Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Explainable Machine Learning for Predicting the Splitting Tensile Strength of SCM-Blended Concrete Using Optimized Gaussian Process Regression

Ali Alnaqbi
University of Sharjah
Abdalla Mahmoud Salim
University of Sharjah
Mohammad Al-Zu'bi
m.alzubi@aut.edu.jo

Aqaba University of Technology
Ibrahim Elnaml
University of Illinois Urbana-Champaign
Abdulaziz Alotaibi
Islamic University of Madinah
Ahmed Mohammad Youssef
New Mansoura University

Research Article

Keywords: Splitting tensile strength, SCM-blended concrete, Machine learning, Bayesian optimization, Partial dependence plots, Explainable artificial intelligence

Posted Date: September 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10989183/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Explainable Machine Learning for Predicting the Splitting Tensile Strength of SCM-Blended Concrete Using Optimized Gaussian Process Regression

Ali Alnaqbi [1], Abdalla Mahmoud Salim [2], Mohammad Al-Zu'bi [3*], Ibrahim Elnaml [4], Abdulaziz Alotaibi [5], Ahmed Mohammad Youssef [6]

1. **Department of Civil and Environmental Engineering, University of Sharjah, Sharjah, United Arab Emirates; E-mail: U21102866@sharjah.ac.ae**
2. **Department of Industrial Engineering and Engineering Management, University of Sharjah, Sharjah, United Arab Emirates; E-mail: U17104768@sharjah.ac.ae**
3. **Department of Civil Engineering, Munib and Angela Masri Faculty of Engineering, Aqaba University of Technology, Aqaba 11947, Jordan; m.alzubi@aut.edu.jo**
4. **Department of Civil and Environmental Engineering, University of Illinois Urbana-Champaign, Urbana, USA. E-mail: elnaml@illinois.edu**
5. **Department of Civil Engineering, Faculty of Engineering, Islamic University of Madinah, Madinah, 42351, Saudi Arabia. Email: abdulaziz.alotaibi@iu.edu.sa**
6. **Civil Engineering Department, Faculty of Engineering, New Mansoura University, New Mansoura, Egypt. E-mail: ahmedyoussef@nmu.edu.eg**

[*]**Correspondence: m.alzubi@aut.edu.jo**

Abstract: This study develops an explainable machine-learning framework to predict the splitting tensile strength of concrete incorporating supplementary cementitious materials (SCMs), which are widely used to reduce reliance on Portland cement. The analysis used 395 valid mixtures extracted from a global SCM-blended concrete database. Cement, water, coarse aggregate, fine aggregate, fly ash, silica fume, ground granulated blast-furnace slag, and superplasticizer were used as predictors. Predictor relevance was assessed using mRMR, F-test, RReliefF, and Random Forest importance. A broad set of regression algorithms was evaluated in MATLAB Regression Learner using 5-fold cross-validation. Rational Quadratic Gaussian Process Regression (GPR) was the best baseline model, with an RMSE of 0.51 MPa, MAE of 0.34 MPa, MAPE of 10.02%, and R^2 of 0.85. The GPR model was then optimized using Bayesian optimization, random search, and grid search. All three strategies produced comparable predictive accuracy. Bayesian optimization required the shortest training time and was therefore selected for the final model. The Bayesian-optimized GPR achieved an RMSE of 0.41898 MPa, MSE of 0.17554 MPa^2, MAE of 0.29095 MPa, MAPE of 8.9%, and R^2 of 0.89. Relative to the baseline model, RMSE and MSE decreased by 17.85% and 32.48%, respectively. Partial dependence plots and SHAP analysis identified water and cement as the dominant predictors. Higher water content generally reduced predicted tensile strength, whereas higher cement content increased it. The findings demonstrate that optimized and explainable GPR provides a reliable data-driven framework for predicting the tensile performance of SCM-blended concrete and supporting performance-informed design of more sustainable cementitious materials.

Keywords: Splitting tensile strength; SCM-blended concrete; Machine learning; Bayesian optimization; Partial dependence plots; Explainable artificial intelligence.

1. Introduction

Concrete is widely used because of its availability, versatility, and structural performance. Its environmental burden, however, is strongly linked to Portland cement production. This has intensified efforts to reduce clinker consumption and develop lower-carbon cementitious systems. Supplementary cementitious materials (SCMs), such as fly ash (FA), silica fume (SF), and ground granulated blast-furnace slag (GGBFS), are established options for partially replacing Portland cement [1–4]. Their effects depend on chemical composition, fineness, replacement level, and curing conditions. SCMs can participate in pozzolanic or latent hydraulic reactions, refine pore structure, and alter hydration. These mechanisms can improve mechanical and durability performance. When suitable mixture proportions are used, SCMs can also reduce the environmental impact associated with cement consumption [1–4].

The mechanical response of SCM-blended concrete depends on interactions among cement, water, aggregates, chemical admixtures, SCM type, and SCM dosage. These interactions are particularly important for tensile behavior. Splitting tensile strength (STS) is a key property because concrete is weak in tension and tensile stresses contribute directly to crack initiation and propagation. Compared with compressive strength, STS has received less attention in both experimental and data-driven studies [5–9]. It is also more difficult to estimate reliably because concrete is heterogeneous and sensitive to mixture composition and internal microstructure. Water, cement, aggregates, admixtures, curing conditions, and SCM incorporation can therefore affect tensile resistance through interacting and nonlinear mechanisms.

Experimental testing remains essential for characterizing concrete properties. Large laboratory programs, however, require substantial material, time, and testing resources. Empirical equations offer convenient estimates, but their applicability can be limited for mixtures containing several SCMs over wide dosage ranges. Machine-learning (ML) methods provide an alternative by learning relationships directly from experimental data without imposing a predefined analytical form. Previous studies have used artificial neural networks, decision trees, support vector machines, ensemble learners, Gaussian process regression, and other data-driven methods to predict concrete properties [10–19]. These approaches have been applied to fly-ash concrete, high-performance concrete containing FA and SF, and high-strength self-compacting concrete with multiple SCMs [10–12,17–21]. Their results show that ML can capture nonlinear relationships between mixture composition and strength. Recent reviews also document the growing use of ML in mixture design, mechanical-property prediction, durability assessment, and optimization [13,15].

Most ML studies on cementitious materials still focus on compressive strength. Splitting tensile strength has received considerably less attention. Li et al. [7], for example, used ML to predict the STS of metakaolin concrete. Other studies have examined tensile strength in high-performance, recycled-aggregate, fiber-reinforced, and SCM-containing concretes [6,8,9,14,19]. These studies demonstrate the potential of ML for tensile-property prediction, but many rely on a single concrete type or a narrowly defined experimental program. Their models may therefore not capture the variability found across different regions, material sources, and published studies. Liu et al. [5] recently compiled a global database containing 1,456 SCM-blended concrete mixtures from 136 publications in 27 countries. The database includes FA-, SF-, and GGBFS-containing concretes and reports several mechanical and durability properties. It provides a broader basis for investigating STS prediction across diverse mixture designs.

Gaussian process regression (GPR) is well suited to this problem because it provides a flexible, nonparametric probabilistic framework for nonlinear relationships [22]. GPR has shown strong performance in predicting the compressive strength of high-performance concrete and has been extended to other concrete-property prediction tasks [23]. Its performance, however, can be sensitive to the selected kernel and associated hyperparameters [24,25]. Different kernel structures can change the smoothness and complexity of the fitted response surface. Systematic hyperparameter optimization is therefore important when GPR is applied to heterogeneous concrete datasets.

High predictive accuracy alone does not provide sufficient engineering insight. Many advanced ML models operate as black boxes, making it difficult to determine how individual mixture constituents affect a prediction. This limitation has encouraged the use of explainable artificial intelligence (XAI) in concrete research. SHapley Additive exPlanations (SHAP), introduced by Lundberg and Lee [26], assign feature contributions to individual predictions and can also summarize global predictor influence. SHAP has been applied to concrete-strength prediction to quantify both feature importance and the direction of feature effects [27–32]. These studies show that SHAP can improve model transparency while revealing relationships that are consistent with established concrete behavior.

Partial dependence plots (PDPs) provide a complementary form of model interpretation. They show the average marginal response of a model as selected predictors vary [33–35]. Used together, PDPs and SHAP can separate overall response trends from observation-specific feature contributions. This distinction is useful in concrete applications, where a model should be both accurate and physically interpretable. Previous studies have used SHAP, PDPs, and related XAI methods to interpret predictions of concrete strength, durability, and time-dependent behavior [27–35].

This study develops an explainable ML framework for predicting the splitting tensile strength of SCM-blended concrete using the global database reported by Liu et al. [5]. The analysis includes 395 mixtures with available STS measurements. Eight mixture variables are used as predictors: cement, water, coarse aggregate, fine aggregate, fly ash, silica fume, GGBFS, and superplasticizer. The study compares a broad range of regression models under a consistent cross-validation framework. These include linear regression, regression trees, support vector machines, ensemble models, GPR, neural networks, and kernel-based methods. Predictor relevance is also evaluated using mRMR, F-test, RReliefF, and Random Forest importance.

The best-performing model family is then optimized using Bayesian optimization, random search, and grid search. Predictive performance and computational efficiency are compared across the three strategies. The final model is interpreted using PDP and SHAP analyses. These methods are used to identify the dominant mixture constituents, determine the direction of their effects on predicted STS, and examine nonlinear behavior and interactions. The overall aim is to develop an accurate and interpretable framework for predicting the tensile performance of sustainable SCM-blended concrete.

2. Research Objectives

The primary objective is to develop an accurate and interpretable machine-learning framework for predicting the splitting tensile strength of SCM-blended concrete from mixture-composition variables. The specific objectives are to:

1. **Characterize the dataset** through descriptive statistical analysis and correlation assessment to understand the distribution and relationships among the input variables and splitting tensile strength.
2. **Evaluate and rank predictor importance** using multiple feature-selection techniques, including mRMR, F-test, RReliefF, and Random Forest importance.
3. **Develop and compare a broad range of machine-learning regression models** using a consistent cross-validation framework and multiple performance metrics, including RMSE, MSE, R^2, MAE, and MAPE.
4. **Identify and optimize the best-performing model** using Bayesian optimization, random search, and grid search, and compare the resulting predictive accuracy and computational efficiency.
5. **Quantify the improvement achieved through hyperparameter optimization** relative to the corresponding baseline model in terms of prediction error, goodness of fit, training time, and prediction speed.
6. **Interpret the final optimized model using partial dependence plots (PDPs) and SHAP. The analysis identifies the dominant predictors and explains their effects on predicted splitting tensile strength from an engineering perspective.**

3. Methodology

Figure 1 summarizes the methodology used in this study. The workflow was designed to develop, optimize, and interpret ML regression models for predicting the splitting tensile strength (STS) of SCM-blended concrete. It comprised dataset acquisition, data preparation, exploratory analysis, feature ranking, model development, baseline model selection, hyperparameter optimization, model interpretation, and final engineering assessment.

The dataset was obtained from the publicly available Mendeley Data repository developed by Liu, Mo, and Lee (2025), entitled A global dataset of SCM-blended concrete and its properties. The original database contains 1,456 SCM-blended concrete mixtures compiled from 136 publications across 27 countries. It includes mixture parameters and several mechanical and durability properties. For this study, records with valid splitting tensile strength measurements were extracted, yielding 395 observations. Eight mixture variables were used as predictors: cement, water, coarse aggregate, fine aggregate, fly ash (FA), silica fume (SF), ground granulated blast-furnace slag (GGBFS), and superplasticizer (SP). Splitting tensile strength, expressed in MPa, was the regression target. Table 1 summarizes the variables, symbols, roles, units, and descriptions [36].

The data were imported into MATLAB and checked for missing values, consistency, and numerical formatting. Predictors and the response variable were then organized for regression analysis. Exploratory data analysis was used to characterize the dataset. It included descriptive statistics, correlation analysis, and examination of pairwise relationships, as shown in Figure 1. Four feature-ranking methods were used to assess predictor relevance: minimum redundancy maximum relevance (mRMR), F-test, RReliefF, and Random Forest importance. The methods provide complementary views based on information relevance, statistical association, neighborhood sensitivity, and model-based importance. Their combined rankings were used to obtain a broader assessment of predictor relevance before model interpretation. Regression modeling was conducted in MATLAB Regression Learner. The evaluated algorithms included linear regression, regression trees, support vector machines (SVM), ensemble methods, Gaussian process regression

(GPR), neural networks, and kernel-based models. All models were assessed using 5-fold cross-validation to provide a consistent basis for comparison. Performance was measured using root mean squared error (RMSE), mean squared error (MSE), coefficient of determination (R^2), mean absolute error (MAE), and mean absolute percentage error (MAPE).

After the baseline comparison, the best-performing model family was selected for hyperparameter optimization. Three search strategies were examined: Bayesian optimization, random search, and grid search. The optimized parameters were the basis function, kernel function, kernel scale, noise parameter σ, and predictor standardization. The same general search domain was used for all three strategies to support a consistent comparison of predictive performance and computational requirements. The optimized models were evaluated using the same validation framework and performance metrics as the baseline models. Training time and prediction speed were also considered when selecting the final model. Model diagnostics included observed-versus-predicted plots and residual plots. These plots were used to assess prediction agreement and residual behavior. Explainable artificial intelligence methods were then applied to interpret the selected model. Partial dependence plots (PDPs) were used to examine the average marginal effect of each predictor on predicted splitting tensile strength. SHapley Additive exPlanations (SHAP) quantified global feature importance and individual feature contributions. SHAP dependence plots were also used to examine nonlinear effects and potential interactions among the most influential variables.

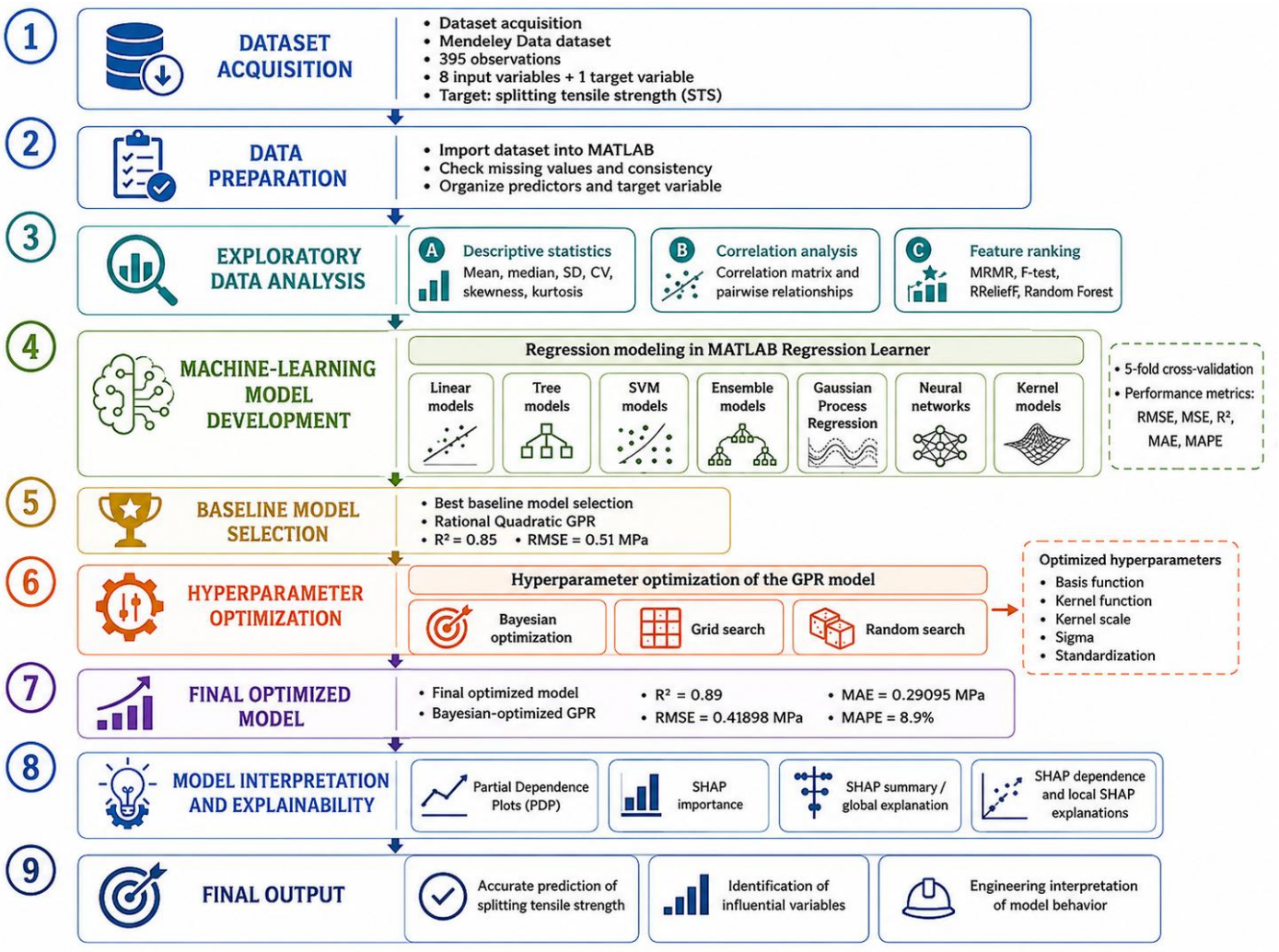


Figure 1. Methodology Framework

Table 1. Description of the dataset variables used for machine-learning prediction

No.	Variable	Symbol	Role	Unit	Description
1	Cement content	C	Input	kg/m³	Mass of cement used per unit volume of the concrete mixture.
2	Water content	W	Input	kg/m³	Mass of mixing water used per unit volume of concrete.

3	Coarse aggregate content	CA	Input	kg/m³	Mass of coarse aggregate incorporated in the concrete mixture.
4	Fine aggregate content	Fine Agg.	Input	kg/m³	Mass of fine aggregate used per unit volume of concrete.
5	Fly ash content	FA	Input	kg/m³	Amount of fly ash used as a supplementary cementitious material in the mixture.
6	Silica fume content	SF	Input	kg/m³	Amount of silica fume incorporated as a supplementary cementitious material.
7	Ground granulated blast-furnace slag content	GGBFS	Input	kg/m³	Amount of ground granulated blast-furnace slag used as a supplementary cementitious material.
8	Superplasticizer content	SP	Input	kg/m³	Dosage of superplasticizer added to improve workability and mixture performance.
9	Splitting tensile strength	STS	**Output**	MPa	Measured splitting tensile strength of concrete used as the target variable for machine-learning regression.

Table 2 summarizes the descriptive statistics of the eight predictors and the splitting tensile strength target. The mixture compositions cover a wide range. Cement content varied from 19.00 to 712.50 kg/m³, and water content ranged from 88.20 to 350.10 kg/m³. Coarse and fine aggregate contents also showed broad variation. The SCM-related variables had much larger coefficients of variation. These were 142.42% for silica fume, 146.87% for GGBFS, 109.81% for fly ash, and 109.54% for superplasticizer. The high variability reflects the presence of zero values in many mixtures and relatively high dosages in others. FA, SF, GGBFS, and SP were positively skewed, indicating a concentration of observations at lower dosage levels. Coarse and fine aggregates were negatively skewed. Splitting tensile strength ranged from 0.96 to 8.26 MPa, with a mean of 3.87 MPa and a median of 3.70 MPa. Its standard deviation was 1.29 MPa, and its coefficient of variation was 33.39%. The moderate positive skewness of 0.86 indicates adequate variability for regression modeling.

Table 2. Descriptive statistical summary of the input variables and target variable used in the study

Variable	Minimum	Q1	Median	Mean	Q3	Maximum	Range	IQR	Std. Dev.	Variance	Std. Error	CV (%)	Skewness	Kurtosis
Cement (kg/m³)	19.00	227.50	300.00	307.66	384.93	712.50	693.50	157.43	123.11	15154.90	6.19	40.01	0.58	0.49
Water (kg/m³)	88.20	155.00	171.00	175.71	197.00	350.10	261.90	42.00	33.22	1103.43	1.67	18.90	1.06	3.30
Coarse aggregate (kg/m³)	115.00	785.00	1069.00	994.46	1139.10	1587.00	1472.00	354.10	222.98	49718.68	11.22	22.42	−0.44	0.13
Fine aggregate (kg/m³)	0.00	647.00	731.20	731.52	837.20	993.00	993.00	190.20	159.26	25362.67	8.01	21.77	−1.06	3.40
Fly ash (FA) (kg/m³)	0.00	0.00	71.25	75.84	126.90	390.00	390.00	126.90	83.28	6935.53	4.19	109.81	1.14	1.13
Silica fume (SF) (kg/m³)	0.00	0.00	0.00	22.64	38.00	187.50	187.50	38.00	32.24	1039.65	1.62	142.42	1.76	3.35
GGBFS (kg/m³)	0.00	0.00	0.00	61.67	122.50	370.00	370.00	122.50	90.57	8203.52	4.56	146.87	1.27	0.50
Superplasticizer (SP) (kg/m³)	0.00	0.00	3.50	4.89	8.23	22.02	22.02	8.23	5.36	28.69	0.27	109.54	1.26	0.85
Splitting tensile strength (MPa)	**0.96**	**2.97**	**3.70**	**3.87**	**4.48**	**8.26**	**7.30**	**1.51**	**1.29**	**1.67**	**0.07**	**33.39**	**0.86**	**0.92**

Figure 2 shows pairwise correlations between the eight predictors and splitting tensile strength (STS). Cement had the strongest positive correlation with STS (r = 0.50), followed by silica fume (r = 0.43) and superplasticizer (r = 0.41). Water and fine aggregate were negatively correlated with STS (r = −0.33 and −0.28), while coarse aggregate showed a weak positive correlation (r = 0.23). The strongest predictor correlations were coarse–fine aggregate (r = −0.69) and water–coarse aggregate (r = −0.58). Most other correlations were weak to moderate. The scatter plots also showed nonlinear patterns and substantial dispersion.

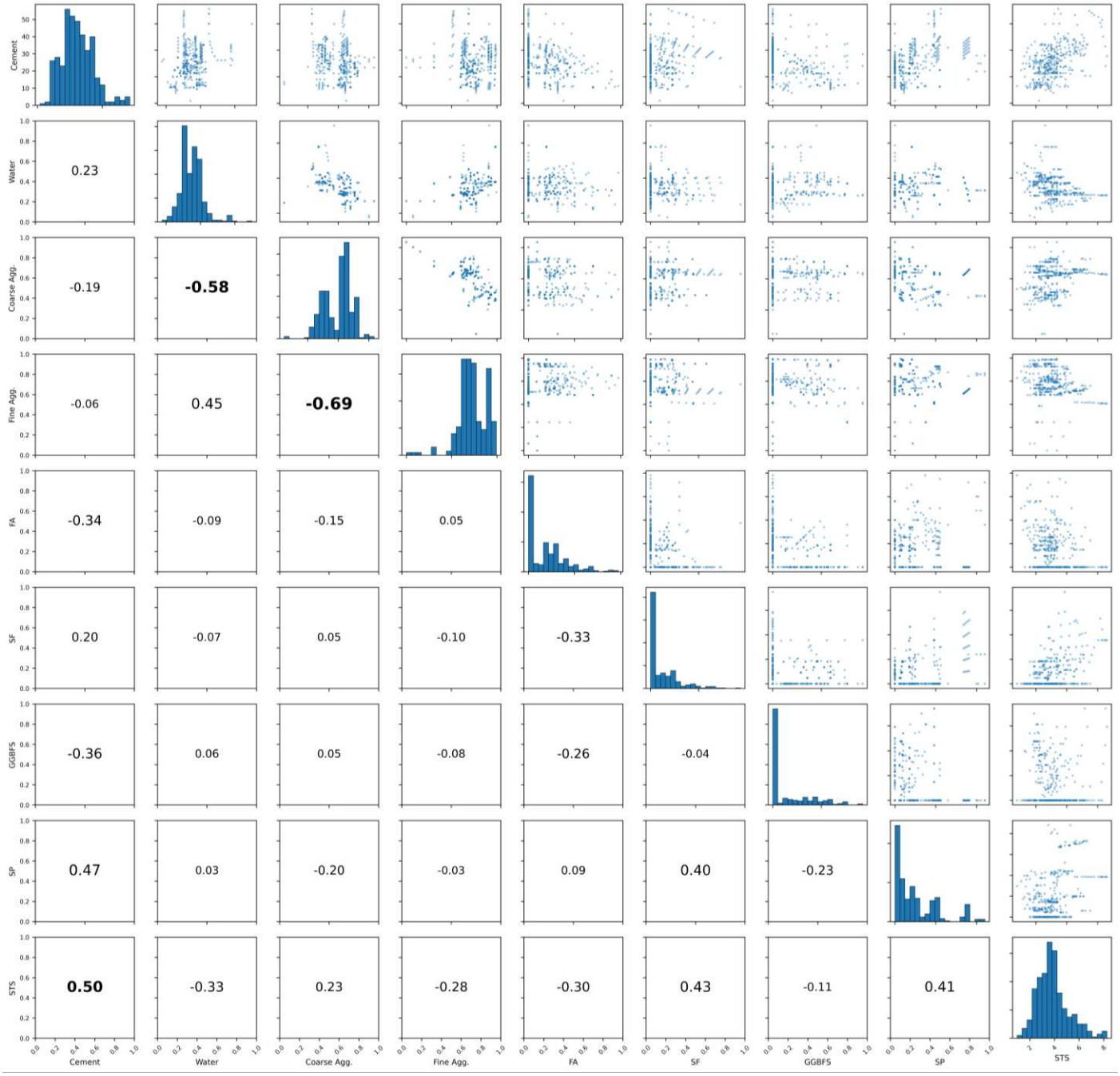


Figure 2. Correlation matrix and pairwise relationships among input variables and the target variable

Figure 3 and Table 3 show that predictor rankings differed across the four feature-selection methods. This variation reflects the different principles used by each method. mRMR ranked SP first, followed by GGBFS and SF. The F-test ranked cement first, followed by water and SF. RReliefF again placed SP first, with water and fine aggregate next. Random Forest importance ranked water first, followed by SP and FA. To obtain an overall assessment, the individual ranks were averaged. SP had the best mean rank (2.50), followed by water (2.75) and cement (3.75). Coarse and fine aggregates occupied intermediate positions. SF, FA, and GGBFS ranked lower overall. The differences among methods show that predictor importance depends on the ranking criterion. Combining several methods therefore provides a more balanced basis for subsequent interpretation.

Table 3. Feature ranking based on different feature selection methods

Variable	mRMR Rank	F-test Rank	RReliefF Rank	Random Forest Rank	Mean Rank	Overall Rank

SP	1	6	1	2	**2.5**	**1**
Water	6	2	2	1	**2.75**	**2**
Cement	4	1	5	5	**3.75**	**3**
Coarse aggregate	5	4	4	6	**4.75**	**4**
Fine aggregate	8	5	3	4	**5**	**5**
Silica fume (SF)	3	3	8	8	**5.5**	**6**
Fly ash (FA)	7	7	6	3	**5.75**	**7**
GGBFS	2	8	7	7	**6**	**8**

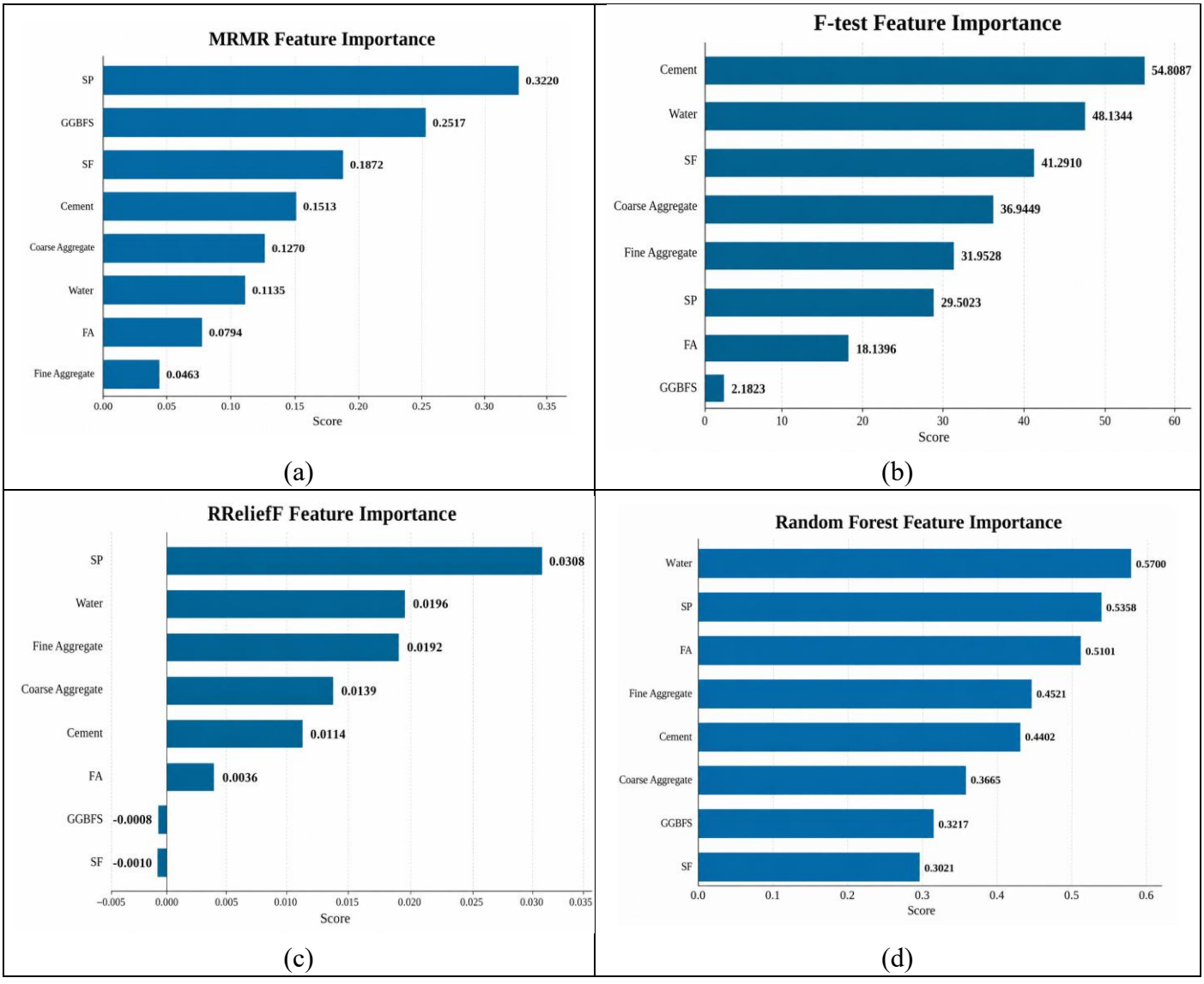


Figure 3. Feature importance analysis of input variables using (a) mRMR, (b) F-test, (c) RReliefF and (d) Random Forest methods.

Table 4 lists the regression models and hyperparameter settings evaluated in this study. The model set was selected to represent both linear and nonlinear relationships between mixture composition and splitting tensile strength. It included conventional and robust linear regression, regression trees, linear and nonlinear SVMs, efficient linear learners, ensemble trees, four GPR kernels, neural networks, and kernel-based regression models. Most presets followed the standard configurations available in MATLAB Regression Learner. Model-specific settings included kernel function,

kernel scale, minimum leaf size, number of learners, learning rate, network architecture, regularization, and data standardization. This broad model set provided a consistent basis for selecting the strongest baseline model before hyperparameter optimization.

Table 4. Summary of machine learning models, presets, and corresponding hyperparameters used in the study.

Model Type	Preset	Hyperparameters
Linear Regression	Linear	Terms: Linear; Robust option: Off
	Interactions Linear	Terms: Interactions; Robust option: Off
	Robust Linear	Terms: Linear; Robust option: On
	Stepwise Linear	Initial terms: Linear; Upper bound on terms: Interactions; Maximum number of steps: 1000
Tree	Fine Tree	Minimum leaf size: 4; Surrogate decision splits: Off
	Medium Tree	Minimum leaf size: 12; Surrogate decision splits: Off
	Coarse Tree	Minimum leaf size: 36; Surrogate decision splits: Off
SVM	Linear SVM	Kernel function: Linear; Kernel scale: Automatic; Box constraint: Automatic; Epsilon: Auto; Standardize data: Yes
	Quadratic SVM	Kernel function: Quadratic; Kernel scale: Automatic; Box constraint: Automatic; Epsilon: Auto; Standardize data: Yes
	Cubic SVM	Kernel function: Cubic; Kernel scale: Automatic; Box constraint: Automatic; Epsilon: Auto; Standardize data: Yes
	Fine Gaussian SVM	Kernel function: Gaussian; Kernel scale: 0.79; Box constraint: Automatic; Epsilon: Auto; Standardize data: Yes
	Medium Gaussian SVM	Kernel function: Gaussian; Kernel scale: 3.2; Box constraint: Automatic; Epsilon: Auto; Standardize data: Yes
	Coarse Gaussian SVM	Kernel function: Gaussian; Kernel scale: 13; Box constraint: Automatic; Epsilon: Auto; Standardize data: Yes
Efficient Linear	Efficient Linear Least Squares	Learner: Least squares; Solver: Auto; Regularization: Auto; Regularization strength (Lambda): Auto; Relative coefficient tolerance (Beta tolerance): 0.0001;
	Efficient Linear SVM	Learner: SVM; Solver: Auto; Regularization: Auto; Regularization strength (Lambda): Auto; Relative coefficient tolerance (Beta tolerance): 0.0001; Epsilon: Auto
Ensemble	Boosted Trees	Minimum leaf size: 8; Number of learners: 30; Learning rate: 0.1; Number of predictors to sample: Select All
	Bagged Trees	Minimum leaf size: 8; Number of learners: 30; Number of predictors to sample: Select All
Gaussian Process Regression	Squared Exponential GPR	Basis function: Constant; Kernel function: Squared Exponential; Use isotropic kernel: Yes; Kernel scale: Automatic; Signal standard deviation: Automatic; Sigma: Automatic; Standardize data: Yes; Optimize numeric parameters: Yes
	Matérn 5/2 GPR	Basis function: Constant; Kernel function: Matérn 5/2; Use isotropic kernel: Yes; Kernel scale: Automatic; Signal standard deviation: Automatic; Sigma: Automatic; Standardize data: Yes; Optimize numeric parameters: Yes
	Exponential GPR	Basis function: Constant; Kernel function: Exponential; Use isotropic kernel: Yes; Kernel scale: Automatic; Signal standard deviation: Automatic; Sigma: Automatic; Standardize data: Yes; Optimize numeric parameters: Yes
	Rational Quadratic GPR	Basis function: Constant; Kernel function: Rational Quadratic; Use isotropic kernel: Yes; Kernel scale: Automatic; Signal standard deviation: Automatic; Sigma: Automatic; Standardize data: Yes; Optimize numeric parameters: Yes
Neural Network	Narrow Neural Network	Number of fully connected layers: 1; First layer size: 10; Activation: ReLU; Iteration limit: 1000; Regularization strength (Lambda): 0; Standardize data: Yes
	Medium Neural Network	Number of fully connected layers: 1; First layer size: 25; Activation: ReLU; Iteration limit: 1000; Regularization strength (Lambda): 0; Standardize data: Yes
	Wide Neural Network	Number of fully connected layers: 1; First layer size: 100; Activation: ReLU; Iteration limit: 1000; Regularization strength (Lambda): 0; Standardize data: Yes
	Bilayered Neural Network	Number of fully connected layers: 2; First layer size: 10; Second layer size: 10; Activation: ReLU; Iteration limit: 1000; Regularization strength (Lambda): 0; Standardize data: Yes

	Trilayered Neural Network	Number of fully connected layers: 3; First layer size: 10; Second layer size: 10; Third layer size: 10; Activation: ReLU; Iteration limit: 1000; Regularization strength (Lambda): 0; Standardize data: Yes
Kernel	SVM Kernel	Learner: SVM; Number of expansion dimensions: Auto; Regularization strength (Lambda): Auto; Kernel scale: Auto; Epsilon: Auto; Standardize data: Yes; Iteration limit: 1000
	Least Squares Regression Kernel	Learner: Least Squares Kernel; Number of expansion dimensions: Auto; Regularization strength (Lambda): Auto; Kernel scale: Auto; Standardize data: Yes; Iteration limit: 1000

4. Results and Discussion

Table 5 compares all regression models using MAPE, RMSE, MSE, R^2, MAE, prediction speed, and training time. Performance varied substantially across model families. The conventional linear models achieved moderate accuracy, with R^2 values from 0.56 to 0.67. Tree-based models showed similar or lower performance, depending on tree complexity. Among the SVM variants, Medium Gaussian SVM performed best. It achieved an RMSE of 0.68 MPa, R^2 of 0.73, MAE of 0.45 MPa, and MAPE of 13.87%. The ensemble models also performed well. Boosted Trees achieved an RMSE of 0.66 MPa and R^2 of 0.74.

The GPR models produced the strongest baseline results. Squared Exponential GPR achieved an RMSE of 0.55 MPa and R^2 of 0.82. The Matérn 5/2, Exponential, and Rational Quadratic kernels further reduced prediction error. Rational Quadratic GPR was the best baseline model overall. It achieved the lowest MAPE (10.02%), an RMSE of 0.51 MPa, an MSE of 0.26 MPa^2, R^2 of 0.85, and an MAE of 0.34 MPa. Exponential GPR produced a similar RMSE and MSE, but its R^2 was slightly lower and its MAPE slightly higher. These results indicate that GPR captured the nonlinear structure of the dataset more effectively than the other tested model families.

The neural-network models showed intermediate performance, with R^2 values between 0.67 and 0.72. The bilayered network had the lowest RMSE in this family (0.69 MPa). The wide network produced the lowest MAPE (13.70%). Some neural-network configurations offered high prediction speeds, but their accuracy remained below that of the GPR models. Efficient linear and kernel-based models also produced larger errors and lower R^2 values. Their performance was therefore less suitable for the present regression problem.

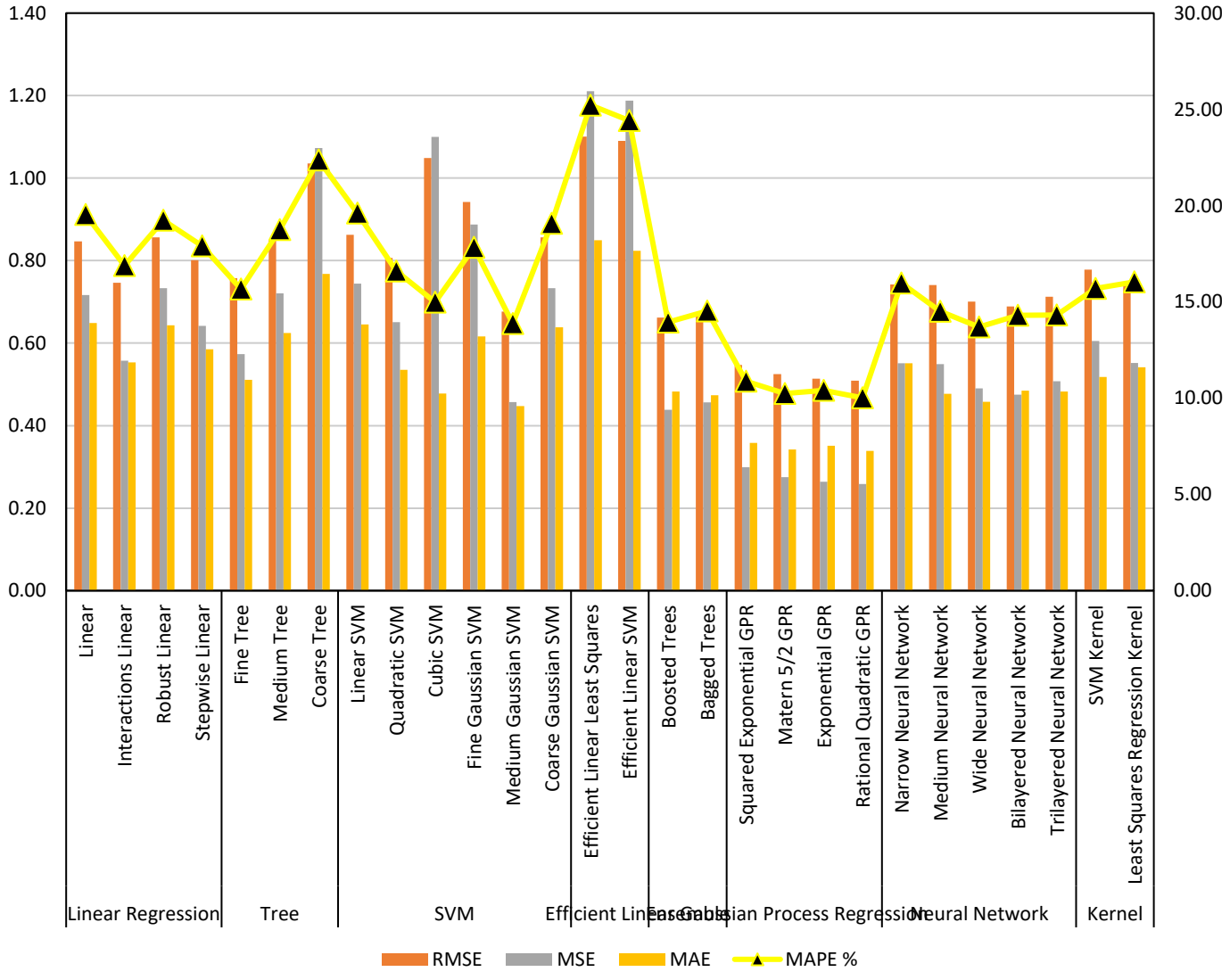
Computational performance also varied across the models. Prediction speed ranged from about 4,527 to 57,738 observations/s, while training time ranged from 4.90 to 27.20 s. The fastest models were not necessarily the most accurate. Model selection therefore required a balance between predictive performance and computational efficiency. Rational Quadratic GPR provided the best overall baseline balance and was selected for hyperparameter optimization. Figure 4 summarizes the comparative results.

Table 5. Performance comparison of machine learning models using validation metrics, including RMSE, MSE, R^2, MAE, prediction speed, and training time.

Model Type	Preset	MAPE %	RMSE	MSE	R^2	MAE	Prediction Speed (obs/sec)	Training time (s)
Linear Regression	Linear	19.54	0.85	0.72	0.57	0.65	4779.18	9.60
	Interactions Linear	16.88	0.75	0.56	0.67	0.55	4526.80	4.90
	Robust Linear	19.24	0.86	0.73	0.56	0.64	28188.71	10.55
	Stepwise Linear	17.91	0.80	0.64	0.62	0.59	19071.16	21.52
Tree	Fine Tree	15.67	0.76	0.57	0.66	0.51	16080.31	9.00
	Medium Tree	18.75	0.85	0.72	0.57	0.62	34096.11	8.29
	Coarse Tree	22.37	1.04	1.07	0.36	0.77	22182.41	7.02
SVM	Linear SVM	19.61	0.86	0.74	0.55	0.65	16993.41	14.10
	Quadratic SVM	16.61	0.81	0.65	0.61	0.54	26329.12	13.82
	Cubic SVM	14.99	1.05	1.10	0.34	0.48	19949.09	13.54
	Fine Gaussian SVM	17.83	0.94	0.89	0.47	0.62	19731.26	13.05
	Medium Gaussian SVM	13.87	0.68	0.46	0.73	0.45	23051.16	12.43

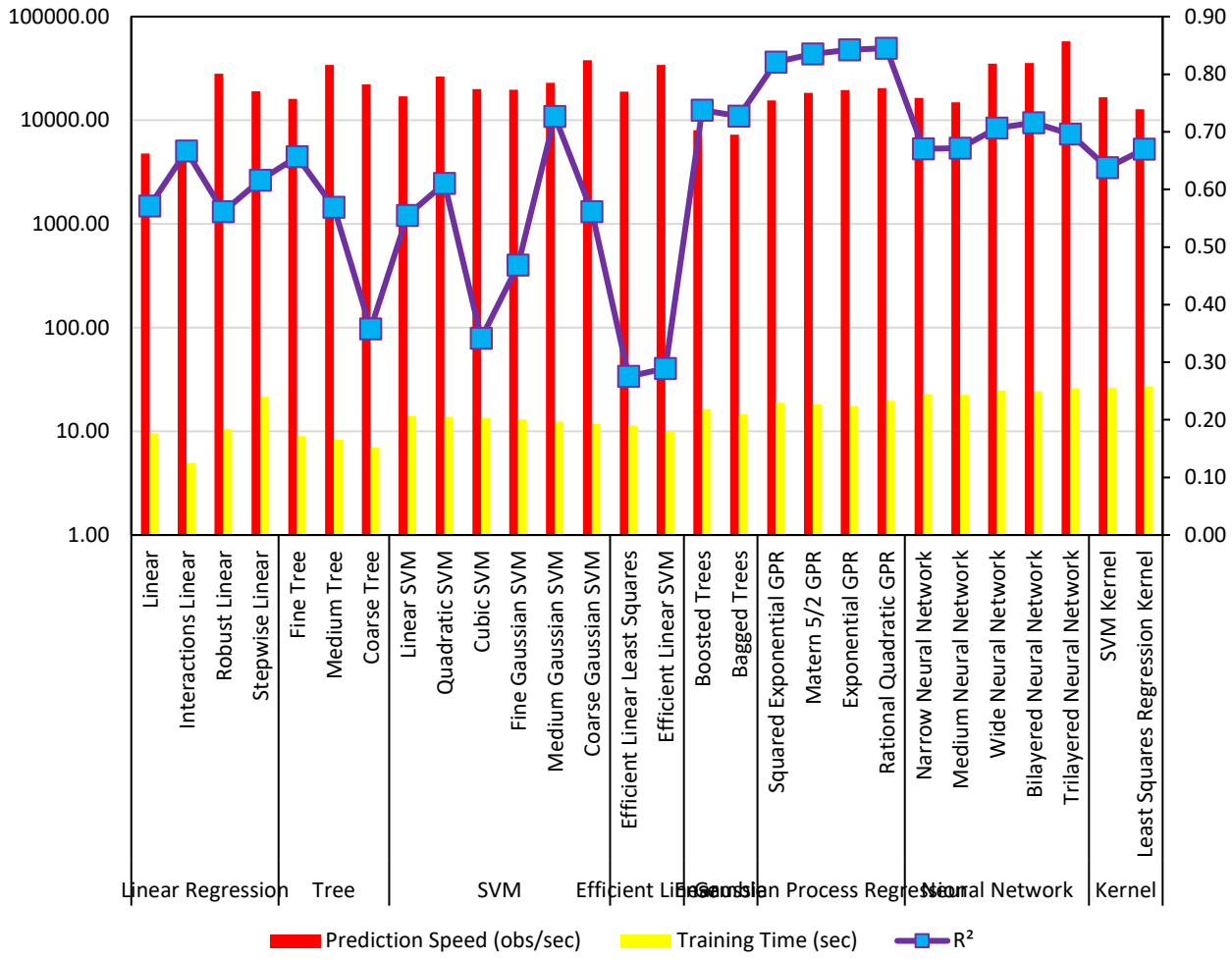
	Coarse Gaussian SVM	19.08	0.86	0.73	0.56	0.64	38074.49	11.85
Efficient Linear	Efficient Linear Least Squares	25.22	1.10	1.21	0.28	0.85	18928.32	11.38
	Efficient Linear SVM	24.42	1.09	1.19	0.29	0.82	34303.38	9.99
Ensemble	Boosted Trees	13.95	0.66	0.44	0.74	0.48	7988.76	16.45
	Bagged Trees	14.53	0.68	0.46	0.73	0.47	7267.12	14.56
Gaussian Process Regression	Squared Exponential GPR	10.88	0.55	0.30	0.82	0.36	15519.41	18.84
	Matérn 5/2 GPR	10.24	0.52	0.28	0.84	0.34	18335.59	18.22
	Exponential GPR	10.40	0.51	0.26	0.84	0.35	19550.29	17.44
	Rational Quadratic GPR	10.02	0.51	0.26	0.85	0.34	20389.31	19.69
Neural Network	Narrow Neural Network	15.98	0.74	0.55	0.67	0.55	16489.39	22.74
	Medium Neural Network	14.51	0.74	0.55	0.67	0.48	14859.62	22.43
	Wide Neural Network	13.70	0.70	0.49	0.71	0.46	35168.63	24.56
	Bilayered Neural Network	14.31	0.69	0.47	0.72	0.48	35743.05	24.33
	Trilayered Neural Network	14.31	0.71	0.51	0.70	0.48	57737.56	26.04
Kernel	SVM Kernel	15.69	0.78	0.61	0.64	0.52	16673.63	26.49
	Least Squares Regression Kernel	16.04	0.74	0.55	0.67	0.54	12754.24	27.20

243



244

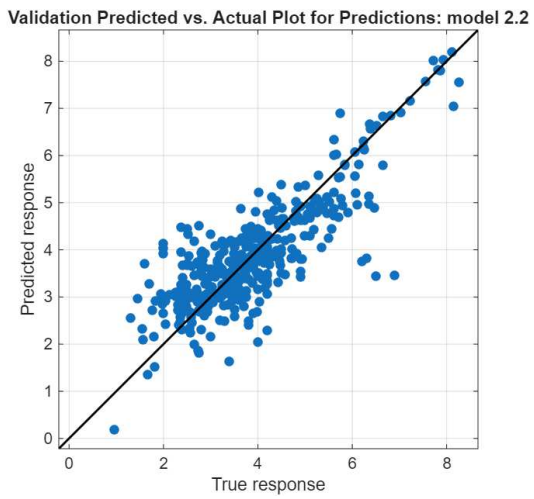
(a)



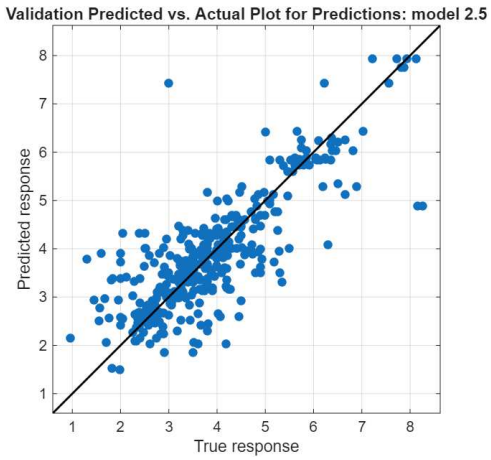
(b)

Figure 4. Machine Learning Results

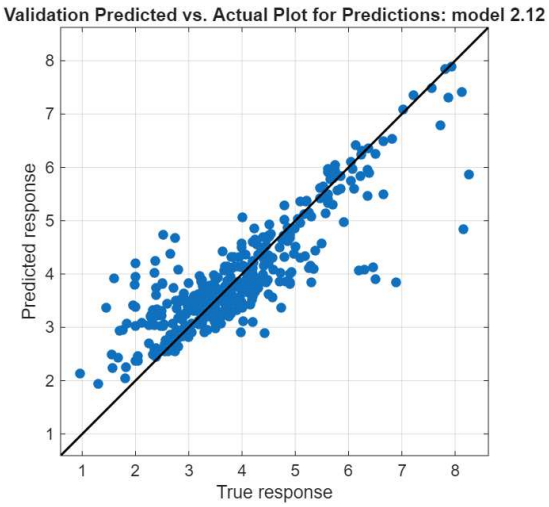
Figure 5 compares validation predictions with observed values for representative models from each major ML family. The 45° line represents perfect agreement. Predictions that cluster near this line indicate better accuracy. Interactions Linear, Fine Tree, Medium Gaussian SVM, Boosted Trees, Bilayered Neural Network, and Least Squares Regression Kernel showed varying levels of scatter. Deviations were more evident at the lower and upper STS ranges for several models. Rational Quadratic GPR showed the closest overall agreement with the 1:1 line and less scatter across most of the response range. This visual pattern is consistent with the quantitative results in Table 5 and supports the selection of GPR for optimization.



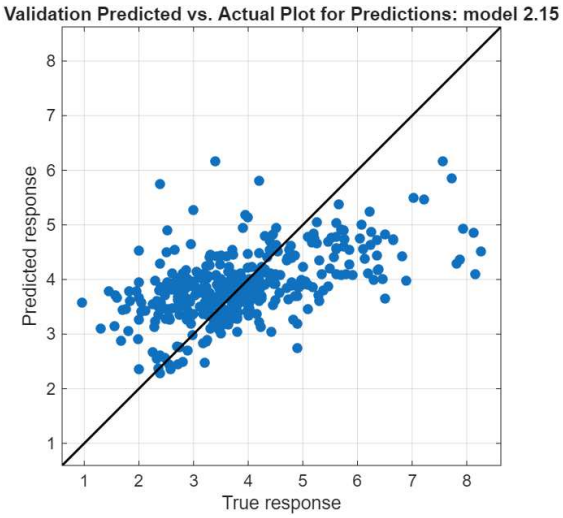
(a) Linear Regression – Interactions Linear



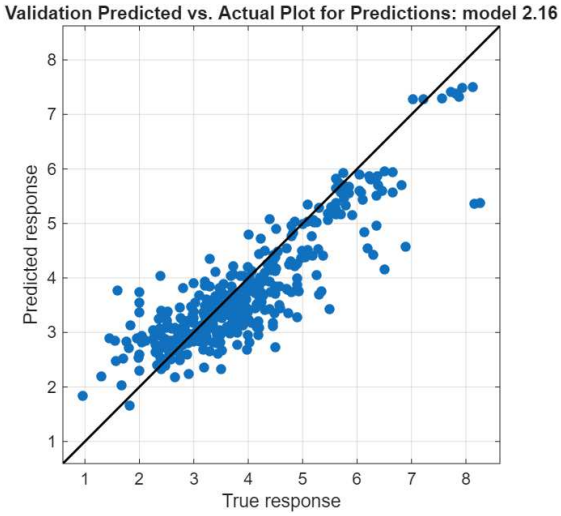
(b) Tree - Fine Tree



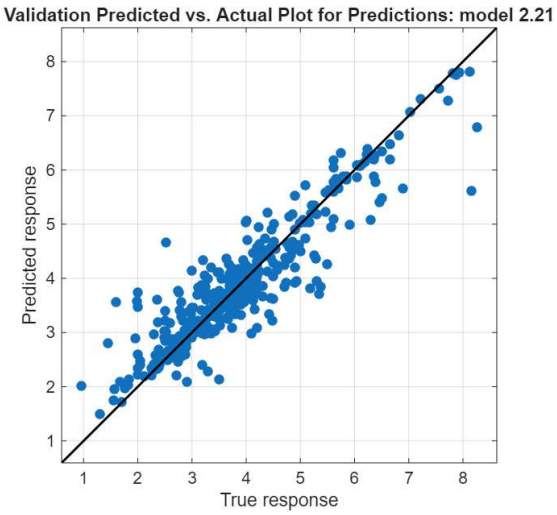
(c) SVM – Medium Gaussian SVM



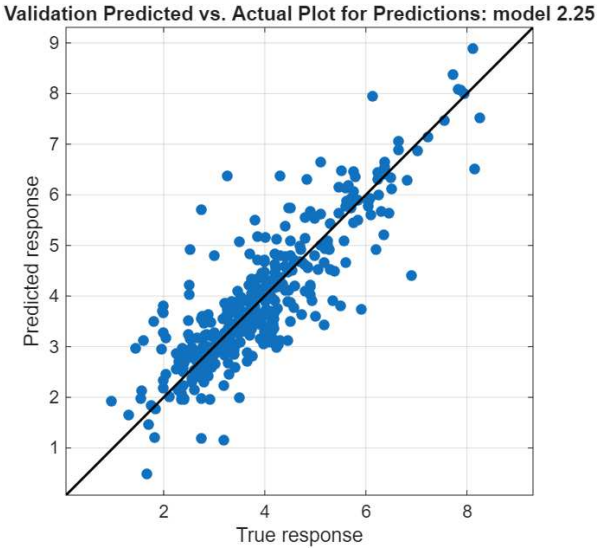
(d) Efficient Linear -Efficient Linear SVM



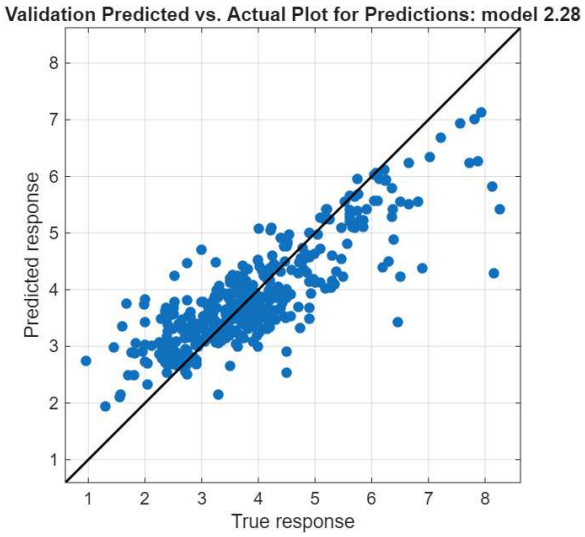
(e) Ensemble - Boosted Trees



(f) GPR -Rational Quadratic GPR



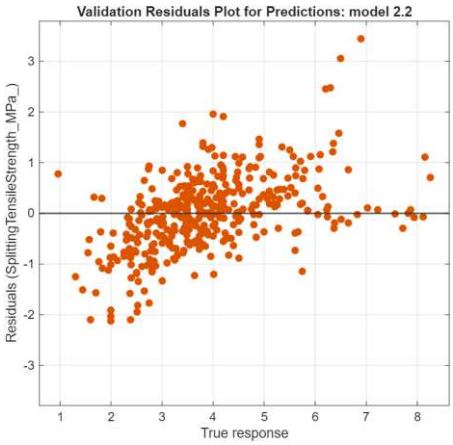
(g) Neural Network -Bilayered Neural Network



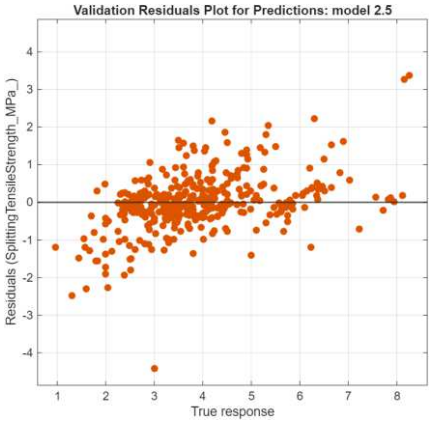
(h) Kernel - Least Squares Regression Kernel

Figure 5. Validation predicted-versus-observed splitting tensile strength for representative machine-learning models: (a) Interactions Linear Regression, (b) Fine Tree, (c) Medium Gaussian SVM, (d) Efficient Linear SVM, (e) Boosted Trees, (f) Rational Quadratic GPR, (g) Bilayered Neural Network, and (h) Least Squares Regression Kernel. The diagonal line represents perfect agreement between observed and predicted values.

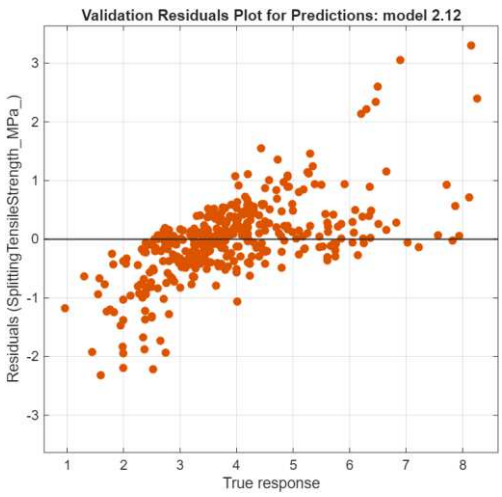
Figure 6 compares residual patterns for representative models across the observed STS range. Ideally, residuals should be randomly distributed around zero with no clear trend or major change in variance. Efficient Linear SVM and Least Squares Regression Kernel showed structured residual patterns and increasing positive residuals at higher response values. This indicates systematic prediction bias. Fine Tree, Medium Gaussian SVM, Boosted Trees, and the neural-network model showed better residual behavior but retained some dispersion and localized trends. Rational Quadratic GPR produced a more compact residual distribution around zero and showed less systematic structure. This finding is consistent with the validation metrics in Table 5 and further supports GPR as the baseline model for optimization.



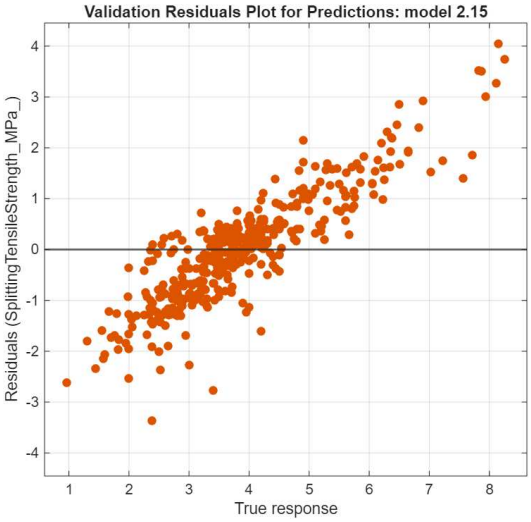
(a) Linear Regression – Interactions Linear



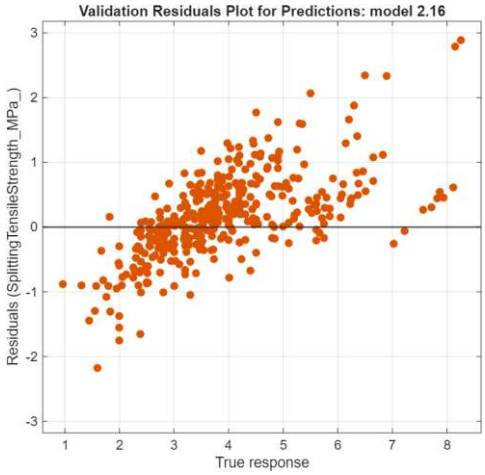
(b) Tree - Fine Tree



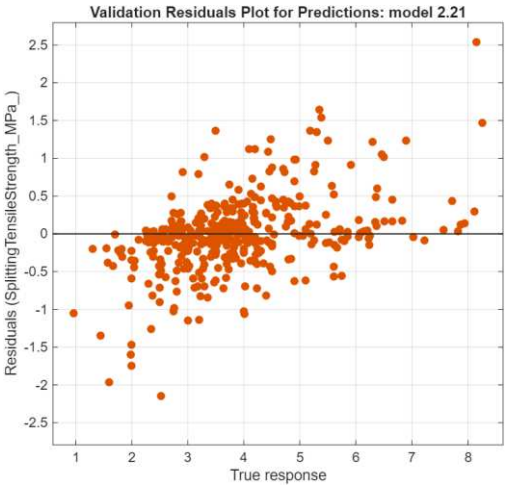
(c) SVM – Medium Gaussian SVM



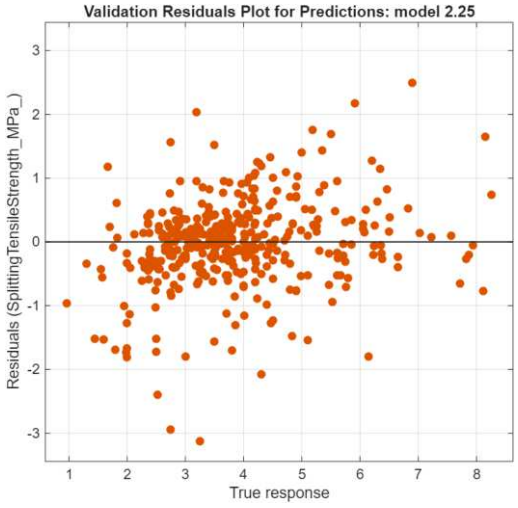
(d) Efficient Linear -Efficient Linear SVM



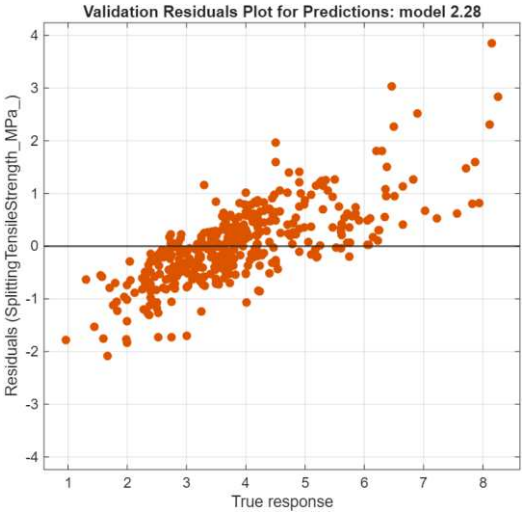
(e) Ensemble - Boosted Trees



(f) GPR -Rational Quadratic GPR



(g) Neural Network -Bilayered Neural Network



(h) Kernel - Least Squares Regression Kernel

Figure 6. Validation residual plots for representative machine-learning models: (a) Interactions Linear Regression, (b) Fine Tree, (c) Medium Gaussian SVM, (d) Efficient Linear SVM, (e) Boosted Trees, (f) Rational Quadratic GPR, (g) Bilayered Neural Network, and (h) Least Squares Regression Kernel.

Table 6 summarizes the GPR hyperparameter search space and optimization settings. The search included both categorical and continuous parameters. Basis function was varied among constant, zero, and linear forms. Ten isotropic and nonisotropic kernel functions were considered. Kernel scale ranged from 0.001 to 1000, while sigma ranged from 0.0001 to 12.9309. Predictor standardization was treated as a Boolean option. Bayesian optimization and random search were each run for 30 iterations. Grid search used 10 grid divisions. For Bayesian optimization, the expected improvement per second plus acquisition function guided the search toward promising configurations while accounting for computational cost.

Table 6. Hyperparameter search space and optimization settings used for GPR model tuning

Category	Hyperparameter / Setting	Search Space / Value	Type / Method
GPR hyperparameters	Basis function	Constant, Zero, Linear	Categorical
	Kernel function	Nonisotropic Exponential; Nonisotropic Matérn 3/2; Nonisotropic Matérn 5/2; Nonisotropic Rational Quadratic; Nonisotropic Squared Exponential; Isotropic Exponential; Isotropic Matérn 3/2; Isotropic Matérn 5/2; Isotropic Rational Quadratic; Isotropic Squared Exponential	Categorical
	Kernel scale	0.001–1000	Continuous
	Sigma	0.0001–12.9309	Continuous
	Standardization	True, False	Boolean
Bayesian optimization	Bayesian iterations	**30**	Optimization setting
	Bayesian acquisition function	**Expected improvement per second plus**	Acquisition strategy
Random search	Random-search iterations	**30**	Optimization setting
Grid search	Grid divisions	**10**	Optimization setting

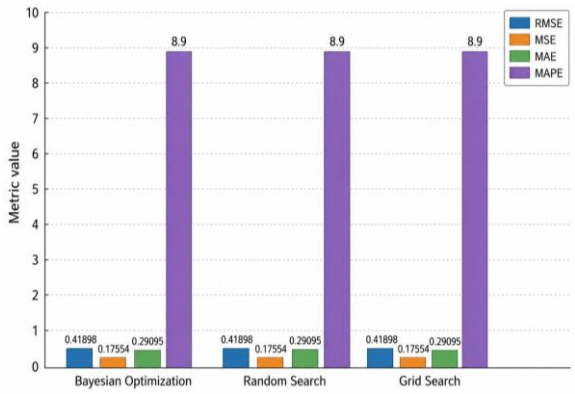
Table 7 compares the optimal GPR configurations obtained by Bayesian optimization, random search, and grid search. All three methods selected a linear basis function and a nonisotropic exponential kernel. They differed, however, in kernel scale, sigma, and standardization. Despite these differences, the three optimized models produced the same reported validation metrics: RMSE = 0.41898 MPa, MSE = 0.17554 MPa^2, MAE = 0.29095 MPa, MAPE = 8.9%, and R^2 = 0.89. Their main difference was computational cost. Bayesian optimization required 89.72 s of training, compared with 187.81 s for random search and 316.86 s for grid search. It was therefore selected as the most computationally efficient solution among the three.

Table 7. Optimized hyperparameters and predictive performance obtained using different optimization methods

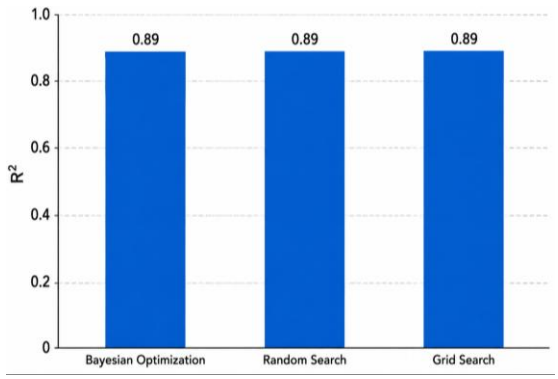
Parameter / Metric	Bayesian Optimization	Random Search	Grid Search
Basis function	Linear	Linear	Linear
Kernel function	Nonisotropic Exponential	Nonisotropic Exponential	Nonisotropic Exponential
Kernel scale	982.1375	1.7799	10
Sigma	12.1345	4.5101	0.0013675

Standardization	No	Yes	No
RMSE (MPa)	**0.41898**	**0.41898**	**0.41898**
MSE (MPa²)	**0.17554**	**0.17554**	**0.17554**
MAE (MPa)	**0.29095**	**0.29095**	**0.29095**
MAPE (%)	**8.9**	**8.9**	**8.9**
R^2	**0.89**	**0.89**	**0.89**
Training time (s)	**89.72**	187.81	316.86

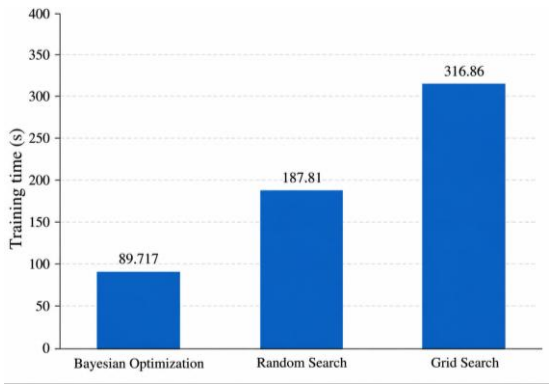
Figure 7 compares predictive performance and computational characteristics across the three GPR optimization strategies. Bayesian optimization, random search, and grid search produced the same reported error metrics and R^2, as shown in Figure 7(a) and (b). Their computational performance differed more clearly. Bayesian optimization had the shortest training time (89.72 s), followed by random search (187.81 s) and grid search (316.86 s). Grid search achieved the highest prediction speed at about 43,000 observations/s. Bayesian optimization followed at about 41,000 observations/s, while random search reached about 30,000 observations/s. Overall, Bayesian optimization offered the best balance between predictive performance and computational cost and was selected for the final GPR model.



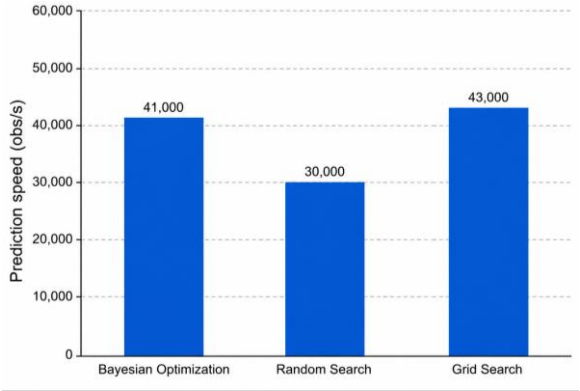
(a)



(b)



(c)



(d)

Figure 7. Comparison of GPR optimization strategies in terms of (a) predictive error metrics, (b) coefficient of determination (R^2), (c) computational training time, and (d) prediction speed.

Figure 8 shows the evolution of minimum mean squared error (MSE) during the three hyperparameter searches. Bayesian optimization remained near an MSE of 0.21 during the early iterations. Around iteration 20, it identified a better solution and reduced the minimum MSE to about 0.176. Random search reached a similar minimum earlier, at about iteration 9, and showed no further improvement. Grid search remained near 0.205 for most of the search and reached the lower MSE only near iteration 26. Thus, the three strategies converged to comparable minimum-error solutions, but they reached their best configurations at different stages of the search.

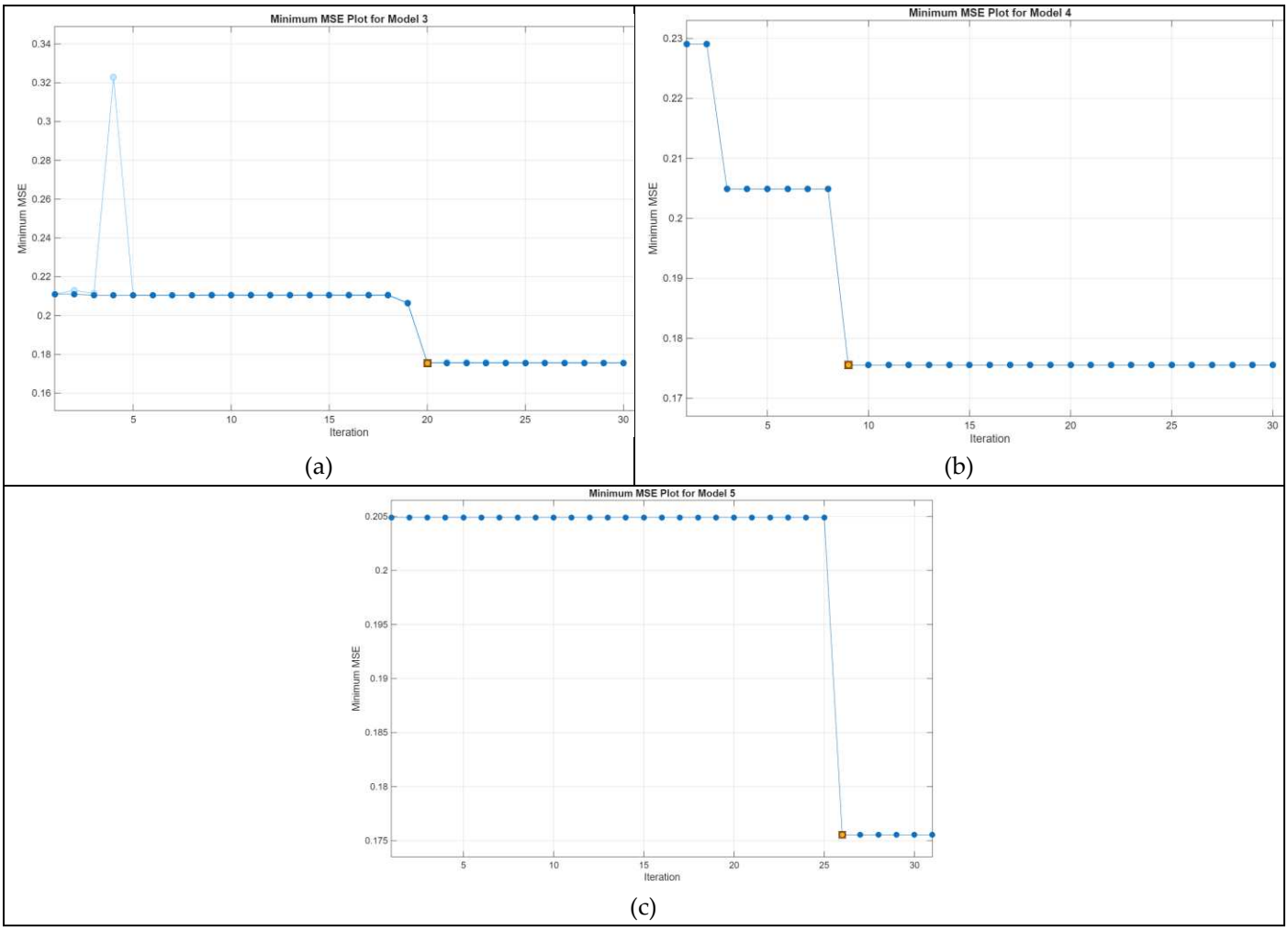


Figure 8. Evolution of the minimum mean squared error (MSE) across optimization iterations for Bayesian optimization, random search, and grid search.

Figure 9 assesses the Bayesian-optimized GPR model using observed-versus-predicted and residual plots. In Figure 9(a), most predictions lie close to the 1:1 reference line. This indicates good agreement between measured and predicted STS across the investigated range. Some dispersion remains at the lower and upper response levels. Figure 9(b) shows residuals distributed around zero without a pronounced systematic pattern. A few larger residuals are present, but most values remain close to zero. The diagnostic plots therefore indicate limited overall bias and support the predictive stability of the optimized model.

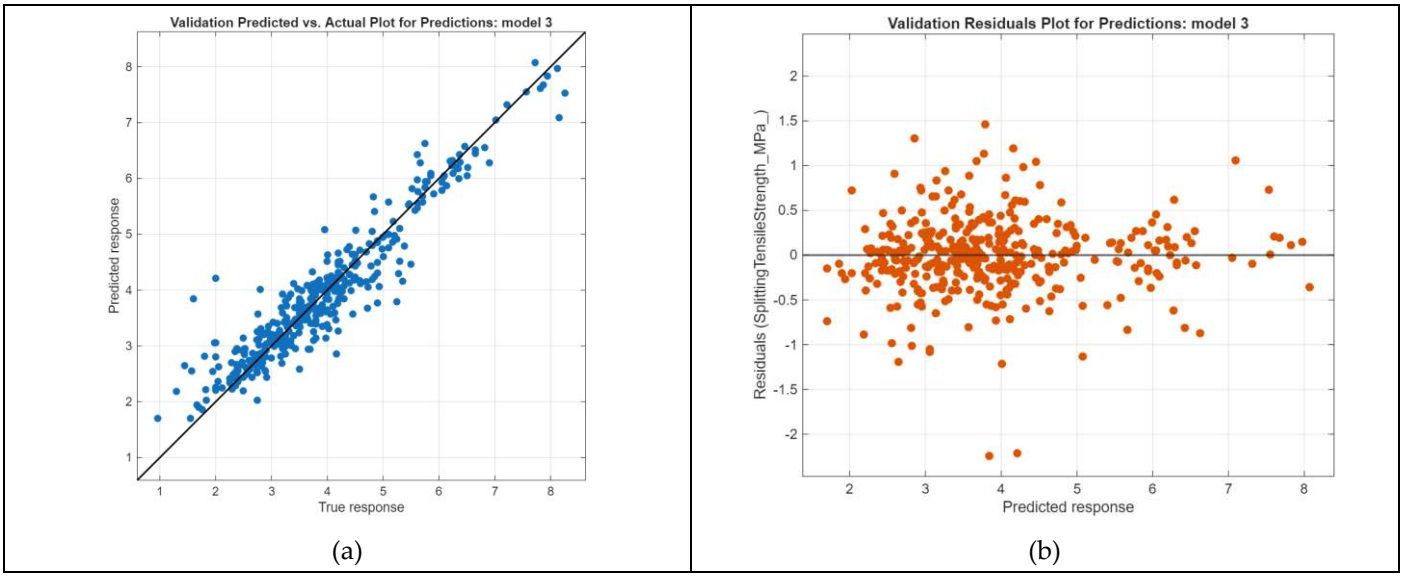


Figure 9. Performance assessment of the best optimized GPR model: (a) observed versus predicted splitting tensile strength values, and (b) residuals versus predicted values.

Table 8 compares the Bayesian-optimized GPR with the baseline Rational Quadratic GPR. RMSE decreased from 0.51 to 0.41898 MPa, a reduction of 17.85%. MSE decreased from 0.26 to 0.17554 MPa², corresponding to a 32.48% reduction. MAE and MAPE decreased by 14.43% and 11.18%, respectively. R^2 increased from 0.85 to 0.89, equivalent to a 4.71% increase. Prediction speed approximately doubled, from 20,389 to about 41,000 observations/s. This gain required more training time, which increased from 19.69 to 89.72 s. Overall, Bayesian optimization improved prediction accuracy and inference speed at the cost of greater training effort.

Table 8. Performance improvement after optimization

Performance Metric	Baseline Rational Quadratic GPR	Bayesian-Optimized GPR	Absolute Change	Relative Change (%)
RMSE (MPa)	0.51	**0.41898**	−0.09102	**17.85% reduction**
MSE (MPa²)	0.26	**0.17554**	−0.08446	**32.48% reduction**
MAE (MPa)	0.34	**0.29095**	−0.04905	**14.43% reduction**
MAPE (%)	10.02	**8.9**	−1.12	**11.18% reduction**
R^2	0.85	**0.89**	0.04	**4.71% increase**
Prediction speed (obs/s)	20,389.31	**41,000**	20,610.69	**101.09% increase**
Training time (s)	19.69	89.72	70.03	**355.65% increase**

Figure 10 presents the partial dependence plots (PDPs) for the Bayesian-optimized GPR model. Cement content shows a clear positive relationship with predicted STS, whereas water content has a pronounced negative effect. Coarse aggregate shows a nonlinear increasing trend with local fluctuations at higher contents. Fine aggregate has a more complex response, with an initial increase, an intermediate peak, a subsequent decline, and partial recovery. Fly ash generally has a negative effect as its content increases. In contrast, silica fume, GGBFS, and superplasticizer show predominantly positive trends over their observed ranges. The PDPs therefore indicate both monotonic and nonlinear relationships between mixture composition and predicted STS.

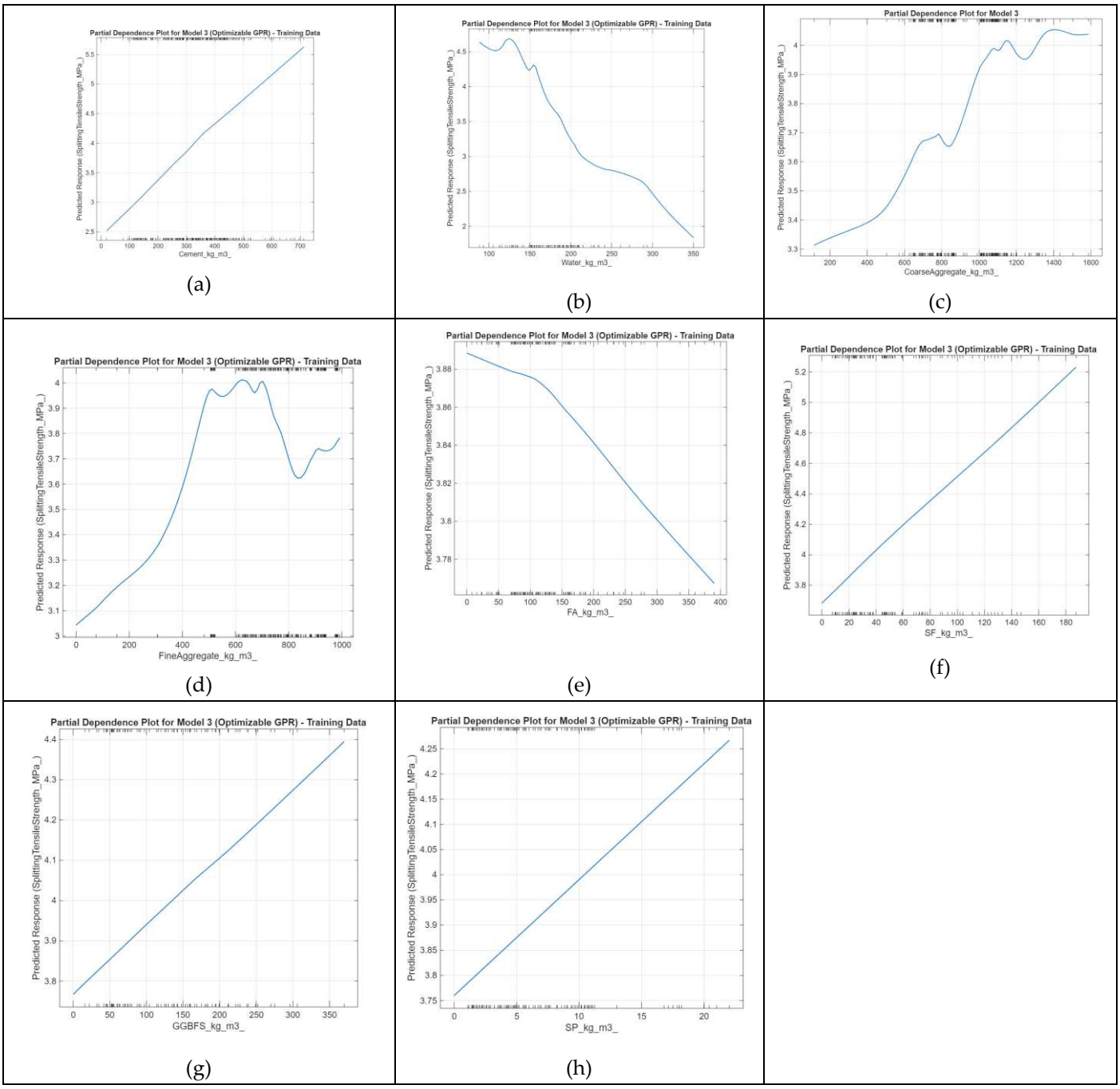


Figure 10. Partial dependence plots (PDPs) of the Bayesian-optimized GPR model showing the effects of (a) cement content, (b) water content, (c) coarse aggregate, (d) fine aggregate, (e) fly ash (FA), (f) silica fume (SF), (g) ground granulated blast-furnace slag (GGBFS), and (h) superplasticizer (SP) on the predicted splitting tensile strength.

Table 9 summarizes the SHAP settings used to interpret the Bayesian-optimized GPR model. The interventional Shapley method was applied to 200 query points. The analysis used 100 background observations and a maximum of 256 predictor subsets across the eight input variables. Splitting tensile strength was the target response. Figure 11 shows that water had the largest mean absolute Shapley value and was therefore the most influential predictor. Cement ranked second. Fine aggregate and silica fume showed intermediate importance, followed closely by coarse

aggregate. GGBFS and superplasticizer made smaller contributions, while fly ash had the lowest global SHAP importance. The model therefore relied most strongly on water and cement when predicting STS.

Table 9. SHAP analysis configuration used for interpretation of the Bayesian-optimized GPR model

SHAP Parameter	Setting
Final model	Bayesian-optimized GPR
Query dataset	Training data
Number of query points	200
Shapley method	Interventional
Number of background observations	100
Maximum predictor subsets	256
Number of predictors	8
Target variable	Splitting tensile strength (MPa)

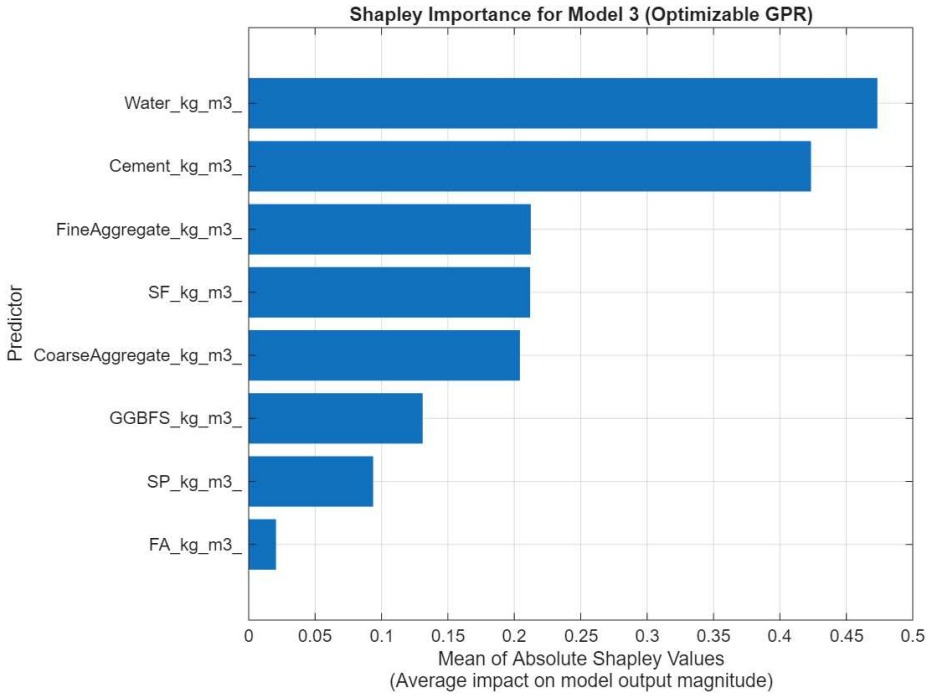


Figure 11. Global SHAP feature importance for the Bayesian-optimized GPR model based on the mean absolute Shapley values of the input variables.

Figure 12 shows the SHAP summary for the Bayesian-optimized GPR model. Water has the widest SHAP-value distribution and the largest overall influence. Higher water contents are mainly associated with negative SHAP values, whereas lower contents tend to increase the predicted STS. Cement shows the opposite pattern. Higher cement contents generally make positive contributions, while lower contents tend to reduce the model output. Higher silica fume contents also show a clear positive effect. GGBFS and superplasticizer have smaller but generally positive contributions as their contents increase. Fine and coarse aggregates show more dispersed and nonlinear effects. Fly ash remains concentrated near zero, indicating limited global influence. Overall, the summary plot confirms water and cement as the dominant predictors and shows the direction of their effects.

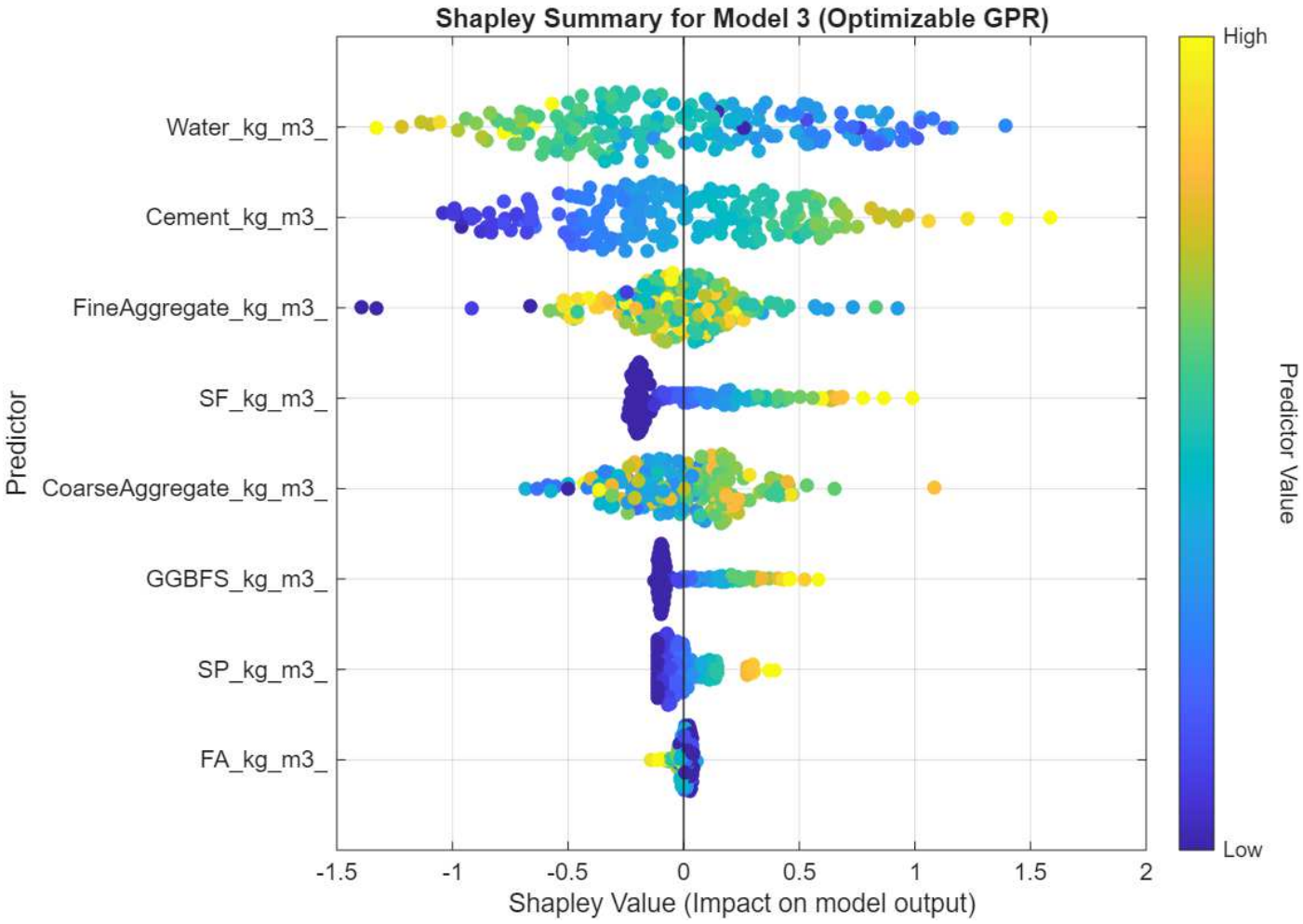


Figure 12. SHAP summary plot

Figure 13 presents SHAP dependence plots for the four most influential predictors. Water in Figure 13(a) shows a clear negative relationship with its SHAP contribution. Lower water contents generally increase predicted STS, whereas higher contents reduce it. Cement in Figure 13(b) shows a strong, nearly monotonic positive relationship. Fine aggregate in Figure 13(c) displays a more dispersed and nonlinear pattern, indicating greater dependence on mixture composition and interaction effects. Silica fume in Figure 13(d) shows an overall positive trend, with higher contents producing more positive SHAP values. The color gradients also indicate interactions with the corresponding secondary predictors. The effect of each principal variable is therefore not fully independent of the other mixture constituents.

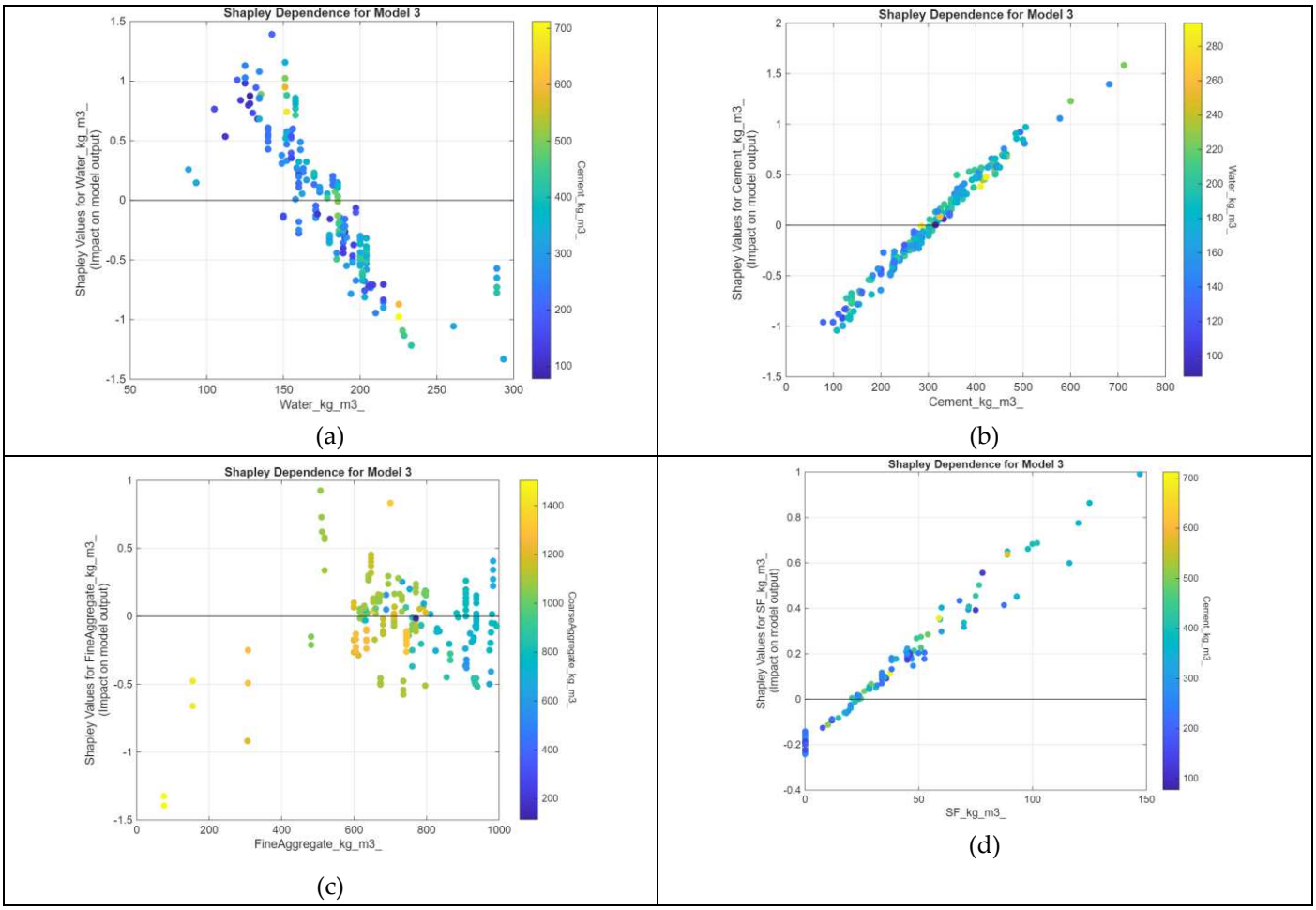


Figure 13. SHAP dependence plots for the four most influential predictors of the Bayesian-optimized GPR model: (a) water content, colored by cement content; (b) cement content, colored by water content; (c) fine aggregate content, colored by coarse aggregate content; and (d) silica fume content, colored by cement content.

5. Conclusion

This study developed an explainable machine-learning framework for predicting the splitting tensile strength of SCM-blended concrete using a global dataset. The analysis included 395 valid observations and eight mixture-related predictors: cement, water, coarse aggregate, fine aggregate, fly ash, silica fume, GGBFS, and superplasticizer. The modeling framework combined descriptive statistics, correlation analysis, multiple feature-ranking methods, broad model screening, hyperparameter optimization, and XAI-based interpretation.

Predictor rankings varied across mRMR, F-test, RReliefF, and Random Forest importance. This variation supports the use of complementary ranking methods rather than a single criterion. Based on the mean rank, superplasticizer, water, and cement were the three most influential variables. Correlation analysis also showed that cement had the strongest positive linear association with STS, whereas water had a moderate negative association.

A broad range of regression algorithms was evaluated. Rational Quadratic GPR provided the best baseline performance, with an RMSE of 0.51 MPa, MSE of 0.26 MPa², MAE of 0.34 MPa, MAPE of 10.02%, and R^2 of 0.85. GPR was therefore selected for further optimization using Bayesian optimization, random search, and grid search.

The three optimization strategies produced the same reported predictive accuracy: RMSE = 0.41898 MPa, MSE = 0.17554 MPa², MAE = 0.29095 MPa, MAPE = 8.9%, and R^2 = 0.89. Their training times differed substantially. Bayesian optimization required 89.72 s, compared with 187.81 s for random search and 316.86 s for grid search. Relative to the baseline model, the Bayesian-optimized GPR reduced RMSE, MSE, MAE, and MAPE by 17.85%, 32.48%, 14.43%, and 11.18%, respectively. R^2 increased by 4.71%, and prediction speed increased by about 101%. These gains were achieved at the cost of a longer training time than the baseline model.

The XAI results clarified how the optimized GPR used the mixture variables. PDP analysis showed generally positive effects for cement, silica fume, GGBFS, and superplasticizer, while water had a pronounced negative effect. Coarse and fine aggregates showed more complex nonlinear responses. Global SHAP analysis ranked water first and cement second. Fine aggregate, silica fume, and coarse aggregate also made meaningful contributions. SHAP summary and dependence plots showed that higher water contents generally reduced predicted STS, whereas higher cement and silica fume contents increased it. These patterns were consistent with the PDP results and provided an engineering interpretation of the model behavior.

The Bayesian-optimized GPR framework provided accurate and interpretable predictions of STS using mixture-composition variables alone. Combining model optimization with PDP and SHAP analysis improved predictive performance and clarified the relationships learned by the model. The findings are limited to the range and composition of the compiled dataset. Future work should therefore include external validation with independent experimental data and study-group-based validation across different data sources. Additional variables, such as curing conditions, specimen characteristics, SCM replacement ratios, and material properties, should also be considered. These extensions would strengthen the robustness and practical applicability of data-driven models for sustainable concrete mixture design.

Funding

This research received no external funding.

Ethics, Consent to Participate, and Consent to Publish declarations: not applicable.

References

[1] Scrivener, K. L., Lothenbach, B., De Belie, N., Gruyaert, E., Skibsted, J., Snellings, R., & Vollpracht, A. (2015). Recent advances in understanding the role of supplementary cementitious materials in concrete. Cement and Concrete Research, 78, 71–80. doi:10.1016/j.cemconres.2015.03.018.

[2] Samad, S., & Shah, A. (2017). Role of binary cement including supplementary cementitious material (SCM), in production of environmentally sustainable concrete: A critical review. International Journal of Sustainable Built Environment, 6(2), 663–674. doi:10.1016/j.ijsbe.2017.07.003.

[3] De Belie, N., Soutsos, M., & Gruyaert, E. (Eds.). (2018). Properties of Fresh and Hardened Concrete Containing Supplementary Cementitious Materials: State-of-the-Art Report of the RILEM Technical Committee 238-SCM, Working Group 4. Springer. doi:10.1007/978-3-319-70606-1.

[4] Ahmed, A. (2024). Assessing the effects of supplementary cementitious materials on concrete properties: A review. Discover Civil Engineering, 1, 145. doi:10.1007/s44290-024-00154-z.

[5] Liu, W., Mo, H., & Lee, C. K. (2025). Comprehensive dataset for performance evaluation of SCM-blended concrete. Data in Brief, 60, 111616. doi:10.1016/j.dib.2025.111616.

[6] Al-Shamasneh, A. R., Mahmoodzadeh, A., Kewalramani, M., Alghamdi, A., Alnahas, J., Sulaiman, M., Ghazouani, N., & Albaijan, I. (2025). Hybrid machine learning models for predicting the tensile strength of reinforced concrete incorporating nano-engineered and sustainable supplementary cementitious materials. Scientific Reports, 15, 35805. doi:10.1038/s41598-025-19741-w.

[7] Li, Q., Ren, G., Wang, H., Xu, Q., Zhao, J., Wang, H., & Ding, Y. (2023). Splitting tensile strength prediction of metakaolin concrete using machine learning techniques. Scientific Reports, 13, 20102. doi:10.1038/s41598-023-47196-4.

[8] Salim, A. M., & Salim, A. M. (2024). Mechanical properties of different types of fiber-reinforced concrete: Benefits and limitations. Proceedings of International Structural Engineering and Construction, 11(2), MAT-09. doi:10.14455/ISEC.2024.11(2).MAT-09.

[9] Hoang, N.-D. (2024). Predicting tensile strength of steel fiber-reinforced concrete based on a novel differential evolution-optimized extreme gradient boosting machine. Neural Computing and Applications, 36, 22653–22676. doi:10.1007/s00521-024-10458-x.

[10] Song, H., Ahmad, A., Farooq, F., Ostrowski, K. A., Maślak, M., Czarnecki, S., & Aslam, F. (2021). Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms. Construction and Building Materials, 308, 125021. doi:10.1016/j.conbuildmat.2021.125021.

[11] Li, Q.-F., & Song, Z.-M. (2022). High-performance concrete strength prediction based on ensemble learning. Construction and Building Materials, 324, 126694. doi:10.1016/j.conbuildmat.2022.126694.

[12] Sobuz, M. H. R., Aditto, F. S., Datta, S. D., Kabbo, M. K. I., Jabin, J. A., Hasan, N. M. S., Khan, M. M. H., Rahman, S. M. A., Raazi, M., & Zaman, A. A. U. (2024). High-strength self-compacting concrete production incorporating supplementary cementitious materials: Experimental evaluations and machine learning modelling. International Journal of Concrete Structures and Materials, 18, 67. doi:10.1186/s40069-024-00707-7.

[13] Luo, D., Wang, K., Wang, D., Sharma, A., Li, W., & Cho, I. H. (2025). Artificial intelligence in the design, optimization, and performance prediction of concrete materials: A comprehensive review. npj Materials Sustainability, 3, 14. doi:10.1038/s44296-025-00058-8.

[14] Alkharisi, M. K., & Dahish, H. A. (2025). Tensile strength of RA concrete containing supplementary cementitious materials and polypropylene fibers utilizing machine learning with GUI. Buildings, 15(24), 4473. doi:10.3390/buildings15244473.
[15] Ben Chaabene, W., Flah, M., & Nehdi, M. L. (2020). Machine learning prediction of mechanical properties of concrete: Critical review. Construction and Building Materials, 260, 119889. doi:10.1016/j.conbuildmat.2020.119889.
[16] Feng, D.-C., Liu, Z.-T., Wang, X.-D., Chen, Y., Chang, J.-Q., Wei, D.-F., & Jiang, Z.-M. (2020). Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach. Construction and Building Materials, 230, 117000. doi:10.1016/j.conbuildmat.2019.117000.
[17] Ahmad, A., Farooq, F., Niewiadomski, P., Ostrowski, K. A., Akbar, A., Aslam, F., & Alyousef, R. (2021). Prediction of compressive strength of fly ash based concrete using individual and ensemble algorithm. Materials, 14(4), 794. doi:10.3390/ma14040794.
[18] Xu, Y., Ahmad, W., Ahmad, A., Ostrowski, K. A., Dudek, M., Aslam, F., & Joyklad, P. (2021). Computation of high-performance concrete compressive strength using standalone and ensembled machine learning techniques. Materials, 14(22), 7034. doi:10.3390/ma14227034.
[19] Bui, D.-K., Nguyen, T. N., Chou, J.-S., Nguyen-Xuan, H., & Ngo, T. D. (2018). A modified firefly algorithm-artificial neural network expert system for predicting compressive and tensile strength of high-performance concrete. Construction and Building Materials, 180, 320–333. doi:10.1016/j.conbuildmat.2018.05.201.
[20] Rathakrishnan, V., Bt. Beddu, S., & Ahmed, A. N. (2022). Predicting compressive strength of high-performance concrete with high volume ground granulated blast-furnace slag replacement using boosting machine learning algorithms. Scientific Reports, 12, 9539. doi:10.1038/s41598-022-12890-2.
[21] Amin, M. N., Khan, K., Abu Arab, A. M., Farooq, F., Eldin, S. M., & Javed, M. F. (2023). Prediction of sustainable concrete utilizing rice husk ash (RHA) as supplementary cementitious material (SCM): Optimization and hyper-tuning. *Journal of Materials Research and Technology*, *25*, 1495–1536. doi:10.1016/j.jmrt.2023.06.006.
[22] Rasmussen, C. E., & Williams, C. K. I. (2006). Gaussian Processes for Machine Learning. MIT Press, Cambridge, MA.
[23] Hoang, N.-D., Pham, A.-D., Nguyen, Q.-L., & Pham, Q.-N. (2016). Estimating compressive strength of high performance concrete with Gaussian process regression model. Advances in Civil Engineering, 2016, 2861380. doi:10.1155/2016/2861380.
[24] Xie, L., He, J., Wang, C., Zhu, S., Zhang, H., & Yang, X. (2025). Enhanced Bayesian Gaussian process regression for compressive strength prediction of multi-binder concrete. Journal of Building Engineering, 111, 113308. doi:10.1016/j.jobe.2025.113308.
[25] Yu, Y., Wang, G., Huseien, G. F., Zou, Z., Ding, Z., & Zhang, C. (2025). Intelligent prediction of compressive strength of self-compacting concrete incorporating silica fume using hybrid IWOA-GPR model. Materials Today Communications, 45, 112282. doi:10.1016/j.mtcomm.2025.112282.
[26] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774.
[27] Ekanayake, I. U., Meddage, D. P. P., & Rathnayake, U. (2022). A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP). Case Studies in Construction Materials, 16, e01059. doi:10.1016/j.cscm.2022.e01059.
[28] Kulasooriya, W. K. V. J. B., Ranasinghe, R. S. S., Perera, U. S., Thisovithan, P., Ekanayake, I. U., & Meddage, D. P. P. (2023). Modeling strength characteristics of basalt fiber reinforced concrete using multiple explainable machine learning with a graphical user interface. Scientific Reports, 13, 13138. doi:10.1038/s41598-023-40513-x.
[29] Alabdullah, A. A., Iqbal, M., Zahid, M., Khan, K., Amin, M. N., & Jalal, F. E. (2022). Prediction of rapid chloride penetration resistance of metakaolin based high strength concrete using light GBM and XGBoost models by incorporating SHAP analysis. Construction and Building Materials, 345, 128296. doi:10.1016/j.conbuildmat.2022.128296.
[30] Elhishi, S., Elashry, A. M., & El-Metwally, S. (2023). Unboxing machine learning models for concrete strength prediction using XAI. Scientific Reports, 13, 19892. doi:10.1038/s41598-023-47169-7.
[31] Wang, X., Chen, A., & Liu, Y. (2022). Explainable ensemble learning model for predicting steel section-concrete bond strength. Construction and Building Materials, 356, 129239. doi:10.1016/j.conbuildmat.2022.129239.
[32] Liang, M., Chang, Z., Wan, Z., Gan, Y., Schlangen, E., & Šavija, B. (2022). Interpretable ensemble-machine-learning models for predicting creep behavior of concrete. Cement and Concrete Composites, 125, 104295. doi:10.1016/j.cemconcomp.2021.104295.
[33] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. The Annals of Statistics, 29(5), 1189–1232. doi:10.1214/aos/1013203451.
[34] Kashem, A., Karim, R., Datta, S. D., & Alharthai, M. (2024). Hybrid data-driven approaches to predicting the compressive strength of ultra-high-performance concrete using SHAP and PDP analyses. Case Studies in Construction Materials, 20, e02991. doi:10.1016/j.cscm.2024.e02991.
[35] Nguyen, N.-H., Abellán-García, J., Lee, S., Garcia-Castano, E., & Vo, T. P. (2022). Efficient estimating compressive strength of ultra-high performance concrete using XGBoost model. Journal of Building Engineering, 52, 104302. doi:10.1016/j.jobe.2022.104302.
[36] Liu, W., Mo, H., & Lee, C. K. (2025). A global dataset of SCM-blended concrete and its properties (Version 2) [Dataset]. Mendeley Data. doi:10.17632/npwgdyy4np.2.