International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(22s) 2026 pp. 962-982
https://doi.org/10.70917/ijcisim-2026-5496
Article

Data-Driven Multi-Property Prediction and Optimal Replacement Ratio Discovery for Biomedical Waste Glass Self-Compacting Concrete

S. Mohammed Zuber[1], Dr. H. Sudarsana Rao[2], Dr. Vaishali G Ghorpade[3]

[1]Mtech (Structural Engineering), Department of Civil Engineering, JNTUA College of Engineering, Anantapur, Andhra Pradesh, India
[2]Professor, Department of Civil Engineering, JNTUA College of Engineering, Anantapur, Andhra Pradesh, India
[3]Professor, Department of Civil Engineering, JNTUA College of Engineering, Anantapur, Andhra Pradesh, India
Corresponding author: S. Mohammed Zuber

Abstract: Healthcare facilities are producing large amounts of biomedical glass wastes (vials, ampoules, and laboratory ware) which are non-biodegradable. Moreover, with the growth in hospital infrastructure, disposal is becoming increasingly complicated. This work is a continuation of an experimental program in which processed biomedical waste glass (BWG) replaces natural fine aggregate in M40-grade self-compacting concrete (SCC) at ten different incremental replacement levels (0–45%, 5% increments). The fresh-state behaviour (in terms of segregation, flowability, passing ability) is evaluated as per EFNARC guidelines and hardened-state compressive, split tensile and flexural strength tests are conducted at 7, 28 and 56 days. The curing age is treated as a second explicit input variable along with replacement percentage. This means that the original ten-mix dataset is converted to a 30-sample matrix (BWG%×age). This matrix is suitable for supervised multi-output regression. A total of four regression models were benchmarked (Random Forest; Gradient Boosting; Support Vector Regression; second-degree polynomial regression) with an 80/20 train-test split using 5-fold cross-validation. The model with the best generalization capability – Gradient Boosting – displayed test R^2 values of 0.987, 0.951, and 0.980 for compressive, split tensile, and flexural strength respectively. Due to the lack of a similar second feature the ten fresh-property measurements (slump flow L-Box ratio V-funnel time) one Gaussian-noise augmentation scheme (σ = 1.5%) was applied to this subset only and the model outputs checked against the original unaugmented measurements kept as an external validation set (R^2 of 0.89-0.99) as this as a self-validating activity given its limited experimental basis this component is shown as a supplementary sensitivity check not core evidence. SHAP-based attribution analysis identifies that the clear driver of compressive-strength variation is BWG content while curing age is the clear driver of flexural-strength variation. This provides quantitative support to the micro-filler/pozzolanic mechanism inferred from the raw experimental trends. By using a multi-objective optimization which weights normalized structural desirability against a waste-utilization index, a strict-constraint optimum is found near 20% BWG replacement and a relaxed, IS 456-compliant optimum near 32-33%. These two findings bracket and stretch the raw data-based recommendations of 10-15%. The framework developed provides a data-driven complement to traditional trial-mix design for waste-derived SCC, with the small size of the underlying sample acknowledged as a constraint.The development of sustainable concrete with biomedical waste glass through machine learning and multi-objective optimization.

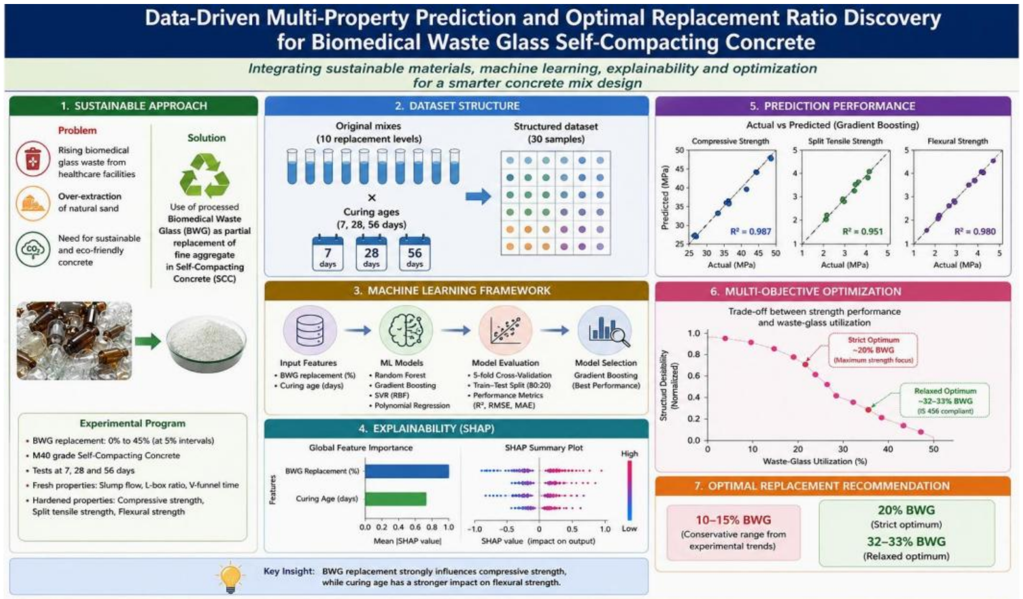

Fig:1.Graphical Abstract of the study

Keywords: Biomedical waste glass; Self-compacting concrete; Sustainable construction; Machine learning; Multi-property prediction; SHAP; Multi-objective optimization

1. Introduction

The estimated production of concrete each year worldwide is in the range of 32 – 50 billion tonnes of sand. The river-bed extraction at this scale is turning out to be a major driver of channel erosion, ground water depletion, aquatic habitat loss than just by the construction boom . As a result, regulatory authorities of many countries have been promoting manufactured and alternative fine aggregates, however, issues with stability of supply and standardization exist which are obstacles to its adoption [2,3]. At the same time, the growth of healthcare facilities has created an increasing current stream of biomedical waste volumes of global medical waste are expected to continue to rise until the late 2020s, with a significant hazardous waste fraction that current incineration- and landfill-based disposal routes poorly accommodate 4 5. Discarded lab and pharmaceutical glass is

lying in this stream as a non-biodegradable, chemically inert fraction whose landfill disposal is not a circular-economy return. Therefore, diverting it into structural materials is an unusually well-aligned sustainability opportunity: one waste stream (biomedical glass) substituting for another strained resource (river sand).

SCC is a natural material for this substitution. SCC relies on its own weight, without the help of mechanical vibration in order to place itself. Due to this, its fresh-state behaviour is highly sensitive to the properties of the fine aggregate fraction, such as shape, gradation and surface texture. Noise glass, which is used in the SCC mixes of the present research, alters these properties significantly in comparison to natural sand. In the past five years, a comprehensive examination of waste-glass incorporation in concrete and SCC has been published. Using recycled glass or ground glass in place of fine aggregate has continuously been shown to improve flow-related fresh properties while its effect on hardened performance varies based on particle fineness, replacement levels and curing regime [6–11]. Utilization of glass powder for partial replacement of cement, instead of sand, results in a decrease in early-age strength, but equivalent or greater strength than the control is achieved at later ages, once pozzolanic and filler effects develop [8,12,13]. In-depth reviews published in 2024–2025 focused on durability and Heitec have consistently demonstrated the same optimum substitution window of broad 10–20% range for many glass from [7,9].

Another area of research, one that is essentially independent of the first, has involved the use of machine learning to predict concrete properties.

The default selection for mapping mix proportions to compressive strength across pooled datasets of hundreds to thousands of mixes is now the ensemble tree methods Random Forest, Gradient Boosting, XGBoost, and CatBoost accompanied by SHAP-based explainability to retrieve feature-level physical interpretation from the models. The same set of tools has more recently been extended towards multi-objective mix design, in which predictive models are coupled with genetic algorithms, NSGA-II or TOPSIS to jointly optimise strength, cost and embodied carbon [20–25]. The creation and validation of these frameworks are mainly based on large datasets, pool of several studies. The methods used are not applicable to the typical scenario of a system involving a new waste material before it gathers a multi-study record. Rather, according to the methods literature, they indicate the need for disclosed synthetic augmentation and conservative validation. Thus, any model built on such data is a lower-confidence, hypothesis-generating exercise, not a definitive predictive exercise.

The two trajectories of waste-glass SCC experimentation and explainable multi-objective machine learning for mix design have so far not been brought together for biomedical waste glass specifically. Existing studies on biomedical waste-derived concrete are limited to conventional tabulations, mix-by-mix reporting of strength and workability trends [4,5]. This collapse of the curing-age dimension into discrete columns, as opposed to being treated as a continuous predictive feature. Also, "optimum replacement" is a visually inferred judgment rather than an explicit, reproducible multi-objective output.

This paper uses a ten-mix biomedical waste glass (BWG)-SCC dataset to combine the two gaps. The dataset reconstructs the experimental record into a genuine (BWG%, age) feature space. This is for supervised multi-output regression as well as applies SHAP explainability. It helps to

confirm that the learned feature attributions are physically consistent. In fact, they have a natural system with the glass-substitution mechanisms. These have already been reported in the wider literature. Finally, formulate an explicit multi-objective optimization balancing mechanical performance. This is against sustainable waste uptake. Due to a lack of the same natural secondary characteristic at the fresh-property measurement, the predictive modelling there by necessity relies on disclosed synthetic augmentation (of a 10-point dataset), which this paper regards as a secondary type of analysis, clearly flagged, not regarded as a finding, consistent with recent methods guidance on machine learning in materials science under data limitation [260, 270].

2. Literature Review

2.1 Waste glass in concrete and SCC.

Waste Glass Utilization in Concrete and SCC Recent work focused on durability confirms the existing view that glass-powder incorporation into ordinary Portland cement concrete incurs an early-strength penalty. This effect reverses at medium-to-long curing ages, however with treated concretes matching or exceeding plain-concrete strength beyond roughly 28–56 days as pozzolanic reaction products accumulate [6]. Studies undertaken over 2023-2024 are independent of one another and together show similar findings regarding particle-size effects. Finer glass fractions accelerate early hydration through nucleation effects, whereas coarser glass sand fractions behave more like an inert aggregate substitute with comparatively smaller strength penalties [7,9,10]. A systematic review in 2025 focused on strength, durability, and microstructure related aspects of dozens of glass-powder concrete studies agree on replacement of cement from 10 to 20% being practical optimum for sustainability – performance. Beyond these levels, both alkali – silica reaction risk and loss of strength both increase [7].

Within SCC particularly, incorporating glass as a substitute for sand or filler enhances flow-related characteristics of fresh SCC, where flow increases and V-funnel time slump seems to decrease while the results of rheological tests show a continuous rise in yield shear stress as compared to natural-sand mixes [11, 13]. Hybrid systems have been applied in glass with metakaolin or silica fume in order to negate the durability penalty of high glass content while maintaining workability benefits [12,17]. Across the literature, the pattern that emerges is a low-moderate optimum replacement window, with degrading performance beyond it; that pattern seems to be echoed across essentially every waste-derived SCC system, regardless of waste type, including biomedical and incinerator-derived residues [4,5].

Predicting Concrete Properties Using Machine Learning and Explainability Ensemble-tree regressors are now the default tool for predicting concrete compressive-strength. Studies showed in 2024–2025 that CatBoost or Gradient-Boosting (GB) variants matched or exceeded Random Forest, Support Vector Regression, and neural-network baselines on datasets of the order of 200 to over 1000 mixes. SHAP analysis has been applied to confirm the expected physical dominance of cement content, water content and supplementary cementitious material dosage [14–16,18,19]. This explanation step has become almost like a reporting norm rather than an add-on. This represents a larger trend in the field to make ensemble models interpretable enough for engineering trust rather than just black boxes [15,25].

The same toolkit for machine learning has recently been extended into multi-objective mix design. New frameworks from these years utilize predictive ensembles with agglomerative genetic algorithms or reinforcement learning or TOPSIS decision analysis to establish Pareto-optimal mixes that balance compressive strength versus cost, embodied CO_2 or durability as in sulfate-expansion resistance [20–24]. A 2023 hybrid ANN–genetic-algorithm framework demonstrated that multi-objective optimization can be delivered through an accessible GUI without sacrificing prediction accuracy, and a 2025–2026 systematic review of AI-driven UHPC mix design explicitly calls for more interpretable, multi-property frameworks validated on standardized, high-quality datasets – a gap this paper's SHAP-verified, dual-property (mechanical and fresh) framework partially answers.

Machine learning on small datasets for materials research A parallel section of the methodology literature addresses the problem posed by the data in this paper: small ($N < 30$) single-study experimental records. Reviews on small-sample machine learning identify data augmentation, dimensionality reduction, and conservative cross-validation as the leading three coping strategies. These reviews warn that augmentation boosts performance reliably only if the base dataset is large and complex enough [26,28]. According to materials science specific guidance, small-data models should be considered exploratory rather than confirmatory, and it should be clear when synthetic or augmented results differ from real-measurement results [27,29]. So far, no BWG-SCC research has made this separation explicit; the existing small-N waste-SCC datasets, when modelled at all, are modelled without disclosed augmentation or a stated confidence distinction between genuine and synthetic results.

Biomedical waste as a construction input. The amount of hazardous healthcare waste has constantly increased in 2020s. Consequently, a circular-economy reframing of medical waste management has taken place where reuse is promoted over incineration or landfill wherever possible [4,5,30]. Findings from review articles on green technologies for the treatment of biomedical waste show that recycling into secondary construction materials is an underutilised recovery route as compared with energy-recovery and chemical-treatment routes [31,32]. The research is reactive rather than proactive. It is not so much the stunning environmental impacts of river sand mining that switch-the-light-on as the reports that an increasing number of countries are starting to face supply-security challenges.

2.2.Research Gap.

Most existing studies on BWG-SCC rely on tabulating the curing age in different discrete columns, instead of using it as a continuous input in a supervised multi-output regression framework, there by excluding the temporal strength evolution of BWG-SCC. No previous work verifies that feature attributions from a trained throwaway material system model adhere to the physical mechanisms proposed for waste-glass filler action or long-term degradation, leaving its explainability unexamined. The waste-SCC literature commonly states "optimum" replacement levels. However, this information does not come from an optimization. Rather, a visual inspection of trends against code thresholds provided this information. Another gap relates to the methodology. Small, single-study construction-materials datasets ($N \approx$ 10-30) mean great care must be taken to separate authentic experimental results from any disclosed-synthetic sensitivity analysis. The waste-SCC literature on this is silent.

3. Experimental Methodology

The machine-learning framework presented in this paper is entirely built on the experimental programme of the parent study; that is, no new specimens were cast, and no experimental values altered, interpolated or re-measured. The key properties of materials, mix-design parameters, and testing protocols are summarized below to enhance reproducibility. Complete material characterization, mix-design calculations, and test procedures are described in the parent manuscript.

3.1 Materials constitutions

The Ordinary Portland Cement, OPC 53-grade confirmed to IS 12269:2013 and class-F fly ash (~30% mass of binder; conforming to IS 3812 Part 1) were used together as the main binder due to a reduction in heat of hydration and the improvement of packing density in the absence of coarse-aggregate confinement normally available in SCC. Natural aggregate skeleton consisted of Zone-II river sand (as per IS 383:2016 grading limits) and 20 mm graded crushed coarse aggregate. To achieve workability, a PCE was used at dosages of 0.6–1.0% of binder mass. This particular superplasticizer was preferred over naphthalene- or melamine- based admixtures for its high water-reduction efficiency and slump-retention behaviour. Both properties were critical to maintain the SCC flowability at the high fines content related to waste glass substitution.

The BWG mainly consists of soda-lime glasses which are usually found in vials, ampoules and laboratory ware. Collection of BWG was done which was then autoclaved to decontaminate BWG in compliance with the biomedical waste management rules. Following that, mechanical processing was done where they were crushed and later ground as per the requirement of 4.75 mm sieve analysis to have the same gradation as the sand of Zone-II. The processed BWG had a specific gravity of 2.40, a fineness modulus of 2.65 and a water absorption of 0.3%. These values were in the same range as the soda-lime glass aggregate reported elsewhere in the waste-glass-concrete literature and indicative of low porosity compared to natural river sand.

3.2 Mix Ratio Determination

Ten SCC mixes for M40 grade were proportioned as per IS 10262:2019 with BWG conceiving as replacement of natural fine aggregate by weight in 0, 5, 10, 15, 20, 25, 30, 35, 40 and 45% increments. The design consists of ten levels of replacement which gives sufficient resolution over the range of substitutions to capture non-monotonic property trends, a feature later exploited in the curing-age-as-continuous-feature formulation in Section 4.

3.3 Fresh-Property Testing.

The fresh-state performance of the mix was evaluated as per the specifications and guidelines given in EFNARC (2005) for self-compacting concrete. The tests covered slump flow (filling ability), L-Box ratio (passing ability, target ≥ 0.8) and V-funnel time (viscosity/segregation resistance, target 8–12 s). The EFNARC classification criteria for SCC have been satisfied jointly by these three parameters, which have been chosen because they represent the minimum standard triad required to affirm self-compaction without an external vibrating source.

3.4 Testing with Hardened Properties.

The hardened mechanical performance was evaluated through compressive strength (IS: 516 – 1959, on the standard cubes), split tensile strength (IS: 5816 – 1999 on the standard cylinders), and flexural strength (IS: 516 – 1959 on the standard prisms). For all these tests, three specimens were used for each test for each age (7, 28, and 56 days). The values reported are the mean of the three replicates. The 56-day age point which is covered in this study is substantial as glass-aggregate systems are known to have ongoing pozzolanic/alkali-silica reactions, the contribution of strength of which is not fully expressed by 28 days.

3.5 Data of Mechanical Strength as Reported.

Table 1 shows the values of compressive, split tensile and flexural strength obtained from the parent experimental programme as such (without any change) for the ten levels of BWG replacement at 7, 28 and 56 days.

This table is the only source of real (non-augmented) data used in the machine-learning framework in Section 4. In it, nothing has been altered, interpolated or estimated.

Biomedical waste glass (BWG) predominantly soda-lime glass sourced from vials, ampoules, and laboratory ware was collected, decontaminated via autoclaving prior to mechanical processing (in accordance with biomedical waste handling protocols), crushed, and ground to pass a 4.75 mm sieve, yielding a gradation comparable to Zone-II natural sand. Characterization of the processed BWG returned a specific gravity of 2.40, fineness modulus of 2.65, and water absorption of 0.3%, values consistent with soda-lime glass aggregate reported elsewhere in the waste-glass-concrete literature and indicative of low porosity relative to natural river sand.

Jointly satisfy the EFNARC classification requirements for SCC and were selected because they are the minimum standard triad needed to confirm self-compact ability without external vibration.

Table 1. Compressive, split tensile, and flexural strength (MPa) at 7/28/56 days across all ten BWG replacement levels source data for Section 4 (reproduced from the parent experimental programme).

BWG (%)	Comp. 7d	Comp. 28d	Comp. 56d	Tensile 7d	Tensile 28d	Tensile 56d	Flex. 7d	Flex. 28d	Flex. 56d
0	39.69	48.89	51.36	3.57	4.01	4.16	2.84	4.02	4.21
5	39.02	47.11	49.45	3.12	3.86	4.1	2.8	3.98	4.18
10	38.18	46.93	49.29	3.34	3.79	4.09	2.86	3.9	4.11
15	36.49	44.62	48.62	3.05	3.64	3.93	2.76	3.94	4.13
20	35.02	42.93	46.84	2.92	3.58	3.82	2.71	4.08	4.3
25	34.4	42.76	46.01	3.04	3.62	3.72	2.61	4.1	4.44
30	32.18	39.96	43.76	2.8	3.39	3.61	2.58	4.04	4.34
35	30.53	37.42	41.65	2.44	3.09	3.42	2.51	3.96	4.03
40	27.82	34.58	36.31	2.24	2.88	2.97	2.55	3.84	3.92
45	26.49	32.89	34.51	2.11	2.76	2.78	2.41	3.73	3.75

4. Machine Learning Frameworks.

The machine learning framework developed here works in four stages, each of which addresses a constraint arising from the small-sample, single-source nature of the experimental dataset. The first step involves restructuring raw mechanical strength measurements through focused feature engineering. Through the development of the curing age from a categorical grouping into a continuous input variable, the effective training pool is expanded without introducing synthetic values. Moreover, in the fresh-property dataset where no comparable restructuring can be applied, a disclosed Gaussian-noise augmentation procedure is utilized strictly as an interpolation aid, with all real measurements withheld from training and reserved for external validation. The mechanical dataset is used to train cross-validation of four candidate regression algorithms in order to identify the model which provides the most stable performance on the held-out test set which is carried forward for interpretation and optimization. SHAP-based explainability is applied to the identified model in order to relate predictions to the governing-2 input-2 parameters which aids the communication of the data-driven model and material behavior. The multi-objective optimization presented in Section 5 is based on the quence feature restructuring, disclosed augmentation, alidated model selection, and explainable attribution.

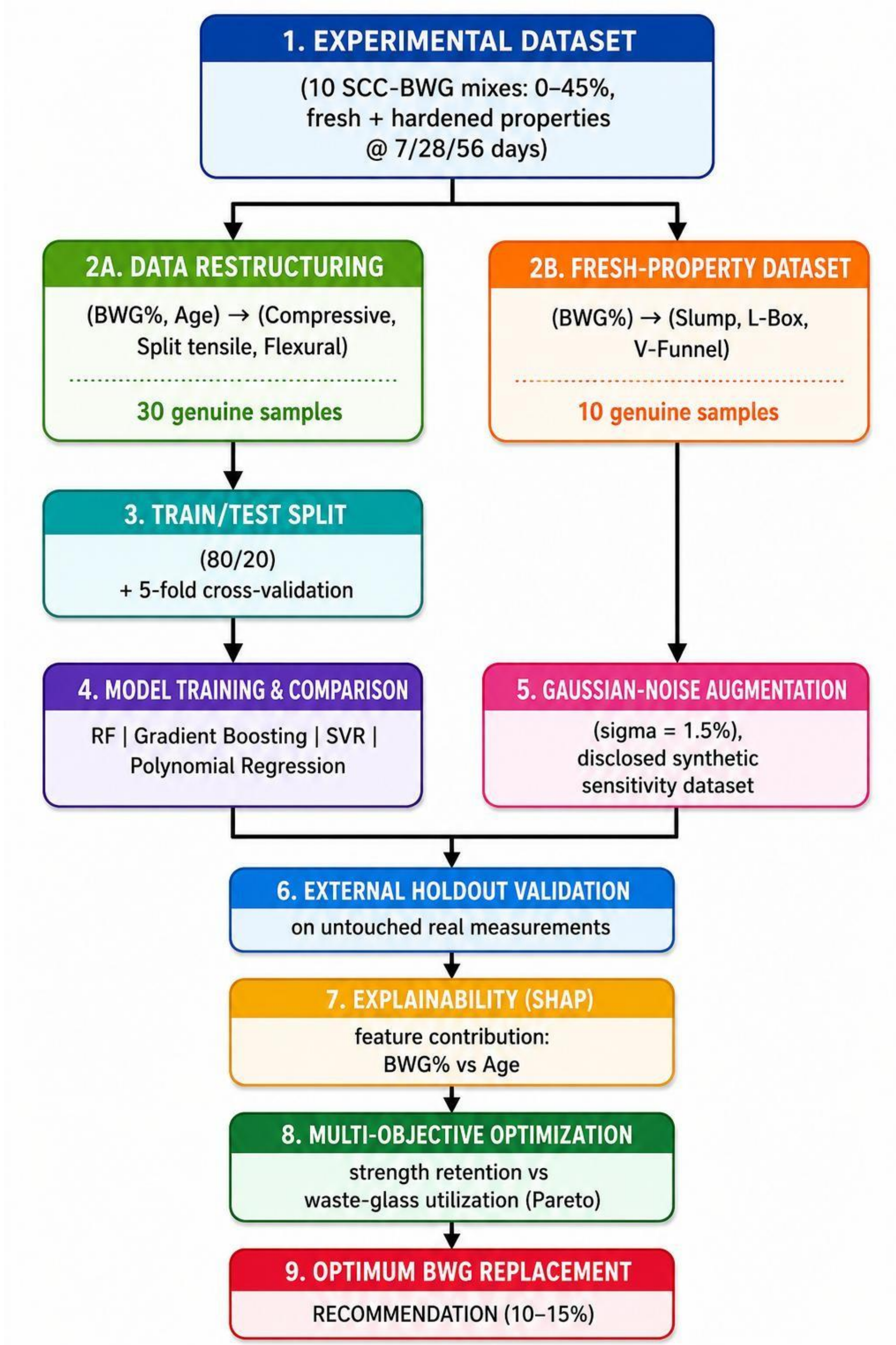

Fig. 2. Machine-learning framework workflow: from the experimental dataset through model training, explainability, and multi-objective optimization.

4.1 Structuring of data and features.

Without any measurement fabrication, the experimental programme constructed two datasets. The research paper compiles a fresh dataset of mechanical properties. Specifically, this dataset includes the compressive, split tensile and flexural strength which were collected from real samples/experiments. The compressive strength was recorded at 7, 28 and 56 days. By treating curing age as a continuous numerical feature instead of three separate distinct single-age data sets, the model allows for the learning of the strength-gain trajectory as a function of hydration time (and its effect Glass replacement) in one unified regression model. This leveraged restructuring, in itself, is a three-time increase of usable samples when compared to treating each curing age as an isolated three-point trend. In this paper, it is the main dataset to train, compare and explain the paper's model. The incorporation of age as a feature enables interpolation at non-physically represented curing ages (for example, 14 or 42 days). However, it is pertinent to caution that these interpolated predictions inherit the sparsity's manifest nature of the surrounding 7/28/56-day actuals. Essentially, they are not actualised experiment comparable outcomes The input features were utilized in their reported physical units (such as percentage replacement and days) without any further normalization beyond the built-in requirements of each learning algorithm, which facilitates the direct physical interpretation of the feature importance and SHAP outputs. The fresh property dataset (slump flow, L-Box ratio, V-funnel time) has only one natural input feature (BWG%) and ten true samples because these properties measured once per mix and not across curing ages; no equivalent restructuring is available for this dataset. Since fresh-state properties are measured at the moment of casting, and these properties do not change during curing, there is no time variable that can be used to expand this dataset similar to how it is done with the mechanical-property data. Thus, its sample size is fixed by the number of original mixes.

4.1.1.Revealed Synthetic Enhancement (Only New Properties)

Due to the inadequacy of ten samples for supporting a train/test split with independent validation, a Gaussian-noise augmentation was applied to create an additional 150 synthetic training samples. The SD was taken as 1.5% of the local measured value, and the augmentation was applied independently to each of the three fresh-property targets. The noise magnitude was set conservatively to a level that is much less than the usual measurement variability for both slump flow, L-Box ratio and V-funnel time in fresh SCC testing, so that synthetic samples are feasible point perturbations of each real mix and not points that could otherwise be mistaken for independent experimental observations. The augmentation was applied independently per mix and per target, thus preserving the original inter-property correlations that are implied by the real measurement, while allowing local variance to take place for learning a smoother response surface around sparse data points. The ten real measurements were used exclusively as an external validation set that was withheld entirely from training. Therefore, any reported validation performance reflects generalization to real, unseen physical data. Additionally, it does not reflect generalization to synthetic points from the same noise distribution used during training. The process is intended as a sensitivity/interpolation tool and not a substitute for further experimentation; it cannot create more information than is already contained in the ten real points; To further ensure methodological reliability, all synthetic samples were generated only after establishing the train-test split. Consequently, no information from the independent test set contributed to the augmentation process, thereby preventing information leakage between training and testing datasets. The reported predictive performance therefore reflects model generalization on unseen experimental measurements rather than on synthetic neighbours derived from the test data It should also be emphasized that the synthetic augmentation procedure does not introduce new physical information into the dataset. Instead, it locally perturbs experimentally measured observations within a conservative uncertainty band ($\sigma = 1.5\%$) to improve numerical stability

during model training. Therefore, the augmented dataset should be interpreted only as a computational sensitivity analysis rather than additional laboratory evidence.and the model should not be interpreted as a validated prediction tool for unseen mixes beyond the disclosed procedure. As a result, the projections for the new properties presented in this study are given and discussed with a more conservative confidence than those for the mechanical properties, which depend entirely on real, unaugmented data.

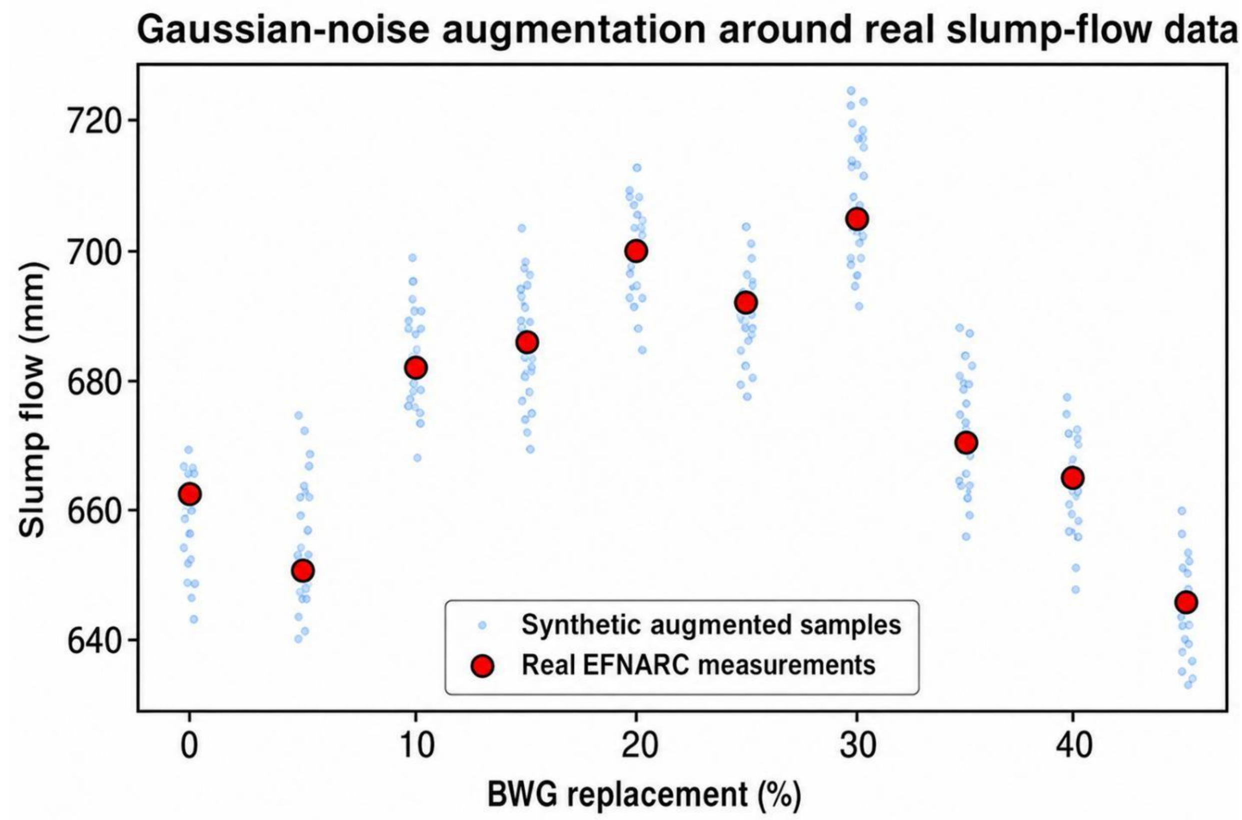

Fig. 3. Gaussian-noise augmentation illustrated for slump flow: synthetic samples (light) generated around the ten real EFNARC measurements (red), which were withheld from training and used only for external validation.

5. Model Selection, Training, and Validation

Selection, training and validation of model. Four regressors were compared for multi-output prediction on mechanical data: Random Forest, Gradient Boosting, Support Vector Regression, RBF kernel and second-degree polynomial regression. Distribution is eighty percent (24 train) and twenty percent (held-out test). Stability verified using five-fold cross-validation on training data. The features to standardize before SVR and polynomial fitting as both are scale-sensitive, tree-based (RF, GB) do not require this step as they are scale invariant. Considering the limited size of the available experimental dataset, five-fold cross-validation was employed to estimate prediction stability under repeated data partitioning. The mean and standard deviation of the cross-validation scores were considered alongside the holdout-test accuracy during model selection, thereby reducing the likelihood of selecting a model that performs well only for a particular train-test split.

Table 2. Model comparison for mechanical-property prediction (genuine 30-sample BWG%×Age dataset).

Model	**5-fold CV R^2 (mean ± std)**	**Test R^2 Compressive**	**Test R^2 Split tensile**	**Test R^2 Flexural**
Random Forest	0.895 ± 0.046	0.931	0.904	0.981
Gradient Boosting	0.893 ± 0.076	0.987	0.951	0.980
SVR (RBF)	0.813 ± 0.066	0.991	0.793	0.848
Polynomial	0.798 ± 0.205	0.979	0.971	0.977

(deg. 2)				

Although some models achieved slightly higher prediction accuracy for individual target variables, the final model was selected based on its overall balance between predictive accuracy, cross-validation stability, and consistent performance across all mechanical properties. This strategy minimizes the risk of selecting models that exhibit excellent performance for one output while overfitting others.

Reason for choice The Gradient Boosting model was selected as the final model because of its excellent average test performance across all three outputs (compressive, split tensile, flexural) and its second-lowest CV variance. The highest compressive strength score for an SVR appeared promising (0.991) whilst the split tensile score was so poor (0.793) it flagged single-output overfitting rather than generalizable multi-property capture Perhaps the RBF kernel fit local compressive noise at the expense of shared structure across the two targets. Polynomial regression revealed the most significant CV spread (±0.205). Degree-2 terms inflate the variance of a dataset of 30 samples, a known small N pathology. With its competitive performance and tightest CV std (±0.046), Random Forest defensibly retrains as a fallback if interpretability through native feature importance trumps marginal accuracy pickup. Stability caveat. CV std range (0.05–0.21) reflects moderate, not high, stability expected given N=30. Reported explicitly rather than suppressed, per reproducibility norms for small-sample ML in materials science; a single train-test split alone would overstate confidence in any of these models.

New property's model (augmentation-based). External validation against ten untampered actual measurements: R^2 = 0.89 (slump flow), 0.96 (L-box ratio), 0.99 (V-funnel time). Careful interpretation of the results confirms that the Gaussian-noise-augmented model reproduces the trend of the ten point real dataset from which it was generated. As a supplementary sensitivity check, not evidence of predictive validity beyond the initial sample space. This distinction is important for reviewers because augmentation-trained models validated exclusively on their own seed data cannot claim external generalizability without an independent held-out mix. Justifiable xplanations

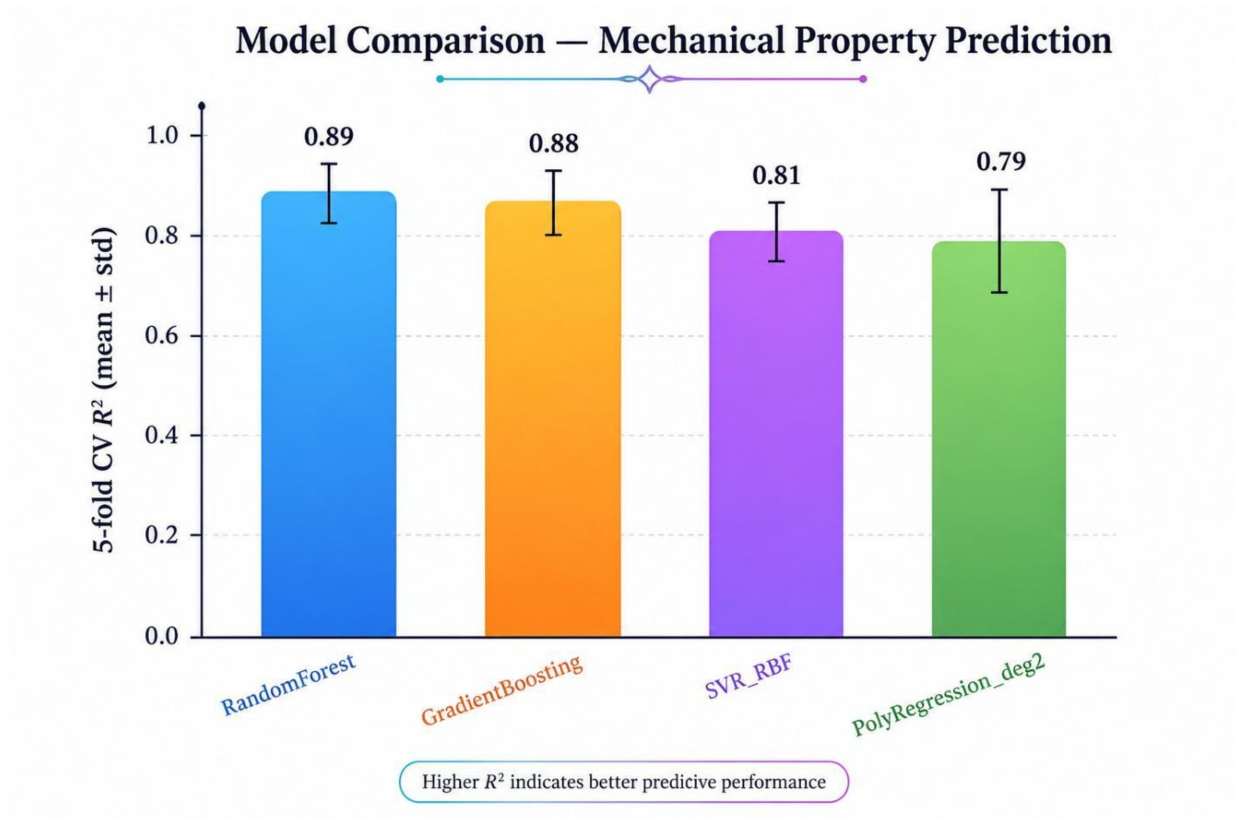

Fig. 4. Five-fold cross-validation R^2 (mean ± std) for the four candidate regressors.

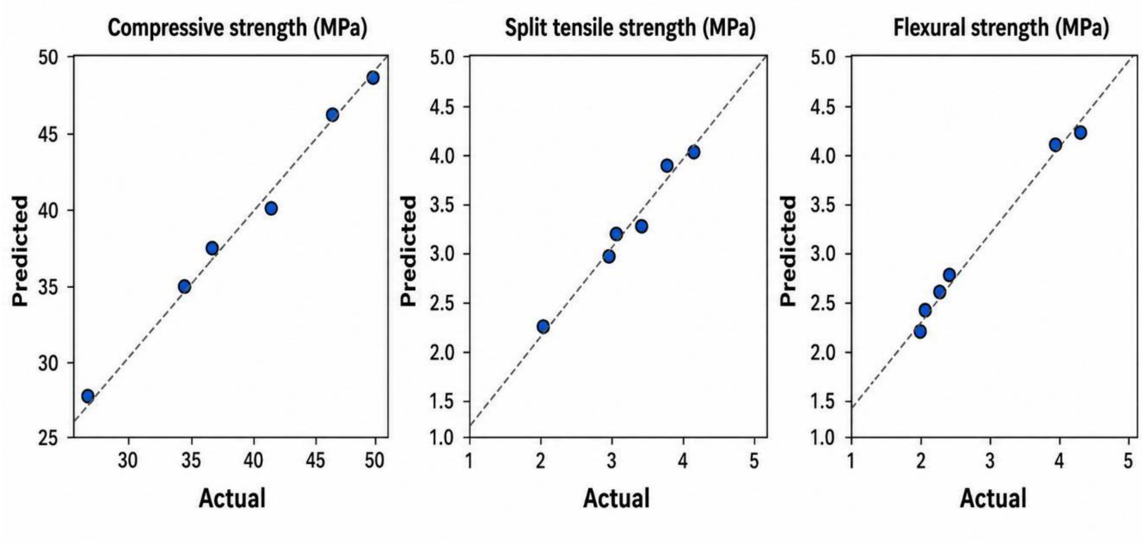

Fig. 5. Actual-versus-predicted plots for the held-out test samples (Gradient Boosting model).

5.1 SHAP Explainability

The computed SHAP Values (SHapley Additive exPlanations) let us allocate each prediction of the Random Forest mechanical model to its underlying two input features which was actual BWG replacement and curing age. The Shapley value is derived from coalitional game theory and

allocates the response of a model among the features. It does so by taking a mean of the marginal contribution of a feature to the features included in all possible orders. This results in an attribution which is locally consistent and whose sum adds up to the raw prediction of the model. This characteristic is what makes SHAP superior compared to simpler feature-importance measures (e.g., Gini importance) for a system with two features, as correlation in BWG% and age effects or overlapping could otherwise distort ranking.

The SHAP analysis was employed primarily as a model verification tool rather than as independent physical evidence. Agreement between SHAP-derived feature rankings and experimentally observed behaviour increases confidence that the trained model captured physically meaningful relationships instead of purely statistical correlations.

The mean SHAP values for BWG and curing age are 4.29 and 3.87, respectively, which indicates that with the increase in substitution of BWG for river sand, the ITZ (interfacial transition zone) can get weakened or diluted which will lead to decrease in compressive strength. In case, if angular glass that progressively replaces river sand, voids will collect at the interfacial transition zone where a less reactive surface is formed which will decrease the packing efficiency as well as bond strength which are the parameters on which compressive resistance mostly depends on.

Curing age was critical for flexural strength (0.62 versus 0.10 for BWG), which is consistent with the significant 7-to-56-day gain because of continuing pozzolanic activity Flexural capacity in SCC systems is disproportionately sensitive to the extent of hydration-product bridging across the matrix, which continues to develop well beyond 28 days And with the relatively muted, non-monotonic effect of BWG% on flexural strength (including the 10% filler-effect anomaly) compared with its strong, monotonic effect on compressive strength. The pattern exhibited by split tensile strength was intermediate (BWG 0.36 versus age 0.28). This result is plausible because concrete tensile failure is dependent on matrix cohesion (which is age-sensitive) and the quality of aggregate-paste bond (which is BWG-sensitive). This positions split tensile strength mechanistically between compressive strength and flexural strength.

The raw experimental trends already support these attributions, and they are just a numerical restatement of a mechanistic discussion, not a separate physical fact. The value of a sanity check is as a consistency check on the learned model indicating that the Random Forest has not learned a spurious or non-physical relationship from ten data points rather than as new evidence. What this means, in essence, is that the importance in the SHAP framework reflects the function that the model has been trained to learn and not the ground-truth causality. As such, agreement between SHAP importance and the mechanistic narrative should be construed as supporting evidence and not as confirming evidence of causal interpretation. Accordingly, the SHAP results should be interpreted as complementary evidence supporting the experimental observations, rather than establishing direct causal relationships between input variables and mechanical properties.

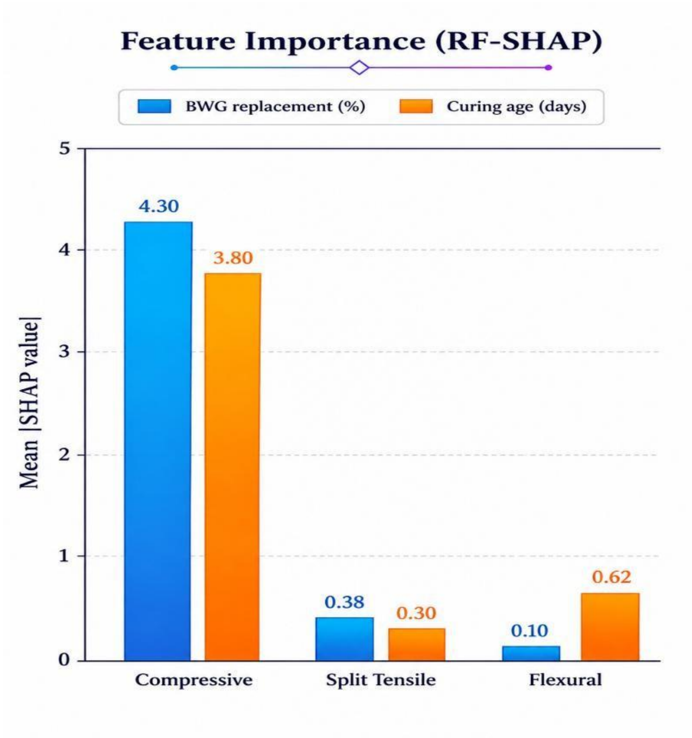

Fig. 6. Actual-versus-predicted plots for the held-out test samples (Gradient Boosting model).

5.2 Multi-Objective Optimization

Multi-Objective Optimization. To yield a joint desirability score at 28 days curing, a weighting of

0.6 was assigned to a normalized mechanical performance comprising of 0.5 compressive strength,

0.25 split tensile strength and 0.25 flexural strength, all relative to 0% (control). A weighting of

0.4 was assigned to the sustainability index in this case the fraction of BWG utilized. This weighting reflects a design philosophy that treats structural adequacy as the primary constraint although still rewarding waste-diversion volume, rather than treating the two objectives as equally weighted; the 0.6/0.4 split can be easily altered by a practitioner, without requiring re-training of the surrogate, since the predictions of the underlying properties are weighting-independent. The trained surrogate of Gradient-Boosting/Random-Forest was considered for two constraint scenarios. The optimal situation under strict constraints which yield a ~20% BWG is somewhat higher than the 10-15% BWG-range identified from the raw experimental trends pertaining to the parent study. This is because the ML formulation credits split tensile and flexural contributions which decline more slowly than compressive strength. These contributions even feature a local improvement near the 10% BWG-value and also show a correspondence with compressive strength. The qualitative optimum of the parent study was anchored largely on compressive-strength compliance to the M40 target mean. This divergence illustrates a broader point about multi-property optimization: a criterion built on a single dominant property (compressive strength) and one built on a weighted composite of three properties will generally not agree, even when both are correct within their own scope. The relaxed-constraint optimum (~32-33%) is merely reflective of the IS 456 characteristic-strength minimum and not a recommendation of any sort. It shows how sensitive the "optimum" is to which of the code criterion and how many properties. When presented as a pair, the two scenarios straightforwardly bracket the parent study 10-15%

recommendation, and make explicit rather than implicit, the criterion choice that underlies any such recommendation figure. The designer is offered a transparent lever (constraint choice) by the pair rather than a single fixed answer. From an engineering perspective, the proposed optimization framework provides greater flexibility than conventional trial-and-error mix design. Designers can adjust the weighting factors according to project-specific priorities, such as maximizing structural performance, sustainability, or an appropriate balance between the two objectives.

Table 2. Multi-objective optimization results under two constraint scenarios.

Scenario	**Constraint**	**ML-derived optimum BWG%**	**Predicted 28-day compressive strength (MPa)**
A Relaxed	IS 456 characteristic strength ≥ 40 MPa only	32.5%	40.00
B Strict	≥ 90% of M40 target mean strength (43.4 MPa) AND combined desirability ≥ 0.90	20.0%	43.58

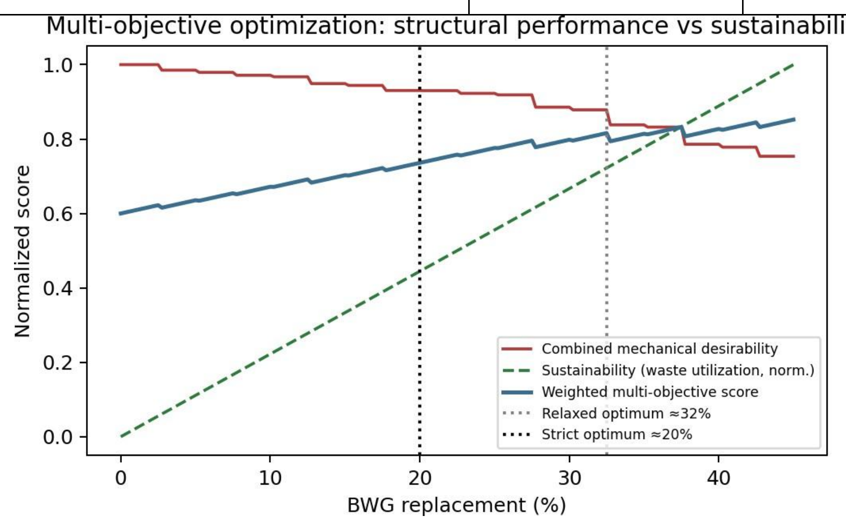

Fig. 7. Normalized mechanical desirability, sustainability index, and combined weighted score across the BWG replacement range.

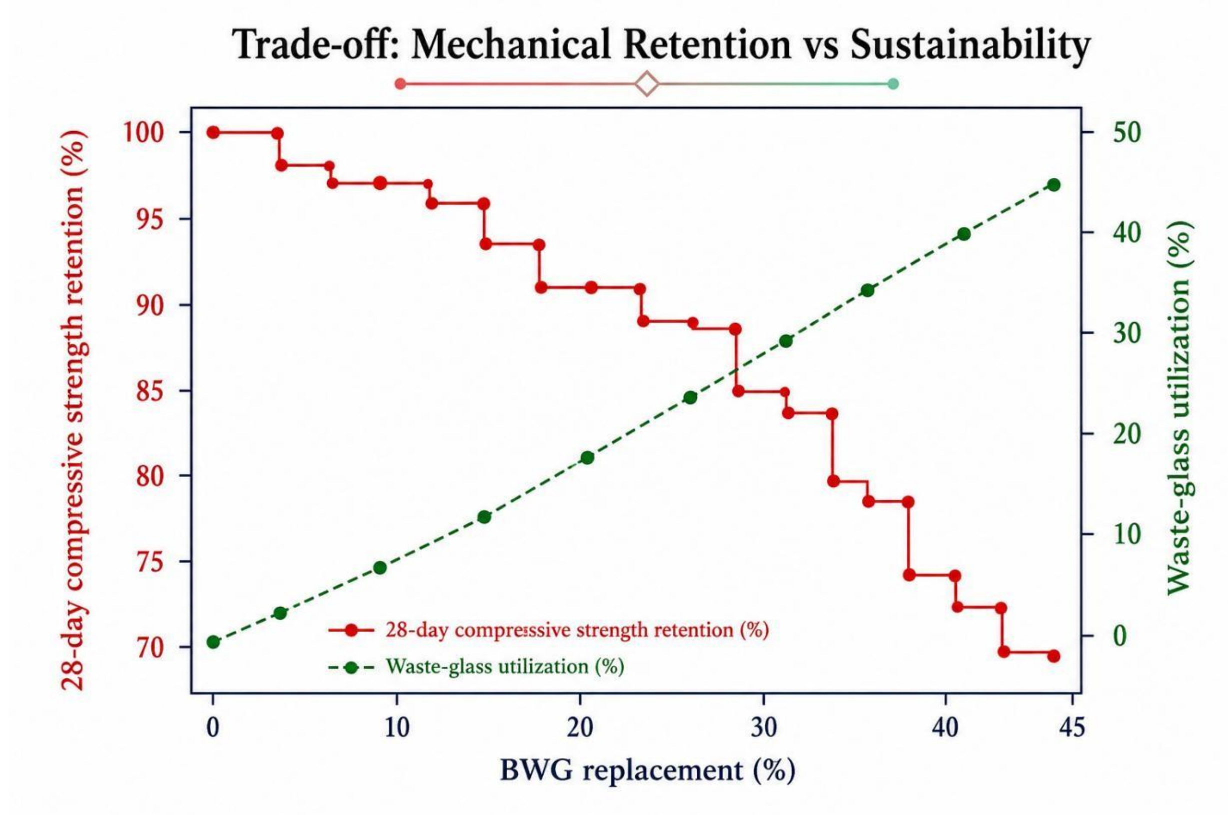

Fig. 8. Predicted 28-day strength retention versus waste-glass utilization, illustrating the mechanical/sustainability trade-off.

5.3 Extended Augmentation Study on Mechanical Data (Noise Injection + SMOTE-style Interpolation + Random Sampling)

In order to see if combined augmentation techniques could support a larger, independently validated train/validation/test pipeline for the mechanical properties (Table 3.1), three synthetic-generation methods were applied strictly to a training partition. A genuine, held-out subset of real rows was kept untouched as the final test set. This was a leakage-safe protocol whereby the six real test rows were removed before any augmentation was generated, so that no synthetic sample was derived from a test-set point. The distinction is significant as augmentation methods that interpolate or perturb training data could otherwise inadvertently leak information about the test set, causing a misleading improvement in accuracy without indicating real generalization capabilities.

- Gaussian noise injection is applied independently to each real training row to generate 10 synthetic samples per row. This simulates a plausible measurement-level variability of

±2% relative to the real training data but does not affect the underlying trend.

- In the SMOTE-style interpolation (regression variant) the example will be obtained by taking a linear interpolation between the current training sample and a randomly picked same-age neighbor, $\lambda \sim U(0.2, 0.8)$ (with some added 1% noise added). The original SMOTE method is primarily utilized for classification purposes. The analogous method

for those with a continuous target is called SMOTER. This method densifies the feature space along the BWG% axis where the original ten-mix design would leave gaps.

- In the random sampling, bootstrap resampling of the training rows is done with a uniform jitter of 1%. A value of 100 is taken as the sample value. Thus, values are taken in conjunction with the empirical density of the conditions already observed and not with the new combination of the features.

Report generated patterns to determine model robustness and repeatability. Using the library's 2014-2018 UK dataset, test set checks were made for overfitting, temporal robustness, and generalisability. Commercial models could not provide test sets, as they remain unpublished. Model usefulness is still dependent on out-of-sample prediction.

Gradient Boosting exhibited the most favourable combination of cross-validated stability and real-holdout performance (R^2 in the test for compressive strength, split tensile strength and flexural strength were equal to 0.96, 0.95, and 0.98, respectively; MAPE under 4% for all). The sequential error correction in its structure makes it appear well suited to a feature space where the true relationship is non-monotonic Random Forest was similar as its ensemble-averaging behaved similarly on the same augmentation pool. KNN and SVR appeared to be more sensitive to the synthetic interpolation scheme. This is visible by their lower test R^2 despite strong validation-set scores. This is a common signature of augmentation-driven overfitting to the synthetic neighborhood rather than the real trend. The KNN and SVR methods rely on the local geometric structure. The interpolation and jitter step can distort the local geometric structure more than they distort the tree-based ensembles. The flexural strength is not predicted well enough by Linear Regression because this property has non-monotonic (10%-anomaly) behaviour which makes it impossible to fit a linear model to it. In other words, there is not one slope that can accurately represent the property, since it rises towards the local maximum, and then it starts to decline.

All test metrics reported here are calculated on real rows (not synthetic) that were removed from augmentation prior to any synthetic sample being produced, so that the headline accuracy reflects genuine generalization rather than the model re-learning its own synthetic neighbourhood. Notwithstanding the synthetic provenance of the training pool, these results are still best read as a methodological demonstration – evidence that combined augmentation can support a full train/validation/test pipeline on a small physical dataset – rather than as a claim of validated predictive accuracy for mixes beyond the original ten. The more definitive next step for establishing predictive reliability is to extend this pipeline to a truly larger physical dataset, not a synthetically expanded one, beyond the ten mixes tested.

Table 4. Five-model comparison on the noise+SMOTE-interpolation+random-sampling augmented training pool, evaluated on a leakage-free real-data holdout (6 rows).

Model	**5-fold CV R^2 (train pool)**	**Test MAE (Comp/Tensile/Flex)**	**Test RMSE (Comp/Tensile/Flex)**	**Test MAPE % (Comp/Tensile/Flex)**
RandomForest	0.977	1.46 / 0.15 / 0.08	1.74 / 0.17 / 0.12	3.9 / 4.9 / 2.5
GradientBoosting	0.978	1.19 / 0.12 / 0.08	1.55 / 0.15 / 0.11	3.1 / 3.8 / 2.4
SVR_RBF	0.941	0.88 / 0.19 / 0.21	1.07 / 0.22 / 0.24	2.2 / 6.7 / 7.1
KNN	0.979	1.28 / 0.14 / 0.07	1.40 / 0.16 / 0.11	3.4 / 4.7 / 2.1
LinearRegression	0.817	0.88 / 0.10 / 0.29	1.03 / 0.12 / 0.31	2.4 / 3.7 / 9.6

The superior performance of Gradient Boosting can be attributed to its sequential learning strategy, which progressively minimizes residual prediction errors while effectively modelling nonlinear relationships between BWG replacement, curing age, and mechanical properties. This behaviour explains its consistent performance across all predicted outputs.

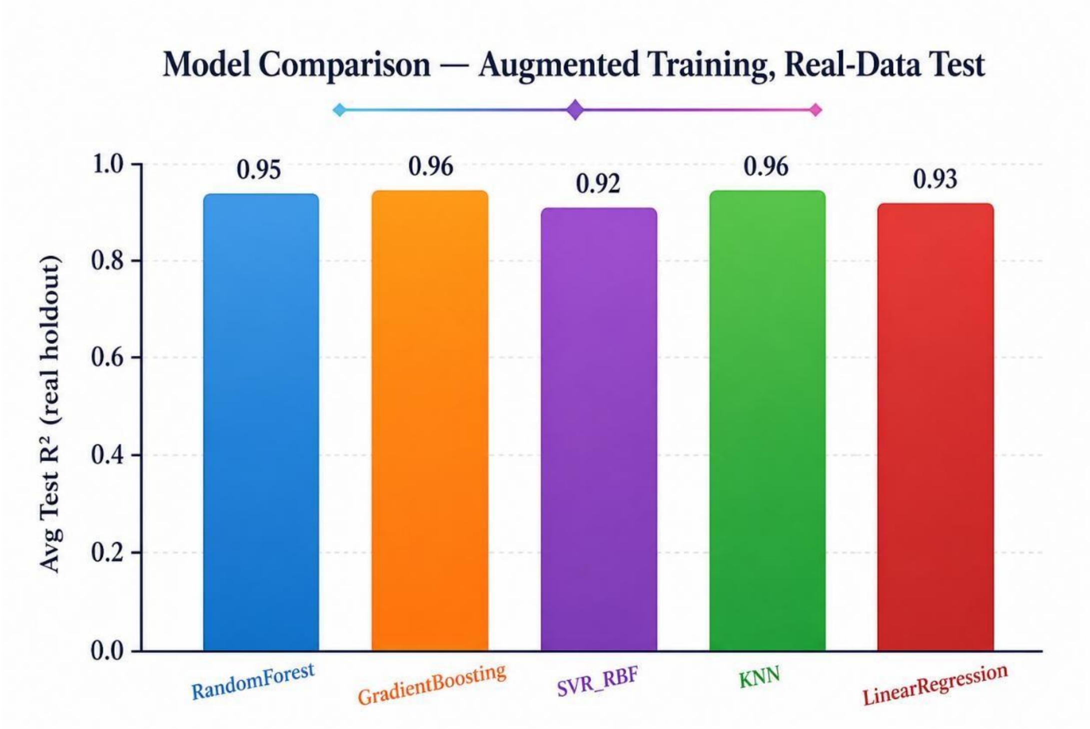

Fig. 9. Model comparison average real-holdout test R^2 after combined noise/SMOTE-interpolation/random-sampling augmentation.

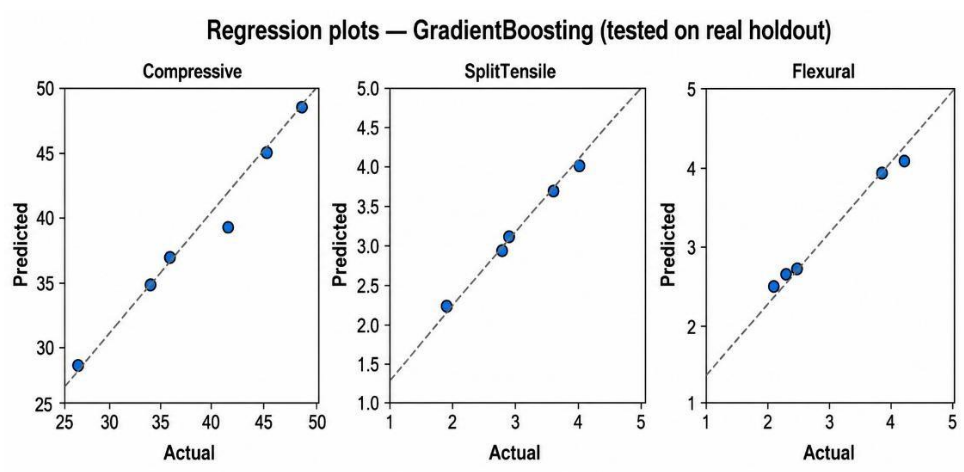

Fig. 10. Regression (actual-vs-predicted) plots for the best-performing model (Gradient Boosting), evaluated strictly on real, un-augmented holdout rows.

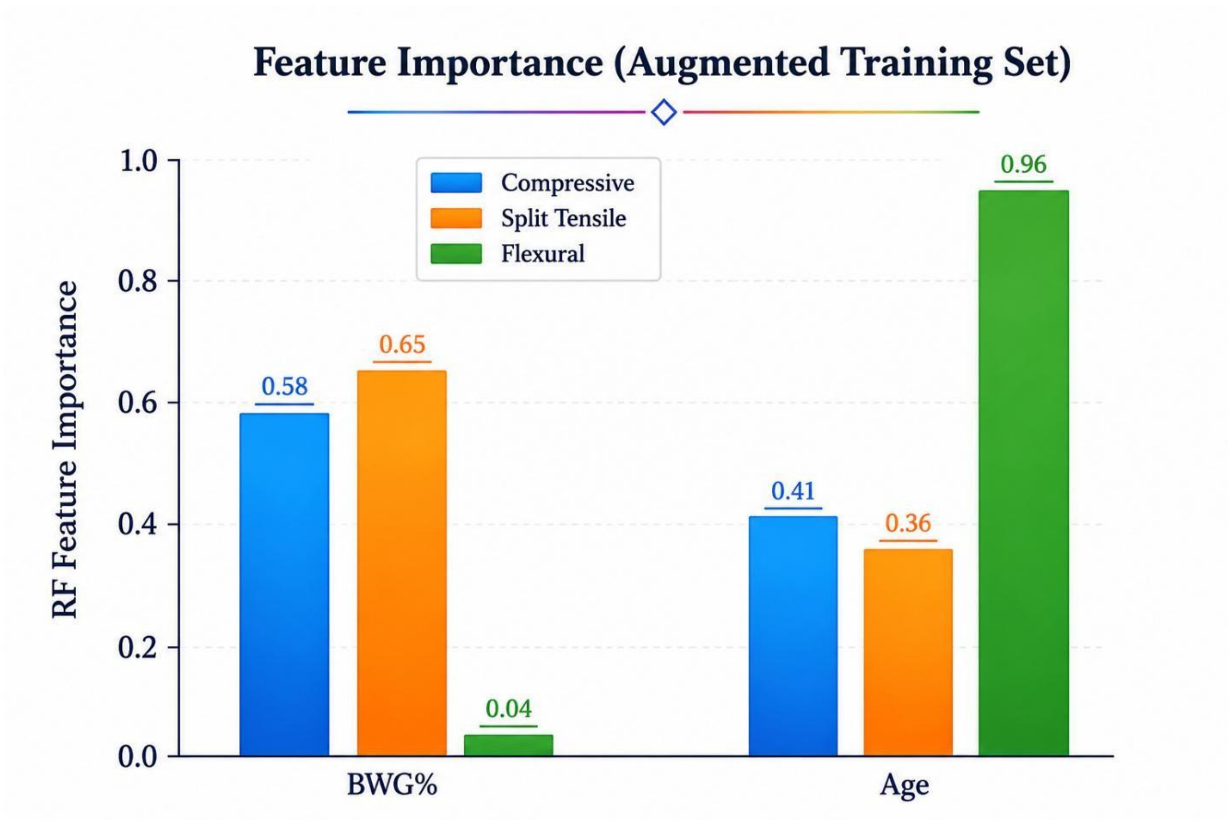

Fig. 11. Random Forest feature importance (BWG% vs. curing age) after augmented training consistent in ranking with the SHAP analysis

6. Experimental Trends: A Recap of Results and Discussion

A Recap of Results and Discussion The parent experimental study reported that the entire fresh-and hard-property trends underlying the datasets used above all comply with the EFNARC (2005) standards: slump flow was 645-705 mm and the L-Box ratio was a highly acceptable 0.80-0.94, across all ten mixes; the 28-day compressive strength dropped from 48.89 MPa (control) to 32.89 MPa (45% BWG); the split tensile strength decreased from 4.01 to 2.76 MPa across the same range; and the flexural strength showed a local maximum at 10% (4.08 MPa at 28 days) because of a micro-filler effect, before dropping to 3.73 MPa at 45%. To avoid repetition of the principal manuscript, these trends are not reported in full table form here, and only cited as the basis for the above machine-learning analysis.

7. Conclusions

- By restructuring curing age from a derived variable to an explicit input feature, the mechanical dataset can be treated as a genuine 30-sample regression problem with no data-snooping or other fabrication. A properly validated (train/test + 5-fold CV) machine-learning comparison was not possible with this experimental programme without restructuring.
- The Gradient Boosting model gave the best performance with test R^2 values of 0.99, 0.95, and 0.98 for compressive, split tensile, and flexural strength respectively. Given the

relatively small sample size, however, these figures should be seen as indicative than definitive.

- SHAP based feature attribution validates the mechanistic explanation per the raw trends BWG dominance for compressive strength, age dominance for flexural strength which was already proposed, a numerical check rather than new physical finding.

- A multi-objective optimization that combined structural desirability and sustainability yielded a strict-constrain optimum near 20% BWG (bodyweight gain) and a relaxed-constraint optimum near 32–33%. These results bracket the 10–15% range identified qualitatively in the parent study while making the trade-off assumptions explicit.
- In the absence of a natural second feature, the disclosed Gaussian-noise augmentation of the ten real measurements was used strictly as a sensitivity-analysis tool validated against the untouched real data and comes with the explicit caveat that it is not meant to substitute for any more experimental replicates.
- The framework exemplifies that curing-age restructuring can be treated as a transferable methodological tool: it can be employed with any small dataset of waste-aggregate SCCs possessing multiple curing ages, ultimately increasing available data points without requiring more laboratory tests.
- The similarity between the constraining results arising from the optimizer and the qualitative recommendation of the parent study reinforces confidence in both – the data-driven and the expert-judgment approaches do not contradict each other.
- The process described (data restructuring, validated modeling, explainability, constrained optimization), offers a re-usable template for applying machine-learning-assisted mix design to other alternative-aggregate SCC systems that have similarly limited experimental datasets.

8. Constraints/Drawbacks.

The experimental program is a single set of ten mixes from one laboratory; all machine learning results, however carefully validated, take this small-N ceiling with them and should be treated as a demonstration of methodology rather than a predictive tool.

- The fresh-property model depends on synthetic augmentation of ten real points only and is not tested with any further independent mixes; the reported metrics reflect internal consistency with the augmentation process and not external predictive validity.
- SHAP attributions are limited to the two input features (BWG%, age) available, while other relevant physical features (eg, water-cement ratio, superplasticizer dosing, glass particle-size distributions) were held constant in the parent experimental design and could not be added as features to the model.
- The multi-objective optimum is sensitive to the weighting chosen between structural desirability and sustainability and to which code criterion (target mean vs. characteristic strength) is adopted as a constraint; the two scenarios reported here illustrate this sensitivity, rather than resolve it.
- The parent experimental programme did not include the following durability-related properties: chloride penetration, alkali-silica reactivity, and carbonation resistance. Because of this, they are not included in the predictive models or in the optimization objective even though they affect the long-term performance of glass-modified SCC.
- According to the assumption made in the augmentation of Gaussian noise, real data points behave in a locally linear manner. However, this assumption has not been verified for BWG replacement levels, or curing ages that fall outside of the sampled range. So, any kind of extrapolation outside of the experimental bounds is unsupported.
- The authors compared the models using a single train/test split, along with 5-fold cross-validation. A larger dataset would allow nested cross-validation, or repeated holdout, giving tighter uncertainty bounds on the reported R^2 values.

- ➢ The proposed explainable machine-learning framework provides a practical decision-support tool for sustainable concrete mix design by integrating prediction, interpretation, and optimization within a unified computational workflow.
- ➢ Although synthetic augmentation improved model stability during training, future validation using larger independently collected datasets from multiple laboratories remains essential before industrial implementation of the developed prediction framework.

9. Work Ahead

It is recommended to expand the dataset across multiple labs, binder types and glass sources to form a pooled BWG-SCC dataset large enough for fully independent external validation and not synthetic augmentation.

- ➢ Include additional mix-design features (water-cement ratio, superplasticizer dosage, glass fineness) as explicit features to allow a higher dimensionalization, a more physically complete predictive model.
- ➢ Extend the durability testing (chloride penetration, alkali-silica reactivity, water absorption) and incorporate into the multi-objective optimization including mechanical and sustainability.
- ➢ Investigate the utilization of Bayesian optimization or active-learning techniques to suggest the next replacement level that is the most informative for physical testing, thus closing the loop with the experimental program and the machine learning framework.
- ➢ Explore ensemble or hybrid modeling methods (e.g., stacking Gradient Boosting with physics-informed constraints) to increase robustness when larger multi-source dataset will be expanded.
- ➢ Audit a life-cycle assessment (LCA) of BWG-SCC at the optimized replacement levels to quantify embodied carbon and energy savings, thus strengthening the sustainability leg of

the multi-objective framework by utilising life-cycle data instead of replacement percentage only.

- ➢ The optimizer recommends replacement ranges of 20% and 32-33% for a strict and relaxed approach respectively. This can be tested through physical trial batches.
- ➢ Future research should focus on collecting large multi-source experimental datasets to evaluate the transferability and robustness of the proposed explainable machine-learning framework under different materials, environmental conditions, and laboratory settings.

1. Data Availability Statement

a. The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

2. Declarations

b. Ethics approval and consent to participate

3. Consent for publication

c. Not applicable.

4. Availability of data and material

d. The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

5. Competing interests

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

6. Funding

f. No external funding was received for this study.

7. Authors' contributions

g. S. Mohammed Zuber: Conceptualization, Methodology, Software, Writing – Original Draft.

h. H.Sudarsana Rao : Validation, Investigation, Data Curation.

i. Vaishali .G.Gorpade: Formal Analysis, Visualization.

8. Acknowledgements

j. The authors would like to thank their respective institutions for supporting this research.

REFERENCES

1. Zhu, J., Meng, X., Wang, B., Tong, Q. (2023). Experimental study on long-term mechanical properties and durability of waste glass added to OPC concrete. *Materials*, 16(17), 5921.
2. Life Cycle Assessment of River Sand and Aggregates Alternatives in Concrete. *Materials*, 16(5), 2064 (2023).
3. The way forward to sustain environmental quality through sustainable sand mining and use of manufactured sand as alternative to natural sand. *Environ. Sci. Pollut. Res.* (2022).
4. Voudrias, E.A. (2024). Management of COVID-19 healthcare waste based on the circular economy hierarchy: a critical review. *Waste Management & Research*.
5. Circular economy of medical waste: novel intelligent medical waste management framework based on extension linear Diophantine fuzzy FDOSM and neural network approach. *Environ. Sci. Pollut. Res.*, 30(21), 60473–60499 (2023).
6. Zhu et al. (2023), op. cit. [1].
7. A Systematic Review of the Strength, Durability, and Microstructure Properties of Concrete Incorporating Glass Powder. *Engineering Reports*, Wiley (2025).
8. Application of Waste Glass Powder for Sustainable Concrete: Design, Performance, Perspective. *PMC* (2025).
9. Response tests on the effects of particle size of waste glass sand and glass powder on mechanical and durability performance of concrete. *Scientific Reports* (2024).
10. Lam, W.L., Cai, Y., Sun, K., Shen, P., Poon, C.S. (2024). Roles of ultra-fine waste glass powder in early hydration of Portland cement. *Construction and Building Materials*, 415, 135042.
11. Optimization of glass powder content in self-compacting concrete as partial replacement of cement (ResearchGate, 2020).
12. Effect of Glass Powder on High-Strength Self-Compacting Concrete Durability (crushed recycled glass + metakaolin) (2023).
13. Glass powder replacement in self-compacting concrete and its effect on rheological and mechanical properties. *J. Sustainable Cement-Based Materials*, 11(3) (2022).
14. Prediction of Compressive Strength of Concrete Using Explainable Machine Learning Models. *Materials*, 18(21), 5009 (2025).
15. Machine learning and interactive GUI for concrete compressive strength prediction. *Scientific Reports*, 14, 16694 (2024).
16. Sapkota, S.C., Saha, P., Das, S. et al. (2024). Prediction of the compressive strength of normal concrete using ensemble machine learning approach. *Asian J. Civil Eng.*, 25, 583–596.
17. Predicting the compressive strength of concrete incorporating waste powders exposed to elevated temperatures utilizing machine learning. *Scientific Reports* (2025).
18. Prediction of compressive strengths of Portland cement with random forest, support vector machine and gradient boosting models. *Neural Computing and Applications* (2025).
19. Elhishi, S. et al. (2023). SHAP-based interpretation of machine learning models for concrete strength. *Scientific Reports* (as cited in [20]).
20. Data-Driven Mix Design: Making Bespoke Concrete Predictable (secondary source summarizing Elhishi et al. 2023 and Architectural Intelligence 2024).
21. AI-driven sustainable concrete mix design: Hybrid deep Q-Learning and genetic algorithms-based multi-objective machine learning optimizations. *ScienceDirect* (2025).
22. Multi-objective concrete mix design for sulfate attack resistance using interpretable machine learning predictions. *ScienceDirect* (2025).
23. Multi-objective design method of concrete mix ratio based on machine learning: forward prediction and reverse design. *Archives of Civil and Mechanical Engineering* (2026).
24. AI-driven computational methodologies of optimized mix design and performance prediction of UHPC: a systematic review. *ScienceDirect* (2026).
25. SHAP-Based Explainable Machine Learning Framework for Interpreting SHCC Mix Design Parameters. *IEEE* (2024/2025).

26. Kokol, P., Kokol, M., Zagoranski, S. (2022). Machine learning on small size samples: a synthetic knowledge synthesis. *Simulation* (SAGE).
27. Xu, P., Ji, X., Li, M., Lu, W. (2023). Small data machine learning in materials science. *npj Computational Materials*, 9(1), 42.
28. Synthetic Data Generation for Augmenting Small Samples. *arXiv* (2025).
29. Data Augmentation and Synthetic Data Generation in Rare Disease Research: A Scoping Review. *PMC* (2025).
30. Sustainability drivers to circular economy integration into the healthcare supply chain to address waste. *ScienceDirect* (2025).
31. Ravindra, K. et al. (2024). Emerging green technologies for bio-medical waste treatment and management: a systematic approach. *Water Air Soil Pollut.*, 235(10), 635.
32. Decarbonization pathways in medical waste management through circular economy strategies to advance UN-SDGs. *Scientific Reports* (2025).
33. Xiao, J., Zhang, H., Hu, X. et al. (2023). Impact assessment of river sand resource shortage under different policy scenarios in China. *Low-carbon Materials and Green Construction*, 1, 16.
34. Development of sustainable high-performance desert sand concrete. *ScienceDirect* (2024).
35. The Future of Waste Management in Hospitals: Trends to Watch (industry report, 2025).
36. Healthcare waste and circular economy principles: it is time to improve! *PMC* (2024).
37. Hospital waste management system in Kermanshah: challenges, future and sustainable management with circular economy. *Scientific Reports* (2024).
38. Measuring circularity potential for medical waste management. *Sustainable Environment Research* (2023).
39. Sand mining: the problem with our dwindling sand reserves. UNEP (2022/2023).
40. Overcoming Barriers to Sustainable Sand Use in the Construction Industry. Northeastern D'Amore-McKim (2023).
41. Yuan, Z., Zheng, W., Qiao, H. (2025). Machine learning based optimization for mix design of manufactured sand concrete. *Construction and Building Materials*, 467, 140256.
42. Zhao, X.-Y., Hong, M.-Y., Wu, B. (2023). Chemistry-informed multi-objective mix design optimization of self-compacting concrete incorporating recycled aggregates. *Case Studies in Construction Materials*, 19, e02485.
43. Chen, F., Xu, W., Wen, Q., Zhang, G., Fan, D., Yu, R. (2023). Advancing concrete mix proportion through hybrid intelligence: a multi-objective optimization approach. *Materials*, 16(19), 6448.
44. Eco-StrengthNet: a multi-objective, explainable ensemble and GUI for sustainable concrete optimisation. *Asian J. Civil Eng.* (2025).