buildings

MDPI

Article

Predicting the Workability of PCE-Modified Cement–Fly Ash Pastes with Explainable Machine Learning

Veysel Gider [1], Serdar Ekinci [2], Davut Izci [3], Burcu Bektaş Güneş [4], Cafer Budak [5], Selin Özteber [6] and Ali Mardani [6,*]

[1] Distance Education Application and Research Center, Rectorate, Batman University, Batman 72100, Türkiye; veysel.gider@batman.edu.tr
[2] Department of Computer Engineering, Bitlis Eren University, Bitlis 13100, Türkiye; sekinci@beu.edu.tr
[3] Department of Electrical and Electronic Engineering, Bursa Uludag University, Bursa 16059, Türkiye; davutizci@uludag.edu.tr
[4] Department of Computer Engineering, Istanbul Gedik University, Istanbul 34876, Türkiye; burcu.gunes@gedik.edu.tr
[5] Department of Electric-Electronic Engineering, Dicle University, Diyarbakır 21280, Türkiye; cafer.budak@dicle.edu.tr
[6] Department of Civil Engineering, Bursa Uludag University, Bursa 16059, Türkiye; ozteberselin@uludag.edu.tr
* Correspondence: alimardani@uludag.edu.tr

Abstract

Reliable assessment of polycarboxylate ether (PCE)–binder combinations requires predictive models whose interpolation performance is distinguished from their performance when an entire formulation is absent from training. This study examined 616 cement–fly ash pastes prepared using twenty-two in-house PCE formulations, four fly ash replacement levels (0, 15, 30, and 45 wt %), seven PCE dosages (0.50–2.00 wt % of binder), and a fixed water-to-binder ratio of 0.35. Marsh funnel flow time was measured for all mixtures, whereas mini-slump measurements were available for 252 mixtures comprising nine PCE formulations. Ten regression algorithms were evaluated using 3 × 5-fold repeated cross-validation and nine non-algebraically redundant input variables. XGBoost achieved R^2 = 0.946 ± 0.027 and RMSE = 8.40 s for flow time, and R^2 = 0.778 ± 0.061 and RMSE = 0.471 cm for mini-slump. These random-resampling results describe interpolation among formulations represented in the training folds. When each PCE chemistry was withheld in turn, the mean R^2 was 0.865 ± 0.198 for flow time and 0.034 ± 0.530 for mini-slump. Performance also decreased when the boundary fly ash levels were withheld. SHAP and permutation analyses identified PCE dose and fly ash replacement as the strongest predictors, while the formulation-level descriptors made smaller contributions. These findings represent associations learned from the present dataset and do not constitute direct evidence of adsorption or dispersion mechanisms. At a nominal 90% level, split-conformal intervals achieved empirical coverages of 85.5% for flow time—moderately below the nominal level, within finite-sample binomial fluctuation—and 90.2% for mini-slump on a random test partition; across fifty repeated train–calibration–test splits, the mean coverages were 89.4 ± 3.6% and 90.1 ± 5.8%. The flow-time model is suitable for preliminary screening within the investigated factor ranges, whereas the mini-slump model should be restricted to interpolation among the sampled formulations.

Keywords: polycarboxylate ether; cement–fly ash paste; Marsh funnel; mini-slump; gradient boosting; model interpretation

Check for updates

Academic Editor: Geo Paul

Received: 8 June 2026
Revised: 18 July 2026
Accepted: 20 July 2026
Published: 25 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Polycarboxylate ether (PCE) superplasticisers have become the principal class of high-range water-reducing admixtures in modern concrete technology because their comb-shaped architecture combines a polycarboxylic backbone with grafted poly(ethylene oxide) side chains, providing dispersion through both electrostatic repulsion and steric hindrance [1–6]. The fluidising performance of a given PCE in a particular cementitious system depends, however, on a multi-dimensional set of factors. These include the following: (i) the molar fraction and spatial distribution of acid groups (i.e., the backbone charge density); (ii) the length and density of poly(ethylene oxide) side chains and the macromonomer-to-acid molar ratio; (iii) the weight-average molecular weight (Mw) and the polydispersity index (PDI); (iv) the surface chemistry, particle-size distribution, and reactivity of the binder; and (v) the dosage of the polymer expressed as a fraction of binder mass [4,7–11]. The interplay of these variables is highly non-linear; small variations in side-chain ratio or in the C_3A content of the cement can shift the dominant adsorption mechanism between calcium-mediated bridging and direct electrostatic anchoring [3,8,12,13].

Phenomenological constitutive models—Bingham, Herschel–Bulkley, and the Krieger–Dougherty correction for solid fraction [14–18]—have been used extensively to summarise the rheological response of PCE-modified pastes, and they remain useful for back-extrapolation and quality control in routine practice. They require, however, empirical parameters that need to be re-fitted whenever the binder, the PCE chemistry or the supplementary cementitious material (SCM) changes. Moreover, single-parameter rheological summaries—yield stress, plastic viscosity, or flow time—are themselves the macroscopic outcome of multiple coupled mechanisms (adsorption, depletion, particle bridging, hydration retardation, and calcium chelation), so the inverse problem of identifying which molecular descriptors govern a particular response is not directly tractable from constitutive models alone.

Machine learning (ML) surrogates offer a complementary route for handling the multi-parameter nature of cementitious composites, whose performance can be affected by mixture composition, reinforcement strategy, and exposure conditions [19–22]. Tree-based ensembles, support-vector regression, neural networks, and Gaussian processes have all been applied to the prediction of compressive strength, modulus of elasticity, or chloride diffusivity of conventional, high-performance, and SCM-modified concretes [23–29]. Interpretability of such models has been advanced by Shapley additive explanations (SHAP) and other model-agnostic frameworks that decompose individual predictions into additive feature contributions [30–33]. These methods have rarely been applied to the fresh-state behaviour of cementitious systems, particularly to cement–fly ash pastes containing several PCE formulations examined under the same experimental conditions. Existing studies typically use one or two PCEs and treat their effect through a single dosage variable, which conflates dose, chemistry, and binder interactions. A recent comprehensive review of concrete material variability and ML model performance [34] further emphasises that dataset heterogeneity and validation design—rather than algorithm choice alone—dominate the accuracy differences reported across studies, a conclusion that directly motivates the grouped validation protocols adopted here.

The present study examines the prediction of Marsh funnel flow time and mini-slump in cement–fly ash pastes using mixture variables and formulation-level PCE descriptors. The experimental programme comprised twenty-two PCE formulations, four fly ash replacement levels, and seven dosage levels, yielding 616 flow-time observations and 252 mini-slump measurements. Ten regression algorithms were evaluated using identical repeated cross-validation partitions. Model performance was also examined by excluding complete PCE formulations, PCE families, and fly ash levels from training, thereby distin-

guishing interpolation within the measured formulation set from prediction when an entire material group was absent from model development.

SHAP analysis, individual and grouped permutation importance, partial-dependence analysis, and feature-group ablation were used to examine the variables associated with the predictions. Bootstrap model spread and split-conformal prediction intervals were evaluated separately. Since adsorption, zeta potential, and rheological flow curves were not measured, the interpretation was confined to relationships represented in the experimental dataset rather than direct attribution of molecular or interparticle mechanisms.

The remainder of the paper is organised as follows. Section 2 surveys the prior literature and locates the present contribution within it. Section 3 describes the materials and the experimental programme. Section 4 details the ML pipeline. Section 5 reports the predictive performance and interpretability outputs, and Section 6 discusses the interpretation, practical scope, and limitations of the results. Section 7 presents the principal conclusions. The corrected dataset and the analysis code used to reproduce the reported numerical results, tables, and figures are provided as Supplementary Material.

2. Related Work

Mechanistic studies of PCE adsorption and dispersion have established the qualitative principles of comb-polymer action on cementitious systems. Yamada et al. [1] showed that the dispersion efficiency of comb-type PCEs is dominated by the side-chain length and by the molar acid-to-side-chain ratio rather than by Mw alone. Sakai et al. [2] subsequently derived the structure–property relations that underlie commercial Japanese PCEs, and Plank & Hirsch [3] linked adsorption to the zeta potential of early hydration phases (notably ettringite and C3A·CaSO4 adducts). Plank et al. [4] reviewed the impact of chemical admixtures on cement microstructure and durability, emphasising that the so-called "compatibility window" of a binder–PCE pair is sensitive both to chemistry and to the physical properties of the polymer. Marchon et al. [5] extended that picture to environmentally optimised comb dispersants, and Houst et al. [7] reported on the dedicated SUPERPLAST programme with novel, custom-synthesised PCEs.

Zingg et al. [8] demonstrated that polyelectrolyte adsorption shifts the zeta potential of cement and individual hydrate phases by tens of millivolts, with direct consequences for interparticle forces. Flatt & Houst [9] proposed a simplified energetic view of the chemical effects perturbing superplasticiser action, while Bessaies-Bey et al. [12] and Ma et al. [13] discussed the trade-offs between molecular weight, viscosity, and cement-phase reactivity that constrain practical PCE design.

The rheological behaviour of cementitious suspensions is commonly described in terms of yield stress and plastic viscosity [14–17]. Marsh funnel flow time is an empirical index that permits comparison of paste flowability when the apparatus geometry, sample volume, timing, and discharge criterion are kept constant. It should not be interpreted as a direct measurement of either yield stress or plastic viscosity. Mini-slump is likewise an empirical spread test whose relation to yield stress depends on mould geometry, mixing history and the interval between mixing and measurement. Kantro [35] introduced the miniature-slump method for evaluating the effect of water-reducing admixtures on cement paste consistency, while Tan et al. [36] demonstrated that reproducibility depends on strict control of the preparation and testing procedure. Computational analysis has also shown that the interpretation of single-point flow tests depends on apparatus geometry and on the rheological range examined [37].

At a fixed water-to-binder ratio and total binder mass, replacement of cement with fly ash changes the relative volume fractions, particle-size distribution, surface area, and early-reactive clinker fraction of the suspension. Depending on its fineness, morphology,

unburned carbon content, and surface chemistry, fly ash may either decrease or increase water and admixture demand. Its interaction with PCE is, therefore, dependent on the characteristics of both the fly ash and the cement and cannot be inferred from replacement level alone [3,4,8,13,18,38–40].

On the methodological side, Yeh [23] introduced artificial neural networks for the prediction of high-performance concrete compressive strength, after which increasingly diverse machine learning approaches were applied to concrete property prediction [24–27,41]. The development of tree-based SHAP methods [32] extended these applications from prediction to feature-level interpretation, while Naser [31] emphasised that model explanations should not be treated as evidence of causality. Recent studies have applied ensemble learning and feature-attribution methods to fresh cementitious properties [42,43]. Tree-based ensembles also remain strong reference methods for small and medium-sized tabular datasets [44–47]. The scikit-learn implementation used to support the preprocessing and regression pipeline is described in [48].

Table 1 summarises representative studies concerning PCE structure, fresh-state property prediction and model interpretation. Previous studies have generally examined PCE adsorption and dispersion, prediction of cementitious material properties or interpretation of trained models separately. The present study combines formulation-level PCE descriptors with two fresh-state responses obtained from one experimental programme and evaluates random-resampling performance separately from group-wise withholding.

Table 1. Comparison of representative prior studies and the present contribution, highlighting methodological and practical differences.

Reference	Material/Dat	Method	Main Contribution	Principal Limitation	Relevance to the Present Study
[1]	Cement paste containing four comb-type PCEs	Adsorption and dispersion experiments	Demonstrated the importance of side-chain and acid-group configuration	One binder; no fly ash or predictive model	Supports the selection of molecular PCE descriptors
[4]	Cementitious systems containing chemical admixtures	Critical literature review	Summarised PCE–binder interactions and compatibility	Qualitative synthesis rather than a predictive study	Provides the physicochemical background
[23]	High-performance concrete dataset	Artificial neural network	Demonstrated the feasibility of data-driven prediction of concrete properties	Hardened-state response; no admixture-chemistry descriptors	Provides an early ML benchmark
[42,43]	Fresh cementitious mixtures	Ensemble ML and model interpretation	Predicted fresh-state properties and examined variable importance	No validation with complete PCE formulations excluded.	Provides direct comparison with fresh-state ML studies
[31]	Civil-engineering ML literature	Review of interpretable ML	Discussed model interpretation and the limitations of feature attribution	No new experimental dataset	Supports cautious interpretation of SHAP results
Present study	616 cement–fly ash–PCE pastes	Ten-model benchmark, grouped validation, and model interpretation	Combines PCE descriptors, flow time, mini-slump, and grouped withholding	One cement, one fly ash, and one water-to-binder ratio	Defines the scope and limitations of the present contribution

3. Materials and Experimental Methods

3.1. Binder Constituents

All paste mixtures were prepared using one CEM I 42.5 R Portland cement and one low-calcium fly ash, each obtained from a single production lot. Their chemical, mineralogical, and physical properties are summarised in Table 2. The cement contained 47.60% C_3S, 20.22% C_2S, 3.60% C_3A, and 16.04% C_4AF and had a Blaine specific surface area of 3754 cm^2/g. The fly ash contained 59.22% SiO_2, 22.86% Al_2O_3, 6.31% Fe_2O_3, and 3.09% CaO and had a Blaine specific surface area of 4300 cm^2/g.

Table 2. Summary of binder constituents and superplasticisers used in this study.

Component	Designation	Key Descriptors
Cement	CEM I 42.5 R	C_3S 47.6%, C_2S 20.22%, C_3A 3.6%, C4AF 16.04%; SSA 3754 cm^2/g; SG 3.20
Fly ash	Low-calcium fly ash	SiO2 59.22%, Al2O3 22.86%, Fe2O3 6.31%, CaO 3.09%; SSA 4300 cm^2/g; SG 2.31
Water	Tap water	245 g per 700 g reference binder (w/b = 0.35)
PCE	K, K-ratio, F, and S series (22 formulations)	Solid content 39.95–40.98%; weight-average molecular weight (Mw) 52,000–65,000 g/mol; PDI 2.0–2.4; Brookfield viscosity 134–192 mPa·s. Bulk density and pH were excluded because the corresponding measurement records could not be verified.

3.2. Polycarboxylate Ether Superplasticisers

Twenty-two aqueous PCE formulations were produced by redox-initiated free-radical copolymerisation of acrylic acid (AA) and a polyethylene glycol-based macromonomer according to the synthesis route reported in [6,11,38]. 3-Mercaptopropionic acid (MPA), ammonium persulfate (AP), and sodium formaldehyde sulfoxylate (SFS) served as the chain-transfer agent, oxidising initiator and reducing activator, respectively. The macromonomer was completely dissolved in water at 40 °C. The solution was subsequently cooled to 35 °C, after which AP was introduced. Aqueous solutions of AA, SFS, and MPA were then added to the reaction medium. Following a total reaction period of 6 h, the pH was adjusted to 8 using aqueous NaOH. The synthesis sequence was kept constant for all formulations, whereas the acid-to-macromonomer ratio and nominal anionic charge level were varied as reported in Table 3.

The formulations were classified into four groups. The K group comprised the base K formulation and the ratio-varied formulations K2.3:1, K3:1, and K4:1. The F series (F1, F3, F5, F7, F9, and F20) and S series (S1, S3, S5, S7, S9, and S20) contained nominal anionic charge levels of 1, 3, 5, 7, 9, and 20 mol % at the base acid-to-macromonomer ratio of 2.3:1. The ratio was additionally varied for F9-2.3:1, F9-3:1, F9-4:1, S5-2.3:1, S5-3:1, and S5-4:1. Together with K2.3:1, K3:1, and K4:1, these formulations constituted the nine-formulation mini-slump subset. The family, charge level, acid-to-macromonomer ratio, and measured formulation properties of all twenty-two PCEs are presented in Table 3.

Weight-average molecular weight and PDI were determined using a Malvern Viscotek GPCmax system (Malvern Panalytical, Malvern, UK). PCE solutions were prepared at a concentration of 2.0 mg/mL and analysed at a flow rate of 0.7 mL/min using an aqueous 0.1% $NaNO_3$ solution as the eluent. Brookfield viscosity and oven-dried solid content were obtained from the formulation-characterisation records [6,11].

During revision, the bulk density and pH entries of the six base S-series formulations could not be traced to complete measurement and instrument records. The recorded density values of 3210–3800 kg/m^3 were also incompatible with aqueous formulations containing

approximately 40% solids. Bulk density and pH were, therefore, removed from both the input matrix and Table 3. No value was reconstructed retrospectively, and all modelling and interpretation analyses were repeated without these variables.

Table 3. Formulation-level properties and response availability of the twenty-two PCEs.

PCE ID	Family	Charge (mol %)	A:M Ratio	Mw (g/mol)	PDI	Viscosity (mPa·s)	Solids (%)	*n* (Flow)	*n* (Slump)
F1	F	1	2.3	58,000	2.2	158	40.16	28	0
F3	F	3	2.3	65,000	2.4	170	40.55	28	0
F5	F	5	2.3	61,000	2.3	156	39.97	28	0
F7	F	7	2.3	56,100	2.1	180	40.85	28	0
F9	F	9	2.3	53,000	2.3	188	40.65	28	0
F9-2.3:1	F	9	2.3	60,000	2.3	134	40.19	28	28
F9-3:1	F	9	3	53,000	2.3	188	40.65	28	28
F9-4:1	F	9	4	54,000	2.4	155	40.14	28	28
F20	F	20	2.3	52,000	2.3	192	40.98	28	0
K	K	0	2.3	55,000	2.1	154	40.78	28	0
K2.3:1	K-ratio	0	2.3	63,000	2.1	178	40.12	28	28
K3:1	K-ratio	0	3.0	55,000	2.1	154	40.78	28	28
K4:1	K-ratio	0	4.0	56,000	2.1	180	40.19	28	28
S1	S	1	2.3	62,100	2.1	168	39.96	28	0
S3	S	3	2.3	63,000	2.2	168	40.15	28	0
S5	S	5	2.3	58,000	2	162	40.56	28	0
S5-2.3:1	S	5	2.3	62,000	2.1	174	39.95	28	28
S5-3:1	S	5	3	58,000	2	162	40.56	28	28
S5-4:1	S	5	4	56,000	2	180	40.07	28	28
S7	S	7	2.3	63,000	2.2	170	39.96	28	0
S9	S	9	2.3	60,000	2.4	162	40.7	28	0
S20	S	20	2.3	61,000	2.1	164	40.15	28	0

The archived synthesis records did not contain independently verifiable values for AP concentration, reagent-feed duration, or reactor volume; these quantities are, therefore, not reported.

3.3. Mix Design and Testing Protocol

All 616 paste mixtures had a fixed water-to-binder ratio, as defined in Equation (1):

$$w/b = \frac{m_{\text{water}}}{m_{\text{cement}} + m_{\text{fly ash}}} = 0.35 \quad (1)$$

Table 4 reports the constituent masses on a reference basis of 700 g total binder. Cement was replaced by fly ash at 0, 15, 30, and 45 wt %, while the water-to-binder ratio remained constant at 0.35. Seven PCE dosage levels, 0.50, 0.75, 1.00, 1.25, 1.50, 1.75, and 2.00 wt % of binder, were applied to every PCE–replacement combination. These levels correspond to 3.5–14.0 g of PCE per 700 g of reference binder. Laboratory batches were prepared at the same constituent ratios in sufficient quantity for the fresh-state tests, giving $22 \times 4 \times 7 = 616$ formulation–mixture combinations.

Paste preparation followed the mixing procedure used in the related experimental studies [6,11,38]. All materials were conditioned at 22 ± 2 °C for at least 24 h before mixing. Water and PCE were first mixed at 800 rpm for 30 s. Cement and fly ash were then added, and mixing continued at 800 rpm for a further 30 s. Mixing was stopped briefly to scrape material adhering to the bowl wall, after which the paste was mixed at 1600 rpm for 120 s. The nominal mixing duration was, therefore, 180 s, excluding the short intervals required for scraping and transfer. Testing began immediately after mixing, and no deliberate resting

period was imposed. Relative humidity was not recorded during the original experimental campaign and is, therefore, not reported retrospectively.

Table 4. Mix-design matrix normalised to 700 g of reference binder; laboratory batches were prepared at the same constituent ratios.

Series	Cement/Fly Ash (wt %)	Cement/Fly Ash (g)	Water (g)	PCE Dose, g (wt % of Binder)
UK0	100/0	700/0	245	3.5–14.0 (0.5–2.0 wt %)
UK15	85/15	595/105	245	3.5–14.0 (0.5–2.0 wt %)
UK30	70/30	490/210	245	3.5–14.0 (0.5–2.0 wt %)
UK45	55/45	385/315	245	3.5–14.0 (0.5–2.0 wt %)

Marsh funnel flow time was measured using a funnel with a top diameter of 150 mm, an overall length of 355 mm, a discharge-tube length of 50 mm and an internal outlet diameter of 5 mm [6,11]. Approximately 1.2 L of freshly mixed paste was introduced into the funnel with the outlet closed. Flow time was defined as the elapsed time required for 700 mL of paste to discharge and was measured using a stopwatch. The funnel was checked for unobstructed discharge before testing. All 616 mixtures discharged within the observation period; no flow-time result was censored or excluded.

Mini-slump was measured using a truncated-cone mould with an internal top diameter of 19.0 mm, an internal bottom diameter of 38.1 mm, and a height of 57.2 mm [6,11,35,36]. The mould was placed on a level glass plate, filled with freshly mixed paste, and lifted vertically. Spread was recorded as the mean of two perpendicular diameters read to the nearest 1 mm. The mould and glass plate were cleaned and dried before each measurement. The miniature-slump method was originally developed for assessing the influence of water-reducing admixtures on cement paste consistency [35], and the timing and execution of the test were controlled in accordance with the considerations reported by Tan et al. [36].

The machine learning dataset contains one reported response value for each PCE–fly ash dosage combination. Replicate-level measurements were not entered as separate observations. The experimental unit used in the analysis was, therefore, one formulation–mixture combination. The surviving records did not permit technical repetitions to be distinguished reliably from separately prepared repetitions; consequently, no replicate count is assigned retrospectively. The earlier statement that each result represented three consecutive measurements on the same batch was removed.

Mini-slump was measured for K2.3:1, K3:1, K4:1, F9-2.3:1, F9-3:1, F9-4:1, S5-2.3:1, S5-3:1, and S5-4:1, yielding $9 \times 4 \times 7 = 252$ observations. No mini-slump result was discarded.

4. Methodological Framework

The experimental and modelling procedure is summarised in Figure 1. The workflow comprises material selection, mixture preparation, fresh-state testing, input variable definition, preprocessing, model comparison, statistical evaluation, grouped validation, uncertainty analysis, and model interpretation.

4.1. Feature Engineering and Preprocessing

Binder chemistry, Blaine fineness, binder specific gravity, and total water proportion were constant throughout the experimental programme and were, therefore, not used as input variables. The mixture identifier was parsed to obtain PCE family, nominal anionic charge level, acid-to-macromonomer ratio, and PCE dose. PCE dose was expressed as a percentage of binder mass according to Equation (2):

$$\text{PCE dose (wt \%)} = \frac{m_{\text{PCE}}}{m_{\text{cement}} + m_{\text{fly ash}}} \times 100 \quad (2)$$

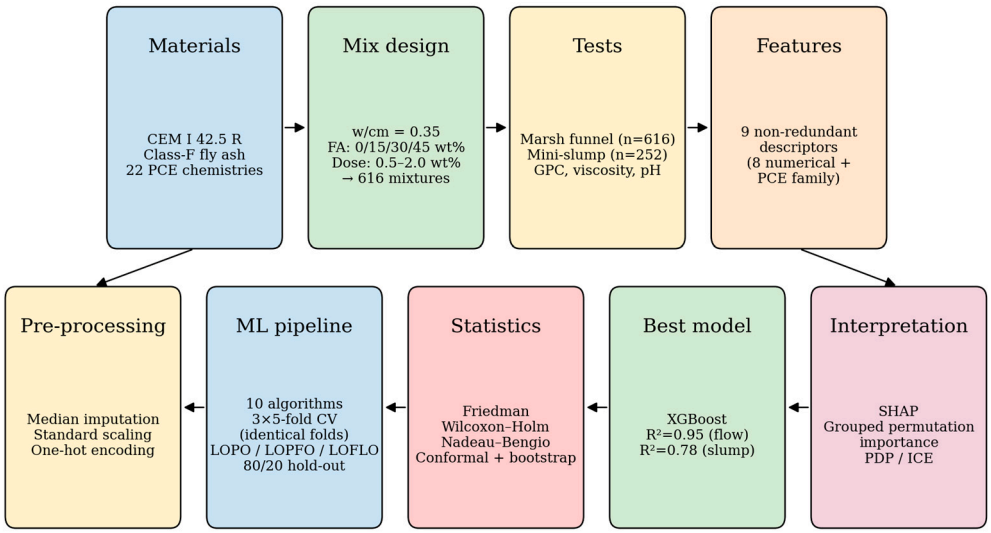


Figure 1. Experimental programme and modelling workflow used in the study.

The initial matrix contained several algebraically duplicated variables. At a constant binder basis and water-to-binder ratio, cement mass, fly ash mass, fly ash replacement, and fly ash-to-binder ratio encode the same compositional change. Similarly, PCE mass, PCE dose as a percentage of binder, and PCE-to-binder ratio are exact transformations of one another. Water-to-cement ratio is also determined by fly ash replacement under the fixed water and binder proportions. Retaining these variables simultaneously can divide or reassign importance among equivalent inputs and obscure the interpretation of SHAP and permutation results.

The revised input matrix, therefore, contained nine non-algebraically redundant variables: PCE dose (wt % of binder), fly ash replacement (wt %), nominal anionic charge level (mol %), acid-to-macromonomer ratio, Mw, PDI, Brookfield viscosity, solid content, and PCE family. Bulk density and pH were excluded for the reasons given in Section 3.2. The retained formulation descriptors are not statistically independent; several remain associated because they originate from a finite set of synthesised formulations. Their attribution values were, therefore, interpreted as associations within the trained model rather than isolated molecular effects.

Numerical variables were median-imputed and standardised to zero mean and unit variance. PCE family was one-hot encoded. All preprocessing operations were fitted separately within each training fold and then applied to the corresponding validation fold, preventing information leakage [48]. The categorical encoder was configured with handle_unknown = "ignore" so that a PCE family excluded during leave-one-family-out validation was not introduced during model training.

4.2. Regression Algorithms

Ten regression algorithms constitute the primary benchmark, chosen to span the relevant inductive biases. Five linear models (ordinary least-squares Linear Regression, Ridge, Lasso, ElasticNet, and Bayesian Ridge) bracket the family of L1/L2 regularisation. The kernel-based support-vector regression with radial basis function (SVR-RBF) and the instance-based k-nearest neighbours (k-NNs, k = 5) provide non-parametric baselines. Three tree-based ensembles (Random Forest, Gradient Boosting, and eXtreme Gradient Boosting, XGBoost [44,45]) represent the state of the art on tabular regression tasks. Two

further gradient-boosting libraries, CatBoost and LightGBM, are evaluated as a supplementary sensitivity analysis under the identical protocol and identical fold assignments (Section 5.10); they are not part of the primary benchmark. Neural network models were not included in the primary comparison because the dataset is small and low-dimensional, particularly for mini-slump, and a balanced neural-model assessment would require a broader architecture and regularisation search than was applied to the other algorithms. Gradient-boosted trees were retained as established baselines for tabular regression [46,47]. The exclusion of neural networks defines the scope of the present model comparison and does not imply that neural architectures are intrinsically unsuitable for these responses. The objective function minimised during XGBoost training is given in Equation (3):

$$L(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (3)$$

where l(·) is a differentiable loss function (squared error in this study), f_k are individual decision trees, and Ω(·) is the regularisation term that penalises tree complexity. Hyperparameters for the primary benchmark were fixed at literature-recommended values (for the boosters: n_estimators = 400, max_depth = 4, learning_rate = 0.05, subsample = 0.9, and colsample_bytree = 0.9; Random Forest: 400 trees; all random seeds = 42). The ten-model comparison of Section 5.2 is, therefore, a fixed-configuration benchmark: every algorithm is evaluated at a single, literature-recommended setting, and no per-model tuning is performed within the primary protocol, so the reported scores should not be read as each algorithm's attainable optimum. Hyperparameter sensitivity is examined separately in Section 5.9, where a nested cross-validation (an explicit grid of 27 configurations scored by inner-loop RMSE) is applied to the selected XGBoost surrogate, and the same nested protocol is additionally applied to the principal competitors (Gradient Boosting, Random Forest, and SVR-RBF) with comparable grids. All experiments used Python 3.10 with scikit-learn 1.7.2, XGBoost 2.1.4, LightGBM 4.6.0, CatBoost 1.2.10, and SHAP 0.49.1; the complete pipeline, including fold generation, is released as Supplementary Material.

4.3. Evaluation Protocol and Statistical Testing

Predictive performance under random resampling was estimated by 3 × 5-fold repeated cross-validation. RMSE, MAE, and R^2 were reported as the mean and standard deviation across the fifteen validation folds. The partitions were generated once using RepeatedKFold with seed 42 and were reused for every algorithm. Because observations belonging to the same PCE formulation could occur in both training and validation folds, this procedure measured interpolation within the sampled formulation set. An 80/20 random hold-out partition was additionally used as an internal interpolation diagnostic rather than as an independent validation dataset. Performance under complete group exclusion was examined using leave-one-PCE-formulation-out, leave-one-PCE-family-out and leave-one-fly-ash-level-out validation. Group-based validation provides a more stringent assessment when related material formulations share common descriptors [49]. RMSE and R^2 are defined in Equations (4) and (5), respectively.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (4)$$

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

Model rankings were examined using the Friedman test followed by Holm-corrected Wilcoxon signed-rank comparisons. Since repeated cross-validation folds share training observations, the resulting fold-level scores are dependent; the corresponding *p*-values

were, therefore, treated as exploratory paired summaries. The prespecified comparison between XGBoost and Gradient Boosting was additionally examined using the Nadeau–Bengio corrected resampled t-test [50].

Uncertainty was quantified in two clearly separated layers. First, 80 bootstrap replicates of the training set were used to retrain the surrogate, and the 2.5th–97.5th percentile spread of the 80 predictions on each hold-out point is reported as a model-uncertainty (epistemic) band; this band does not include residual (measurement) variance and is, therefore, not a prediction interval for new observations. Second, split-conformal prediction [51] (Section 5.8) supplies distribution-free prediction intervals with marginal coverage guarantees under exchangeability. Interval widths, split sizes, and seeds are reported alongside the coverages.

4.4. Interpretability Tools

4.4.1. SHAP Attributions

SHAP additive explanations were computed with TreeExplainer where applicable and with the Permutation explainer otherwise. The local SHAP decomposition of an individual prediction f(x) is given in Equation (6):

$$f(x) = \phi_0 + \sum_{i=1}^{M} \phi_i(x) \quad (6)$$

where ϕ_0 is the expected model output over the training set and $\phi_i(x)$ is the additive contribution of feature i to the specific prediction [32].

4.4.2. Permutation Importance

Model-agnostic permutation feature importance was computed with twenty random permutations per feature on the hold-out partition, scoring on R^2. The permutation importance I_i of feature i is given in Equation (7):

$$I_i = R^2_{\text{full}} - \frac{1}{K}\sum_{k=1}^{K} R^2_{\text{perm(i),k}} \quad (7)$$

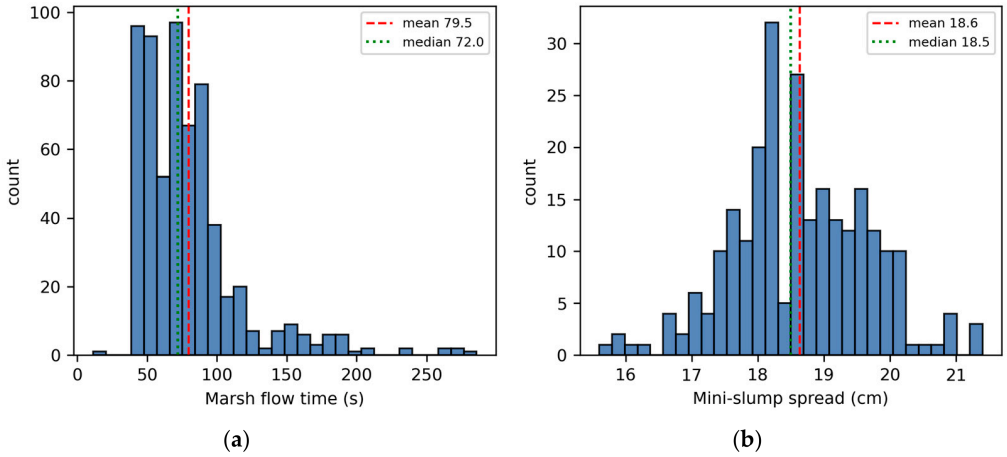
4.4.3. Partial Dependence and ICE

Partial-dependence and individual conditional expectation (ICE) curves were generated for the six most influential numerical features to inspect non-linearity and interaction. Two-dimensional partial-dependence surfaces (Section 5.5) were computed for the highest-ranked feature pairs.

5. Results

5.1. Distribution of the Rheological Responses

Flow time spans 11.2 to 285.8 s with a strong right-skew due to the high-viscosity tail of low-dose mixtures (Figure 2a). The mini-slump occupies a comparatively narrow window of 15.6–21.4 cm with a roughly Gaussian shape (Figure 2b). The two responses are moderately anti-correlated overall (Spearman $\rho \approx -0.57$ on the 252 paired observations; Figure 3) and weakly correlated with each individual molecular descriptor in isolation, supporting the need for multivariate, non-linear models.



(**a**) (**b**)

Figure 2. Empirical distributions of (**a**) Marsh funnel flow time (n = 616) and (**b**) mini-slump spread (n = 252). Vertical lines denote the corresponding sample means and medians.

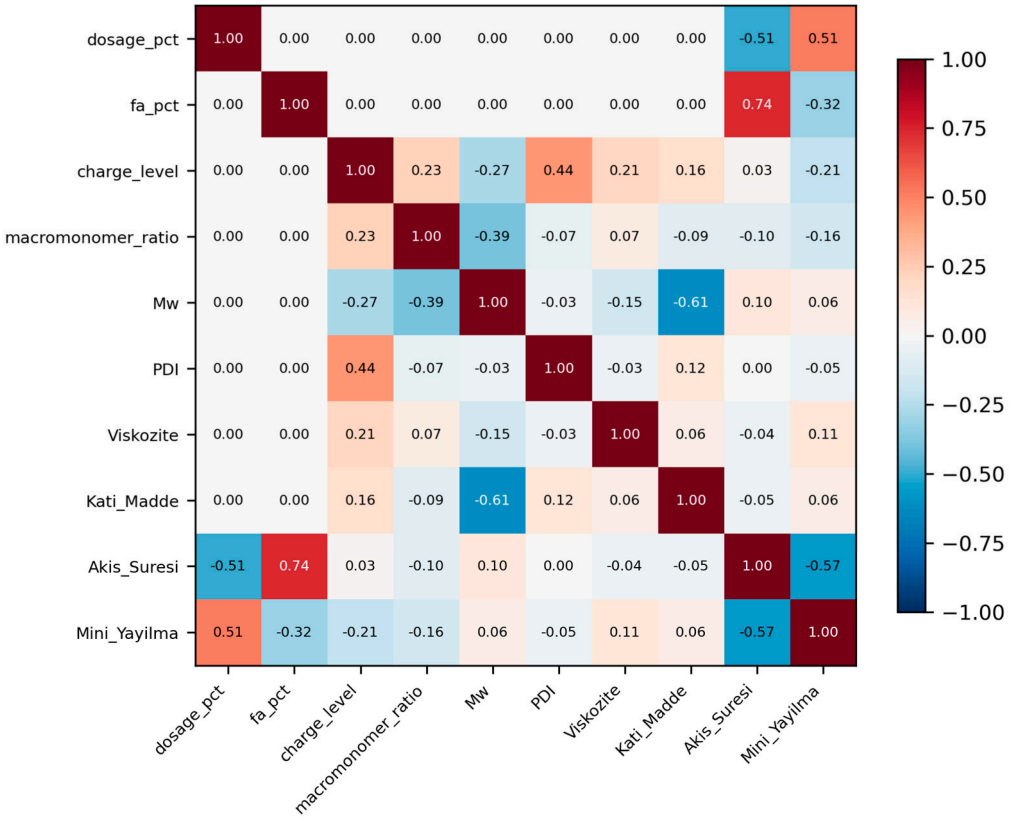


Figure 3. Spearman rank-correlation matrix of the final input variables and measured responses. Correlations involving mini-slump were calculated using the 252 observations for which both responses were available.

5.2. Cross-Validated Performance and Statistical Comparison

Tables 5 and 6 summarise the RMSE, MAE, and R^2 values of all ten regression algorithms on the two responses under the reduced nine-feature matrix. For flow time, the gradient-boosted ensembles share the top tier (XGBoost $R^2 = 0.946 \pm 0.027$ and Gradient Boosting 0.942 ± 0.030), followed by Random Forest (0.883 ± 0.053). The instance-based and kernel baselines reach R^2 of 0.633 (k-NN) and 0.617 (SVR-RBF), while the five linear models saturate near $R^2 \approx 0.60$. The same hierarchy is observed for the mini-slump response with a smaller absolute gap: XGBoost attains $R^2 = 0.778 \pm 0.061$, Gradient Boosting 0.742 ± 0.057, Random Forest 0.728 ± 0.044, and SVR-RBF 0.714 ± 0.055, with the linear

baselines at $R^2 \approx 0.37$. The near-identical scores obtained with the previous, redundant seventeen-feature matrix (R^2 = 0.943 and 0.772) indicate that removal of the algebraically duplicated inputs did not materially reduce predictive performance.

Table 5. Repeated 3 × 5-fold cross-validation results for flow time (n = 616), reduced nine-feature set. Values are mean ± standard deviation across 15 folds.

Model	RMSE (s)	MAE (s)	R^2
XGBoost	8.40 ± 2.71	3.46 ± 0.56	0.946 ± 0.027
Gradient Boosting	8.68 ± 2.71	3.54 ± 0.54	0.942 ± 0.030
Random Forest	12.47 ± 3.74	4.84 ± 0.95	0.883 ± 0.053
k-NN	22.40 ± 3.96	12.06 ± 1.73	0.633 ± 0.065
SVR-RBF	23.04 ± 5.02	9.73 ± 1.74	0.617 ± 0.080
Bayesian Ridge	23.55 ± 4.11	15.69 ± 1.28	0.598 ± 0.050
ElasticNet	23.56 ± 4.08	15.76 ± 1.27	0.597 ± 0.049
Lasso	23.57 ± 4.06	15.81 ± 1.27	0.597 ± 0.049
Ridge	23.57 ± 4.07	15.80 ± 1.27	0.597 ± 0.049
Linear Regression	23.58 ± 4.06	15.82 ± 1.27	0.596 ± 0.049

Table 6. Repeated 3 × 5-fold cross-validation results for mini-slump (n = 252), reduced nine-feature set. Values are mean ± standard deviation across 15 folds.

Model	RMSE (cm)	MAE (cm)	R^2
XGBoost	0.471 ± 0.059	0.363 ± 0.048	0.778 ± 0.061
Gradient Boosting	0.508 ± 0.045	0.403 ± 0.042	0.742 ± 0.057
Random Forest	0.525 ± 0.044	0.404 ± 0.042	0.728 ± 0.044
SVR-RBF	0.536 ± 0.036	0.414 ± 0.032	0.714 ± 0.054
k-NN	0.669 ± 0.061	0.515 ± 0.031	0.560 ± 0.058
Bayesian Ridge	0.797 ± 0.057	0.630 ± 0.060	0.374 ± 0.083
Ridge	0.797 ± 0.061	0.632 ± 0.066	0.374 ± 0.091
ElasticNet	0.798 ± 0.059	0.632 ± 0.063	0.373 ± 0.085
Linear Regression	0.798 ± 0.064	0.634 ± 0.068	0.371 ± 0.094
Lasso	0.800 ± 0.057	0.633 ± 0.061	0.370 ± 0.082

The Friedman test rejects the null hypothesis of equal model performance for both responses (χ^2 = 99.51, p = 2.0 × 10^{-17} for flow time; χ^2 = 115.86, p = 9.4 × 10^{-21} for mini-slump). Holm-corrected pairwise Wilcoxon comparisons place XGBoost and Gradient Boosting in a single statistically indistinguishable group at the top, both significantly above all linear, kernel and instance-based models for the two responses. As anticipated in Section 4.3, these *p*-values should be read as anti-conservative because the fifteen-fold scores share training data. The variance-corrected Nadeau–Bengio t-test applied to the top pair (XGBoost versus Gradient Boosting) yields $p = 0.62$ for flow time and $p = 0.16$ for mini-slump, i.e., the two boosters cannot be separated once the fold dependence is accounted for, whereas the ensemble-versus-linear gap (≈0.35 R^2 points) is far beyond any correction. Figure 4 visualises the distribution of RMSE values across the fifteen cross-validation folds.

5.3. Hold-Out Diagnostic and Bootstrap Model-Uncertainty Bands

On the 20% random hold-out partition (124 mixtures for flow time and 51 for mini-slump), the XGBoost surrogate attained R^2 = 0.953 and RMSE = 8.57 s for the flow-time response and R^2 = 0.768 and RMSE = 0.497 cm for the mini-slump response, in line with the cross-validated estimates. The random hold-out partition is an internal interpolation diagnostic rather than an external validation dataset. Figure 5 plots the measured–predicted scatter together with the 1:1 line and the ±10% envelope: 91.1% of the flow-time hold-out

points and 100% of the mini-slump points lie within the envelope (both values are reported in the figure insets).

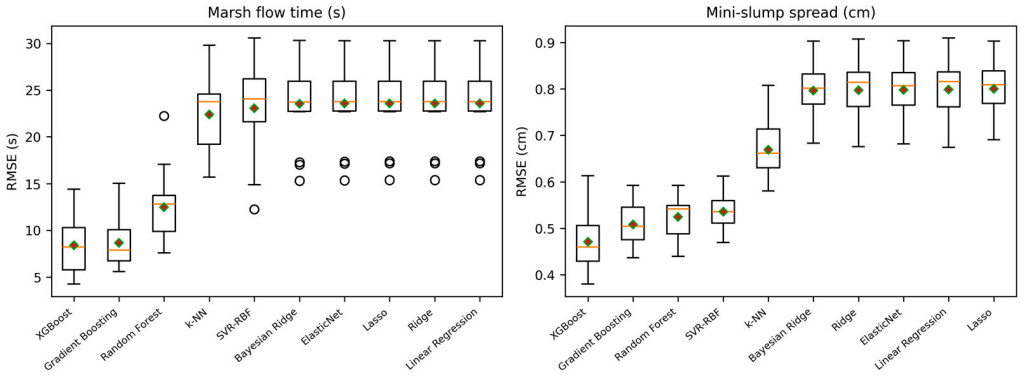


Figure 4. Distribution of RMSE values across the fifteen cross-validation folds, ordered by mean RMSE. Diamonds mark the per-model mean. The horizontal line marks the median; whiskers show the typical range, with circles marking outliers. Tree-based ensembles form a clearly separated top tier in both panels.

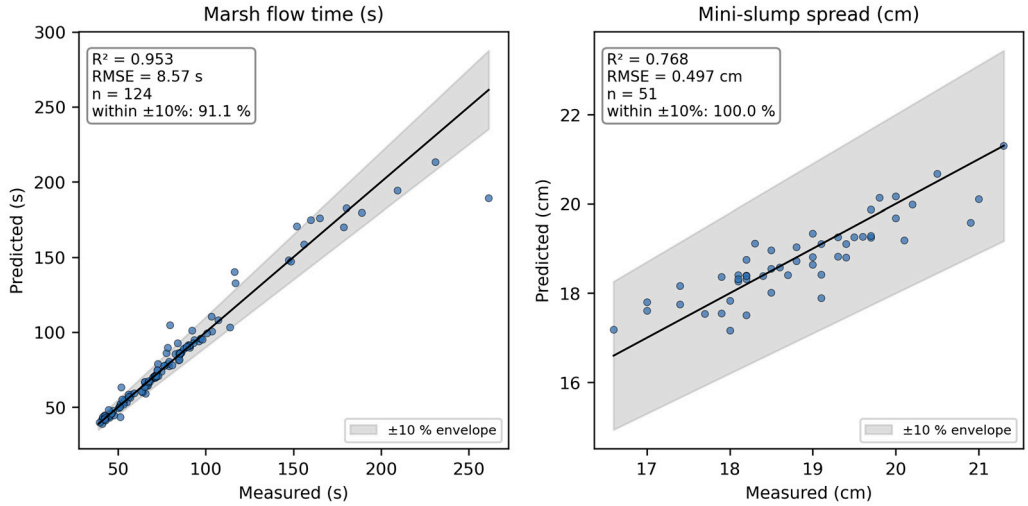


Figure 5. Measured versus predicted on the 20% random hold-out partition (interpolation diagnostic) for the XGBoost surrogate. Solid line: 1:1 reference; shaded band: ±10% envelope. Insets report R^2, RMSE, sample size, and the share of points inside the envelope.

Residual diagnostics in Figure 6 show no pronounced trend across most of the prediction range. The residuals are centred near zero, and the Q–Q plots indicate approximately normal distributions for both responses, although the flow-time model exhibits heavier tails at the upper end of the measured range, particularly for mixtures with low PCE dosage and high fly ash replacement.

5.4. Feature Importance and SHAP Analysis

Individual permutation importance, grouped permutation importance, and SHAP rankings agree closely on both responses (Tables 7 and 8, Figures 7 and 8). For Marsh flow time, PCE dose and fly ash replacement dominate: together they account for approximately 80% of the absolute SHAP mass (41.1% and 39.2%, respectively), followed by Mw (6.5%) and Brookfield viscosity (5.0%), while the PCE family categorical is informative only at the margin. For mini-slump, the same two mix variables lead (31.2% and 27.0%), but the molecular descriptors gain substantial relative weight: anionic charge level, Mw, viscosity, and macromonomer ratio each contribute $\approx$ 8% of the SHAP mass. Because even the

reduced feature set retains some correlation among the molecular descriptors, the grouped permutation scheme (Figure 7c,d)—in which physically related variables are permuted jointly (dose; fly ash replacement; anionic charge; macromonomer ratio; molecular size Mw + PDI; formulation viscosity + solids; family one-hots)—was added as a robustness check. It reproduces the individual rankings, with ΔR^2 of 1.16 (dose) and 0.67 (fly ash) for flow time, and 0.84 (fly ash), 0.41 (dose), 0.17 (formulation), 0.15 (molecular size), and 0.14 (macromonomer ratio) for mini-slump. It's shown that, when stratified by PCE family, F-series and S-series polymers exhibit distinct trends with respect to anionic charge level for both responses. Throughout this section, the rankings describe associations learned by the model on this dataset; they are not, by themselves, evidence of a physical mechanism (Section 6).

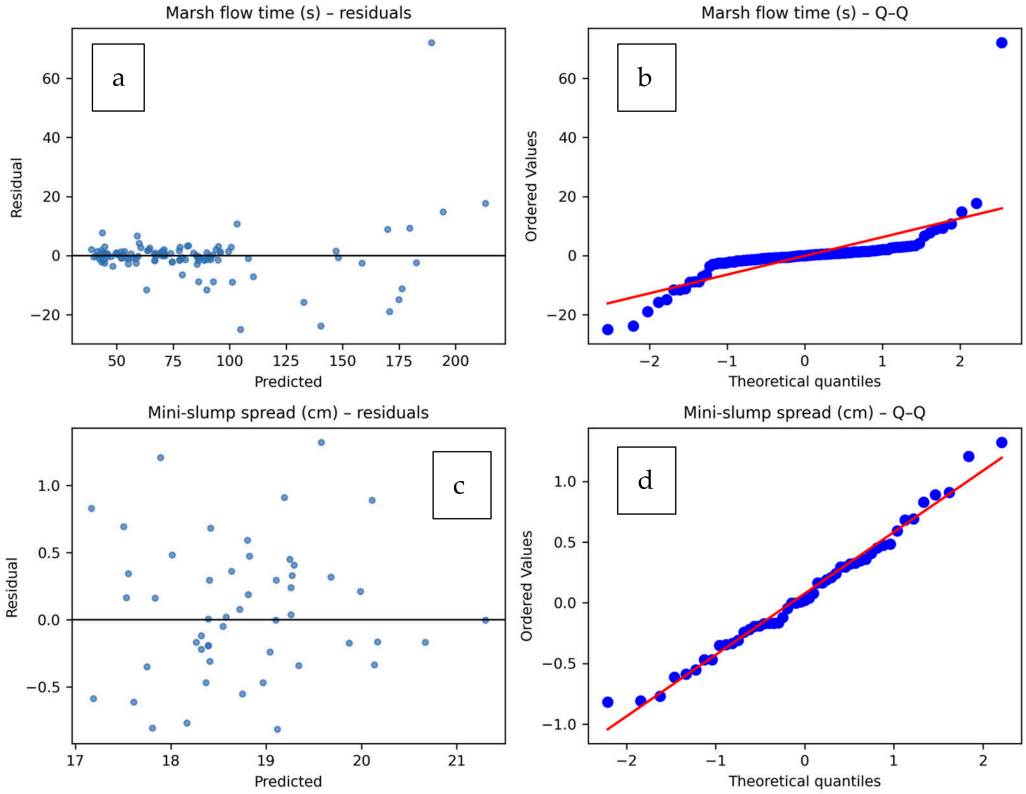


Figure 6. Residual diagnostics for the XGBoost surrogate. (**a**,**c**) Residual versus predicted; (**b**,**d**) normal Q–Q plot. The absence of trend in (**a**,**c**) supports the assumption that the model captures the deterministic part of the variance.

Table 7. Feature ranking for flow time (XGBoost, reduced feature set): mean |SHAP| share and permutation ΔR^2.

Rank	Feature	Mean |SHAP| Share (%)	Permutation ΔR^2
1	PCE dose (wt % binder)	41.1	1.139
2	Fly ash replacement (wt %)	39.2	0.633
3	PCE Mw (g/mol)	6.5	0.048
4	PCE viscosity (mPa·s)	5.0	0.131
5	Solid content (%)	3.5	0.033
6	Anionic charge (mol %)	1.5	0.002
7	PCE family = F	0.9	0.002
8	PCE PDI	0.9	0.005
9	PCE family = S	0.6	0.002
10	Macromonomer ratio	0.4	0.001

Table 8. Feature ranking for mini-slump (XGBoost, reduced feature set): mean |SHAP| share and permutation ΔR^2.

Rank	Feature	Mean |SHAP| share (%)	Permutation ΔR^2
1	PCE dose (wt % binder)	31.2	0.415
2	Fly ash replacement (wt %)	27.0	0.789
3	Anionic charge (mol %)	8.3	0.048
4	PCE Mw (g/mol)	8.3	0.114
5	PCE viscosity (cps)	8.2	0.131
6	Macromonomer ratio	7.9	0.142
7	Solid content (%)	4.1	0.016
8	PCE PDI	3.2	0.008
9	PCE family = K-ratio	0.9	0.006
10	PCE family = S	0.9	0.006

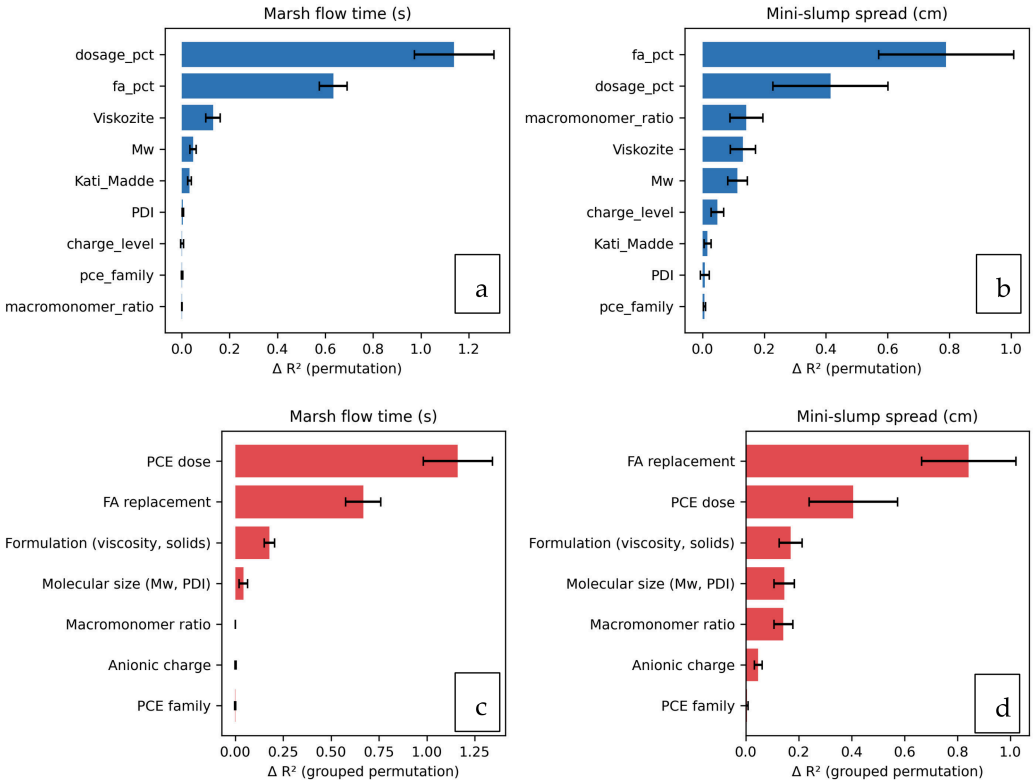


Figure 7. Feature importance for the XGBoost surrogate on the reduced nine-feature set. (**a**,**b**) Individual permutation importance (20 random permutations per feature on the hold-out partition; bars: mean ΔR^2; whiskers: standard deviation). (**c**,**d**) Grouped permutation importance in which physically related features are permuted jointly.

The 80-replicate bootstrap ensemble shown in Figure 9 was used to quantify model uncertainty. The 95% bootstrap model-spread band covered 83.1% of the observed flow-time values, with mean and median widths of 11.8 and 5.3 s, respectively, and 64.7% of the mini-slump values, with a mean width of 1.02 cm. These bands represent variation among retrained models and do not include residual or measurement variance; they are, therefore, not interpreted as prediction intervals. Split-conformal prediction intervals for in-distribution observations are reported separately in Section 5.8.

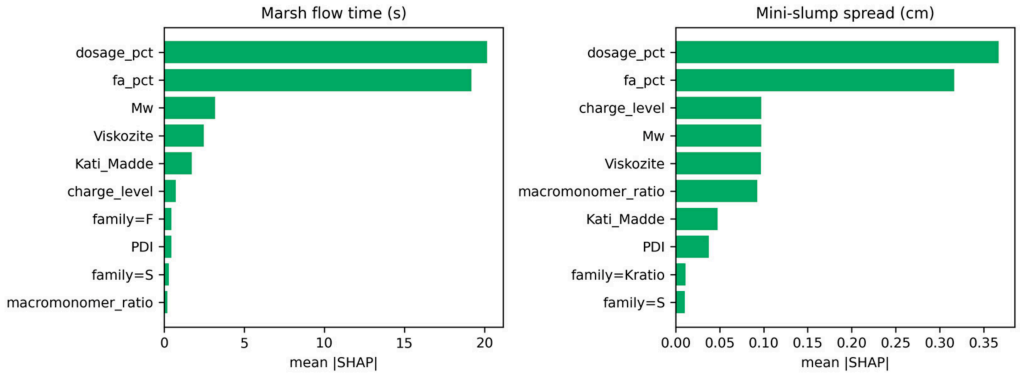


Figure 8. Mean absolute SHAP values for the same XGBoost surrogate as Figure 7. The ordering is broadly consistent with permutation importance.

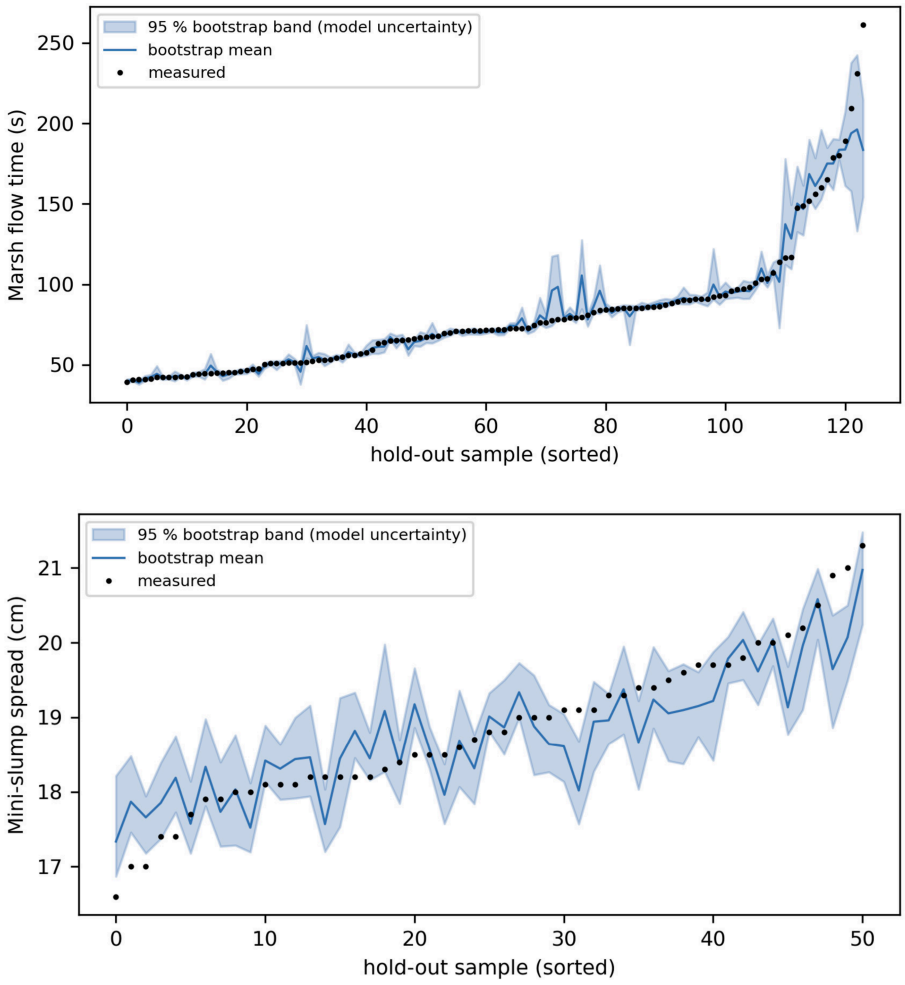


Figure 9. Bootstrap model-uncertainty bands over 80 resamples on the hold-out partition ((**top**): flow time; (**bottom**): mini-slump). Filled band: 2.5th–97.5th percentile spread of the 80 retrained models (epistemic uncertainty only—not a prediction interval); coloured curve: bootstrap mean; black markers: measured values, sorted by magnitude.

5.5. Model-Derived Interpretation by PCE Family

Figure 10 stratifies the two rheological responses by PCE family and by anionic charge level. The K, K-ratio, F, and S families occupy clearly separated regions of the flow-time distribution (panel a), but their median mini-slump values overlap substantially (panel b), suggesting that the family-level effect on the two responses is partly decoupled. The F-series exhibits a near-monotonic decrease in flow time with anionic charge level up to ≈ 9 mol %,

beyond which a plateau emerges (panel c). The S-series shows a non-monotonic trend with a minimum around 5 mol % charge—a trend that may reflect formulation-level differences among the S-series polymers [4,5,8]. Figure 11 reports the corresponding dose-response curves stratified by fly ash replacement level.

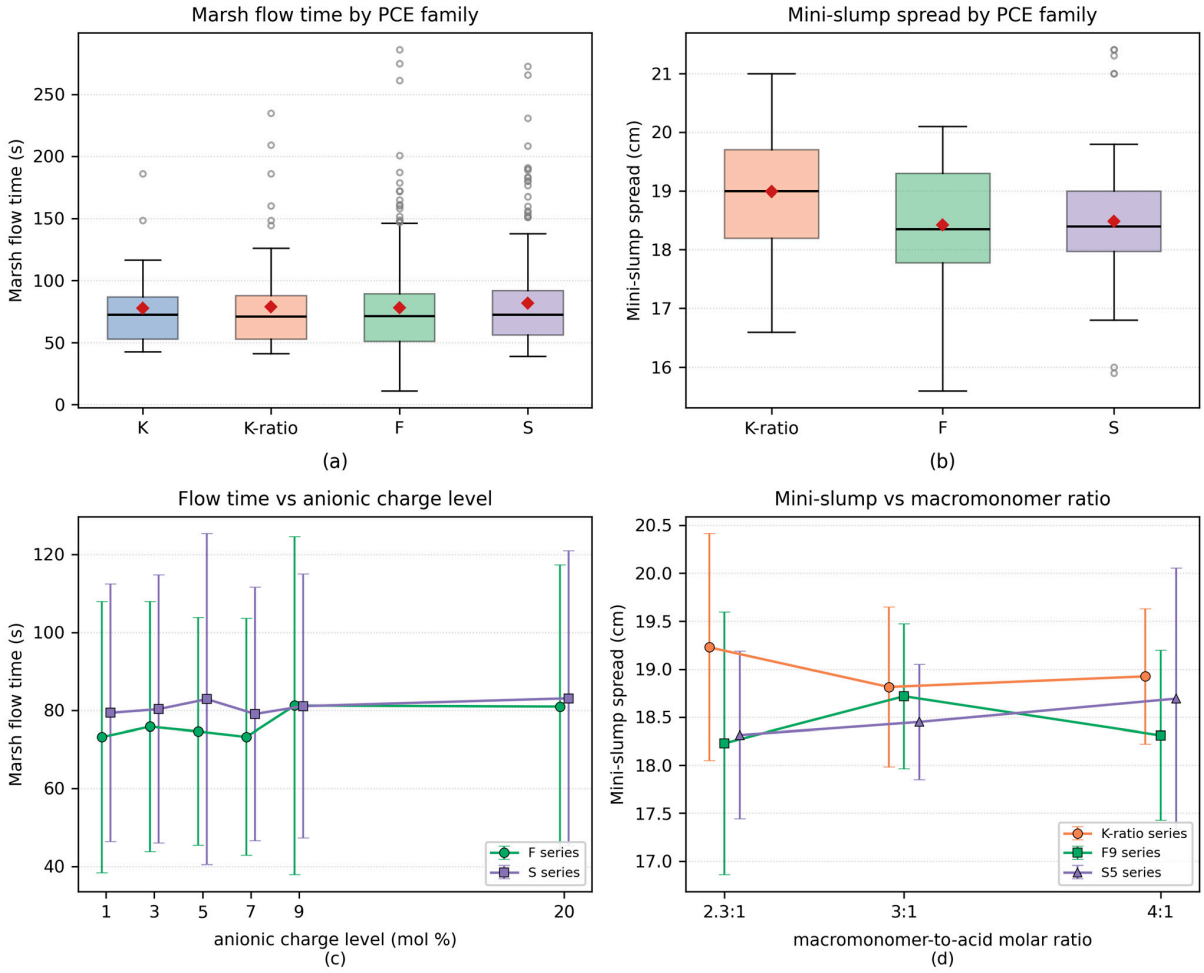


Figure 10. Family- and charge-level effects on the two rheological responses. (**a**,**b**) Box plots stratified by PCE family. (**c**,**d**) Mean ± standard deviation as a function of the anionic charge level for F and S families. Error bars correspond to one standard deviation across dosage levels and replacement ratios.

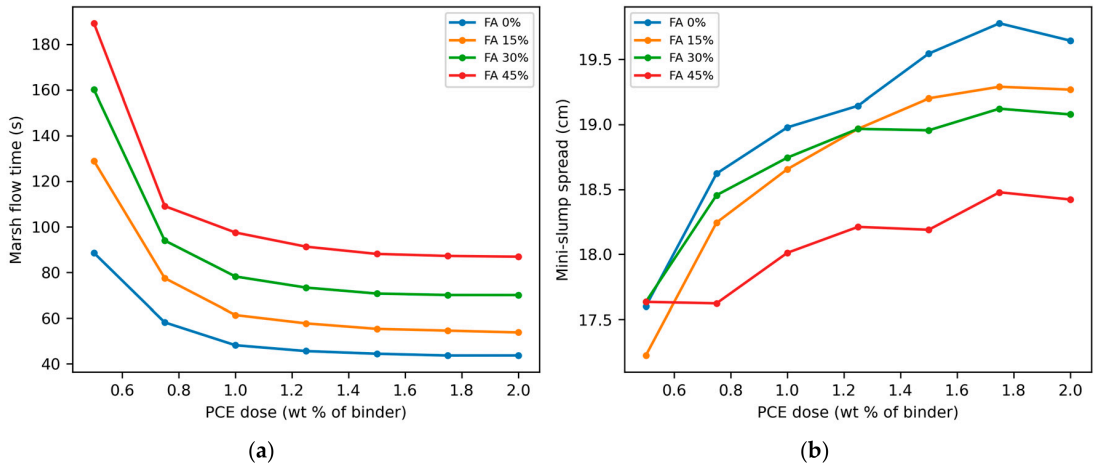


Figure 11. Dose-response curves stratified by fly ash replacement level. Marsh funnel flow time decreases steeply between 0.5 and 1.25 wt % PCE before reaching a plateau (**a**). Mini-slump shows a smoother increase up to about 1.5 wt % with a mild ceiling thereafter (**b**).

5.6. Grouped Generalisation Tests (Primary Validation)

The LOPO, LOPFO, and LOFLO results are summarised in Figure 12.

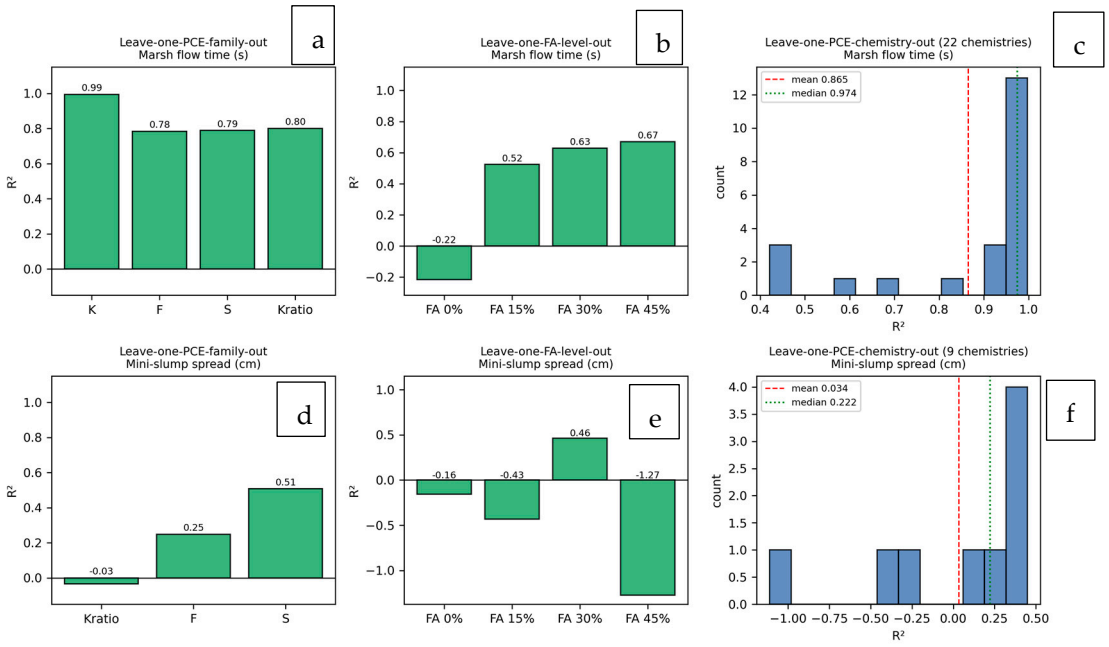


Figure 12. Grouped generalisation tests (primary validation). (**a**,**d**) Leave-one-PCE-family-out R^2 for Marsh flow time and mini-slump. (**b**,**e**) Leave-one-fly-ash-level-out R^2. (**c**,**f**) Leave-one-PCE-chemistry-out R^2 distribution across 22 held-out chemistries (Marsh) and 9 (mini-slump). Mean and median are marked with red and green dashed lines.

Flow-time prediction retained substantially higher performance than mini-slump prediction when complete formulation groups were excluded. Across the twenty-two LOPO folds, flow time produced a mean R^2 of 0.865 $\pm$ 0.198 and a median R^2 of 0.974. The corresponding LOPFO values were 0.995 for K, 0.801 for K-ratio, 0.784 for F, and 0.788 for S. Mini-slump showed markedly weaker group-exclusion performance, with a LOPO mean R^2 of 0.034 $\pm$ 0.530 and LOPFO values not exceeding 0.51. Prediction accuracy also decreased when the boundary fly ash levels were withheld; flow-time R^2 was -0.22 when the 0% fly ash level was excluded, while mini-slump produced negative R^2 values at three of the four excluded levels. These results restrict the use of the flow-time model to the formulation and compositional domain represented in the present dataset and restrict the mini-slump model to interpolation among the sampled formulations.

5.7. Feature-Group Ablation

The relative contributions of mixture variables and polymer descriptors were examined by retraining the XGBoost model under five feature scenarios: (i) mixture variables only, (ii) polymer descriptors including PCE family, (iii) polymer descriptors excluding PCE family, (iv) the complete nine-input set, and (v) the complete set excluding PCE family. Each scenario was evaluated using the same 3 $\times$ 5-fold repeated cross-validation partitions as the primary comparison. The resulting R^2 values are shown in Figure 13.

Mix variables alone explain 78.1% of the Marsh-flow variance and 33.4% of the mini-slump variance, which sets a non-trivial floor and is consistent with the comparatively large contribution of dose and binder composition. Polymer descriptors alone produce $R^2 \approx 0$ for both responses, indicating that PCE chemistry is informative only when contextualised by the dose at which it is applied. The combination of the two feature groups recovers the full reported accuracy (R^2 = 0.946 for flow time and 0.778 for mini-slump). The PCE

family categorical contribution contributes less than 0.002 to either score, meaning that the continuous polymer descriptors (charge level, macromonomer ratio, Mw, PDI, viscosity, and solids) already encode the family-level differences relevant to the prediction task.

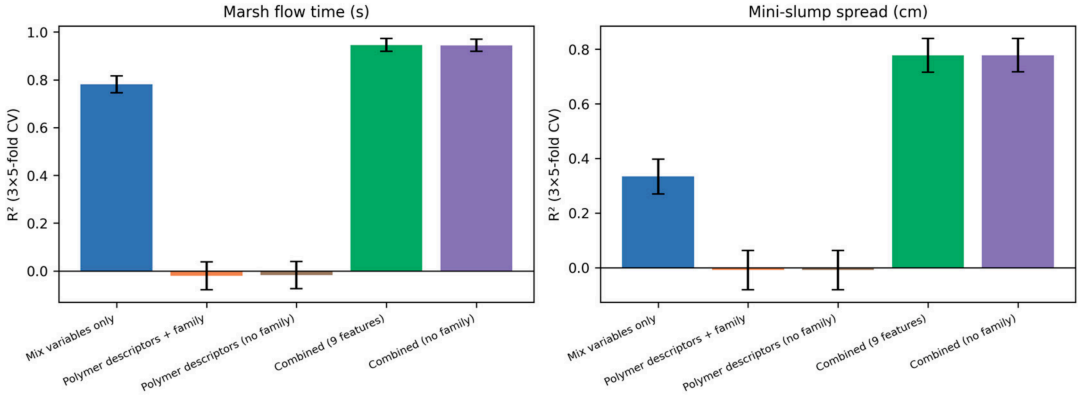


Figure 13. Feature-group ablation under 3 × 5-fold repeated CV. Bars represent R^2 mean; error bars show one standard deviation across the fifteen folds. The combined feature set (green) is essential to recover the full predictive accuracy.

5.8. Split-Conformal Prediction Intervals

Split-conformal prediction was applied using a 60/20/20 train–calibration–test division with seed 42. The corresponding split sizes were 369/123/124 for flow time and 150/51/51 for mini-slump. The conformal quantile was calculated from the absolute calibration residuals and applied to the independent test partition. At a nominal coverage level of 90%, the empirical coverage on this single split was 85.5% for flow time (half-width ±7.2 s) and 90.2% for mini-slump (half-width ±0.83 cm). The flow-time value is below the nominal level and is described here as moderate undercoverage; with 124 test observations, the 95% binomial interval around the empirical value is 79.3–91.7%, so the shortfall is compatible with finite-sample fluctuation. To assess calibration beyond a single partition, the full conformal procedure (model refit, calibration, and evaluation) was repeated over fifty random train–calibration–test splits (seeds 0–49): the mean empirical coverage at the 90% level was 89.4 ± 3.6% for flow time (range 80.6–96.0%) and 90.1 ± 5.8% for mini-slump (range 74.5–100%), indicating that the procedure is well calibrated on average and that the single-split value reflects sampling variability rather than systematic miscalibration. The empirical and nominal coverage levels are compared in Figure 14. These results apply to random, in-distribution observations and do not establish interval coverage when an entire PCE formulation or material family is absent from training. The bootstrap bands reported in Section 5.3 quantify model spread and are not prediction intervals; consequently, their coverages are not directly comparable with the split-conformal coverages.

5.9. Nested Cross-Validation

Nested cross-validation was used to examine whether the reported XGBoost scores were materially affected by the tested parameter grid. A five-fold outer loop and a three-fold inner loop were used. The inner grid contained 27 combinations of maximum depth {3, 4, 5}, learning rate {0.03, 0.05, 0.10}, and number of estimators {200, 300, 400}, with RMSE as the selection criterion. The resulting outer-fold estimates were $R^2 = 0.947 \pm 0.021$ for flow time and $R^2 = 0.784 \pm 0.039$ for mini-slump, as shown in Figure 15. Their agreement with the fixed-parameter estimates indicates that the tested XGBoost grid did not materially alter the reported scores. It does not establish that overfitting was absent. To examine whether the fixed-configuration benchmark disadvantaged the principal competing algorithms, the identical nested protocol (5-fold outer × 3-fold inner, inner-loop

RMSE) was additionally applied to Gradient Boosting (same grid as XGBoost), Random Forest (n_estimators ∈ {200, 400}; max_depth ∈ {none, 10, 20}; max_features ∈ {$\sqrt{p}$, p}), and SVR-RBF (C ∈ {1, 10, 100}; ε ∈ {0.05, 0.1, 0.2}; γ ∈ {scale, 0.1}). The tuned nested estimates are R^2 = 0.944 ± 0.021 (Gradient Boosting), 0.886 ± 0.026 (Random Forest), and 0.787 ± 0.064 (SVR-RBF) for flow time, and 0.752 ± 0.066, 0.735 ± 0.027, and 0.755 ± 0.047 for mini-slump. Tuning benefits SVR-RBF most (its flow-time R^2 rises from 0.617 at the fixed configuration to 0.787), confirming that the fixed-configuration scores understate the attainable accuracy of some competitors. Nevertheless, the ranking relevant to model selection is unchanged—the boosted-tree ensembles remain the top tier on both responses, and XGBoost remains the selected surrogate.

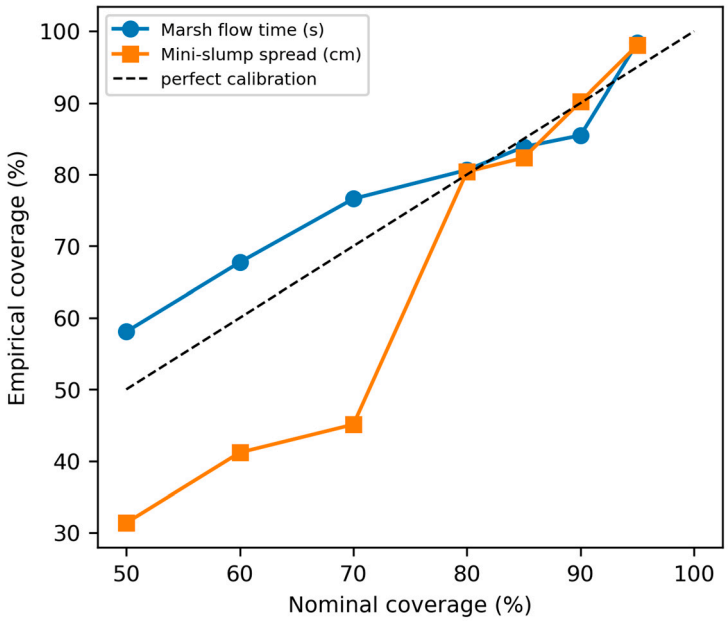


Figure 14. Inductive conformal-prediction calibration curves for both responses. The dashed line denotes perfect calibration; the curves correspond to the single seed-42 split, whose deviations from the nominal level lie within finite-sample (binomial) fluctuation; over fifty repeated splits, the mean coverage at the 90% level is 89.4% (flow time) and 90.1% (mini-slump).

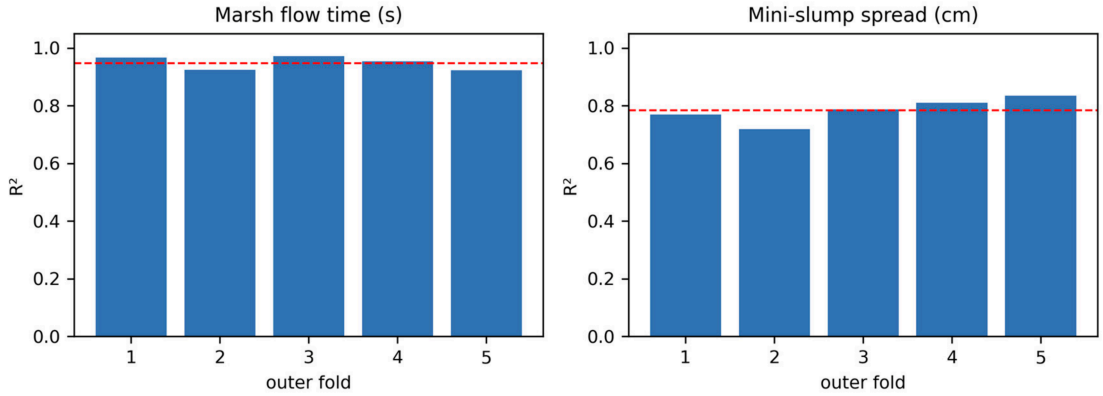


Figure 15. Outer-fold R^2 values from 5-fold outer and 3-fold inner nested cross-validation. The dashed line marks the mean value for each response.

5.10. Modern Gradient-Boosting Comparison (CatBoost/LightGBM/XGBoost)

LightGBM and CatBoost were evaluated under the same preprocessing, 3 × 5-fold repeated cross-validation partitions and fold assignments as XGBoost. The numerical

results are reported in Table 9, and the corresponding mean R^2 values are shown in Figure 16. This analysis is supplementary to the ten-model comparison reported in Tables 5 and 6.

Table 9. Modern gradient-boosting comparison on both responses (3 × 5-fold repeated CV, identical preprocessing and folds).

Target	Model	R^2 (Mean ± Std)	RMSE	MAE
Flow time	XGBoost	0.946 ± 0.027	8.40 s	3.46 s
Flow time	LightGBM	0.934 ± 0.039	9.10 s	4.10 s
Flow time	CatBoost	0.952 ± 0.027	7.81 s	3.25 s
Mini-slump	XGBoost	0.778 ± 0.061	0.471 cm	0.363 cm
Mini-slump	LightGBM	0.731 ± 0.047	0.521 cm	0.404 cm
Mini-slump	CatBoost	0.786 ± 0.038	0.464 cm	0.356 cm

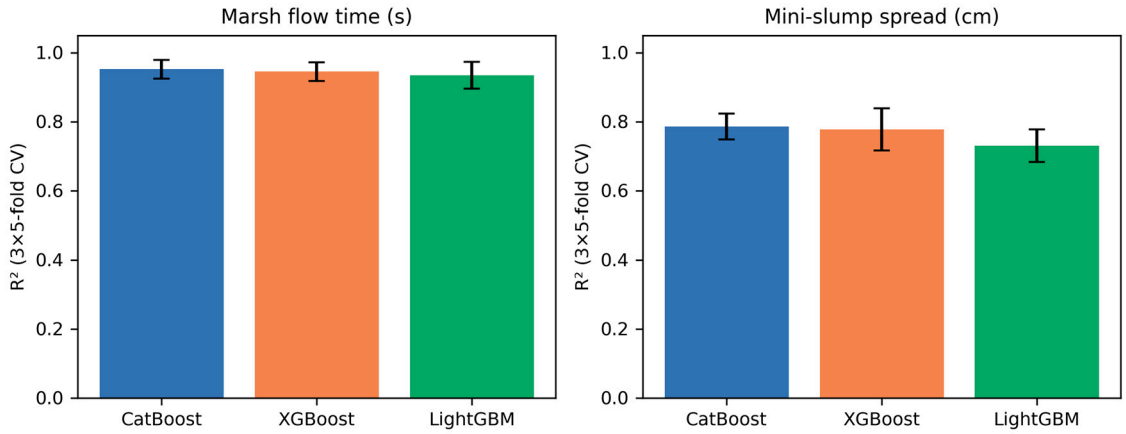


Figure 16. Modern gradient-boosting comparison on the two rheological responses under the identical 3 × 5-fold repeated CV and identical folds. Bars show R^2 mean; whiskers show one standard deviation.

On flow time, CatBoost yields a marginal improvement over XGBoost (R^2 = 0.952 ± 0.027 versus 0.946 ± 0.027; RMSE = 7.81 s versus 8.40 s), with LightGBM third (R^2 = 0.934 ± 0.039). On mini-slump, CatBoost and XGBoost are likewise within one standard deviation of each other (R^2 = 0.786 ± 0.038 versus 0.778 ± 0.061), while LightGBM trails by ≈ 0.05 R^2 points. The differences across the three boosters are within one standard deviation throughout, indicating that the choice of booster family is not a primary driver of predictive accuracy on this dataset. XGBoost is retained as the reference surrogate for three practical reasons: its tree-exact SHAP integration used throughout Section 5.4, its maturity in the materials-informatics literature, and the reproducibility of the full interpretability pipeline. XGBoost was retained as the reference model to maintain one consistent pipeline for SHAP, permutation importance, ablation, bootstrap, and grouped-validation analyses; this choice is not interpreted as evidence of statistical superiority over CatBoost.

6. Discussion

The better performance of the boosted-tree models than the linear models indicates that the measured responses depend on non-linear effects and interactions among PCE dose, fly ash replacement, and formulation descriptors. This result does not identify a particular adsorption or dispersion mechanism, because the models were trained using macroscopic flow-time and spread measurements rather than adsorption, zeta-potential, or rheometer data.

PCE dose and fly ash replacement were the two most influential inputs for both responses, as shown by Tables 7 and 8 and Figures 7 and 8. Their importance is consistent

with the experimental design. Dose defines the quantity of admixture introduced relative to binder mass, whereas fly ash replacement changes the cement-to-fly ash balance, particle population, and early-reactive clinker fraction at constant total binder and water-to-binder ratio. The flattening of the predicted response at higher dosages, visible in Figure 11, is compatible with a reduction in the incremental effect of additional PCE; however, the present measurements do not establish an adsorption-saturation concentration or identify the solid phases on which adsorption occurred. Similarly, the influence of fly ash cannot be attributed solely to an increase in adsorption surface because its fineness, morphology, carbon content, and surface chemistry may affect admixture demand in different directions.

Mw, PDI, viscosity, solid content, nominal anionic charge level, and acid-to-macromonomer ratio supplied additional predictive information. These formulation-level descriptors remain partly correlated with one another and with the PCE family, as indicated by Figure 3. Their SHAP and permutation rankings, therefore, show which variables were used by the trained model but do not isolate the molecular effect of each descriptor. Mw is not a direct measure of side-chain extension, and Brookfield viscosity is not an independent measure of molecular architecture.

Experimental studies on related PCE formulations have shown that main- and side-chain configuration can alter adsorption, hydrodynamic dimensions, and fresh-state performance [6,11]. In those studies, intermediate chain configurations frequently produced better dispersion than the shortest or longest chains, while excessive main-chain length was associated with lower adsorption efficiency and possible bridging behaviour. These findings provide a physicochemical context for the formulation effects observed in the present model, but they cannot be transferred directly because adsorption and hydrodynamic diameter were not measured for all twenty-two formulations.

Marsh funnel time and mini-slump spread are empirical responses obtained under different geometries and deformation histories, but neither test alone separates electrostatic and steric contributions. Mini-slump performance was weaker because measurements were available for only nine formulations and because leave-one-formulation-out validation produced a mean R^2 close to zero with substantial variation among excluded formulations, as shown in Figure 12. The result indicates that the available descriptors are insufficient for formulation-level prediction of mini-slump; it does not demonstrate that mini-slump is governed predominantly by an electrostatic mechanism.

Grouped validation defines the scope of the models more clearly than random cross-validation. The flow-time model retained useful performance when individual formulations were excluded within the present experimental matrix, whereas the mini-slump model did not. Accuracy also decreased when the boundary fly ash levels were withheld. The models should, therefore, be confined to the measured dosage and replacement ranges and to binder materials comparable with those examined. The random hold-out results in Figure 5 and the conformal intervals in Figure 14 apply to mixtures drawn from the same experimental campaign and do not constitute external validation.

The principal limitations are the use of one cement, one fly ash, one water-to-binder ratio, and one laboratory campaign; incomplete mini-slump coverage; absence of time-dependent flow measurements; and lack of adsorption, zeta-potential, and direct rheological characterisation. Bulk density and pH were excluded because their original measurement records were incomplete. The interpretation is consequently restricted to the relationships represented in the present dataset. In addition, the dataset stores a single reported response value for each formulation–mixture combination, and the surviving records do not permit technical repetitions to be distinguished reliably from independently prepared repetitions (Section 3.3); measurement repeatability and experimental uncertainty,

therefore, cannot be separated from the model residual error, and the reported RMSE values should be read as an upper bound that pools both contributions.

7. Conclusions

Within the experimental boundaries described above, the present study supports the following conclusions:

- A 616-mixture experimental dataset with twenty-two in-house PCE chemistries (documented individually in Table 3), four fly ash replacement levels and seven dosage levels at fixed w/b = 0.35 was compiled, audited for internal consistency, and is openly released together with the full analysis code as Supplementary Material. The dataset comprises 616 flow-time and 252 mini-slump measurements (nine macromonomer ratio-varied chemistries).
- Among the ten regression algorithms benchmarked on the reduced nine-feature matrix under repeated cross-validation with identical folds, gradient-boosted tree ensembles provided the best predictive accuracy: XGBoost reached R^2 = 0.946 ± 0.027 (RMSE = 8.40 s) for flow time and R^2 = 0.778 ± 0.061 (RMSE = 0.471 cm) for mini-slump, with CatBoost marginally ahead on flow time (0.952 ± 0.027) in the supplementary booster comparison. Friedman and Holm-corrected Wilcoxon tests confirm the ensemble-versus-linear separation ($p < 10^{-16}$), while the variance-corrected Nadeau–Bengio test shows that the top two boosters are statistically indistinguishable.
- These random-CV figures quantify interpolation. Under the primary grouped assessment, the flow-time surrogate transfers well to unseen PCE chemistries (LOPO mean R^2 = 0.865 ± 0.198, median 0.974; LOPFO 0.78–1.00 across all four families), whereas the mini-slump surrogate does not (LOPO mean R^2 = 0.034 ± 0.530), and both responses degrade when boundary fly ash levels are withheld. The practical scope of the tool is, therefore, screening of new PCE chemistries within the sampled design space for Marsh flow and interpolation only for mini-slump.
- PCE dose and fly ash replacement jointly account for approximately 80% (flow time) and 58% (mini-slump) of the absolute SHAP mass, with molecular descriptors acting as secondary modulators: Mw and viscosity for flow time; anionic charge level, macromonomer ratio, Mw, and PDI for mini-slump. Individual and grouped permutation importance reproduce these rankings on the non-redundant feature set. All such readings are reported as model-derived associations and physics-consistent but not mechanistically verified.
- Feature-group ablation shows that the two mix variables alone explain 78.1% of the flow-time variance and 33.4% of the mini-slump variance, while polymer descriptors alone are uninformative without dose context; only the combined feature set recovers the full predictive accuracy (0.946/0.778).
- Uncertainty is quantified in two separated layers: split-conformal 90% prediction intervals achieve single-split empirical coverages of 85.5% (±7.2 s; moderately below the nominal level, within finite-sample binomial fluctuation) and 90.2% (±0.83 cm), with mean coverages of 89.4 ± 3.6% and 90.1 ± 5.8% over fifty repeated splits, and the 80-replicate bootstrap band quantifies model (epistemic) uncertainty only. Nested cross-validation (R^2 = 0.947 ± 0.021 and 0.784 ± 0.039) indicates that the tested parameter grid did not materially alter the reported XGBoost scores; this agreement does not by itself establish the absence of overfitting.
- The conclusions reported here are dataset-specific. The model is a domain-specific surrogate for the cement, fly ash, and water-to-binder ratio used in the experimental campaign rather than a universal rheology predictor. The random hold-out is not an external validation, and independent multi-laboratory replication on alternative

binders and SCMs is required before transferability can be claimed. Because a single response value is stored per mixture and replicate counts cannot be reconstructed from the surviving records, measurement repeatability is not separable from the model residual error, and the reported error metrics pool both contributions.

- The released surrogate may serve as a screening tool for PCE–binder combinations during early-stage mix design—for unseen PCE chemistries in the case of Marsh flow, and for interpolation among sampled chemistries in the case of mini-slump—and as a starting point for transfer learning to other binder–admixture systems.

Future work will extend the present matrix to other binders, time-resolved rheology, zeta-potential, and FT-IR adsorption-isotherm measurements, oscillatory-rheometer characterisation, verified re-measurement of the excluded density and pH descriptors, and a coordinated multi-laboratory replication.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/buildings16152965/s1. Table S1: Chemical composition (wt %) and physical properties of the CEM I 42.5 R Portland cement; Table S2: Chemical composition (wt %) and physical properties of the class-F fly ash; Table S3: Physico-chemical descriptors of the seventeen PCE formulations; Table S4: Hyperparameter configuration of the ten regression algorithms; Table S5: Per-fold RMSE values (s) for the ten models on Marsh-cone flow time; Table S6: Per-fold RMSE values (cm) for the ten models on mini-slump spread; Table S7: Pairwise Wilcoxon-Holm test—Marsh-cone flow time; Table S8: Pairwise Wilcoxon-Holm test—mini-slump spread; Table S9: Full permutation importance—Marsh-cone flow time; Table S10: Full permutation importance—mini-slump spread; Table S11: Full SHAP ranking—Marsh-cone flow time; Table S12: Full SHAP ranking—mini-slump spread; Table S13: LOPFO results—Marsh-cone flow time; Table S14: LOPFO results—mini-slump spread; Table S15: LOFLO results—Marsh-cone flow time; Table S16: LOFLO results—mini-slump spread; Table S17: LOPO results—Marsh-cone flow time; Table S18: LOPO results—mini-slump spread; Table S19: Feature-group ablation—repeated 3 × 5-fold CV; Table S20: Nested CV—per-outer-fold results and best hyperparameters; Table S21: Conformal calibration—Marsh-cone flow time; Table S22: Conformal calibration—mini-slump spread.

Author Contributions: Conceptualization, V.G., S.E. and A.M.; methodology, V.G., S.E. and D.I.; software, V.G. and B.B.G.; validation, S.E., D.I. and C.B.; formal analysis, V.G. and B.B.G.; investigation, V.G., D.I. and C.B.; resources, S.E., C.B., S.Ö. and A.M.; data curation, V.G. and D.I.; writing—original draft preparation, V.G. and S.Ö.; writing—review and editing, V.G., S.E., D.I., B.B.G., C.B. and A.M.; visualization, V.G. and B.B.G.; supervision, S.E. and A.M.; project administration, V.G.; funding acquisition, S.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Bursa Uludağ University Scientific Research Projects Unit (BAP), grant number FBG-2025-2550 and FGA-2025-2048. The APC was funded by Bursa Uludağ University Scientific Research Projects Unit (BAP) through Grant Nos. FBG-2025-2550, FGA-2025-2048 and FHIZ-2026-2891.

Data Availability Statement: The corrected experimental dataset (PCE3_clean_R1.csv; n = 616 mixtures, including the curated mini-slump subset, n = 252), the complete formulation inventory (Table 3), and the full Python code used for preprocessing, model training, grouped validation, statistical testing, SHAP analysis, uncertainty quantification, and figure generation are openly released with this article as Supplementary Material, and have additionally been deposited in a public Zenodo repository (https://doi.org/10.5281/zenodo.21416492). No restrictions apply.

Acknowledgments: The authors gratefully acknowledge the laboratory staff who supported the flow-time and mini-slump campaigns, the admixture suppliers for providing material safety data and GPC characterisations, and the cement and fly ash producers for supplying the binder lots. Computational work made use of open-source Python libraries (scikit-learn, XGBoost, LightGBM, CatBoost, and SHAP), and the authors thank the maintainers of these projects. The corresponding author further gratefully acknowledges the support of the Turkish Academy of Sciences (TÜBA).

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

Abbreviations

The following abbreviations are used in this manuscript:

PCE	Polycarboxylate ether superplasticiser
SCM	Supplementary cementitious material
ML	Machine learning
XGBoost	eXtreme Gradient Boosting
SHAP	Shapley additive explanations
PDP	Partial-dependence plot
ICE	Individual conditional expectation
SVR-RBF	Support-vector regression with radial basis function kernel
k-NNs	k-nearest neighbours
CV	Cross-validation
RMSE	Root-mean-squared error
MAE	Mean absolute error
GPC	Gel-permeation chromatography
FT-IR	Fourier-transform infrared spectroscopy
BAP	Bilimsel Araştırma Projeleri (Scientific Research Projects Unit)
DOE	Design of experiments
XAI	Explainable artificial intelligence
LOPFO	Leave-one-PCE-family-out (generalisation protocol)
LOFLO	Leave-one-fly-ash-level-out (generalisation protocol)
LOPO	Leave-one-PCE-chemistry-out (generalisation protocol)
SSA	Specific surface area (Blaine fineness)
SG	Specific gravity
SYMBOLS	
M_w	Weight-average molecular weight (kg/mol)
PDI	Polydispersity index (M_w/M_n)
w/b	Water-to-binder (cementitious material) ratio = 0.35 (fixed)
FA/b	Fly ash-to-binder ratio (mass fraction)
R2	Coefficient of determination
φ	Permutation importance score (ΔR^2)
ϕ_0	SHAP base value (expected model output over training set)
ϕ_i	SHAP additive contribution of feature i to a prediction
η	Dynamic viscosity (cps, Brookfield)
ρ	Spearman rank-correlation coefficient
$\Omega(f_k)$	Regularisation term penalising tree complexity (XGBoost objective)
$l(\cdot)$	Differentiable loss function (squared error in this study)
n	Number of observations in the dataset
y_i	Measured response value for observation i
$\hat{y}_i$	Model-predicted response value for observation i
$\bar{y}$	Sample mean of the measured response
χ^2	Chi-squared test statistic (Friedman test)
p	Probability value (statistical significance)
I_i	Permutation importance of feature i (ΔR2 when feature i is shuffled)

References

1. Yamada, K.; Takahashi, T.; Hanehara, S.; Matsuhisa, M. Effects of the chemical structure on the properties of polycarboxylate-type superplasticizer. *Cem. Concr. Res.* **2000**, *30*, 197–207. [CrossRef]
2. Sakai, E.; Yamada, K.; Ohta, A. Molecular structure and dispersion–adsorption mechanisms of comb-type superplasticizers used in Japan. *J. Adv. Concr. Technol.* **2003**, *1*, 16–25. [CrossRef]

3. Plank, J.; Hirsch, C. Impact of zeta potential of early cement hydration phases on superplasticizer adsorption. *Cem. Concr. Res.* **2007**, *37*, 537–542. [CrossRef]
4. Plank, J.; Sakai, E.; Miao, C.; Yu, C.; Hong, J. Chemical admixtures—Chemistry, applications and their impact on concrete microstructure and durability. *Cem. Concr. Res.* **2015**, *78*, 81–99. [CrossRef]
5. Marchon, D.; Sulser, U.; Eberhardt, A.; Flatt, R.J. Molecular design of comb-shaped polycarboxylate dispersants for environmentally friendly concrete. *Soft Matter* **2013**, *9*, 10719–10728. [CrossRef]
6. Kobya, V.; Karakuzu, K.; Mardani, A.; Felekoğlu, B.; Ramyar, K. Effect of polycarboxylate-based water-reducing admixture chains length on Portland cement–admixture compatibility. *J. Sustain. Cem.-Based Mater.* **2023**, *13*, 69–86. [CrossRef]
7. Houst, Y.F.; Bowen, P.; Perche, F.; Kauppi, A.; Borget, P.; Galmiche, L.; Le Meins, J.-F.; Lafuma, F.; Flatt, R.J.; Schober, I.; et al. Design and function of novel superplasticizers for more durable high-performance concrete (SUPERPLAST project). *Cem. Concr. Res.* **2008**, *38*, 1197–1209. [CrossRef]
8. Zingg, A.; Winnefeld, F.; Holzer, L.; Pakusch, J.; Becker, S.; Gauckler, L. Adsorption of polyelectrolytes and its influence on the rheology, zeta potential, and microstructure of various cement and hydrate phases. *J. Colloid Interface Sci.* **2008**, *323*, 301–312. [CrossRef] [PubMed]
9. Flatt, R.J.; Houst, Y.F. A simplified view on chemical effects perturbing the action of superplasticizers. *Cem. Concr. Res.* **2001**, *31*, 1169–1176. [CrossRef]
10. Kobya, V.; Mardani, A. Interaction of polycarboxylate-based water-reducing admixture molecular structure with fly-ash-substituted cementitious systems: Marsh-funnel and mini-slump performance. *Eurasia Proc. Sci. Technol. Eng. Math.* **2023**, *26*, 693–699. [CrossRef]
11. Kobya, V.; Karakuzu, K.; Mardani, A.; Felekoğlu, B.; Ramyar, K. Combined interaction of PCE chain lengths, C_3A and water content in cementitious systems. *Constr. Build. Mater.* **2023**, *378*, 131178. [CrossRef]
12. Bessaies-Bey, H.; Baumann, R.; Schmitz, M.; Radler, M.; Roussel, N. Effect of polyacrylamide on rheology of fresh cement pastes. *Cem. Concr. Res.* **2015**, *76*, 98–106. [CrossRef]
13. Ma, Y.; Sha, S.; Zhou, B.; Lei, F.; Liu, Y.; Xiao, Y.; Shi, C. Adsorption and dispersion capability of polycarboxylate-based superplasticizers: A review. *J. Sustain. Cem.-Based Mater.* **2022**, *11*, 319–344. [CrossRef]
14. Roussel, N. A theoretical frame to study stability of fresh concrete. *Mater. Struct.* **2006**, *39*, 81–91. [CrossRef]
15. Roussel, N.; Lemaître, A.; Flatt, R.J.; Coussot, P. Steady-state flow of cement suspensions: A micromechanical state of the art. *Cem. Concr. Res.* **2010**, *40*, 77–84. [CrossRef]
16. Wallevik, O.H.; Wallevik, J.E. Rheology as a tool in concrete science: The use of rheographs and workability boxes. *Cem. Concr. Res.* **2011**, *41*, 1279–1288. [CrossRef]
17. Pal, R. Non-Newtonian behaviour of suspensions and emulsions: Review of different mechanisms. *Adv. Colloid Interface. Sci.* **2024**, *333*, 103299. [CrossRef]
18. Xu, K.; Yang, J.; He, H.; Wei, J.; Zhu, Y. Influences of additives on the rheological properties of cement composites: A review of material impact. *Mater* **2025**, *18*, 1753. [CrossRef] [PubMed]
19. Kumar, A.; Arora, H.C.; Kapoor, N.R.; Mohammed, M.A.; Kumar, K.; Majumdar, A.; Thinnukool, O. Compressive strength prediction of lightweight concrete: Machine learning model. *Sustainability* **2022**, *14*, 2404. [CrossRef]
20. Ziolkowski, P.; Niedostatkiewicz, M. Machine learning techniques in concrete mix design. *Materials* **2019**, *12*, 1256. [CrossRef] [PubMed]
21. Ünverdi, M.; Kaya, Y.; Mardani, N.; Mardani, A. Investigation of the microstructural and mechanical properties of fiber-reinforced roller-compacted concrete under high-temperature exposure. *Materials* **2025**, *18*, 2430. [CrossRef]
22. İsafça-Kaya, T.; Kaya, Y.; Mardani, N.; Mardani, A.; Doğangün, A. Improving lime-based restoration mortars: Effect of type and utilization rate of binder and aggregate. *Materials* **2025**, *18*, 961. [CrossRef] [PubMed]
23. Yeh, I.-C. Modeling of strength of high-performance concrete using artificial neural network. *Cem. Concr. Res.* **1998**, *28*, 1797–1808. [CrossRef]
24. Ben Chaabene, W.; Flah, M.; Nehdi, M.L. Machine learning prediction of mechanical properties of concrete: Critical revie. *Constr. Build. Mater.* **2020**, *260*, 119889. [CrossRef]
25. Salami, B.A.; Iqbal, M.; Abdulraheem, A.; Jalal, F.E.; Alimi, W.; Jamal, A.; Tafsirojjaman, T.; Liu, Y.; Bardhan, A. Estimating compressive strength of lightweight foamed concrete using neural, genetic and ensemble machine learning approaches. *Cem. Concr. Compos.* **2022**, *133*, 104721. [CrossRef]
26. Tipu, R.K.; Panchal, V.; Pandya, K. An ensemble approach to improve BPNN model precision for predicting compressive strength of high-performance concrete. *Structures* **2022**, *45*, 500–508. [CrossRef]
27. Sobuz, H.R.; Khatun, M.; Kabbo, K.I.; Sutan, N.M. An explainable machine learning model for encompassing the mechanical strength of polymer-modified concrete. *Asian J. Civ. Eng.* **2024**, *26*, 931–954. [CrossRef]
28. Mai, H.-V.T.; Nguyen, T.-A.; Ly, H.-B.; Tran, V.Q. Investigation of ANN model containing one hidden layer for predicting compressive strength of concrete with blast-furnace slag and fly ash. *Adv. Mater. Sci. Eng.* **2021**, *2021*, 5540853. [CrossRef]

29. Unverdi, M.; Kazemi, R.; Kaya, Y.; Mardani, N.; Mardani, A.; Mirjalili, S. Predicting compressive and splitting tensile strength of high-volume fly-ash roller-compacted concrete using ANN and ANN-biogeography-based optimization models. *Sci. Rep.* **2025**, *15*, 21794. [CrossRef] [PubMed]
30. Cakiroglu, C.; Aydın, Y.; Bekdaş, G.; Geem, Z.W. Interpretable predictive modelling of basalt-fiber-reinforced concrete splitting tensile strength using ensemble machine learning methods and SHAP approach. *Materials* **2023**, *16*, 4578. [CrossRef] [PubMed]
31. Naser, M. An engineer's guide to explainable artificial intelligence and interpretable machine learning: Navigating causality, forced goodness, and the false perception of inference. *Autom. Constr.* **2021**, *129*, 103821. [CrossRef]
32. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [CrossRef] [PubMed]
33. Jia, J.-F.; Chen, X.-Z.; Bai, Y.-L.; Li, Y.-L.; Wang, Z.-H. An interpretable ensemble learning method to predict the compressive strength of concrete. *Structures* **2022**, *46*, 201–213. [CrossRef]
34. Bahmani, H.; Mostafaei, H.; Santos, P.; Ferrández, D. Concrete material variability and machine learning model performance: A comprehensive review. *Buildings* **2026**, *16*, 556. [CrossRef]
35. Kantro, D. Influence of water-reducing admixtures on properties of cement paste—A miniature slump test. *Cem. Concr. Aggreg.* **1980**, *2*, 95–102. [CrossRef]
36. Tan, Z.; Bernal, S.A.; Provis, J.L. Reproducible mini-slump test procedure for measuring the yield stress of cementitious pastes. *Mater. Struct.* **2017**, *50*, 235. [CrossRef]
37. Oizuni, A.S.; dos Santos, B.C.R.; Lofrano, F.C.; Romano, R.C.d.O.; Monte, R.; Kurokawa, F.A. Computational fluid dynamics analysis of single-point rheological tests for assessing cementitious material workability. *Constr. Build. Mater.* **2026**, *518*, 145846. [CrossRef]
38. Altun, M.G.; Özen, S.; Mardani-Aghabaglou, A. Effect of side-chain-length change of polycarboxylate-ether-based high-range water-reducing admixture on properties of cementitious systems containing fly ash. *J. Mater. Civ. Eng.* **2021**, *33*, 04021015. [CrossRef]
39. Ahmed, A. Assessing the effects of supplementary cementitious materials on concrete properties: A review. *Discov. Civ. Eng.* **2024**, *1*, 1–47. [CrossRef]
40. Nassiri, S.; Butt, A.A.; Zarei, A.; Roy, S.; Filani, I.; Pandit, G.A.; Mateos, A.; Haider, M.; Harvey, J.T. Opportunities for supplementary cementitious materials from natural sources and industrial byproducts: Literature insights and supply assessment. *Buildings* **2025**, *15*, 3099. [CrossRef]
41. Qing, S.; Li, C. Data-driven prediction on critical mechanical properties of engineered cementitious composites based on machine learning. *Sci. Rep.* **2024**, *14*, 15322. [CrossRef] [PubMed]
42. Khan, M.; Javed, M.F.; Alabduljabbar, H.; Ahmad, F. Explainable machine learning and ensemble models for predicting fresh properties of self-consolidating concret. *Sci. Rep.* **2025**, *15*, 42288. [CrossRef] [PubMed]
43. Liu, Y.; Mohammed, Z.M.A.; Ma, J.; Xia, R.; Fan, D.; Tang, J.; Yuan, Q. Machine learning-driven fluidity and rheological properties prediction of fresh cement-based material. *Materials* **2024**, *17*, 5400. [CrossRef] [PubMed]
44. Adler, A.I.; Painsky, A. Feature importance in gradient boosting trees with cross-validation feature selection. *Entropy* **2022**, *24*, 687. [CrossRef]
45. 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, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [CrossRef]
46. Grinsztajn, L.; Oyallon, E.; Varoquaux, G. Why do tree-based models still outperform deep learning on typical tabular data? *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 507–520. [CrossRef]
47. Shwartz-Ziv, R.; Armon, A. Tabular data: Deep learning is not all you need. *Inf. Fusion* **2022**, *81*, 84–90. [CrossRef]
48. Nelli, F. Machine learning with scikit-learn. In *Python Data Analytics*; Springer Nature: Berlin, Germany, 2023; pp. 259–287. [CrossRef]
49. Durdy, S.; Gaultois, M.W.; Gusev, V.V.; Bollegala, D.; Rosseinsky, M.J. Random projections and kernelised leave-one-cluster-out cross-validation: Universal baselines and evaluation tools for supervised machine learning of material properties. *Digit. Discov.* **2022**, *1*, 763–778. [CrossRef]
50. Nadeau, C.; Bengio, Y. Inference for the generalization error. *Mach. Learn.* **2003**, *52*, 239–281. [CrossRef]
51. Angelopoulos, A.N.; Bates, S. Conformal prediction: A gentle introduction. *Found. Trends Mach. Learn.* **2023**, *16*, 494–591. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.