. Biogeography-based optimization. *IEEE Trans. Evol. Comput.* **12**, 702–713 (2008).
56. Mirjalili, S., Mirjalili, S. M. & Lewis, A. Let a biogeography-based optimizer train your Multi-Layer Perceptron. *Inf. Sci. (Ny)* **269**, 188–209 (2014).
57. Ma, H., Simon, D., Siarry, P., Yang, Z. & Fei, M. Biogeography-based optimization: A 10-year review. *IEEE Trans. Emerg. Top. Comput. Intell.* **1**, 391–407 (2017). https://doi.org/10.1109/TETCI.2017.2739124
58. Raschka, S. Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning (2018).
59. Mehlig, B. Machine learning with neural networks. *Mach. Learn. Neural Netw.* https://doi.org/10.1017/9781108860604 (2021).
60. El-Bakry, M. Y. Feed forward neural networks modeling for K-P interactions. *Chaos Solitons Fractals* **18**, 995–1000 (2003).
61. Haykin, S. *Neural Networks and Learning Machines Third Edition. Pearson Education India* vol. 3 (2009).
62. Adeli, H. & Seon Park, H. Counterpropagation neural networks in structural engineering. *J. Struct. Eng.* **121**, 1205–1212 (1995).
63. Lundberg, S. M. & Lee, S. I. A unified approach to interpreting model predictions. in *Advances in Neural Information Processing Systems* vols 2017-Decem 4766–4775 (2017).

Acknowledgements
During the preparation of the current study, the authors have used AI-based tools for enhancing the fluency, and solving the grammatical/typographical issues of the context. After using these tools/services, the authors have reviewed and edited the content, as needed, and take full responsibility for the content of the publication.

Author contributions
N.M and M.U collected the data and prepared the dataset. R.M developed the AI model and run the experiments. The results are analyzed by all authors. All authors contributed to the writing. A.M and S.M supervised the project.

Funding
Open access funding provided by Óbuda University.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-32972-1.

Correspondence and requests for materials should be addressed to S.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026