ENG. Struct. Civ. Eng. 2026, 20(3): 555–579
https://doi.org/10.1007/s11709-026-1277-9

RESEARCH ARTICLE

Data-driven and experimental synergy: Metaheuristic-optimized extreme gradient boosting modeling of self-compacting concrete compressive strength

Divi SAI VARDHAN[a], Akhilendra SHARMA[a], Rahul BISWAS[a*], Akshay BURA[b]

[a] *Department of Applied Mechanics, Visvesvaraya National Institute of Technology, Nagpur 440022, India*
[b] *Department of Civil Engineering, School of Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra 411038, India*
[*] *Corresponding author. E-mail: rahulbiswas@apm.vnit.ac.in*

ABSTRACT This investigation develops a comprehensive framework that integrates machine learning modeling with experimental validation for predicting the compressive strength of self-compacting concrete (SCC). The predictive models were constructed using a consolidated data set from existing literature and validated through independent laboratory experiments on ternary blends incorporating fly ash and silica fume. Experimental results demonstrated that optimized combinations of 30% fly ash and 5%–7.5% silica fume produced superior fresh-state properties, including enhanced flowability and reduced segregation, while achieving substantial improvements in compressive, tensile, and flexural strengths across all curing ages. Advanced ensemble learning techniques were employed by coupling Extreme Gradient Boosting (XGBoost) with three metaheuristic optimization algorithms: Tuna Swarm Optimization (TSO), Coyote Optimization Algorithm, and Giant Trevally Optimization. The XGB-TSO model demonstrated superior predictive performance, achieving coefficient of determination R^2 = 0.9690, root mean square error of 2.15 MPa, weighted mean absolute percentage error of 2.63%, and Nash-Sutcliffe efficiency of 0.9674. Model interpretability analysis using SHapley Additive exPlanations (SHAP) identified concrete age and cement content as the most influential parameters, providing transparent insights into strength development mechanisms. Taylor diagram analysis confirmed statistical robustness through high correlation coefficients and minimal centered root mean square error. A graphical user interface was developed to enable real-time strength prediction from standard mix parameters, facilitating practical implementation. Experimental validation using laboratory-produced SCC specimens demonstrated excellent agreement with model predictions, achieving validation R^2 = 0.9698 and mean absolute error of 1.83 MPa. The strong correlation between predicted and measured values validates the framework's reliability for engineering applications. This research advances concrete technology by providing a validated, interpretable, and deployable tool that combines data-driven modeling with experimental verification, enabling intelligent mix design and informed decision-making in sustainable concrete construction.

KEYWORDS SCC, XGBoost, metaheuristic optimization, TSO, SHAP analysis, graphical user interface

1 Introduction

Concerns over the concrete industry's influence on the environment have grown in recent years due to the fact that Portland cement accounts for over 8% of CO_2 emissions worldwide and uses around 13% of the world's industrial energy [1]. A significant challenge for civil engineering professionals is to implement projects that align with natural ecosystems through sustainable development, utilizing high-performance, environmentally friendly materials that are cost-effective. For concrete, which is the most common building material, it is important to find alternatives to cement that are less expensive. The use of supplementary cementitious

Article history: Received Sep 1, 2025; Accepted Oct 17, 2025

materials (SCMs), such as fly ash (FA), blast furnace slag, silica fume (SF), metakaolin, rice husk ash, etc., has been shown by numerous researchers in recent years to improve the various properties of concrete in both its fresh and hardened states and to reduce the increase in construction costs [2]. To conserve primary materials, it is essential to reduce the utilization of cement, aggregates, and water, while also incorporating waste materials and industrial by-products. These measures are crucial for diminishing the reliance on non-renewable materials and mitigating environmental damage [3–6]. In this regard, the utilization of SCMs is a prevalent choice, as these materials have the potential to enhance the properties of concrete while simultaneously decreasing its cement consumption [7–10]. In reality, some researchers have developed concretes that achieve compressive strength (CS) ranging from 60 to 120 MPa by substituting Portland cement for up to 75% of SCMs [9,11,12]. Since the late 1980s, researchers have investigated the feasibility of using high volumes of fly ash (HVFA) in place of Portland cement in concrete [13]. Although these levels can go as high as 80% or more [14,15], the optimal mechanical performances are typically achieved at 40 to 60% replacement. Along the same lines, numerous researchers have proved that high volumes of fly ash-self-compacting concrete (HVFA-SCC) is one of the most valuable waste material applications [16–18].

The research and manufacturing of self-compacting concrete (SCC) commenced in Japan in the late 1980s to address the deficiencies of normal vibrated concrete. SCC is a type of concrete that can be positioned and compacted only by its own weight, with minimal to no vibrational assistance, while maintaining sufficient cohesion to prevent segregation or bleeding during handling. It is utilized to promote and guarantee adequate filling of the formwork and superior structural performance in confined spaces and substantially reinforced structural elements. The SCC has rapidly achieved global recognition due to its significant advantages in enhanced mechanical and durability qualities. The numerous advantages associated with the utilization of SCC include a reduction in construction time and labor costs, the elimination of vibration requirements, a decrease in noise pollution, enhanced compact ability even in densely packed structural elements, and ultimately, improved construction quality that ensures superior structural performance [19]. SCC mix often comprises a large quantity of powdered components along with a superplasticizer and viscosity-modifying admixtures. To balance out the high powder concentration, mineral admixtures like slag, FA, etc. are commonly used [20]. In light of its superior workability, effective utilization of industrial by-products, and reduced cement content, HVFA SCC has emerged as a promising candidate among these alternatives. The incorporation of FA enhances rheological properties and mitigates thermally induced cracking in concrete by decreasing the overall heat of hydration [21]. The slump of the concrete mix could be increased without an increase in cost by using mineral admixtures such as FA. Sadrmomtazi et al. [22] demonstrated that a substitution of up to 30% FA in the concrete mix can enhance the slump value by 45%. As per Dave et al. [23], substituting 30% and 50% of ordinary Portland cement (OPC) with FA led to increases in standard consistency of 1.6% and 3.1%, respectively, attributable to the enhanced fineness of the blended mixture. Incorporating FA further reduces the requirement for SCC's cement, fine fillers, and sand [24]. According to Siddique [25], the slump values in HVFA concrete went up by 20, 25, and 35 mm compared to the control group when substitution fractions of 45%, 50%, and 55% FA were used. The inclusion of HVFA is claimed to enhance CS from 28 and 120 d of curing, hence affirming the persistent pozzolanic activity of FA over time [26]. In both the control concrete and HVFA concrete, the CS improved with time as a result of the ongoing hydration of OPC and the enhanced pozzolanic reaction of FA, as reported by Mugahed Amran et al. [27]. Even with these benefits, adding large amounts of FA to concrete can affect its early-age strength, particularly CS, a crucial performance indicator. HVFA concrete has lower strength than OPC concrete because of the low pozzolanic reaction and dilution effect of FA at an early age [2,12,13]. When 50% OPC was replaced with FA, the concrete CS dropped by 57.34%, 41.61%, 37.50%, 29.42%, 28.32%, and 35.52% at 3, 14, 28, 56, and 91 d, respectively [27]. The difficulty of striking a balance between workability and strength, which are frequently impacted by a number of variables, such as FA quality, mix proportions, and curing conditions complicates this task even more. Researchers have explored the incorporation of SF, an exceptionally reactive pozzolanic ingredient, into HVFA-SCC mixtures to mitigate these limitations. The inclusion of SF adversely affects the properties of fresh concrete, specifically leading to a reduction in slump values. Due to the large surface area and adsorption, excessive SF content in concrete mixes may limit fluidity and increase superplasticizer (SP) demand to maintain workability [28]. Incorporating SF with FA and OPC in HVFA-SCC mixtures enhances the strength and permeability resistance of the concrete [29,30]. Wongkeo et al. [31] investigated the effects of incorporating 50%, 60%, and 70% FA and SF as substitutes for Portland cement on the CS and chloride resistance of SCC. FA and SF were found to improve the CS and chloride resistance of SCC when used at high volumes as replacements for Portland cement. The combination of FA and SF in a ternary blend enhances concrete properties, causing SCC more cost-effective and environmentally sustainable [29,32]. The

use of SF greatly enhances the early-age strength and matrix density of concrete, making it more suitable for structural applications. To investigate the strength and durability of SCC mixes made with 10% SF and a 0.28 *w*/*b* ratio, Yazıcı [33] substituted cement with FA between 30% to 60%. The studies showed that SF enhances fresh and hardened properties of HVFA concrete. The combined effects of FA and SF complicate the estimation of mechanical strength, indicating a need for further research. Therefore, the findings of the research we conducted to produce and evaluate the mechanical properties of SCCs consisting of high-volume class F FA and SF are presented in this publication.

Accurately predicting the mechanical properties of SCC, especially its CS, has become more important due to the growing demand for high-performance concrete in contemporary, intricate construction projects. To uphold quality standards and guarantee the longevity and safety of structures, accurate prediction of these attributes is crucial. Application of soft computing and machine learning (ML) techniques to real-world problems has grown substantially because of their exceptional capacity to detect nonlinear and unclear correlations between a data set's input and output variables [34–43]. ML techniques such as artificial neural networks (ANN), decision trees (DT), support vector machines (SVM), genetic engineering programming and deep learning are used to estimate various civil engineering problems [43–46]. One such application is to improve the SCC properties using ML based predictive models. The use of ML for predicting the CS of SCC presents transformative potential for the construction industry ML techniques can be readily applied to quantify the fresh and mechanical properties of SCC using formulated mixture ratios, focusing on various strength predictions [47,48]. Techniques such as neural networks, SVM, and fuzzy logic have demonstrated significant potential for handling complex data sets and providing accurate predictions [49]. Many researchers have proposed multiple ML approaches to estimate the characteristics of cement and concrete, yielding satisfactory outcomes. SVM was used to predict fresh and mechanical properties of FA blended SCC [50]. ML models using Extreme Gradient Boosting (XGBoost) and ANN informational models have been found to accurately predicted the CS of waste marble powder concrete with a high degree of reliability (R^2 > 0.97) [51]. CS of ternary blended SCC with marble powder and rice husk ash has been predicted using multiple ML models [52]. Although the aforementioned ML paradigms achieved satisfactory results in predicting the CS of concrete and cement, they have multiple limitations such as uncertainty, time consumption, and local minima trapping issues. Thus, it is necessary to develop a high-performance ML paradigm, especially for estimating the CS of SCC. Civil engineering optimization commonly uses meta-heuristic algorithms to optimize data mining models for ML algorithm hyperparameters. Thus, training ML models requires choosing the right meta-heuristic algorithm. Finding effective solutions and improving the precision of regression algorithms have been widely achieved in recent years through the use of optimization algorithms based on neural networks [53]. While this study adopts a data-driven approach via metaheuristic-optimized XGBoost, it is important to note that physics-informed frameworks such as the deep energy method and its neural operator extension visual inference and neural network optimization [54,55] provide powerful alternatives for forward and inverse problems, uncertainty quantification, and inverse design. Integration of such methods in future work could enhance the generalizability of SCC strength predictions.

SCC strength predictions have rarely used XGBoost algorithm. In the present study, HVFA-SCC was developed and evaluated. The rheological properties were examined in accordance with the European guidelines for SCC, and the results were found to be within the prescribed acceptable limits. Subsequently, mechanical performance was assessed through CS, split tensile strength, and flexural strength tests. In addition to experimental evaluation, Artificial Intelligence and Machine Learning (AIML) models, namely Adaptive Boosting (AdaBoost), Gradient Boosting (GB), Categorical Boosting (CatBoost) and XGBoost, were employed to predict the mechanical behavior of HVFA-SCC. These models were further enhanced using metaheuristic optimization algorithms, including Giant Trevally Optimization (GTO), Coyote Optimization Algorithm (COA), and Tuna Swarm Optimization (TSO). A graphical user interface (GUI) was also developed to facilitate user interaction and prediction. The outcomes obtained from the AIML models and optimization approaches were compared with the corresponding experimental results to validate the predictive accuracy and robustness of the proposed models. The novelty of this research lies in the integration of newly emerging metaheuristic optimizers with ensemble learning models for the predictive modeling of HVFA-SCC properties, coupled with the development of a user-friendly GUI tool, an approach not widely explored in the context of sustainable SCC systems.

2 Research significance

The framework developed in this study establishes a direct link between advanced ML methodologies and practical engineering practice by providing a robust, deployable tool for the design and quality control of SCC. Metaheuristically optimized XGBoost models were rigorously trained and validated using extensive

laboratory data derived from HVFA and SF ternary blends. A user-oriented graphical interface allows practitioners to specify conventional mix parameters, cement content, SCMs, water, admixtures, and age, and to obtain accurate CS estimates in real time without requiring specialized computational expertise. Model interpretability is ensured through SHapley Additive exPlanations (SHAP) and Taylor diagram analysis, which elucidate the contributions of key variables such as concrete age and cement dosage to strength development. The computational efficiency of the optimized XGBoost algorithm facilitates rapid prediction on standard computing hardware, enabling seamless integration into existing quality management systems and mobile applications for field deployment. By combining empirical validation, algorithmic innovation, and an accessible deployment platform, this end-to-end pipeline advances SCC mix design from a predominantly empirical process to one guided by data-driven insight, thereby promoting enhanced sustainability, cost effectiveness, and consistency in concrete construction worldwide. Figure 1 represents the flowchart of the methodology.

3 Material properties

3.1 Cementitious materials

In this study, OPC 53-grade cement from UltraTech was used as shown in Fig. 2, which is confirming to IS 12269-1987 [56]. The physical tests on cement were done based on IS 4031-1996 [57]. FA which is sampled from Khaperkheda thermal power station, Maharashtra as shown in Fig. 3, was used in this study. The undensified SF with product name ELKEM MICROSILICA ® 920U as shown in Fig. 4, was used in the current study is

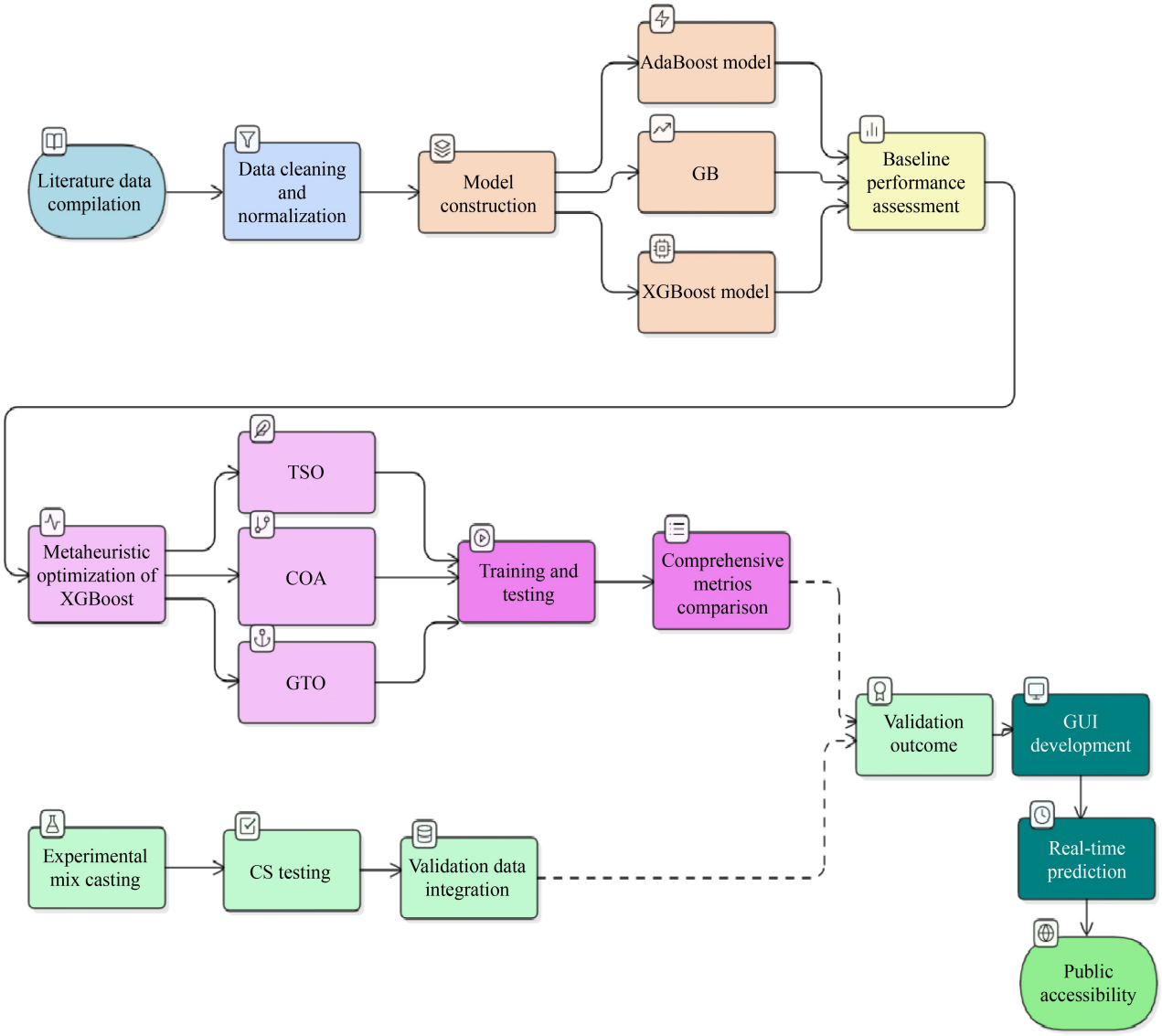

Fig. 1 Flowchart of methodology.

another pozzolanic material which was brought from Elkem south Asia Pvt Ltd. The sample satisfied the requirements prescribed in IS 15388-2003 [58]. The chemical tests like X-ray fluorescence (XRF) and X-ray diffraction (XRD) were conducted for FA, cement, and SF and the results were confirming to the IS 3812-2013 [59].

Fig. 2 OPC 53 cement.

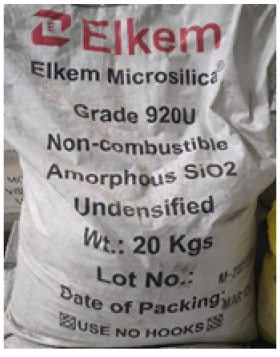

Fig. 3 Undensified SF.

3.2 Aggregates

The natural river sand and coarse aggregate which were procured from the local vendor were used for the study. Followed the standards given in IS 2386-2021 [60] specific gravity tests alongside water absorption assessments on both coarse and fine aggregates were assessed. The coarse aggregate and fine aggregate exhibited a specific gravity measurement of 2.83 and 2.65, respectively. The coarse aggregate and fine aggregate displayed a water absorption rate of 0.8% and 1.1%, respectively. Sieve analysis done on the fine aggregate and coarse aggregate as per IS 383-2016 [61], and the results were confirming to Zone I and 12.5 mm well graded aggregates, respectively. Figure 5 represents the combined grading of the aggregates.

3.3 Chemical admixture

This paper has used the superplasticizer HS PRO

Fig. 4 FA.

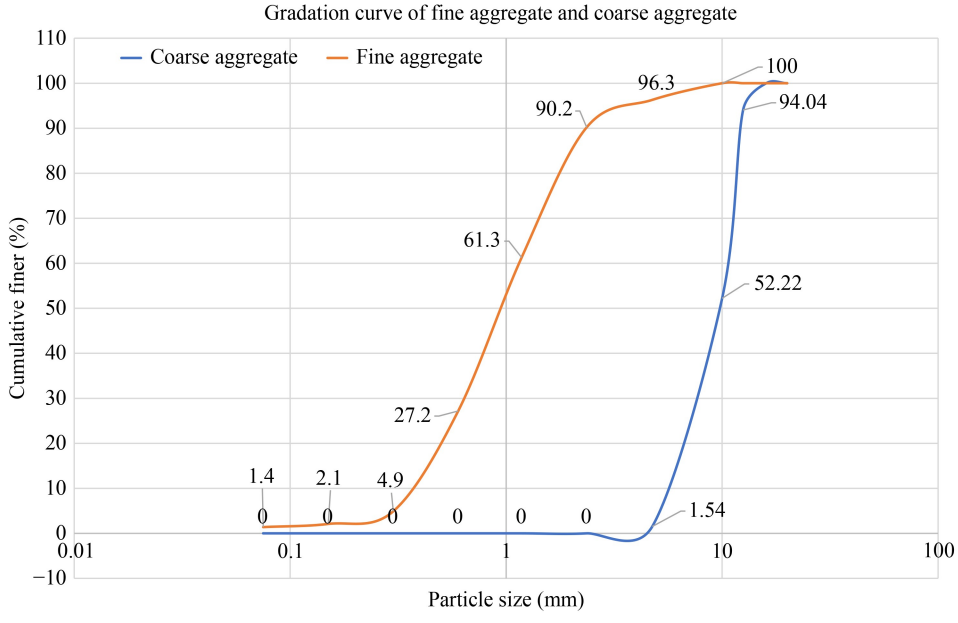

Fig. 5 Combined grading of the aggregates.

HYPERPLAST RT 2000 NSG of Berger Paints to enhance the performance of SCC and its specific gravity is 1.08. It is a highly technical admixture specially designed to increase workability and reduce water requirement, thus helping to realize a highly flowable and at the same time cohesive concrete mix. When used in SCC, HS PRO HYPERPLAST RT 2000 NSG was found to ensure easier placement and compaction, even in highly reinforced structural members, with a reduced potential of segregation or bleeding. Its agreement with numerous cementitious materials and the capacity to hold slump throughout the long-term makes it a substantial asset to the projects where thick and high-quality concrete is needed with a superior finish. The product is in conformity with the appropriate industry requirement, which also helps in enhancing the mechanics properties as well as the durability of concrete.

4 Tests on materials

4.1 X-ray diffraction test

A popular non-destructive method for examining a material's crystallographic structure is XRD test. It is used to measure lattice dimensions, identify distinct phases, and comprehend the overall structure of a crystal by revealing the arrangement of atoms within it. This approach is widely used in disciplines like metallurgy, geology, chemistry, and materials science. The test was conducted using the Rigaku smart laboratory apparatus as shown in Fig. 6 and usually test will be done in 30 to 45 min for one sample. The material is prepared as a fine powder or a thin solid sample. Powder is commonly used because it provides a more uniform diffraction pattern. X-rays are generated by directing a high-energy electron beam at a metal target usually copper or molybdenum. The interaction between the electrons and the metal produces X-rays of a specific wavelength. When X-rays are directed at a crystalline sample in XRD, Bragg's Law states that the atomic planes will diffract the X-rays. The intensity of these diffracted rays is recorded by a detector at different angles, creating a distinct peak pattern that represents the crystal structure of the material. The atomic arrangement, unit cell dimensions, crystallinity, lattice parameters, crystallite size, and internal strain can all be determined in detail by examining the locations and intensities of these peaks. To determine the material and its phase composition, this pattern is frequently compared with reference databases such as the International Centre for Diffraction Data. The output is shown in Fig. 7. XRD traces of FA, OPC 53 cement and undensified SF are showing in Figs. 8–10, respectively, by taking the references from research papers and based on intensity we will give the naming to the different peaks.

Fig. 6 Rigaku smart laboratory setup.

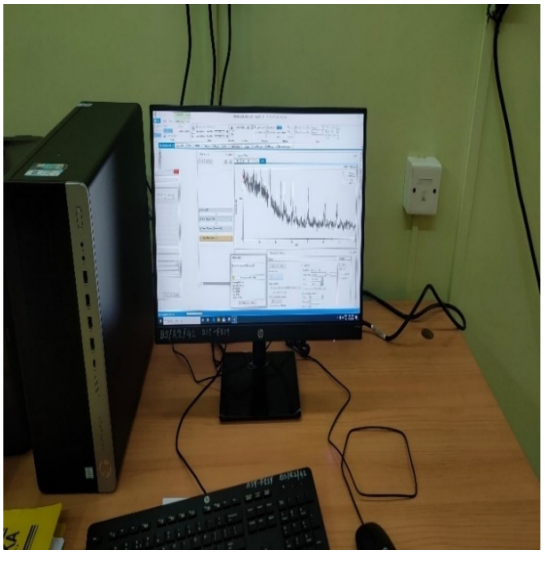

Fig. 7 XRD output in OPC.

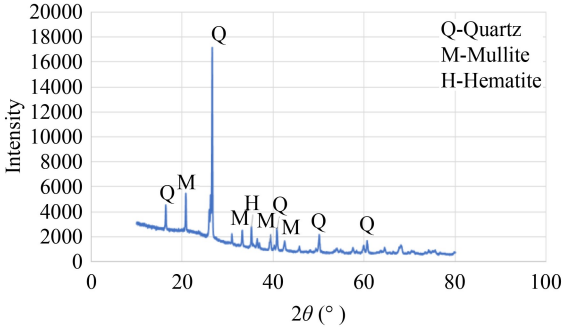

Fig. 8 XRD results of FA.

4.2 X-ray fluorescence analysis

XRF test is a non-destructive method to analyze the

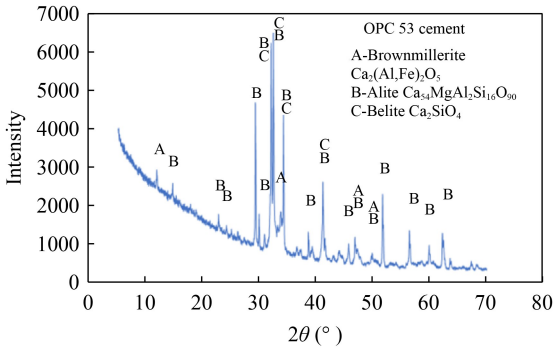

Fig. 9 XRD results of OPC 53 cement.

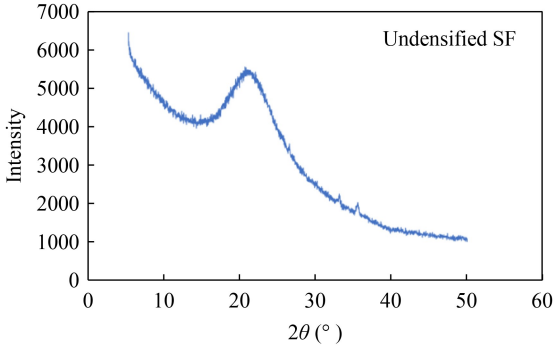

Fig. 10 XRD results of undensified SF.

contents of a material and find out its elemental components. The process begins by exposing high-energy X-rays on a sample which causes the atoms to give off secondary X-rays (fluorescence). X-rays that come from the samples have particular energy levels assigned to specific elements which makes it possible to find and measure the elements involved. The XRF analysis was carried out using Shimadzu EDX7000 XRF machine on the cement, undensified SF and FA samples. Satisfactory results were obtained and matched the requirements in IS 4032-1985 [62], IS 15388-2003 [58], and IS 3812-2013 [59]. The test procedure of conducting XRF is first the powder sample placed in the polypropylene (PP) cap as shown in Fig. 11, then the sample placed in device Shimadzu EDX7000XRF instrument which shown in Fig. 12 and whole XRF setup shown in Fig. 13, then the X-ray source directs a beam of X-rays at the sample, causing atoms in the sample to become excited. When atoms revert to their stable state, they release fluorescent X-rays that are specific to the elements present in the material. These emitted X-rays are detected by the XRF instrument, which measures their energy and intensity. The instrument's software analyses the detected X-ray signals to identify which elements are present and in what quantities as the output as shown in Fig. 14. The results are displayed or printed, showing the elemental composition of the sample. The total test is done in 90 s. Table 1 shows the results of XRF tests for OPC 53 cement, FA, SF and physical test results of materials.

Fig. 11 FA sample placed in PP cap.

Fig. 12 Shimadzu EDX7000 XRF.

Fig. 13 XRF instrument setup.

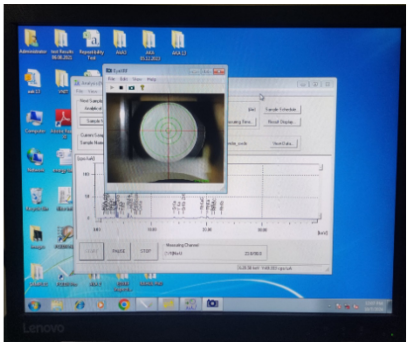

Fig. 14 XRF output in PC.

5 Mixture proportions

The IS 10262-2019 [63] was used for HVFA-SCC mix

Table 1 Chemical properties of different materials form XRF test

Property	Analyte	Measured value for OPC 53	Measured value for FA	Measured value for SF
Chemical property	silica as SiO_2 (%)	27.12	60.871	96.586
	lime as CaO (%)	57.386	0.738	0.512
	alumina as Al_2O_3 (%)	8.012	31.617	–
	sulfate as SO_3 (%)	2.653	–	0.237
	potassium as K_2O (%)	0.484	1.154	0.920
	ferric oxide as Fe_2O_3 (%)	3.552	4.007	1.525
	manganese oxide as MnO (%)	0.060	0.044	0.137
	copper oxide as CuO (%)	0.009	0.009	0.019
	titanium dioxide as TiO_2 (%)	–	1.407	–
Physical property	specific gravity (g/cm^3)	3.15	2.11	2.3
	initial setting time (min)	30	44	36
	final setting time (min)	600	283	218
	specific surface (m^2/g)	0.224	0.369	19.3
	consistency (%)	31	34	33

proportioning and followed the guidelines for typical range of constituents in SCC given by EFNARC-2005 [64]. As several trials were done to find out the optimum SP dosage and correct slump required for flowability properties to meet the acceptance criteria for SCC prescribed in EFNARC. After fixed the optimum SP dosage for control mix i.e. 0% FA mix, the SP dosage was reduced as the FA percentage increasing by referring the research papers and by experience. The mix proportions of all SCC mixes are shown in Table 2. While the control mix did not contain FA, the binary mixes included FA ranging from 0% to 50% and replaced the cement with SF from 2.5% to 10% while maintaining 30% FA as constant for ternary mixes after the binary mix of 30% FA showed significantly good CS and as economical concern 10% and 20% FA mixes for replacing SF were not considered.

6 Test program and procedures

6.1 Flowability tests

Unlike traditional concrete SCC have different type of fresh property tests called flowability tests which represents characteristics of SCC. The tests including slump flow test, T_{500}, J-ring, L-box, and V-funnel test were conducted to check the filling ability, viscosity, flow through confined spaces, passing ability and viscosity of SCC with maximum aggregate size, respectively. Followed the guidelines provided in EFNARC-2005 [64] for the acceptance criteria of flowability properties of SCC for each mix and the results are satisfied. The slump flow test, J-ring test, L-box, and V-funnel test were performed as shown in Figs. 15–18, respectively.

Table 2 Mix proportion of HVFA-SCC

FA (wt.%)	SF (wt.%)	Cement (kg)	Coarse aggregate (kg)	Fine aggregate (kg)	Water (L)	Super plasticizer/cement (wt.%)
0	0	533	829	882	181	1.5
10	0	480	819	872	181	1.4
20	0	427	809	861	181	1.3
30	0	373	799	805	181	1.1
40	0	320	789	840	181	0.8
50	0	267	779	829	181	0.7
30	2.5	360	797	848	181	1.1
30	5	347	795	846	181	1.1
30	7.5	333	795	846	181	0.7
30	10	320	793	844	181	0.5

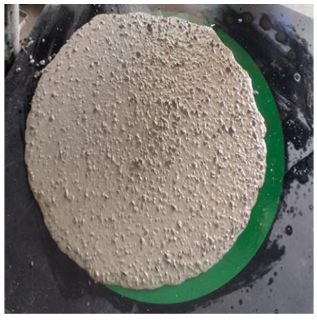

Fig. 15 Slump flow test.

Fig. 16 J-ring test.

Fig. 17 L-box test.

Fig. 18 V-funnel test.

6.2 Strength test

1) The CS test was conducted by performing uniaxial compression testing on cube specimen of 150 mm in CONTROLS AUTOMAX Pro m compression testing machine which have 3000 kN capacity. The test is conducted as per the guidelines given in IS 516-2021 [65]. The samples are tested at 7, 28, 56, and 90 d of curing, minimum three samples were tested for each result.

2) Split tensile strength test was performed on the cylindrical specimens of 150 mm diameter and 300 mm height at 28 d of curing in Compression testing machine by following the guidelines given in IS 516-2021 [65]. The formula for finding split tensile strength is expressed as Eq. (1):

$$\text{Split tensile strength} = \frac{2P}{\pi DL}, \tag{1}$$

where P = maximum applied load on cylinder, L = length of cylinder, D = diameter of cylinder.

3) Flexural strength was performed on the beam sizes of 500 mm length, 100 mm width, and 100 mm height by following procedure prescribed in IS 516-2021 [65]. Two-point loading was adopted and minimum three beams were tested for each result. The formula for two-point bending is expressed as Eq. (2):

$$\text{Flexural strength} = \frac{PL}{bd^2}, \tag{2}$$

where P = maximum applied load (N), L = span length between supports (mm), b = width of the specimen (mm), and d = depth of the specimen (mm).

7 Experimental findings

7.1 Fresh-state properties

The results of flowability tests were shown in the Table 3. In the binary mixes as the increase of FA content the slump value, j-ring flow and L-box ratio is increasing indicating enhanced deformability and passing ability. Meanwhile t_{500} time and V-funnel time is decreasing indicating that faster flow rate and better filling ability. In general, the workability continues to improve as the SF% increases from 0 to 10. Higher SF% values are associated with decreased t_{500} and V-funnel times, as well as increased slump flow diameter, J-ring flow, and L-box H_2/H_1 ratio, much like the FA% variation. This suggests that SF plays a major role in rheology of self-compact concrete, even at comparatively low percentages. SF acts as a lubricant and filler, the ultra-fine SF particles increase the paste's volume and lower inter-particle friction, which improves the mixture's flowability and

Table 3 Flowability tests results of HVFA-SCC

FA (wt.%)	SF (wt.%)	Slump flow diameter (mm)	t_{500} (s)	J-ring flow (mm)	V-funnel time (s)	L-box H_2/H_1
0	0	652	4.3	648	11.8	0.81
10	0	661	3.9	655	11.2	0.85
20	0	665	3.5	658	10.3	0.88
30	0	676	2.8	669	9.2	0.9
40	0	680	2.5	676	8.5	0.95
50	0	702	2	697	8.1	0.97
30	2.5	692	3.5	685	8.9	0.87
30	5	702	3.2	696	8.1	0.94
30	7.5	715	2.9	706	7.5	0.95
30	10	719	2.4	708	6.8	0.97

passing ability. All the flowability results are satisfied the guidelines provided by EFNARC-2005 [64].

7.2 Hardened-state properties

7.2.1 Compressive strength

In the binary mixtures as shown in Fig. 19, the control (0% FA) exhibited consistently high CS at all curing ages, with strength plateauing between 28 and 56 d before increasing by 4 MPa at 90 d. Replacing 10% of cement with FA reduced early-age strength (7 d) relative to the control but yielded a 7 MPa gain by 90 d, reflecting the onset of pozzolanic reactions. As the FA content increased to 40%, mixes gradually approached the target mean strength of 58.25 MPa by 56 d; however, the 50% replacement mix underperformed at all ages, indicating diminishing returns at very high FA levels due to delayed hydration.

For the ternary blends (30% FA plus SF), adding 2.5% SF did not affect the 7 d strength compared to the 30% FA binary mix but produced notable gains at 28, 56, and 90 d. Incorporating 5% SF improved early‐age (7 d) strength, although subsequent age gains were modest. The 7.5% SF blend delivered the highest long-term strengths among all ternary mixes, attributable to the highly reactive SF accelerating the conversion of calcium hydroxide into additional C-S-H gel and densifying the microstructure in the presence of FA. At 10% SF, although CS surpassed that of the plain 30% FA mix, it declined relative to the 7.5% blend, likely due to particle agglomeration and increased water demand at higher SF dosages.

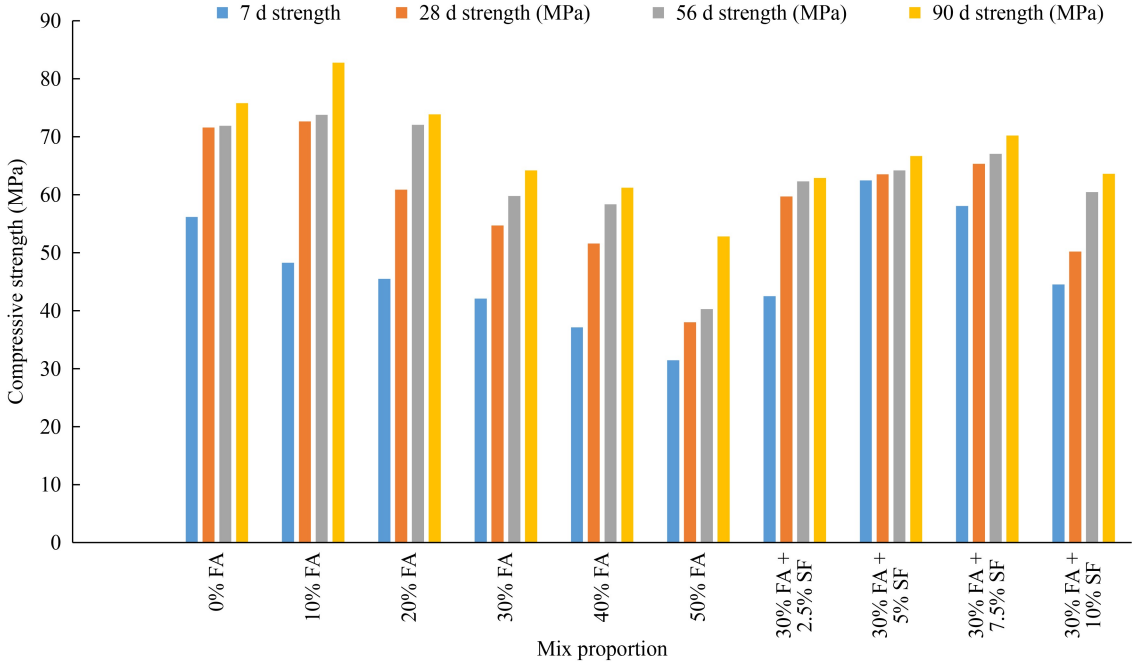

Fig. 19 Compressive strength in 7, 28, 56, and 90 d.

7.2.2 Split tensile strength

The control mixture, containing 0% FA, exhibited the highest split tensile strength at 5.57 MPa as shown in Fig. 20. As the proportion of FA replacement increased, tensile strength progressively decreased, reaching 4.12 MPa at 50% FA content. This reduction is attributable to the delayed pozzolanic reaction and diminished early-age reactivity associated with HVFA incorporation. Conversely, introducing SF at a 5% replacement level alongside 30% FA resulted in a notable enhancement in tensile performance, raising strength to 5.32 MPa. This improvement likely stems from the ultrafine SF particles refining the interfacial transition zone and promoting a denser, more cohesive microstructure, thereby compensating for the strength loss induced by FA. However, increasing SF content beyond 5% led to further strength reductions, with the 10% SF blend attaining only 4.02 MPa. This deterioration may reflect suboptimal dispersion and agglomeration of SF particles, which can introduce localized weaknesses within the concrete matrix.

7.2.3 Flexural strength

The control mixture with 0% FA exhibited a flexural strength of 5.71 MPa at 28 d. Partial replacement of cement with FA produced modest variations in performance: a 10% replacement yielded 5.51 MPa, whereas increasing FA content to 20% and 30% enhanced flexural strength to 5.99 and 5.90 MPa, respectively, indicating that up to 30% substitution can be beneficial as shown in Fig. 21. Beyond this threshold, strength declined to 5.46 and 5.20 MPa at 40% and 50% FA, respectively, reflecting the adverse impact of excessive substitution on matrix cohesion and early-age pozzolanic reaction.

In the ternary blends incorporating 30% FA, the addition of 2.5% SF resulted in a minimum flexural strength of 4.70 MPa, likely due to insufficient filler synergy or suboptimal dispersion. Elevating SF content to 5% restored strength to 5.70 MPa, comparable to the control mix. A peak flexural strength of 6.30 MPa was achieved at 7.5% SF, attributable to the ultra-fine particles and heightened reactivity of SF, which refine the interfacial transition zone, reduce porosity, and densify the microstructure. However, further increasing SF to 10% resulted in a reduction to 5.30 MPa, suggesting that particle agglomeration and dispersion challenges at higher dosages can introduce localized weaknesses within the concrete matrix.

8 Proposed machine learning models

8.1 Adaptive boosting

AdaBoost is a classic boosting algorithm which builds a powerful classifier via sequentially combining many weak learners, decision stumps usually. The procedure starts with all training samples having equal weights. A weak learner is trained at every iteration, and the weights of the incorrectly classified samples are boosted, so that later learners pay more attention to the hard examples.

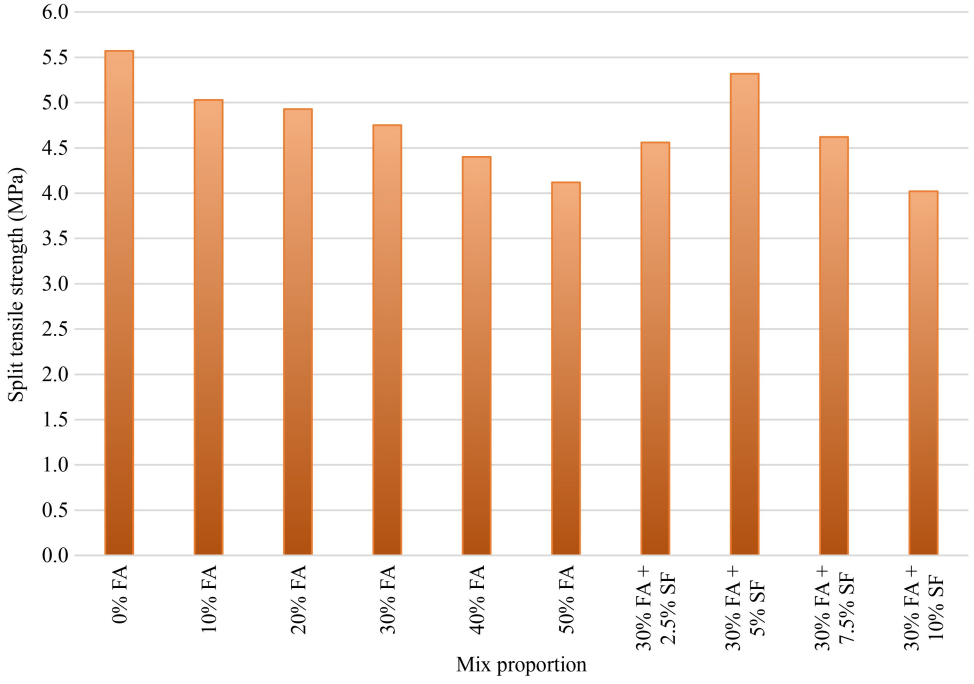

Fig. 20 Split tensile strength in 28 d.

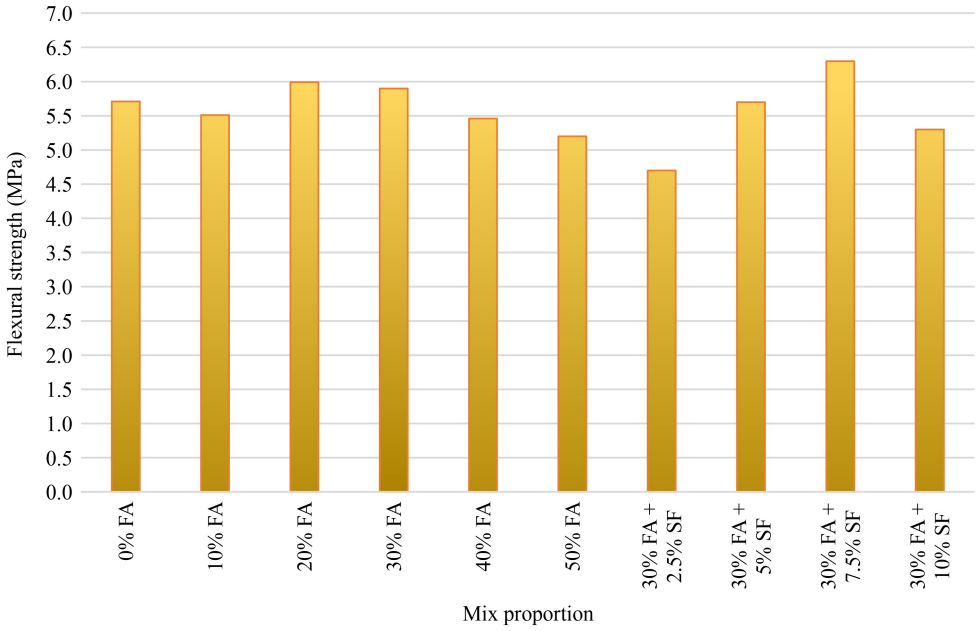

Fig. 21 Flexural strength in 28 d.

This adaptive reweighting is repeated a specified number of times or till the achievement of perfect classification. The last prediction is carried out by doing a weighted majority vote of the output of all weak learners and what is obtained is a model with a much better accuracy and generalization. Theoretical arguments have demonstrated that AdaBoost expands the margins of the training examples and gives a bound on the generalization error, and this is useful in avoiding overfitting despite a large number of boosting rounds [66,67].

8.2 Gradient boosting

GB is a more liberal and articulate boosting architecture which builds an ensemble of weak learners, stage-wise. GB, unlike AdaBoost which is focused on reweighting misclassified instances, fits each new learner to the residual errors (gradients) of the current ensemble of learners' predictions, and is equivalent to minimizing a specified loss function. The method enables the algorithm to be used with both regression and classification tasks as well as to minimize a large variety of loss functions [68,69]. This step-by-step correction of mistakes repeatedly allows GB to create highly accurate predictive models, and its flexibility has caused it to become the method of choice in many real-world ML applications.

8.3 Categorical boosting

CatBoost is a gradient-boosting framework tailored for tabular data with categorical features, addressing key pitfalls of conventional boosting such as target leakage and overfitting through its ordered boosting scheme and permutation-driven target statistics [70]. Instead of requiring one-hot or target encoding as a preprocessing step, CatBoost natively transforms categorical variables using unbiased estimates computed on sequential permutations of the data, thereby stabilizing training and improving generalization. Its use of symmetric (oblivious) DT yields fast, cache-friendly learning and robust regularization, while built-in L2 penalties, shrinkage, and depth constraints further control model complexity. CatBoost is efficient on both Central Processing Unit and Graphics Processing Unit, includes sparsity-aware handling of missing values, supports custom loss functions and constraints, and provides rich interpretability tooling (e.g., SHAP-based feature attributions). In practice, it delivers state-of-the-art results across regression, classification, and ranking tasks on real-world, mixed-type data sets, making it a strong baseline and often a top performer in academic studies and industrial applications.

8.4 Extreme gradient boosting

XGBoost is an extremely efficient and scalable GB framework, which is used to overcome the computational and statistical shortcomings of the conventional boosting algorithms. Some of the main improvements made by XGBoost are regularization (e.g., L1 and L2 penalties) to avoid overfitting and to constrain the complexity of the model, and improved tree learning algorithms, which make parallel and distributed computation possible [71]. It also has a sparsity-conscious algorithm to deal with missing data and can efficiently use memory, which is helpful to large-scale data sets. The flexibility of XGBoost in dealing with different objective functions, allowing the use of custom loss functions, and offering

strong measures of feature importance has made it popular in academic work and industry. In empirical evaluations, XGBoost repeatedly appears to reach state-of-the-art performance on many ML tasks, such as classification, regression, and ranking [72]; it is now a standard tool in data science competitions and industry practice

8.5 Metaheuristic optimization techniques for hybrid extreme gradient boosting

8.5.1 Giant trevally optimization

The GTO is a novel stochastic metaheuristic algorithm inspired by the foraging behavior and strategic hunting patterns of the giant trevally fish, particularly their systematic foraging movements, selection of optimal foraging zones, and dynamic hunting tactics, including leaping from the water to capture prey. The GTO algorithm simulates these behaviors through a multi-phase process: it commences with an extensive search phase, wherein candidate solutions (each representing an individual trevally) explore the solution space using mechanisms like Lévy flights to mimic the long-range, stochastic foraging behavior of the fish and to evade local optima [73]. During the area selection phase, each trevally evaluates and is drawn to regions inside the search space that have exhibited superior potential, informed by both the best-known solution and the population's average knowledge thus far. Finally, in the attack or exploitation phase, the algorithm refines solutions by navigating around the best-known solution, emulating the trevally's strategic leap to capture prey, and converging toward the global optimum [74]. The GTO framework facilitates robust exploration and exploitation, with research validating its effectiveness and efficiency across diverse optimization and engineering challenges, often demonstrating superiority or parity with other established metaheuristic algorithms [75,76].

8.5.2 Coyote optimization algorithm

The COA is a metaheuristic algorithm which is bio-inspired and is based on the social structure and adaptive behavior of coyotes in the wild. The population in COA is organized in a number of packs, each consisting of a group of coyotes and the optimization process is achieved in four key phases: initialization of coyotes, growth of coyotes, birth and death of pups and migration of coyotes between packs [77]. A candidate solution is represented by each coyote and the adaptability (fitness) of the coyote is determined according to the objective function of the problem. The social interactions are modeled by the algorithm that enables coyotes in a pack to push one another around and the new solutions (pups) are produced by recombining the social characteristics of randomly chosen parents taking into account the environmental characteristics [78,79]. Coyotes also have the tendency of migrating between packs to ensure there is diversity and premature convergence is avoided. COA variants that add a mechanism to explore the space (Levy flights and chaotic maps) have also been shown to increase the robustness and efficiency of the algorithm. COA has shown to compete on a range of optimization problems, maintaining an appropriate level of exploration and exploitation via its social interactions and dynamic adjustments [80].

8.5.3 Tuna swarm optimization

TSO is a population-based metaheuristic algorithm derived from the helical and parabolic foraging behaviors shown by tuna during hunting [81–83]. The method starts with the random initialisation of candidate solutions inside the search space, followed by fitness assessment according to the specified objective function. TSO utilizes two principal strategies: spiral foraging for global exploration and parabolic foraging for local exploitation. During each cycle, people adjust their locations using a probabilistic method that equilibrates both techniques. The optimal solution is preserved, and the procedure continues until convergence or a certain number of repetitions is achieved. TSO has shown efficacy in diverse optimization issues owing to its robust exploration-exploitation equilibrium and resilience against premature convergence [83,84]. Proposals for enhanced versions using chaotic maps or Lévy flights have been made to augment variability and convergence rate [81].

9 Data preparation and machine learning algorithms

A total of 624 samples of concrete mix composition and CS were cleaned up to enhance the quality of the data and predictive modeling. The first data cleaning involved replacing the missing values with the mean and eliminating the duplicate entries to minimize redundancy. Authenticated numerical variables were converted to regular data formats. The outliers were eliminated to have fewer data points that are statistically significant to reduce noise using the Z-score method with the threshold set to stop 3. The range of the continuous input variables in the cleaned data set was normalized with Standard Scaler to enhance gradient-based ML models. Lastly, the data was stratified sampled into 80:20 training and test to maintain the distribution of the target variable. The Bayesian optimization was performed with a search space including (list key parameters: e.g., learning rate 0.001–0.1, max_depth 3–10), 50 iterations, and Gaussian

process priors. This ensured a balanced exploration of hyperparameters rather than only reporting final optimal values.

Preprocessing input features was followed by descriptive statistical analysis to summarize input feature properties in Table 4. Cement content was significantly different with a mean of 318.67 kg/m^3 and a range of 489 kg/m^3. FA, Slag, and SF demonstrated a high coefficient of variation and strong skew especially Slag (skewness = 8.60, kurtosis = 77.40) and SF (skewness = 3.18, kurtosis = 9.46), which are not normally distributed. The water content had a mean value of 182.74 kg/m^3, whereas the fine and coarse aggregates were symmetrically distributed. The age of the concrete samples was between 1 and 182 d with positively skewed distribution (1.59) which means that there is a focus on the early-age testing. The dosage of superplasticizer was skewed to the right (0.97). The CS (F_c) was slightly positively skewed and platykurtic with the mean of 40.88 MPa and standard deviation of 19.38 MPa. Such descriptive observations show the heterogeneity of data and the necessity of using powerful and flexible ML models capable of learning the nonlinear, complex interactions of concrete behavior. The Pearson correlation matrix is shown in Fig. 22 to demonstrate the parametric influence of each individual parameter to the CS of the concrete.

10 Results and discussion

10.1 SHapley additive exPlanations analysis

SHAP is a unified framework for interpreting ML models by assigning each feature a contribution value, called a SHAP value, which quantifies the extent to which that feature raises or lowers the model's baseline prediction. Rooted in cooperative game theory, SHAP values provide fair and consistent feature attribution, offering both global insight, average importance across all predictions, and local explanation, feature contributions for individual predictions, in a single, theoretically grounded approach. Figures 23–25 collectively elucidate the influence of input variables on the XGB-TSO model's predictions for SCC CS. Figure 23 ranks feature by their mean absolute SHAP values and demonstrates that curing age and cement content dominate model behavior, contributing on average approximately 8.7 and 7.9 MPa, respectively. Water and superplasticizer dosage follow with mean contributions near 4.5 and 2.3 MPa, while coarse and fine aggregates exhibit moderate effects of 1.8 and 1.7 MPa. Slag and SF show minimal impact, each below 0.5 MPa. Figure 24 presents SHAP decision-plot trajectories for all 624 instances, each beginning at the model's baseline prediction of 40 MPa and accumulating feature contributions to reach individual outputs. Early-stage shifts driven by age and cement can span ±20 MPa; instances cured beyond 60 d contribute up to 18 MPa and those under 7 d subtract up to 16 MPa. Mid-trajectory effects from water (±8 MPa) and superplasticizer (±5 MPa) further adjust predictions, whereas aggregate and SCMs contributions remain within ±4 MPa. Figure 25's bees warm plot synthesizes these patterns by depicting the distribution of SHAP values for each feature, with horizontal dispersion indicating the range of contributions and color indicating feature magnitude. High values of cement and age consistently drive positive contributions, whereas elevated water content yields negative contributions. The narrow spread of SF and slag contributions underscores their comparatively minor roles. Together, these visualizations confirm that the model's internal logic aligns with established physico-chemical principles governing concrete strength development.

10.2 Performance evaluation of predicting compressive strength

The predictive efficacy of the proposed hybrid models, integrating ML and metaheuristics, was assessed utilizing

Table 4 Descriptive analysis

Parameter	Cement (kg/m^3)	Fly Ash (kg/m^3)	Slag (kg/m^3)	SF (kg/m^3)	Water (kg/m^3)	FA (kg/m^3)	CA (kg/m^3)	SP (kg/m^3)	Age (d)	F_c (MPa)
Count	624	624	624	624	624	624	624	624	624	624
Mean	318.7	153.2	1.6	5.7	182.7	840.7	809.8	4.3	33.9	40.9
Std	107.0	95.4	13.2	17.2	28.6	95.4	111.6	3.4	34.2	19.4
Min	61	0	0	0	124.8	523	540	0	1	4.44
25%	228.8	96.5	0.0	0.0	163.0	786.0	716.1	1.5	7.0	26.1
50%	307.5	160.0	0.0	0.0	180.0	850.0	837.0	3.2	28.0	40.7
75%	388.0	216.0	0.00	0.00	197.6	900.2	875.8	6.50	28.00	55.26
Max	550	378	150	82.2	286	1054	1190	17.1	182	94.6
Range	489	378	150	82.2	161.2	531	650	17.1	181	90.16
Skewness	0.33	−0.02	8.60	3.18	0.67	−0.61	0.28	0.97	1.59	0.15
Kurtosis	−0.60	−0.60	77.40	9.46	0.92	1.26	1.24	0.59	2.71	−0.76

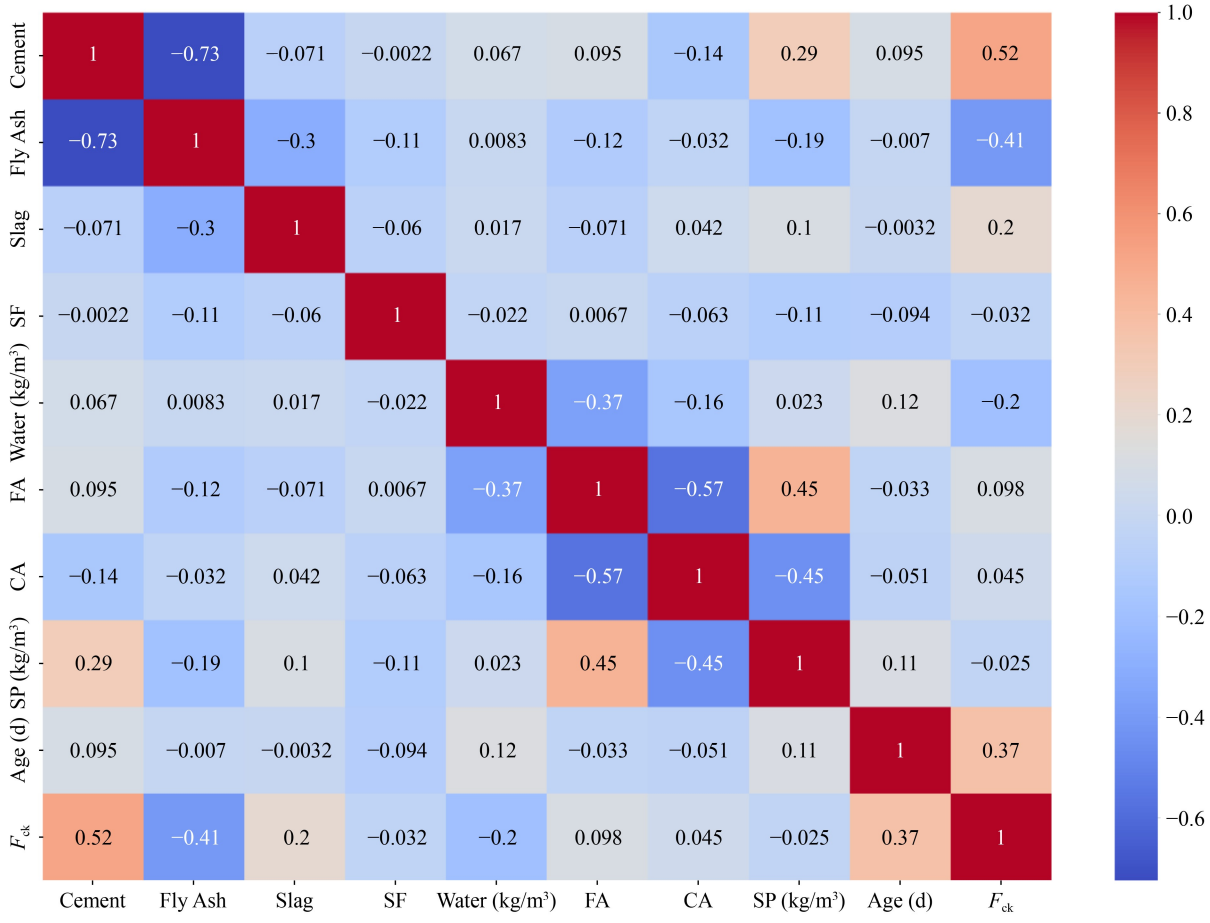

Fig. 22 Pearson correlation.

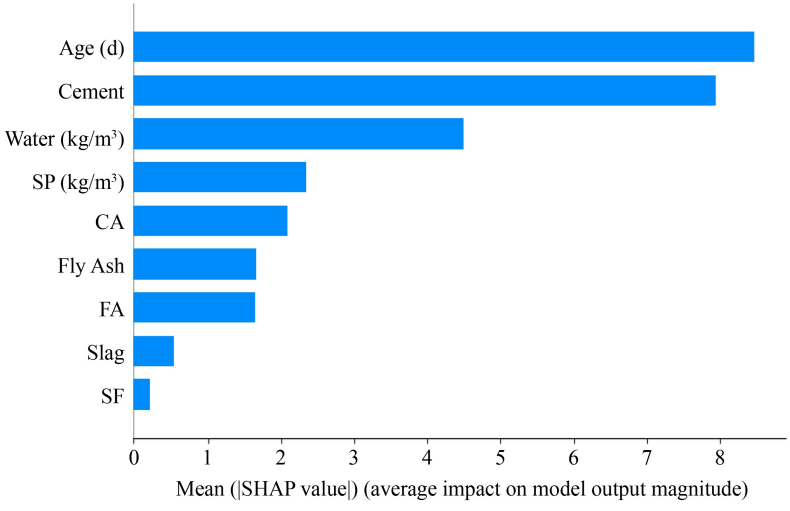

Fig. 23 Ranks feature by their mean absolute SHAP.

a range of standard statistical metrics, including the coefficient of determination (R^2), Root Mean Square Error (*RMSE*), Nash–Sutcliffe Efficiency (*NS*), Weighted Mean Absolute Percentage Error (*WMAPE*), Variance Accounted For (*VAF*), and Willmott's Index of Agreement (*WI*). These variables together assess model accuracy, bias, variation, and prediction consistency in both training and testing data sets [85–87].

The training data set findings (Table 5) indicate that the XGBoost model demonstrated the superior predictive ability among the base learners, with R^2 = 0.9987 and *RMSE* = 0.715. The hybrid versions, namely XGB-TSO and XGB-GTO, exhibited similarly good performance with equal R^2 values (0.9987 for TSO and 0.9968 for

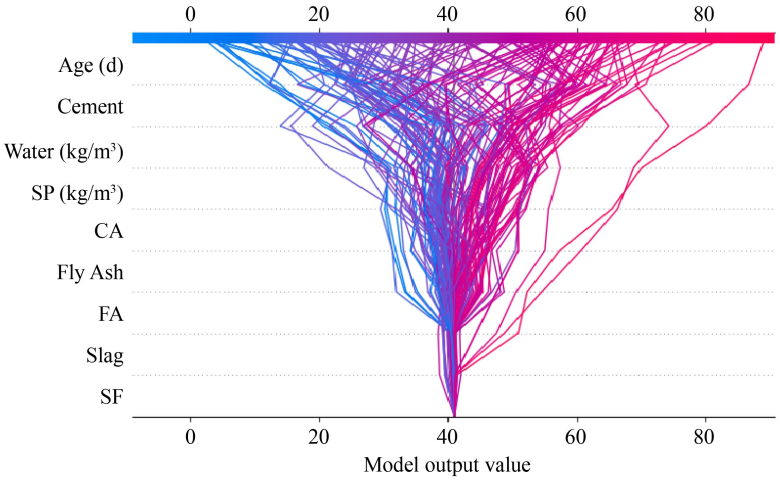

Fig. 24 SHAP decision plot.

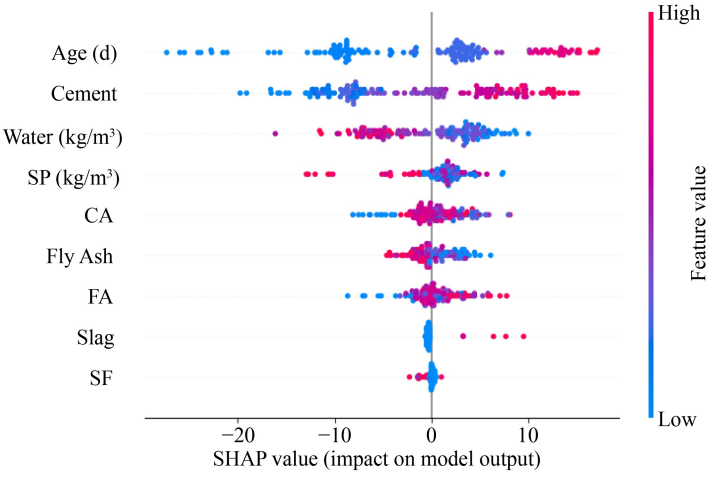

Fig. 25 Bee swarm plot.

Table 5 Model performance in training data set

Parameter	AdaBoost	GB	CatBoost	XGBoost	XGB-GTO	XGB-TSO	XGB-COA
R^2	0.8990	0.9514	0.9831	0.9987	0.9968	0.9987	0.9975
WMAPE	0.150	0.080	0.048	0.010	0.020	0.010	0.017
NS	0.859	0.949	0.983	0.999	0.997	0.999	0.997
RMSE	7.410	4.435	2.560	0.715	1.115	0.715	0.987
VAF	86.09	94.93	98.31	99.87	99.68	99.87	99.75
WI	0.788	0.938	0.996	0.999	0.996	0.999	0.997

GTO) and *RMSE* values of 0.715 and 1.115 MPa, respectively. The second-best performing model is CatBoost with R^2 = 0.9831 and *RMSE* = 2.560. The low *RMSE* values indicate minimal deviation from actual strength values, indicating strong learning and generalisation ability. AdaBoost and GB demonstrated lower accuracy, with R^2 values of 0.8990 and 0.9514, and *RMSE*s of 7.410 and 4.435 MPa, respectively. The associated *WMAPE* data corroborate this: XGBoost versions attained a minimum of 0.01 (1%) in *WMAPE*, but AdaBoost exhibited a comparatively high error of 15%. The findings of the testing data in Table 6 validated the superiority of the hybrid models. XGB-TSO attained the greatest R^2 of 0.9718, closely followed by XGB-GTO at 0.9660, but GB and AdaBoost trailed with R^2 values of 0.9204 and 0.8773, respectively. XGB-TSO demonstrated the lowest *RMSE* (3.525) and *WMAPE* (0.066), confirming its stability and accuracy in the presence of test data. The NS values for the hybrid models varied from 0.959 to 0.972, in sharp contrast to AdaBoost's value of 0.826, so underscoring the superior error resistance of the hybrid algorithms. The *WI* and *VAF*

Table 6 Model performance in testing data set

Parameter	AdaBoost	GB	CatBoost	XGBoost	XGB-GTO	XGB-TSO	XGB-COA
R^2	0.8773	0.9204	0.9300	0.9375	0.9660	0.9718	0.9588
WMAPE	0.179	0.111	0.100	0.099	0.072	0.066	0.080
NS	0.826	0.917	0.930	0.937	0.966	0.972	0.959
RMSE	8.724	6.008	5.532	5.249	3.872	3.525	4.256
VAF	82.61	91.74	93.00	93.75	96.58	97.17	95.87
WI	0.939	0.977	0.982	0.983	0.991	0.993	0.989

values were consistently elevated for the hybrid models ($WI > 0.98$, $VAF > 95\%$), indicating substantial predictive agreement and reduced unexplained variation. These measures together highlight the strength and practical use of XGB-based hybrid models, especially when enhanced by TSO and GTO optimization.

Scatter graphs comparing actual and expected CS values were generated for each model to complement the numerical assessment (Figure: Actual vs expected). Figure 26 illustrate a more compact grouping of points around the 45° line for AdaBoost, GB and XGBoost whereas Figs. 27(a)–27(c) clearly illustrate a more compact grouping for XGBoost and its hybrid models (COA, TSO, and GTO), corroborating reduced prediction error. Conversely, AdaBoost and GB plots exhibit increased dispersion, indicating bigger variances. XGBoost optimized with TSO repeatedly shown to be the most precise and steady predictor of SCC CS across all measurements and visual tools. The integration of optimization techniques, including GTO, TSO, and COA,

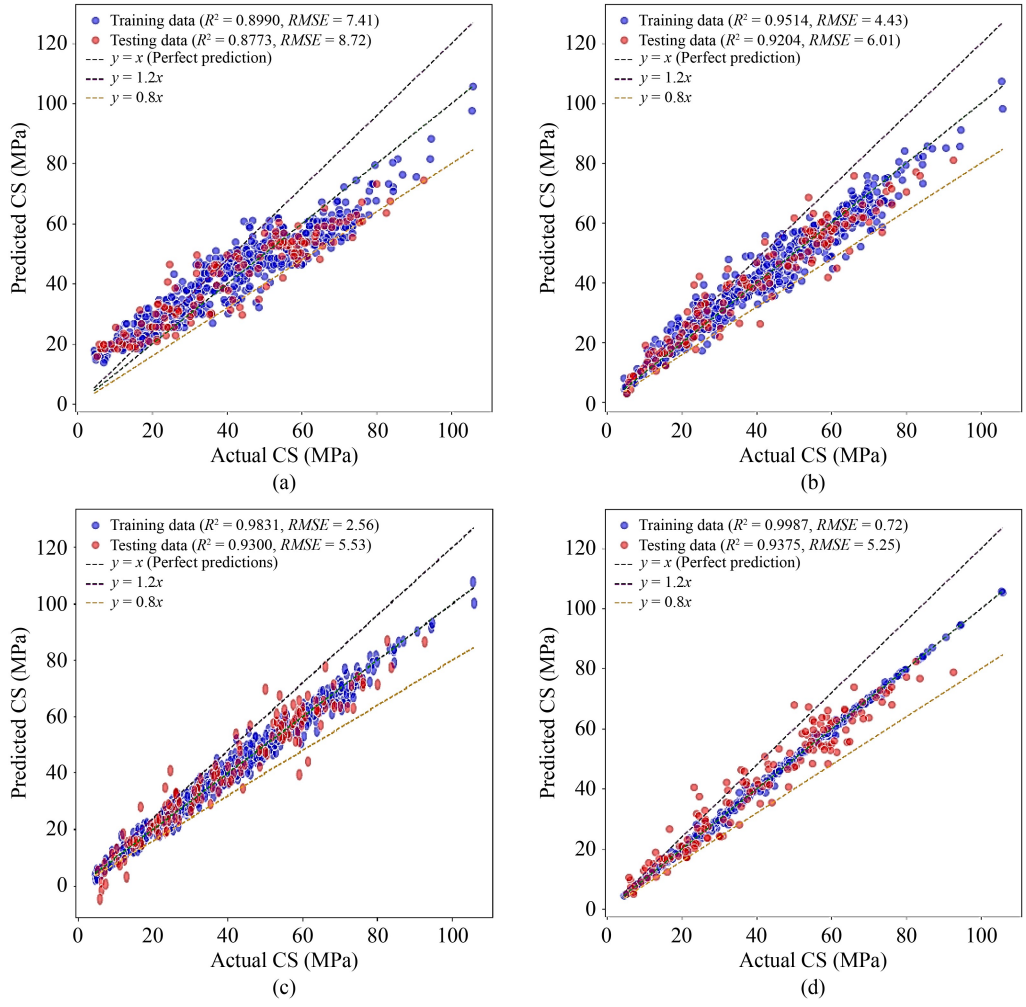

Fig. 26 Actual vs predicted CS for: (a) AdaBoost; (b) GB; (c) CatBoost; (d) XGBoost.

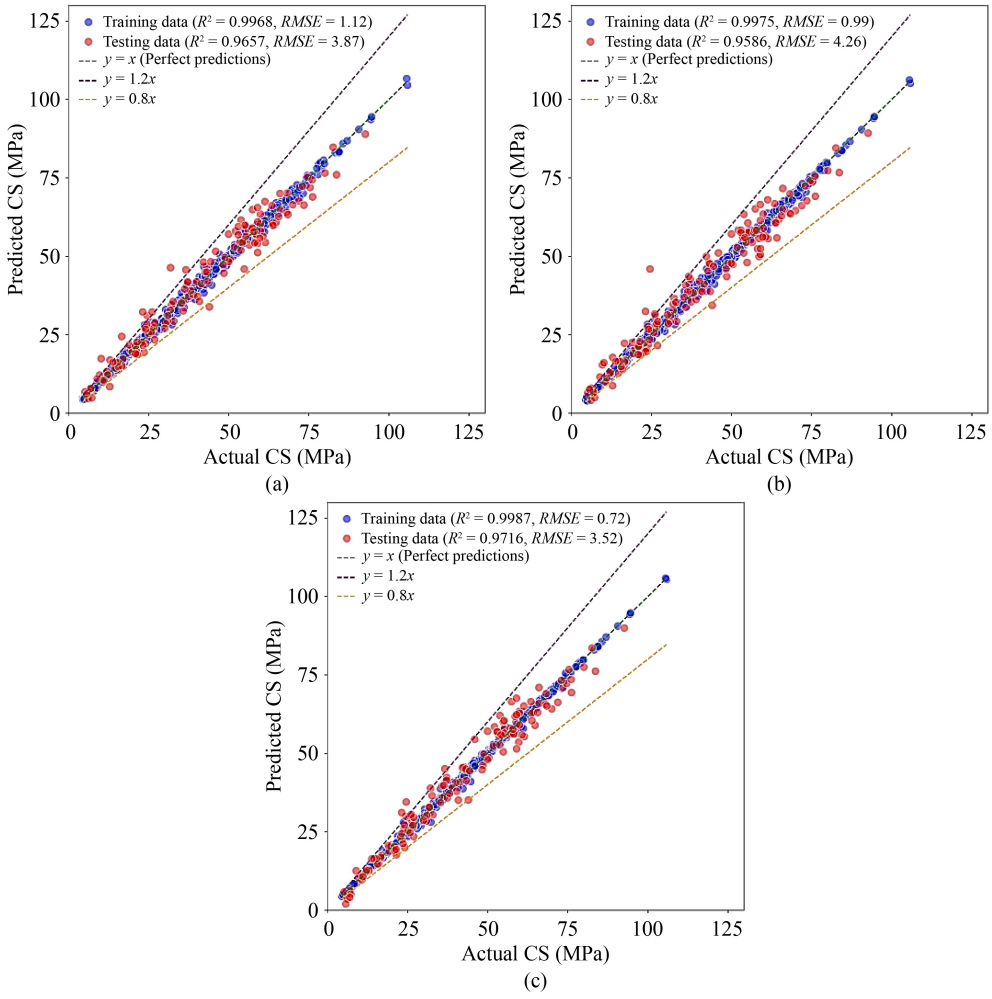

Fig. 27 Actual vs predicted CS for hybridized XGBoost with: (a) GTO; (b) COA; (c) TSO.

significantly improved XGBoost's baseline performance, decreasing *RMSE* by more than 80% relative to unoptimized ensemble learners such as AdaBoost. These results provide strong support for the use of hybrid XGBoost frameworks, especially with TSO, in high-fidelity models for sustainable concrete research. This substantiates the novel use of powerful metaheuristic algorithms into ensemble learning frameworks for predicting concrete strength.

10.3 Taylor diagram presentation for machine learning models

Taylor diagrams are mathematical visualization data sets aimed to graphically summarize the statistical association involving modeled and observed information by concurrently showing 3 commonly used measures of performance: the Pearson correlation coefficient, the centered root-mean-square error (*CRMSE*), and the standard deviation. These diagrams were first proposed by Karl E. Taylor in 2001 and employ polar coordinates in which the correlation coefficient is plotted as the azimuthal angle, the standard deviation as the radial distance to the origin and the *CRMSE* as the distance to the reference point on the *x*-axis [88,89].

By examining the displayed Taylor diagrams of training and testing phases in Figs. 28(a) and 28(b), it can be seen that all of the compared models, namely, ensemble boosting techniques (AdaBoost, GB, CatBoost, XGBoost) and metaheuristic optimization algorithms (GTO, TSO, COA) possession good performance traits, as they have high correlations and relatively low centered *RMSE*. All the models being clustered close to the reference point (correlation = 1.0) show an excellent agreement between the predicted and observed values with standard deviations of 16–20 units. Strikingly, training and testing results are similar, indicating strong generalization ability on all methods, and no considerable indication of overfitting. The close clustering of the optimization algorithms with the classic ensemble approaches shows that the metaheuristic-optimized models are statistically comparable to known ML approaches and therefore are useful in predictive modeling tasks [90].

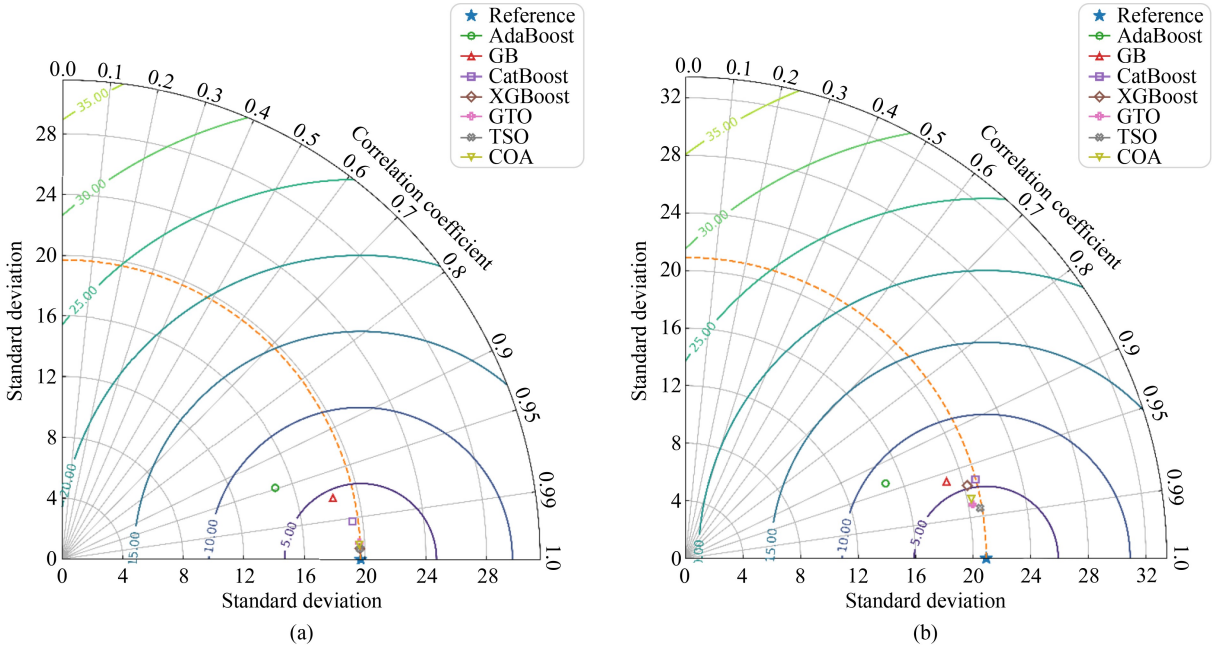

Fig. 28 Taylor diagram in: (a) training; (b) testing.

11 Validation

The validation of the developed hybrid ML models, specifically XGBoost combined with metaheuristic algorithms such as GTO, TSO, and COA was conducted using an independent experimental data set. This Independent experimental data set provided a robust foundation for evaluating the models' generalization and predictive reliability for estimating the CS of SCC. The validation process incorporated a suite of standard performance metrics as mention in Section 9 The results highlighted that all hybrid models demonstrated strong predictive capacity in Table 7, with R^2 values consistently exceeding 0.965, indicating the models could explain the vast majority of variance in observed CS. Low *RMSE* values, typically around 2.15 to 2.25 MPa, indicated highly accurate and precise predictions in alignment with actual test results. *WMAPE* values in the range of 2.6%–2.7% further underscored the models' minimal average prediction error. High *NS* and *WI* values, both approaching unity, confirmed that these models not only captured overall data trends but also replicated the nuances of experimental behavior, supporting their robust generalization and real-world utility. The *VAF* percentages, all above 96.5%, corroborated the high explanatory power of these hybrid approaches.

Table 7 Model performance in testing data set

Parameter	GTO	TSO	COA
R^2	0.9698	0.9690	0.9655
WMAPE	0.0270	0.0263	0.0266
NS	0.9671	0.9674	0.9643
RMSE	2.1617	2.1546	2.2533
VAF	96.96	96.90	96.54
WI	0.9915	0.9916	0.9908

Visual analysis added clarity to these metrics: scatter plots exhibited tight clustering along the refence line (Figs. 29(a)–29(c)) and line diagrams (Fig. 30) showed excellent agreement between predicted and measured strengths, reinforcing the models' precision and reliability throughout the data set. The integration of metaheuristic optimization into the XGBoost model framework noticeably enhanced predictive capability, outperforming base models as well as traditional ensemble learners. Critically, these validation outcomes demonstrate that the established modeling approach is not only robust in theory but also highly effective and reliable for practical engineering applications in SCC strength estimation. To mitigate local minima issues, multiple random initializations were tested, yielding consistent results. Out-of-distribution behavior was assessed through sensitivity checks, clarifying model limits. Additionally, an independent experimental data set was used for validation, achieving $R^2 \approx 0.97$ and $RMSE \approx 2.15$ MPa, confirming the robustness of the proposed hybrid models.

12 Graphical user interface for practical engineering

The increasing reliance on ML in carrying out predictive analysis in concrete technology suggests the need to align the data-driven approach with the classical empirical-based approach, thereby ensuring that both remain applicable to design engineers. A novel prediction model

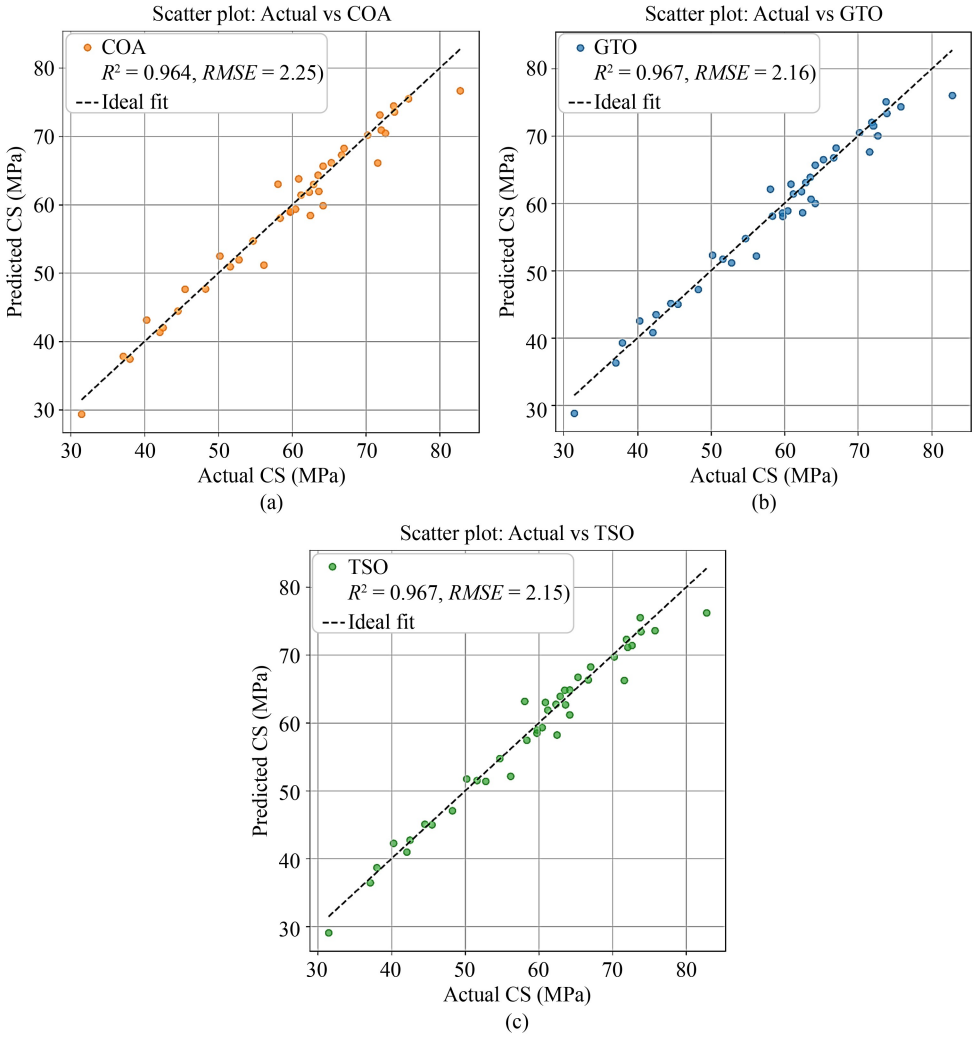

Fig. 29 Actual vs predicted graph for Independent experimental data set validation with: (a) COA; (b) GTO; (c) TSO.

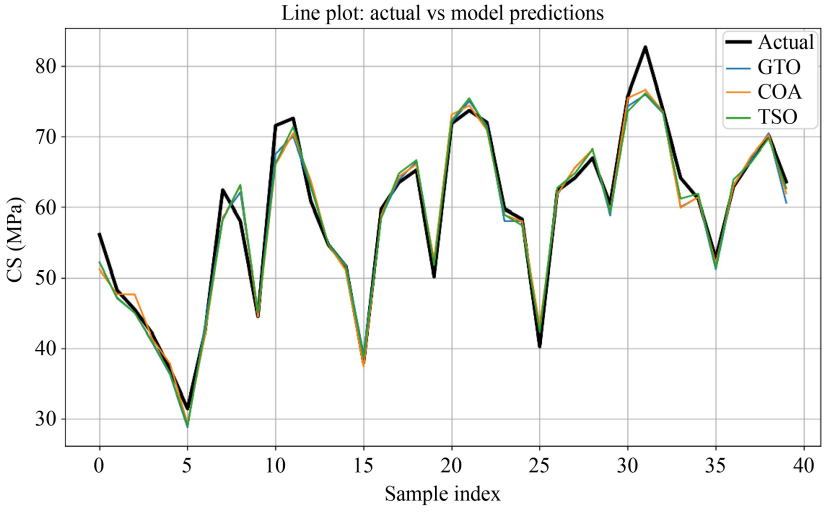

Fig. 30 Actual vs predicted line plot experimental validation.

XGBoost, which is optimized TSO algorithm, is introduced in this paper to predict the CS of SCC. In addition to this optimized model, a user-friendly GUI is developed as shown in Fig. 31 and can be accessed on GitHub. The interface enables the user to input important mix design variables, which are cement, water, superplasticizer, FA, SF, aggregates and concrete age, on a per cubic meter basis. When a value is entered, the

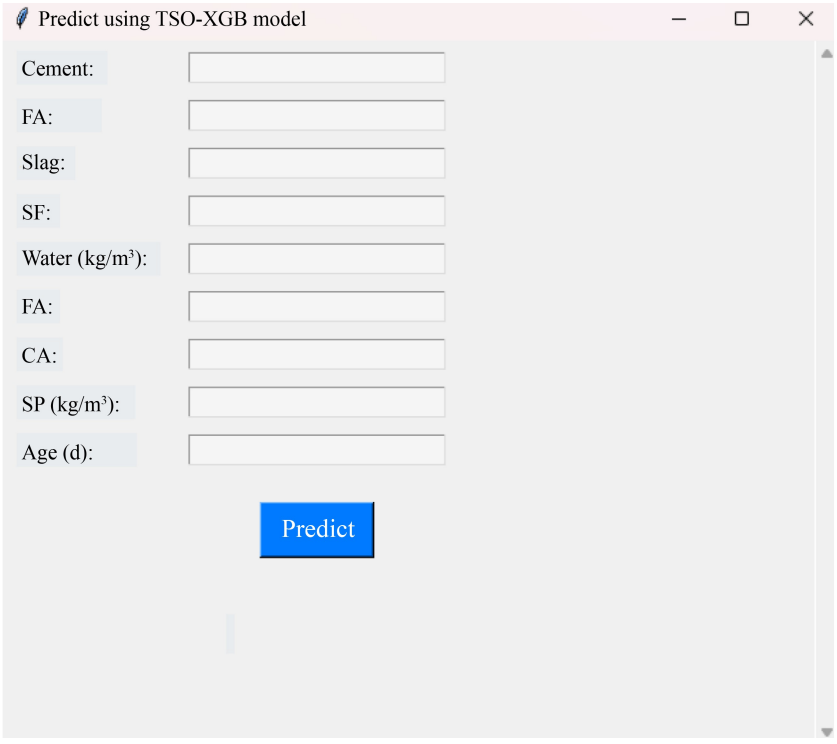

Fig. 31 Developed GUI.

technology will quickly estimate the expected CS and offer quick information concerning the mix feasibility. Its simplified architecture makes it usable by users with a wide range of technical expertise, and removes the need to have programming knowledge, bridging the gap between research level models and field usable applications. The incorporation of the TSO-optimized XGBoost model into this platform signifies a significant progress in the accessibility, accuracy, and efficiency of SCC strength prediction.

13 Conclusions

This study established a unified experimental and ML-driven framework for predicting the CS of SCC. Through the integration of laboratory test data with advanced hybrid ML algorithms, the research offers a bridge between traditional materials science and contemporary predictive analytics in concrete technology. The XGBoost model optimized by XGB-TSO was identified as the most effective predictive tool among those evaluated. The key findings of the study include.

1) The optimal SCC mix, comprising 30% FA and 5%–7.5% SF as cement replacements, resulted in pronounced improvements in fresh-state properties (flowability, reduced segregation) as well as enhanced compressive, tensile, and flexural strengths across all curing periods. This ternary blend provided well-balanced pozzolanic reactivity and particle packing, supporting superior mechanical performance and matrix densification over time. The addition of SF notably accelerated early strength development, while the long-term gains in strength were retained due to the latent hydraulic activity of FA.

2) The XGB-TSO model demonstrated the strongest predictive accuracy, with R^2 = 0.9690, *RMSE* = 2.1546 MPa, *WMAPE* = 2.63%, and *NS* = 0.9674. These results validate the model's suitability and reliability for SCC strength prediction.

3) Model interpretability using SHAP analysis highlighted age and cement content as the most influential predictors, with decision plots illustrating both the cumulative and individual contributions of input features to the predicted CS.

4) The Taylor diagram further confirmed the robustness of the XGB-TSO model, indicating the highest correlation values (> 0.97), the lowest centered *RMSE*, and standard deviation that closely matched the corresponding experimental data.

5) To facilitate practical use, a user-friendly GUI was developed, enabling real-time CS predictions based on user-specified SCC mix inputs. This advancement supports the accessibility and deployment of the developed model in field and design contexts.

6) Independent experimental validation illustrated a strong match between predicted and observed CS (validation R^2 = 0.9698; mean absolute error = 1.83 MPa; maximum absolute error within ±3.5 MPa). Visualizations, including scatter and residual plots, confirmed that predictions clustered closely around the 45° line, with low and uniformly distributed error margins, providing

further assurance of predictive stability and minimal bias.

The novelty of this work is evidenced by its integrated approach of combining metaheuristic optimized ML, parallel experimental evaluation, SHAP-based interpretability, Taylor diagram visualization and practical interface development, thus presenting a comprehensive end-to-end pipeline from laboratory experimentation to real-world deployment.

14 Limitation

Nevertheless, the study acknowledges some limitations. Despite the strong results, the work has constraints, including 1) data set primarily limited to SCC with FA and SF, with certain SCMs (e.g., slag) showing uneven distribution; 2) restricted hyperparameter optimization of ANN; 3) absence of comprehensive global uncertainty analysis; 4) experimental simplifications (controlled curing regimes). These factors will guide the scope of future extensions. In addition, while CatBoost has been incorporated and discussed in detail, further comparisons with emerging models (e.g., Light Gradient Boosting Machine, Transformer-based architectures) could provide deeper benchmarking. Future research may also explore the integration of physics-informed approaches and Artificial Intelligence for partial differential equations to reduce data requirements and improve generalizability. Expanding the framework to include a wider range of SCMs, different curing conditions, and reinforcement parameters will further increase applicability. Enhancing the GUI with dynamic visualization, uncertainty quantification, and database integration is also recommended to strengthen its deployment in practical engineering contexts.

Acknowledgements The authors gratefully acknowledge Berger Paints India Ltd. for kindly providing the superplasticizer HS PRO HYPERPLAST RT 2000 NSG used in this study. We also extend our sincere thanks to Elkem South Asia Pvt. Ltd. for supplying Elkem Microsilica® 920U (silica fume), which was an essential component of the experimental program. Their support in facilitating the procurement of high-quality materials is deeply appreciated.

Competing interests The authors declare that they have no competing interests.

References

1. Ali M B, Saidur R, Hossain M S. A review on emission analysis in cement industries. Renewable & Sustainable Energy Reviews, 2011, 15(5): 2252–2261
2. Dinakar P, Babu K G, Santhanam M. Durability properties of high volume fly ash self compacting concretes, 2008, 30(10): 880–886
3. Liew K M, Sojobi A O, Zhang L W. Green concrete: Prospects and challenges. Construction & Building Materials, 2017, 156: 1063–1095
4. Hayles M, Sanchez L F M, Noël M. Eco-efficient low cement recycled concrete aggregate mixtures for structural applications. Construction & Building Materials, 2018, 169: 724–732
5. Barra Bizinotto M, Faleschini F, Jiménez Fernández C G, Aponte Hernández D F. Effects of chemical admixtures on the rheology of fresh recycled aggregate concretes. Construction & Building Materials, 2017, 151: 353–362
6. Faleschini F, Jiménez C, Barra M, Aponte D, Vázquez E, Pellegrino C. Rheology of fresh concretes with recycled aggregates. Construction & Building Materials, 2014, 73: 407–416
7. Aïtcin P C. Cements of yesterday and today: Concrete of tomorrow. Cement and Concrete Research, 2000, 30(9): 1349–1359
8. Schneider M. Process technology for efficient and sustainable cement production. Cement and Concrete Research, 2015, 78: 14–23
9. Miller S A, Monteiro P J M, Ostertag C P, Horvath A. Concrete mixture proportioning for desired strength and reduced global warming potential. Construction & Building Materials, 2016, 128: 410–421
10. Celik K, Meral C, Petek Gursel A, Mehta P K, Horvath A, Monteiro P J M. Mechanical properties, durability, and life-cycle assessment of self-consolidating concrete mixtures made with blended Portland cements containing fly ash and limestone powder. Cement and Concrete Composites, 2015, 56: 59–72
11. Huang W, Kazemi-Kamyab H, Sun W, Scrivener K. Effect of cement substitution by limestone on the hydration and microstructural development of ultra-high performance concrete (UHPC). Cement and Concrete Composites, 2017, 77: 86–101
12. Sethy K P, Pasla D, Chandra Sahoo U. Utilization of high volume of industrial slag in self compacting concrete. Journal of Cleaner Production, 2016, 112: 581–587
13. Atiş C D. Heat evolution of high-volume fly ash concrete. Cement and Concrete Research, 2002, 32(5): 751–756
14. Huang C H, Lin S K, Chang C S, Chen H J. Mix proportions and mechanical properties of concrete containing very high-volume of Class F fly ash. Construction & Building Materials, 2013, 46: 71–78
15. Yu J, Lu C, Leung C K Y, Li G. Mechanical properties of green structural concrete with ultrahigh-volume fly ash. Construction & Building Materials, 2017, 147: 510–518
16. Kurad R, Silvestre J D, de Brito J, Ahmed H. Effect of incorporation of high volume of recycled concrete aggregates and fly ash on the strength and global warming potential of concrete. Journal of Cleaner Production, 2017, 166: 485–502
17. Wang X Y, Park K B. Analysis of compressive strength development of concrete containing high volume fly ash. Construction & Building Materials, 2015, 98: 810–819
18. Habert G, Roussel N. Study of two concrete mix-design strategies to reach carbon mitigation objectives. Cement and Concrete Composites, 2009, 31(6): 397–402
19. Dinakar P, Babu K G, Santhanam M. Mechanical properties of high-volume fly ash self-compacting concrete mixtures. Structural Concrete, 2008, 9(2): 109–116

20. Kumar A, Sen S, Sinha S. Machine learning based prediction models for the compressive strength of high-volume fly ash concrete reinforced with silica fume. Asian Journal of Civil Engineering, 2025, 26(4): 1683–1701
21. Kumar S, Rai B. Pulse velocity–strength and elasticity relationship of high volume fly ash induced self-compacting concrete. Journal of Structural Integrity and Maintenance, 2019, 4(4): 216–229
22. Sadrmomtazi A, Tahmouresi B, Kohani Khoshkbijari R. Effect of fly ash and silica fume on transition zone, pore structure and permeability of concrete. Magazine of Concrete Research, 2018, 70(10): 519–532
23. Dave N, Misra A K, Srivastava A, Kaushik S K. Experimental analysis of strength and durability properties of quaternary cement binder and mortar. Construction & Building Materials, 2016, 107: 117–124
24. Soleymani Ashtiani M, Scott A N, Dhakal R P. Mechanical and fresh properties of high-strength self-compacting concrete containing class C fly ash. Construction & Building Materials, 2013, 47: 1217–1224
25. Siddique R. Performance characteristics of high-volume class F fly ash concrete. Cement and Concrete Research, 2004, 34(3): 487–493
26. Kumar S, Rai B. Synergetic effect of fly ash and silica fume on the performance of high volume fly ash self-compacting concrete. Journal of Structural Integrity and Maintenance, 2022, 7(1): 61–74
27. Mugahed Amran Y H, Soto M G, Alyousef R, El-Zeadani M, Alabduljabbar H, Aune V. Performance investigation of high-proportion Saudi-fly-ash-based concrete. Results in Engineering, 2020, 6: 100118
28. Choudhary R, Gupta R, Nagar R. Impact on fresh, mechanical, and microstructural properties of high strength self-compacting concrete by marble cutting slurry waste, fly ash, and silica fume. Construction & Building Materials, 2020, 239: 117888
29. Hassan K E, Cabrera J G, Maliehe R S. Effect of mineral admixtures on the properties of high-performance concrete. Cement and Concrete Composites, 2000, 22(4): 267–271
30. Biswas R, Rai B. Efficiency concepts and models that evaluates the strength of concretes containing different supplementary cementitious materials. Civil Engineering Journal, 2019, 5(1): 18–32
31. Wongkeo W, Thongsanitgarn P, Ngamjarurojana A, Chaipanich A. Compressive strength and chloride resistance of self-compacting concrete containing high level fly ash and silica fume. Materials & Design, 2014, 64: 261–269
32. Shi C. Effect of mixing proportions of concrete on its electrical conductivity and the rapid chloride permeability test (ASTM C1202 or ASSHTO T277) results. Cement and Concrete Research, 2004, 34(3): 537–545
33. Yazıcı H. The effect of silica fume and high-volume class C fly ash on mechanical properties, chloride penetration and freeze–thaw resistance of self-compacting concrete. Construction & Building Materials, 2008, 22(4): 456–462
34. Asteris P G, Douvika M G, Karamani C A, Karamani A D, Chlichlia K, Cavaleri L, Daras T, Armaghani D J, Zaoutis T E. A novel heuristic algorithm for the modeling and risk assessment of the COVID-19 pandemic phenomenon. Computer Modeling in Engineering & Sciences, 2020, 125(2): 815–828
35. Bai J, Liu G R, Rabczuk T, Wang Y, Feng X Q, Gu Y T. A robust radial point interpolation method empowered with neural network solvers (RPIM-NNS) for nonlinear solid mechanics. Computer Methods in Applied Mechanics and Engineering, 2024, 429: 117159
36. Sun J, Liu Y, Wang Y, Yao Z, Zheng X. BINN: A deep learning approach for computational mechanics problems based on boundary integral equations. Computer Methods in Applied Mechanics and Engineering, 2023, 410: 116012
37. Wang Y, Sun J, Li W, Lu Z, Liu Y. CENN: Conservative energy method based on neural networks with subdomains for solving variational problems involving heterogeneous and complex geometries. Computer Methods in Applied Mechanics and Engineering, 2022, 400: 115491
38. Wang Y, Zhuang X, Timon R, Liu Y. AI for PDEs in solid mechanics: A review. Advances in Mechanics, 2025, 55(2): 231–287
39. Wang Y, Sun J, Bai J, Anitescu C, Eshaghi M S, Zhuang X, Rabczuk T, Liu Y. A physics-informed deep learning framework for solving forward and inverse problems based on Kolmogorov–Arnold Networks. Computer Methods in Applied Mechanics and Engineering, 2025, 433: 117518
40. Wang Y, Bai J, Lin Z, Wang Q, Anitescu C, Sun J, Sadegh Eshaghi M, Gu Y, Feng X Q, Zhuang X, et al. Artificial intelligence for partial differential equations in computational mechanics: A review. 2024, arXiv: 2410.19843
41. Bai J, Lin Z, Wang Y, Wen J, Liu Y, Rabczuk T, Gu Y T, Feng X Q. Energy-based physics-informed neural network for frictionless contact problems under large deformation. Computer Methods in Applied Mechanics and Engineering, 2025, 437: 117787
42. Wang Y, Bai J, Eshaghi M S, Anitescu C, Zhuang X, Rabczuk T, Liu Y. Transfer learning in physics-informed neurals networks: Full fine-tuning, lightweight fine-tuning, and low-rank adaptation. International Journal of Mechanical System Dynamics, 2025, 5(2): 212–235
43. Wang Y, Sun J, Rabczuk T, Liu Y. DCEM: A deep complementary energy method for linear elasticity. International Journal for Numerical Methods in Engineering, 2024, 125(24): e7585
44. Liu J, Guan J. A continuum and computational framework for viscoelastodynamics: II. Strain-driven and energy–momentum consistent schemes. Computer Methods in Applied Mechanics and Engineering, 2023, 417: 116308
45. Biswas R, Kumar M, Kumar D R, Samui P, T P, Rajak M K, Armaghani D J, Singh S. Application of novel deep neural network on prediction of compressive strength of fly ash based concrete. Nondestructive Testing and Evaluation, 2025, 40(10): 4638–4688
46. Khan K, Biswas R, Gudainiyan J, Amin M N, Qureshi H J, Arab A M A, Iqbal M. PCA-Based hybrid intelligence models for estimating the ultimate bearing capacity of axially loaded concrete-filled steel tubes. Materials, 2022, 15(18): 6477
47. Ahmad W, Ahmad A, Ostrowski K A, Aslam F, Joyklad P, Zajdel P. Application of advanced machine learning approaches to predict the compressive strength of concrete containing supplementary cementitious materials. Materials, 2021, 14(19): 5762
48. Xie T, Yang G, Zhao X, Xu J, Fang C. A unified model for

predicting the compressive strength of recycled aggregate concrete containing supplementary cementitious materials. Journal of Cleaner Production, 2020, 251: 119752
49. Liao J, Asteris P G, Cavaleri L, Mohammed A S, Lemonis M E, Tsoukalas M Z, Skentou A D, Maraveas C, Koopialipoor M, Armaghani D J. Novel fuzzy-based optimization approaches for the prediction of ultimate axial load of circular concrete-filled steel tubes. Buildings, 2021, 11(12): 629
50. Azimi-Pour M, Eskandari-Naddaf H, Pakzad A. Linear and non-linear SVM prediction for fresh properties and compressive strength of high volume fly ash self-compacting concrete. Construction & Building Materials, 2020, 230: 117021
51. Asadi Shamsabadi E, Roshan N, Hadigheh S A, Nehdi M L, Khodabakhshian A, Ghalehnovi M. Machine learning-based compressive strength modelling of concrete incorporating waste marble powder. Construction & Building Materials, 2022, 324: 126592
52. Mahmood M S, Elahi A, Zaid O, Alashker Y, Şerbănoiu A A, Grădinaru C M, Ullah K, Ali T. Enhancing compressive strength prediction in self-compacting concrete using machine learning and deep learning techniques with incorporation of rice husk ash and marble powder. Case Studies in Construction Materials, 2023, 19: e02557
53. Chen H, Asteris P G, Jahed Armaghani D, Gordan B, Pham B T. Assessing dynamic conditions of the retaining wall: Developing two hybrid intelligent models. Applied Sciences, 2019, 9(6): 1042
54. Eshaghi M S, Anitescu C, Thombre M, Wang Y, Zhuang X, Rabczuk T. Variational physics-informed neural operator (VINO) for solving partial differential equations. Computer Methods in Applied Mechanics and Engineering, 2025, 437: 117785
55. Samaniego E, Anitescu C, Goswami S, Nguyen-Thanh V M, Guo H, Hamdia K, Zhuang X, Rabczuk T. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications. Computer Methods in Applied Mechanics and Engineering, 2020, 362: 112790
56. IS 12269-1987. Bureau of Indian Standards. Ordinary Portland Cement, 53-grade—Specification. New Delhi: Bureau of Indian Standards, 1987
57. IS 4031-1996. Methods of Physical Tests for Hydraulic Cement. New Delhi: Bureau of Indian Standards, 1996
58. IS 15388-2003. Silica Fume—Specification. New Delhi: Bureau of Indian Standards, 2003
59. IS 3812-2013. Pulverized Fuel Ash—Specification, Part 1: For Use as Pozzolana in Cement, Cement Mortar and Concrete. New Delhi: Bureau of Indian Standards, 2013
60. IS 2386-2021. Methods of Test for Aggregates for Concrete—Part 3: Specific Gravity, Density, Voids, Absorption and Bulking. New Delhi: Bureau of Indian Standards, 2021
61. IS 383-2016. Coarse and Fine Aggregate for Concrete—Specification. New Delhi: Bureau of Indian Standards, 2016
62. IS 4032-1985. Method of Chemical Analysis of Hydraulic Cement. New Delhi: Bureau of Indian Standards, 1985
63. IS 10262-2019. Concrete Mix Proportioning—Guidelines. New Delhi: Bureau of Indian Standards, 2019
64. EFNARC-2005. Specification and Guidelines for Self-Compacting Concrete. Farnham: European Federation of National Associations Representing for Concrete, 2005
65. IS 516-2021. Methods of Tests for Strength of Concrete—Part 1: Compressive Strength of Hardened Concrete. New Delhi: Bureau of Indian Standards, 2021
66. Wang R. AdaBoost for feature selection, classification and its relation with SVM, a review. Physics Procedia, 2012, 25: 800–807
67. Schapire R E. Empirical Inference. Berlin: Springer, 2013, 37–52
68. Natekin A, Knoll A. Gradient boosting machines, a tutorial. Frontiers in Neurorobotics, 2013, 7(21): 1–21
69. Otchere D A, Ganat T O A, Ojero J O, Tackie-Otoo B N, Taki M Y. Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions. Journal of Petroleum Science Engineering, 2022, 208: 109244
70. Prokhorenkova L, Gusev G, Vorobev A, Dorogush A V, Gulin A. Catboost: Unbiased boosting with categorical features. Advances in neural information processing systems, 2018, 31: 6638–6648
71. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: Association for Computing Machinery, 2016, 785–794
72. Shahani N M, Zheng X, Liu C, Hassan F U, Li P. Developing an XGBoost regression model for predicting Young's modulus of intact sedimentary rocks for the stability of surface and subsurface structures. Frontiers in Earth Science, 2021, 9: 761990
73. Aliyari M. A new meta-heuristic algorithm of giant trevally for solving engineering problems. Journal of Algorithms and Computation, 2023, 55(1): 37–51
74. Bhol S, Sahu N C, Kabat S R. Optimal scheduling of renewable sources based micro grid with PV and battery storage using giant trevally optimizer. Tuijin Jishu/Journal of Propulsion Technology, 2023, 44(4): 5205–5222
75. Hashish M S, Hasanien H M, Ullah Z, Alkuhayli A, Badr A O. Giant trevally optimization approach for probabilistic optimal power flow of power systems including renewable energy systems uncertainty. Sustainability, 2023, 15(18): 13283
76. Sadeeq H T, Abdulazeez A M. Car side impact design optimization problem using giant trevally optimizer. Structures, 2023, 55: 39–45
77. Zhang Q, Bu X, Zhan Z H, Li J, Zhang H. An efficient optimization state-based coyote optimization algorithm and its applications. Applied Soft Computing, 2023, 147: 110827
78. Venkatasatish R, Chittathuru D. Coyote optimization algorithm-based energy management strategy for fuel cell hybrid power systems. Sustainability, 2023, 15(12): 9638
79. Thom de Souza R C, de Macedo C A, dos Santos Coelho L, Pierezan J, Mariani V C. Binary coyote optimization algorithm for feature selection. Pattern Recognition, 2020, 107: 107470
80. Wang Z, Liu S, Liang W, Liu J, Zhou Y, Lei K, Gao Y, Ou W. Predictive modeling of compressive strength in tailings concrete using explainable machine learning approaches. Results in Engineering, 2025, 27: 105516
81. Xie L, Han T, Zhou H, Zhang Z R, Han B, Tang A. Tuna swarm optimization: A novel swarm-based metaheuristic algorithm for global optimization. Computational Intelligence and Neuroscience, 2021, 2021(1): 9210050

82. Wang W, Tian J. An improved nonlinear tuna swarm optimization algorithm based on circle chaos map and levy flight operator. Electronics, 2022, 11(22): 3678
83. Li W, Pandit R K. Data-centric predictive control with tuna swarm optimization-backpropagation neural networks for enhanced wind turbine performance. Renewable Energy, 2024, 237: 121821
84. Wang X, Liu Y, Xue G, Bai F, Huang S. Fast non-dominated sorting tuna swarm optimization algorithm (FNS-TSO): Time-energy-impact multi-objective optimization of underwater manipulator trajectories. Journal of Marine Science and Engineering, 2025, 13(5): 916
85. Tabani A, Sharma A, Biswas R, Sivenas T, Asteris P G. Revealing the nature of ultra-high-performance concrete using computational intelligence. Construction & Building Materials, 2025, 492: 143082
86. Tabani A, Biswas R. Assessment of compressive strength of ultra-high-performance concrete using advanced machine learning models. Structural Concrete, 2025, 70076
87. Kumar D R, Samui P, Burman A, Biswas R, Vanapalli S. A novel approach for assessment of seismic induced liquefaction susceptibility of soil. Journal of Earth System Science, 2024, 133(3): 128
88. Biswas R, Li E, Zhang N, Kumar S, Rai B, Zhou J. Development of hybrid models using metaheuristic optimization techniques to predict the carbonation depth of fly ash concrete. Construction & Building Materials, 2022, 346: 128483
89. Taylor K E. Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research, 2001, 106(D7): 7183–7192
90. Kumar M, Biswas R, Kumar D R, Samui P, Kaloop M R, Eldessouki M. Soft computing-based prediction models for compressive strength of concrete. Case Studies in Construction Materials, 2023, 19: e02321