

Research Article

Ahmed A. Alawi Al-Naghi*, Muhammad Nasir Amin*, Suleman Ayub Khan*, and Muhammad Tahir Qadir

Modeling the strength parameters of agro waste-derived geopolymer concrete using advanced machine intelligence techniques

<https://doi.org/10.1515/rams-2024-0035>
received March 11, 2024; accepted June 04, 2024

Keywords: artificial intelligence, geopolymer concrete, mechanical properties

Abstract: The mechanical strength of geopolymer concrete incorporating corncob ash and slag (SCA-GPC) was estimated by means of three distinct AI methods: a support vector machine (SVM), two ensemble methods called bagging regressor (BR), and random forest regressor (RFR). The developed models were validated using statistical tests, absolute error assessment, and the coefficient of determination (R^2). The importance of various modeling factors was determined by means of interaction diagrams. When estimating the flexural strength and compressive strength of SCA-GPC, R^2 values of over 0.85 were measured between the actual and predicted findings using both individual and ensemble AI models. Statistical testing and k -fold analysis for error evaluation revealed that the RFR model outperformed the SVM and BR models in terms of accuracy. As demonstrated by the interaction graphs, the mechanical characteristics of SCA-GPC were found to be extremely responsive to the mix proportions of ground granulated blast furnace slag, fine aggregate, and corncob ash. This was the case for all three components. This study demonstrated that highly precise estimations of mechanical properties for SCA-GPC can be made using ensemble AI techniques. Improvements in geopolymer concrete performance can be achieved by the implementation of such practices.

Abbreviations

AI	artificial intelligence
BR	bagging regressor
CCA	corncob ash
CS	compressive strength
FA	fine aggregate
FAS	fly ash
FS	flexural strength
GPC	geopolymer concrete
GGBFS	ground granulated blast furnace slag
MAE	mean absolute error
MK	metakaolin
MAPE	mean absolute percentage error
ML	machine learning
PC	Portland cement
RFR	random forest regressor
RHA	rice husk ash
RM	red mud
RMSE	root mean squared error
SCMs	supplementary cementitious materials
SF	silica fume
SVM	support vector machine
SCA-GPC	concrete incorporating corncob ash and slag

* **Corresponding author: Ahmed A. Alawi Al-Naghi**, Civil Engineering Department, University of Ha'il, Ha'il, 55476, Saudi Arabia, e-mail: a.alnaghi@uoh.edu.sa

* **Corresponding author: Muhammad Nasir Amin**, Department of Civil and Environmental Engineering, College of Engineering, King Faisal University, Al-Ahsa 31982, Saudi Arabia, e-mail: mgadir@kfu.edu.sa

* **Corresponding author: Suleman Ayub Khan**, Department of Civil Engineering, COMSATS University Islamabad, Abbottabad 22060, Pakistan, e-mail: sulemanayub@cuiatd.edu.pk

Muhammad Tahir Qadir: Department of Civil and Environmental Engineering, College of Engineering, King Faisal University, Al-Ahsa 31982, Saudi Arabia

1 Introduction

The long history of concrete as an essential building material has highlighted the environmental impact of concrete over years [1]. With the global demand for cement and concrete expected to triple by 2,050, carbon emissions are projected to increase, and biodiversity is likely to decline at a faster rate than previously anticipated [2].

Due to its large energy and carbon footprint, Portland cement (PC) has been the target of researchers seeking to create alternative binders [2]. In the manufacturing process of PC, crucial for concrete binding, approximately 1.80 metric tons of raw materials are utilized, resulting in the emission of 0.8 metric tons of CO₂ [3]. Thus, cement output must be mitigated immediately to reduce environmental change [2]. One methodical and technical approach to ensuring materials' long-term viability is to recycle them into fresh construction materials from agricultural and industrial waste [4]. There are societal, economic, and environmental benefits to producing supplementary cementitious materials from recycled agricultural and industrial waste [5,6]. Using recycled materials in place of PC has been proven to be an efficient, affordable, and long-term strategy for reducing one's carbon footprint [7–9].

Sustainable concrete, also known as geopolymers concrete (GPC), replaces the PC with recycled agro-industrial resources, making a cementitious binder redundant [10,11]. The utilization of alkali hydroxide and alkali silicate appears to be a component of the activation process for raw materials based on the aluminosilicate structure [12]. There is a wide variety of reprocessed agronomic and manufacturing materials that have potential as precursors, including fly ash (FAS), red mud (RM), geopolymers (alumino-silicates), rice husk ash (RHA), ground granulated blast furnace slag (GGBFS), silica fume (SF), and metakaolin (MK) [13–18]. GGBFS in producing GPC presents minimal environmental repercussions alongside favorable cost-effectiveness, heightened rigidity, and exceptional resistance to chemical degradation. Moreover, it holds promise as a key ingredient in eco-friendly and economically viable concrete formulations [19–23]. The corn cob ash (CCA) component, instead, is novel. More traditional pozzolanic components, including FAS and RHA, can be replaced or supplemented with CCA due to its elevated level of silica. The usage of on-the-spot heated GPC is associated with a number of problems; thus, researchers are considering creating this green concrete at room temperature instead. It is also critical to know that there are other criteria for judging performance outside of reaching strength norms. Evaluating a structure's resistance to environmental and other pressures is essential for accurately estimating its lifespan. GPC is a prospective concrete solution that could be used in ecologically sensitive places because of its improved mechanical capabilities and improved resilience [31]. All of the aforementioned factors point to GPC's unique chemical makeup as the source of the material's exceptional mechanical capabilities and endurance [18,24,25]. Using nano-silica and reused plastic particles has allowed GPC to perform better in recent years [26–28]. Waste-based GPC has numerous advantages, as can be seen in Figure 1.



Figure 1: Advantages of waste-derived GPC in construction [29].

Experts in the fields of science, engineering, research, and computer programming are starting to notice that AI is having a major influence on how new products are developed and enhanced. Problems exist in the engineering industry, and there is a high demand for individuals who can find ways to integrate AI into their jobs. Nevertheless, there are still certain downsides and performance issues with AI-based systems, even though the future seems bright. Artificial intelligence programs have formidable obstacles when it comes to tasks that people typically take for granted, such as object identification and natural conversation understanding [30]. This poses a challenge for modern AI in creating appropriate alternatives for training computer perception. AI systems have utilized machine learning (ML) to tackle these issues [30,31]. ML algorithms allow computers to gain the necessary expertise for autonomous action by analyzing a sufficiently large dataset [32,33]. Getting back to the qualities that make the most explicit data is the first step before putting the plan into action. The term “feature extraction” is now used to describe this procedure. Then, ML is used to train sample data, attributes, and pattern separation instructions [30,34,35]. Modern civil engineering research relies on statistical methods and AI to address ever-increasingly complicated issues. Estimating concrete's compressive strength (CS) is a typical use case for these techniques in civil engineering [15,36]. The ability to forecast self-compacting concrete's slump and impact strength [37], varied column axial

bearing [38], and shear behavior of beams in a structure [39], as well as the forecasting of chloride contamination [40] are some of the harder challenges solved utilizing these strategies. These estimates assist in decreasing the number of test configurations for future investigations, shortening their duration and expense. ML approaches such as artificial neural networks, gradient boosting (GB), expression trees (ETs), Gaussian process regression (GPRs), decision trees (DTs), support vector machines (SVMs), and extreme gradient boosting (XGB) may estimate concrete strength [41–43]. The mechanical properties of GPC were better predicted by the individual and ensemble models than by any of the other models.

This study used experimental data and AI algorithms to forecast the mechanical properties of slag and corncob ash-centered geo-polymer concrete (SCA-GPC), a GPC composed of slag and CCA. One standalone ML method and two ensemble ML processes were employed in the study to accomplish its goals. One method to evaluate the models' accuracy involved comparing the predicted and actual outcomes, using statistical tests, and performing *K*-fold analysis. Carrying out experiments is difficult because of the lengthy and complex procedures involved in collecting materials, casting samples, curing them to increase strength, and evaluating them. Modern modeling techniques like ML can significantly aid the construction industry by overcoming these challenges. Conventional testing methods struggle to assess the overall impact of all parameters on SCA-GPC strength. To identify the most important variables, this study employed interaction graphs. Data necessary for ML approaches can be gathered from existing research. The dataset now has a plethora of potential uses, such as in ML algorithms, impact studies, and material property estimations. Utilizing an experimental dataset, this article validates the efficacy of ensemble ML algorithms in predicting SCA-GPC strength. The study's findings might pave the way

for greener construction methods, which would enhance GPC's value to the business.

2 Study methods

2.1 Collecting and evaluating data

The research employed ML models, including SVM, bagging regressor (BR), and random forest regressor (RFR), to predict the CS and FS of SCA-GPC. The experimental investigation yielded a dataset comprising 260 data points [44]. According to the eight input variables (NaOH pellets [SHP], molar concentration [MC], GGBFS, curing day [CD], fine aggregate [FA], water [W], CCA, and concrete grade [CG]), the CS and FS of SCA-GPC were predicted. The data were collected and organized using data preparation. Data preparation for mining data is the standard technique for knowledge discovery from data to minimize major obstacles. Data preparation involves eliminating noise and unnecessary details from the dataset. Descriptive statistics in Table 1 provide a comprehensive summary of key characteristics within the refined dataset, offering valuable insights into its central tendencies, variability, and distribution. These statistics serve as fundamental tools for understanding the dataset's structure, facilitating informed decision-making and hypothesis testing in subsequent analyses. One common way to find parameter dependencies is to use Pearson's correlation coefficient (*r*) [45]. Two Figure 2(a) for CS and 2(b) for FS, show the results of the association map plot for the attributes. The *r*-squared test is useful for demonstrating parameter dependency and multicollinearity [46]. Within the range of -1 to $+1$, a strong negative relationship is provided by -1 , a strong positive link by $+1$, and no correlation at all by 0 for the *r*-value [47]. This correlation

Table 1: Variable descriptions using statistics (parameters similar to [48–50])

Parameters	Mean	Standard error	Median	Mode	Standard deviation	Sample variance	Kurtosis	Skewness	Minimum	Maximum
GGBFS ($\text{kg}\cdot\text{m}^{-3}$)	218.65	9.54	228.00	0.00	153.90	23685.60	−1.12	0.12	0.00	488.00
CCA ($\text{kg}\cdot\text{m}^{-3}$)	215.24	9.48	195.00	0.00	152.78	23341.70	−1.12	0.13	0.00	488.00
FA ($\text{kg}\cdot\text{m}^{-3}$)	818.12	3.53	841.00	841.00	56.90	3237.34	−1.41	−0.14	728.00	899.00
W ($\text{kg}\cdot\text{m}^{-3}$)	35.22	0.13	35.16	37.86	2.15	4.63	−1.53	0.04	32.64	37.86
SHP ($\text{kg}\cdot\text{m}^{-3}$)	23.38	0.13	23.44	20.74	2.15	4.63	−1.53	−0.04	20.74	25.96
CD (days)	45.25	1.93	42.00	7.00	31.20	973.43	−1.34	0.25	7.00	90.00
MC (M)	14.00	0.10	14.00	12.00	1.65	2.72	−1.53	0.00	12.00	16.00
CG (MPa)	34.96	0.31	30.00	30.00	5.01	25.10	−2.02	0.02	30.00	40.00
CS (MPa)	35.90	0.76	36.04	29.04	12.19	148.64	−0.74	0.17	10.67	64.09
FS (MPa)	5.36	0.06	5.47	5.00	1.01	1.03	−0.58	−0.24	2.81	7.45

between the input variables and the output (CS and FS) is displayed in the bottom row of Pearson's array.

2.2 ML modeling

Laboratory studies were used to assess the mechanical properties of SCA-GPC. While CS and FS need ten inputs,

the prediction models were built using just eight of the variable inputs. The SCA-GPC's CS and FS were predicted using advanced ML algorithms that included SVM, BR, and RFR. The study achieved its goals by using Python code in the Spyder environment of Anaconda Navigator (version 5.1.5). Typically, ML algorithms are utilized to compare outputs with inputs throughout the process. Researchers allocated 70% of the data for training ML models, reserving the remaining 30% for testing. Additionally, the R^2 value of the

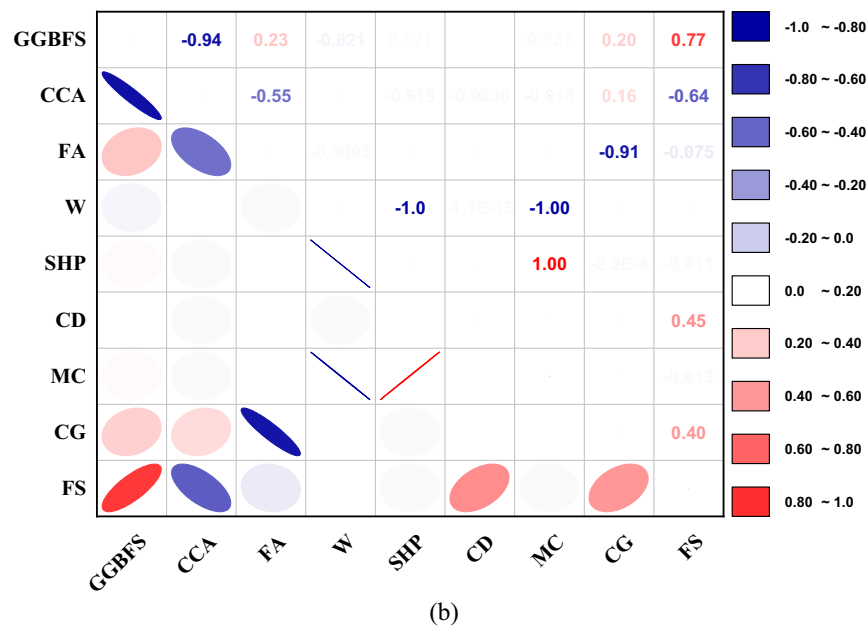
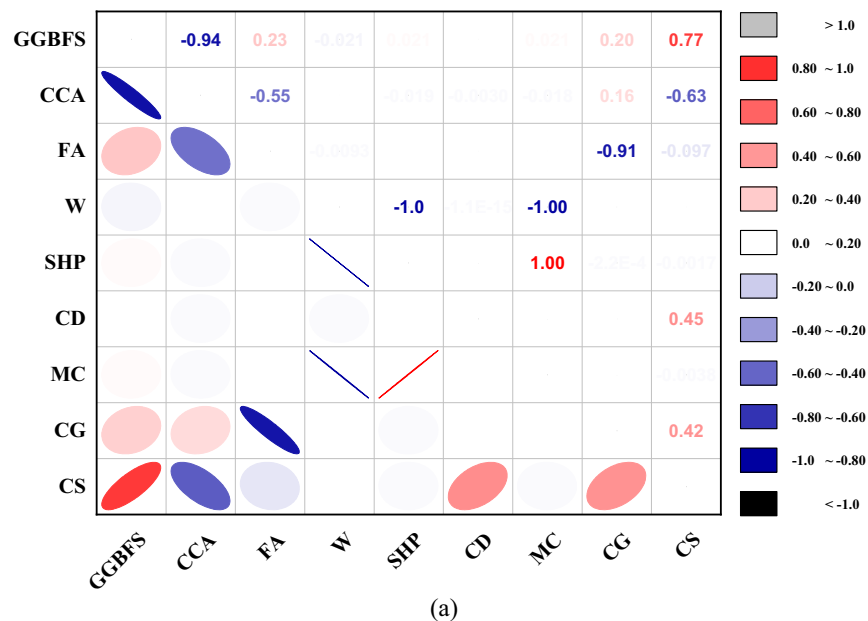


Figure 2: Parameter correlation heat map: (a) CS and (b) FS.

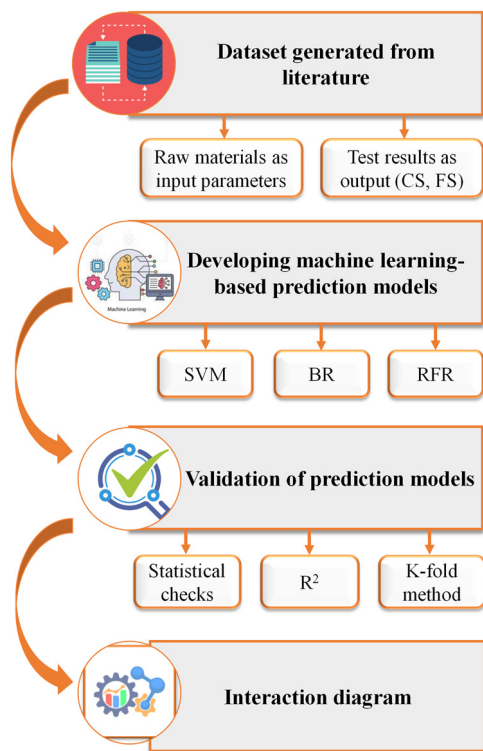


Figure 3: ML process outline.

predicted outcome served as an indicator of the model's reliability. A low R^2 score signifies a significant deviation between predicted and actual outcomes, highlighting substantial discrepancies in the model's predictive accuracy. This metric serves as a crucial indicator of the model's efficacy in capturing the variance within the dataset, with lower scores suggesting a less accurate representation of the observed data [51]. The correctness of the model was validated by a number of analyses, which included statistical examinations and evaluations of errors. A simple graphical

representation of an event model is shown in Figure 3, which may be seen hereunder.

2.2.1 Support vector machine

For supervised ML tasks like data regression and classification evaluation, there is the support vector machine (or SVM). SVM classification systems employ diverse categorization strategies aimed at maximizing the separation between different classes to the greatest extent possible within practical constraints. This approach ensures robust classification performance by effectively delineating boundaries between distinct categories in the feature space. For the purpose of depicting the samples, this method makes use of points on a plane or line. The additional instances are arranged in a manner that corresponds to their orientation along the vector, as shown in Figure 4. Figure 5 delineates the systematic approach for implementing SVM models, meticulously designed to deliver a holistic assessment of material strength considering multiple influential factors. This framework empowers users to fine-tune SVM model parameters using sophisticated optimization techniques, thereby augmenting its predictive precision and utility in material strength analysis.

2.2.2 BR

The BE technique is illustrated in Figure 6, a simplified flow diagram. A comparable ensemble method is the most effective way to describe the steps required to augment the forecast model with additional training data sets. Asymmetric sample statistics are substituted for the original set of statistics. With each new batch of training samples, it is

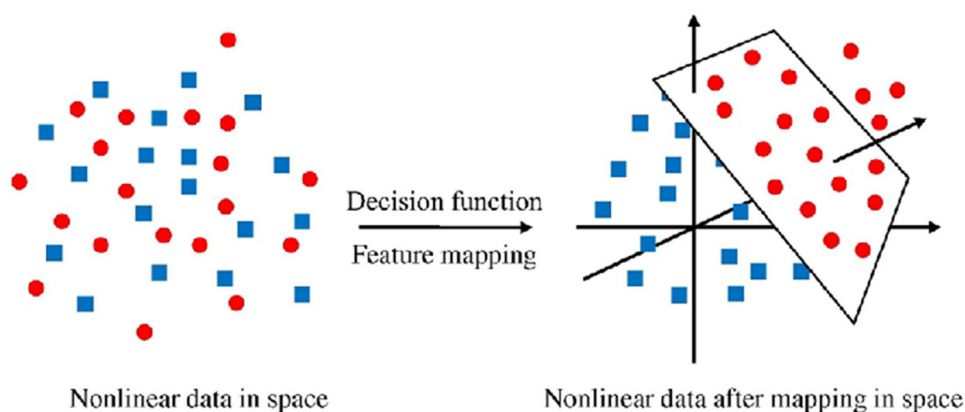


Figure 4: Model mapping SVM [52].

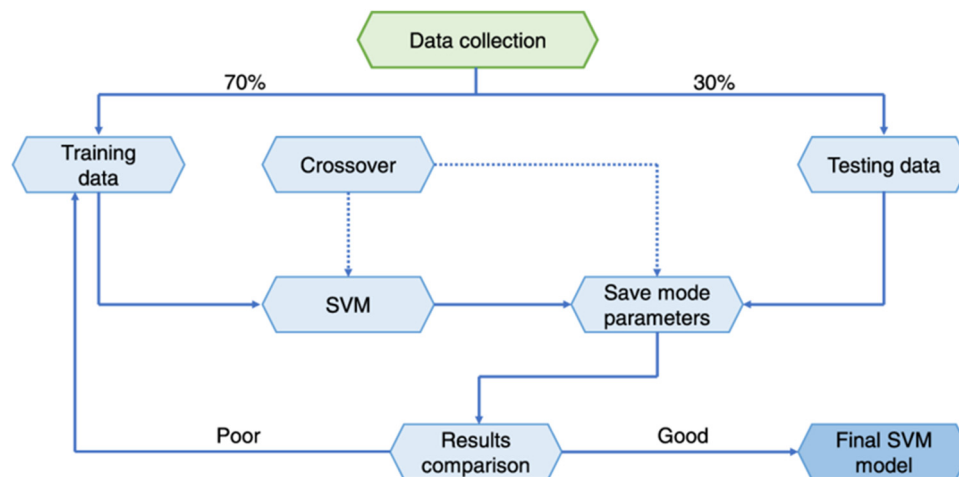


Figure 5: Process flow diagram for SVM [53].

possible to find the same or similar results. Once the bagging procedure is completed, each component has an equal chance of being included in the updated dataset. The overall projection quality remains unaffected by the size of the training dataset. The divergence could be significantly reduced if the target output is more accurately approximated. For each run of the simulations, this ensemble takes the average prediction and uses it. The median prediction from several simulations is utilized in regression [54]. Twenty separate models are used separately to optimize the SVM-based bagging technique and find its best output.

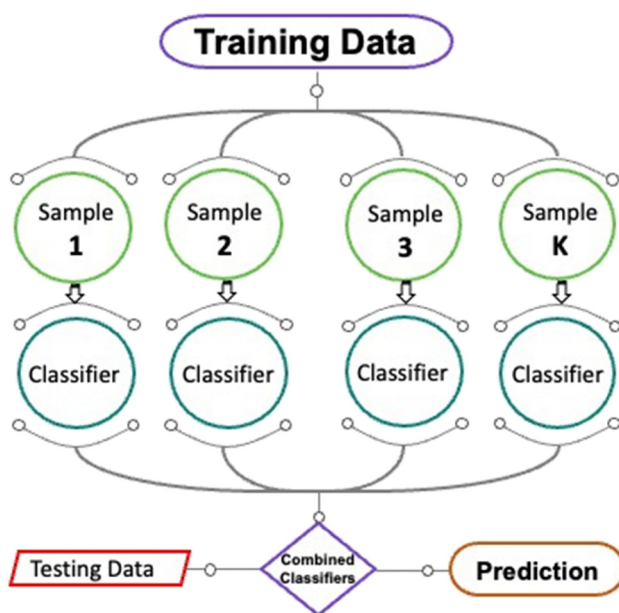


Figure 6: Bagging ensemble model: a schematic explanation [55].

2.2.3 Random forest

Random split selection, in conjunction with bagged decision trees, allows for the attainment of RFR [56]. The assembly and operation of the RFR model are depicted in a simplified diagram, as shown in Figure 7. There is a random selection process for both the training data and the input parameters required to create each branch split in the forest's trees [57]. The natural diversity of the tree is enhanced by the presence of this variable. When it comes to the forest, only completely developed binary trees are there. In the realm of universal regression techniques, the RFR method has been demonstrated to be effective. When the amount of variables exceeds the maximum number of possible clarifications, it has been demonstrated that combining the results of a large number of decision trees that were chosen at random yields more accurate results. It is useful for both planned and unplanned learning activities because the significance of its indications shifts significantly throughout the course of time [56].

2.3 Model's validation

For the purpose of ensuring that the ML models had an accurate representation of the data, a number of distinct mathematical techniques and k -fold procedures were developed and implemented. The k -fold technique is frequently applied for the purpose of determining whether or not a procedure is considered to be effective. This strategy involves arbitrarily dividing the data set into ten different categories [59]. As depicted in Figure 8, ML simulations are trained using nine distinct sets, with only one reserved for validation. ML

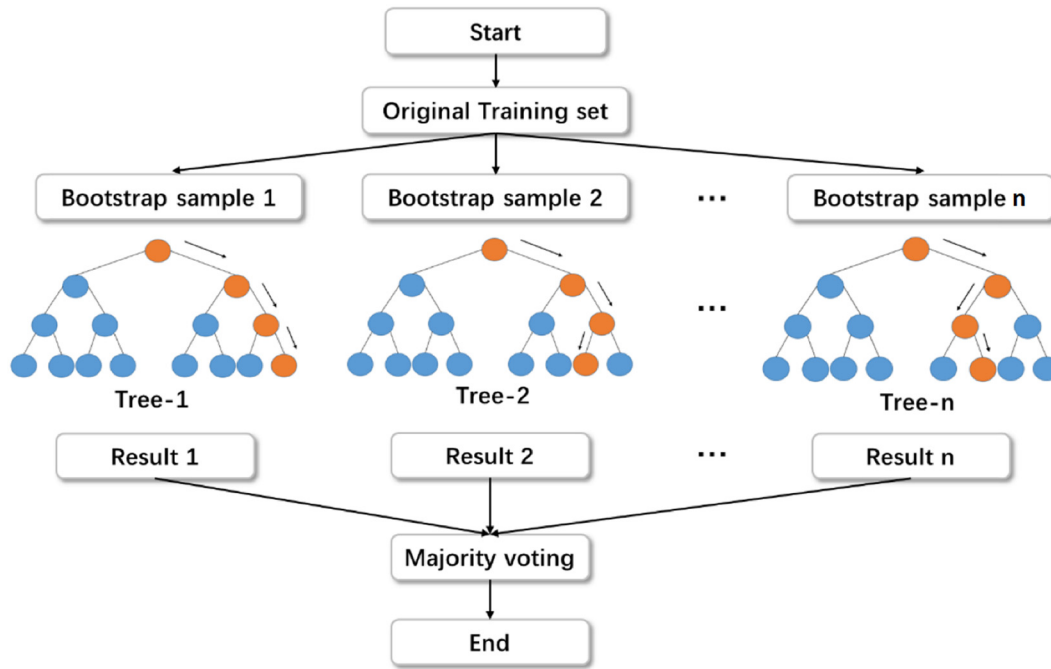


Figure 7: Random forest modeling and structure [58].

methods exhibit good performance in scenarios with low error and high R^2 . Additionally, to yield positive outcomes, the treatment needs to be conducted a total of ten times. The precision of the model, which was previously quite excellent, is greatly improved by this procedure. Various ML approaches were also correlated by employing statistical error evaluation metrics like mean absolute percentage error, mean absolute error (MAE), and root mean squared error.

Eqs. (1) and (3), obtained from previous works, were employed to statistically test the precision of the ML methods' estimates [60,61]

$$\text{RMSE} = \sqrt{\sum \frac{(P_i - T_i)^2}{n}}, \quad (1)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |P_i - T_i|, \quad (2)$$

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \frac{|P_i - T_i|}{T_i}. \quad (3)$$

In this context, n stands for the total number of observations, P_i refers to the anticipated results, and T_i indicates the actual measured values.

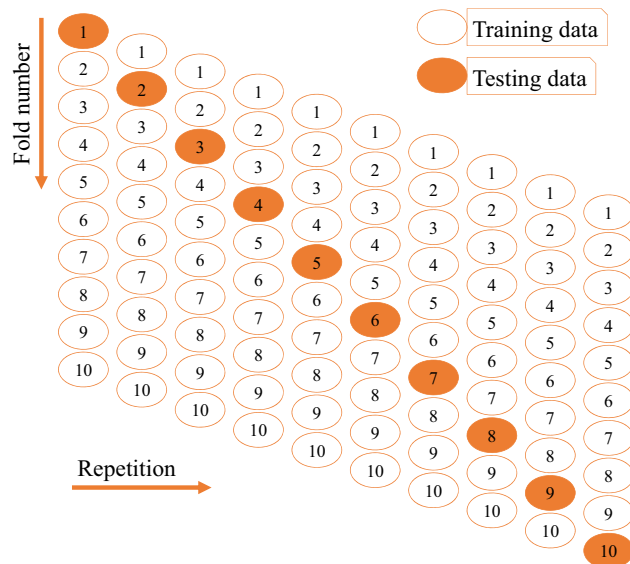


Figure 8: K-method operational procedure [52].

2.4 Input parameter interaction analysis

Python and Jupyter Notebook 6.4.12 simulated input feature interaction. Matplotlib was used to create interaction graphs. Jupyter Notebook enables users to write and share interactive code, graphs, equations, and text documents online [62]. Among the numerous uses for this platform are filtering of data and alterations, mathematical simulation, arithmetical modeling, and data conception, among others [63]. For visualizing two-dimensional data arrays, Matplotlib is one of the Python libraries that is used the

most frequently [64]. Initiating the `plot()` method, preparing the data, and establishing the required dependencies are all prerequisites to starting. In order to display a plot, the `show()` method must be used. Matplotlib is a Python library that uses NumPy, an extension for Python used in numerical mathematics [65]. It includes a variety of graphs: line, bar, scatter, and histogram. The study utilized scatter plots to visually depict the relationship between the input variables, a method commonly employed in numerous comparable investigations [66,67].

3 Results and analysis

3.1 CS models

3.1.1 CS-SVM model

Figure 9 shows the results of estimating the CS of SCA-GPC using the SVM model. Figure 9(a) graphically illustrates the agreement between the anticipated and observed CS. The dotted lines indicate a 20% deviation from the solid black line, which represents a perfect match with the data. Predictions for CS from the SVM model and the measured values were very close. To effectively determine the CS of SCA-GPC, the SVM technique was utilized. The results displayed a notable level of accuracy, with 83% of predictions falling within the 20% criterion and an R^2 value of 0.8745. Figure 9(b) shows the range of differences (errors) between experimental and predicted values using the SVM method is illustrated. The erroneous values exhibited a standard deviation ranging from 0.09 to 9.50 MPa, with an average of approximately 3.42 MPa. Specifically, there were 17 values below 1 MPa, 20 falling between 1 and 3 MPa, and 41 exceeding 3 MPa. Despite the scattered data, the error distribution suggests that the SVM model can effectively predict the CS of SCA-GPC.

3.1.2 CS-BR model

Figure 10 presents the outcomes of estimating the CS of SCA-GPC using the BR model. Figure 10(a) shows a distinct correlation between observed and predicted CS values, where the solid black line represents an ideal fit, and the dotted lines indicate a deviation of up to 20%. The experimental CS values closely align with the predictions from the BR model. The BR technique demonstrates remarkable performance, achieving an R^2 value of 0.9365 and with 97% of predictions falling within the 20% deviation threshold, indicating significant accuracy enhancement. Figure 10(b)

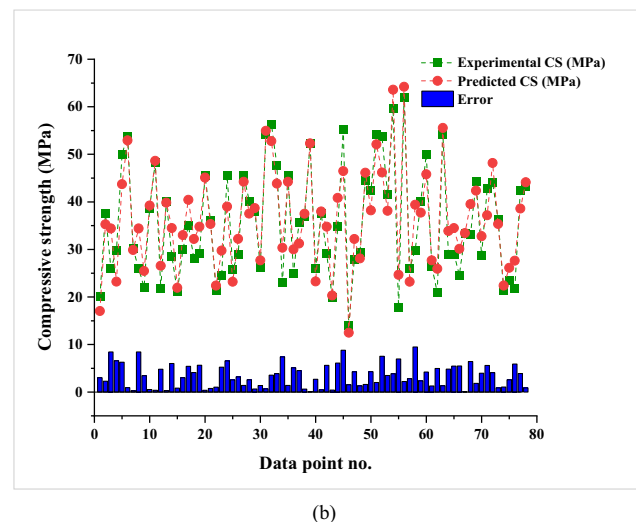
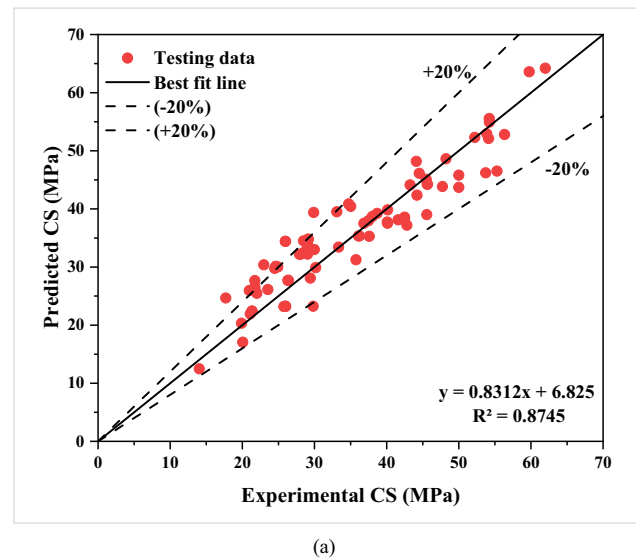


Figure 9: (a) Connection between experimental and predicted CS in the CS-SVM model and (b) scattering of errors and predicted CS.

shows the range of errors between experimental and predicted values using the BR technique, with incorrect results exhibiting a standard deviation ranging from 0.07 to 7.42 MPa and an average of 2.19 MPa. The data are categorized into 19 instances below 1 MPa, 27 falling between 1 and 3 MPa, and 22 exceeding 3 MPa. Error distribution analysis indicates that the BR model provides more precise predictions for SCA-GPC CS compared to the SVM model, albeit with slightly narrower variability in measurements.

3.1.3 CS-RFR model

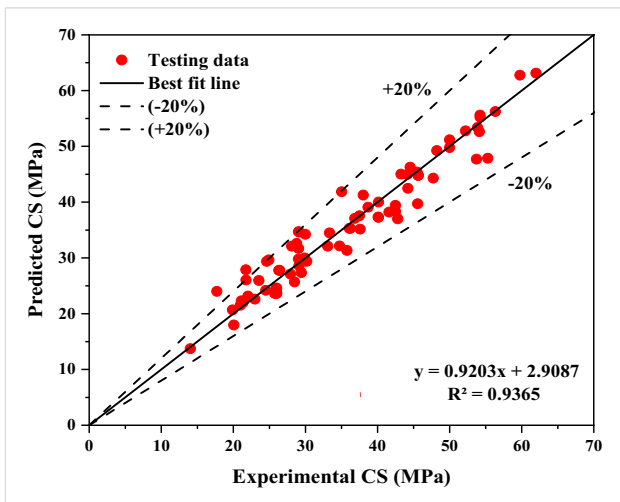
Figure 11 shows the utilization of the RFR model to estimate the CS of SCA-GPC. Figure 11(a) illustrates the agreement between observed and predicted CS values. The CS values

predicted by the RFR model closely resemble those obtained experimentally. The RFR technique exhibits remarkable accuracy in estimating the CS of SCA-GPC, boasting an impressive R^2 value of 0.9688, with 99% of predictions falling below the 20% threshold, as depicted in Figure 11(b). The error range, depicting the differences between experimental and predicted values using the RFR approach, varies with a standard deviation ranging from 0.01 to 11.56 MPa and an average of approximately 1.32 MPa. Further analysis reveals that among the total values, 48 were below 1 MPa, 21 fell between 1 and 3 MPa, and only 9 exceeded 3 MPa. The error distribution underscores the superior accuracy of the RFR model in predicting the CS of SCA-GPC compared to both the SVM and BR models, with significantly reduced error spread.

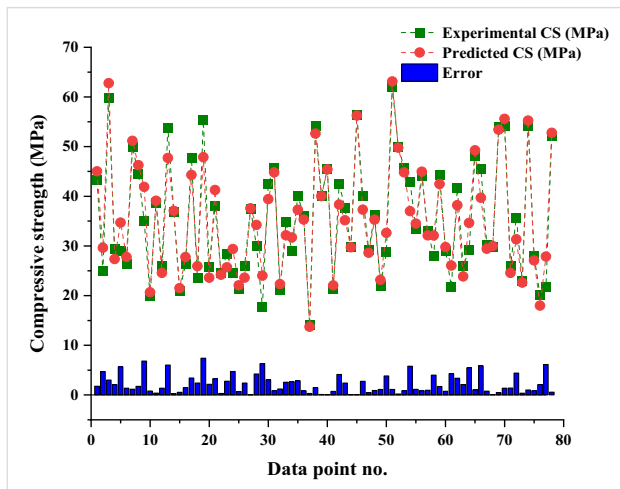
3.2 FS models

3.2.1 FS-SVM model

Figure 12 shows the result of using the SVM model to approximate the FS of SCA-GPC. Figure 12(a) shows the agreement between the expected and observed FS. In terms of FS, the predictions made by the SVM model were pretty comparable to the values that were measured. An effective estimation of the FS of SCA-GPC was achieved through the utilization of the SVM analysis. A high level of accuracy was exhibited by the model, which, similar to the CS-SVM model, had an R^2 value of 0.8853 and had a 100% of its predictions falling below the threshold of 20%. As

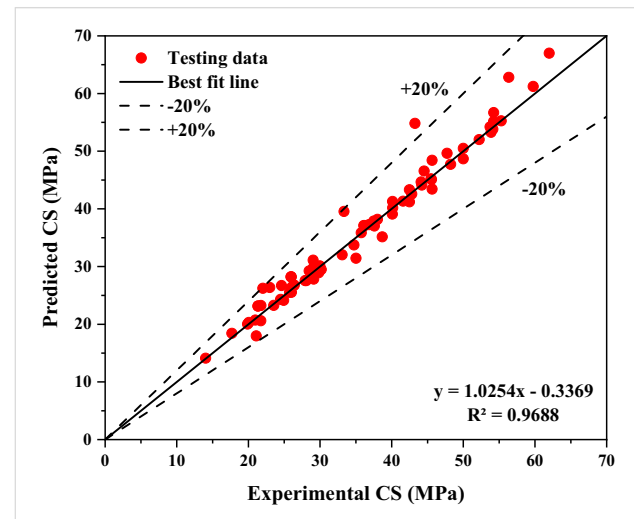


(a)

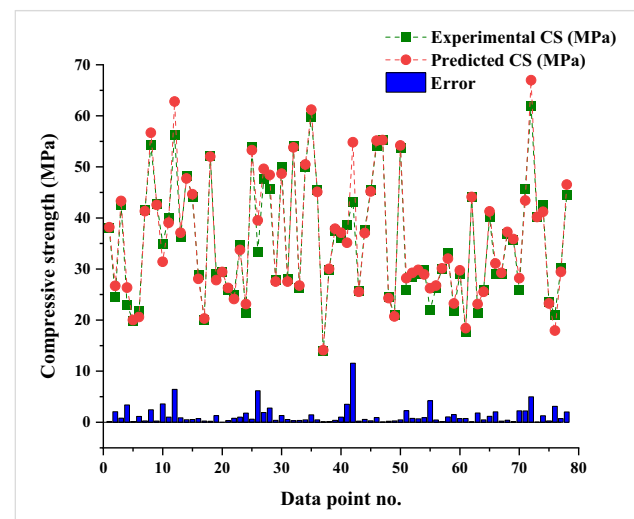


(b)

Figure 10: (a) Connection between experimental and predicted CS in the CS-BR model and (b) scattering of errors and predicted CS.

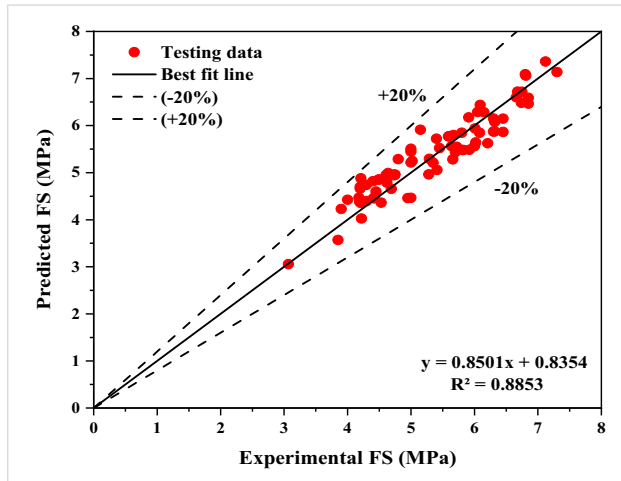


(a)

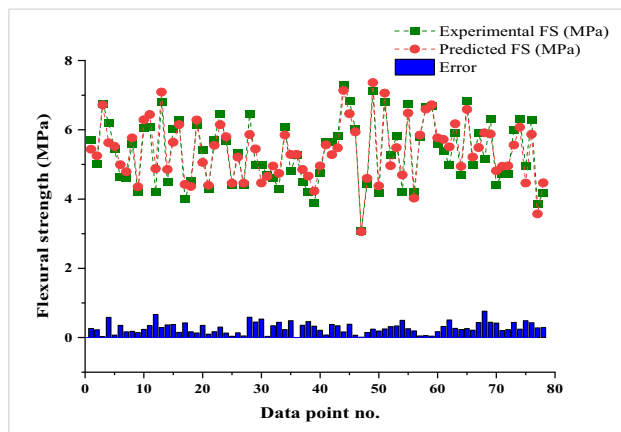


(b)

Figure 11: (a) Connection between experimental and predicted CS in the CS-RFR model and (b) scattering of errors and predicted CS.



(a)



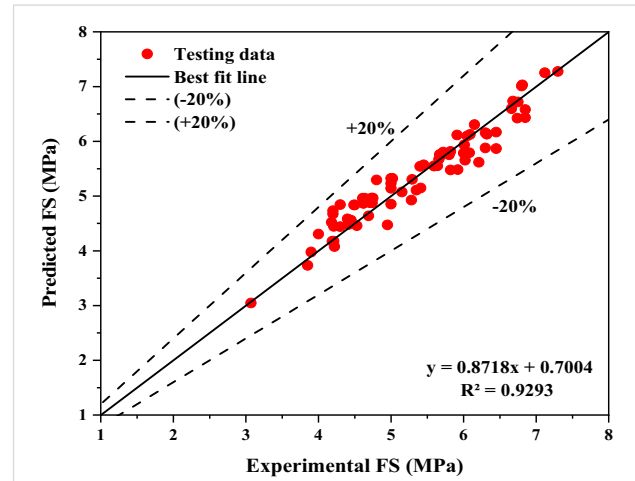
(b)

Figure 12: (a) Connection between experimental and predicted FS in the FS-SVM model and (b) scattering of errors and predicted FS.

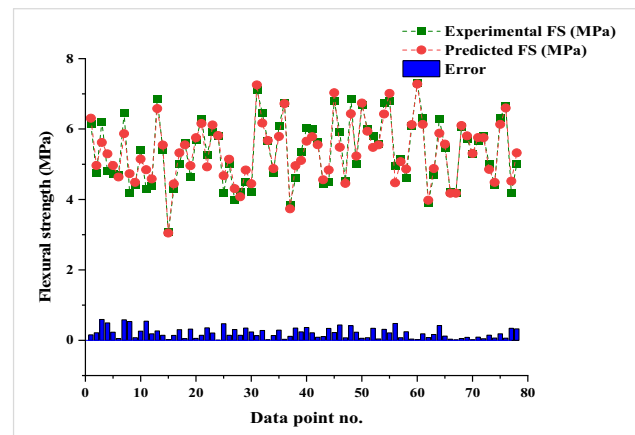
shown in Figure 12(b), the SVM method's projected values differ from the experimental values by a wide variety of margins. The erroneous results had a standard deviation ranging from 0.003 to 0.763 MPa, with an average of approximately 0.276 MPa. Additionally, the analysis unveiled that 45 of the values were below 0.3 MPa, 27 fell within the range of 0.3–0.5 MPa, and 6 were found to exceed 0.5 MPa. It is clear, after examining the distribution of the errors, that the FS of SCA-GPC may be predicted by applying an SVM model despite the fact that its measurements (errors) are widely dispersed.

3.2.2 FS-BR model

For the purpose of approximating the FS of SCA-GPC, the BR model was utilized, as shown in Figure 13. Figure 13(a)



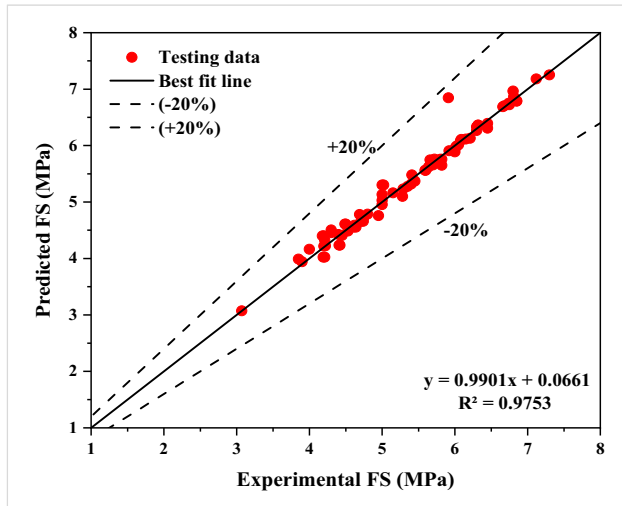
(a)



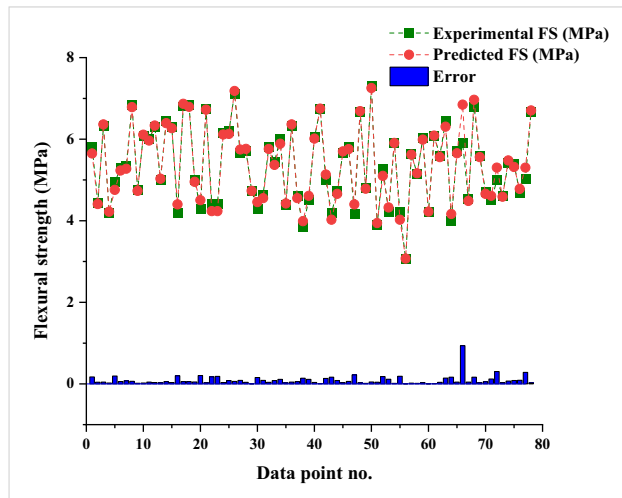
(b)

Figure 13: (a) Connection between experimental and predicted FS in the FS-BR model and (b) scattering of errors and predicted FS.

shows the graphic representation of the agreement between the observed and projected FS. The experimental results for FS were very close to the predictions made by the BR model. A remarkable level of accuracy was achieved when the BR approach was used to efficiently identify the FS of SCA-GPC. The method's R^2 value was 0.9293, and all of its predictions fell inside the 20% criterion. Figure 13(b) illustrates the range of discrepancies (errors) between the BR-predicted and experimental values. The error values, averaging approximately 0.205 MPa, had a standard deviation ranging from 0.003 to 0.591 MPa. Moreover, 56 of the values were below 0.3 MPa, 18 fell between 0.3 and 0.5 MPa, and 4 exceeded 0.5 MPa. The BR model's FS prediction of SCA-GPC was noticeably more accurate than the SVM model's, with slightly lower spread measurements (errors), as can be seen from the distribution of the errors.



(a)



(b)

Figure 14: (a) Connection between experimental and predicted FS in the FS-RFR model and (b) scattering of errors and predicted FS.

3.2.3 FS-RFR model

For the purpose of approximating the FS of SCA-GPC, the RFR model was utilized, as shown in Figure 14. It is possible to notice a graphical representation of the degree of agreement

that exists between the anticipated and observed FS in Figure 14(a). The FS values predicted by the RFR model and those obtained experimentally were very similar. An R^2 value of 0.9753, coupled with all predictions falling within the 20% threshold, highlights significantly improved accuracy in determining the FS of SCA-GPC through the RFR approach. In Figure 14(b), the distribution of errors or discrepancies between the experimental and predicted values using the RFR approach is illustrated. On average, the incorrect readings were around 0.090 MPa, with a standard deviation ranging from 0.002 to 0.936 MPa. Additionally, it was observed that 77 of the values were below 0.3 MPa, there were no values falling between 0.3 and 0.5 MPa, and only 1 value exceeded 0.5 MPa. By looking at the distribution of the errors, it is evident that the prediction of FS of SCA-GPC utilizing the RFR model was expressively more accurate than both the SVM and BR models, with significantly lesser spread measurements (errors).

3.3 Validation of models

Table 2 displays the results of Eqs. (1)–(3) applied to the CS and CS-approximation models in terms of the computed errors (MAE), root-mean-square error (RMSE), and mean absolute percentage error (MAPE). The MAEs for CS predictions using SVM, BR, and RFR were 3.420, 2.190, and 1.320 MPa, respectively. SVM, BR, and RFR all improved performance by an average of 11.10%, 6.90%, and 3.90%, respectively, according to the MAPE metric. Moreover, the RMSE values were calculated as 4.194 MPa for SVM, 2.907 MPa for BR, and 2.211 MPa for RFR. Similar trends were observed in the prediction models for flexural strength (FS) regarding MAE, RMSE, and MAPE, as seen in the CS prediction models. These findings indicate that compared to SVM and BR models, the RFR method offers superior accuracy. Table 3 displays the outcomes of computing R^2 , RMSE, and MAE to validate the K -fold approach, while Figure 15 illustrates the K -fold assessments of various ML techniques for predicting CS and FS. The SVM approach

Table 2: Assessment of errors through statistical methods

ML technique	CS			FS		
	MAPE (%)	RMSE (MPa)	MAE (MPa)	MAPE (%)	RMSE (MPa)	MAE (MPa)
SVM	11.10	4.194	3.420	5.30	0.321	0.277
BR	6.90	2.907	2.190	4.00	0.256	0.205
RFR	3.90	2.211	1.320	1.80	0.148	0.091

Table 3: Accuracy metrics (RMSE, R^2 , and MAE) obtained from k -fold analysis

Property	ML model	Parameters	K-fold									
			1	2	3	4	5	6	7	8	9	10
CS	SVM	MAE	2.43	3.99	3.93	3.12	3.65	5.92	3.78	2.72	4.69	3.11
		RMSE	3.09	5.34	5.81	4.36	4.66	6.09	4.73	3.26	5.38	4.34
		R^2	0.88	0.60	0.69	0.53	0.81	0.74	0.86	0.87	0.72	0.84
	BR	MAE	2.00	3.90	4.76	3.37	4.25	5.05	3.24	2.87	4.61	2.55
		RMSE	3.05	5.16	6.16	4.22	5.21	5.83	4.55	3.16	5.35	4.00
		R^2	0.94	0.61	0.62	0.55	0.78	0.74	0.89	0.92	0.69	0.92
	RFR	MAE	2.87	1.22	1.19	0.97	1.84	4.43	1.58	2.80	1.68	1.25
		RMSE	3.33	1.84	1.31	1.55	2.38	4.49	2.04	3.27	2.10	1.64
		R^2	0.93	0.95	0.98	0.98	0.94	0.85	0.97	0.91	0.96	0.98
FS	SVM	MAE	0.21	0.25	0.24	0.18	0.28	0.35	0.19	0.22	0.22	0.18
		RMSE	0.24	0.32	0.25	0.31	0.32	0.42	0.32	0.27	0.27	0.21
		R^2	0.85	0.84	0.89	0.89	0.87	0.75	0.85	0.83	0.89	0.86
	BR	MAE	0.18	0.14	0.20	0.21	0.32	0.31	0.21	0.22	0.22	0.23
		RMSE	0.14	0.31	0.30	0.24	0.35	0.38	0.26	0.16	0.28	0.27
		R^2	0.93	0.86	0.85	0.87	0.85	0.79	0.93	0.93	0.88	0.93
	RFR	MAE	0.19	0.19	0.13	0.17	0.13	0.26	0.15	0.24	0.18	0.09
		RMSE	0.24	0.23	0.17	0.21	0.14	0.31	0.20	0.34	0.19	0.09
		R^2	0.96	0.92	0.95	0.86	0.96	0.85	0.97	0.82	0.95	0.97

produced CS estimates with MAE values ranging from 2.43 to 5.92 MPa, with an average of 3.73 MPa. The CS-BR model showed an MAE of 3.66 MPa, with a confidence interval from 2.00 to 5.05 MPa. Additionally, RFR had a mean MAE of 1.98 MPa, ranging from 0.97 to 4.43 MPa (minimum to maximum). The mean RMSE was also 4.71 MPa for the CS-SVM method, 4.67 MPa for the CS-BR method, and 2.39 MPa for the CS-RFR method. The highest R^2 values for CS-based SVM, BR, and RFR are, nevertheless, 0.88, 0.94, and 0.98. The FS prediction K -fold analysis showed a significant decrease in MAE and RMSE values from SVM to BR to RFR, with a slight increase in R^2 across the same models. The top-performing RFR model exhibited the highest R^2 and the lowest error rate for predicting both CS and FS of SCA-GPC. Based on K -fold R^2 measurements and error analysis, the RFR model proved to be the most accurate. However, both SVM and BR models also demonstrated strong performance. Consequently, SVM, BR, and RFR models may provide a more precise means of assessing the CS and FS of SCA-GPC.

3.4 Interaction of input parameters

This section analyses the relationship between the input variables and the final product, CS. Figure 16 illustrate the scatter plots comparing the CS of SCA-GPC with different inputs. The scatter plots are accompanied by bar graphs depicting the frequencies of the input and output components. The GGBFS effect and interaction are illustrated in

Figure 16(a), which clearly demonstrates that the mechanical properties of concrete were directly impacted by both inputs. This indicates that the SCA-GPC's strength was linearly proportional to the GGBFS content. The increased silica content in the GGBFS employed in different research may explain the higher quantities of GGBFS [68]. Figure 16(b) shows that the relationship between SCA-mechanical GPC's and CCA's characteristics was indirect. As the CCA concentration increased, the mechanical characteristics of SCA-GPC gradually degraded. Up to $800 \text{ kg}\cdot\text{m}^{-3}$, the strength of FA-input rose as the FA content grew; after that, the strength decreased significantly. Subsequently, when the FA concentration surpassed $850 \text{ kg}\cdot\text{m}^{-3}$, the SCA-GPC once again gained strength. Figure 16(c) shows the relationship between the FA content and the mechanical characteristics of SCA-GPC.

Figure 16(d)–(j) illustrate that factors such as CA, W, SHP, SSG, CD, MC, and CG have a minimal impact on concrete strength due to the low variability in the content of these input factors. The outcomes of the interaction analysis were notably influenced by both the raw material utilized and the size of the data sample under examination. Adjusting the input parameters and sampling frequency yields different results. It is important to note that the inputs and database size used to run the algorithms determined the aforementioned results. Using different databases and input factors can result in different outcomes. Further research is needed to enhance understanding of the relationship between the material's components.

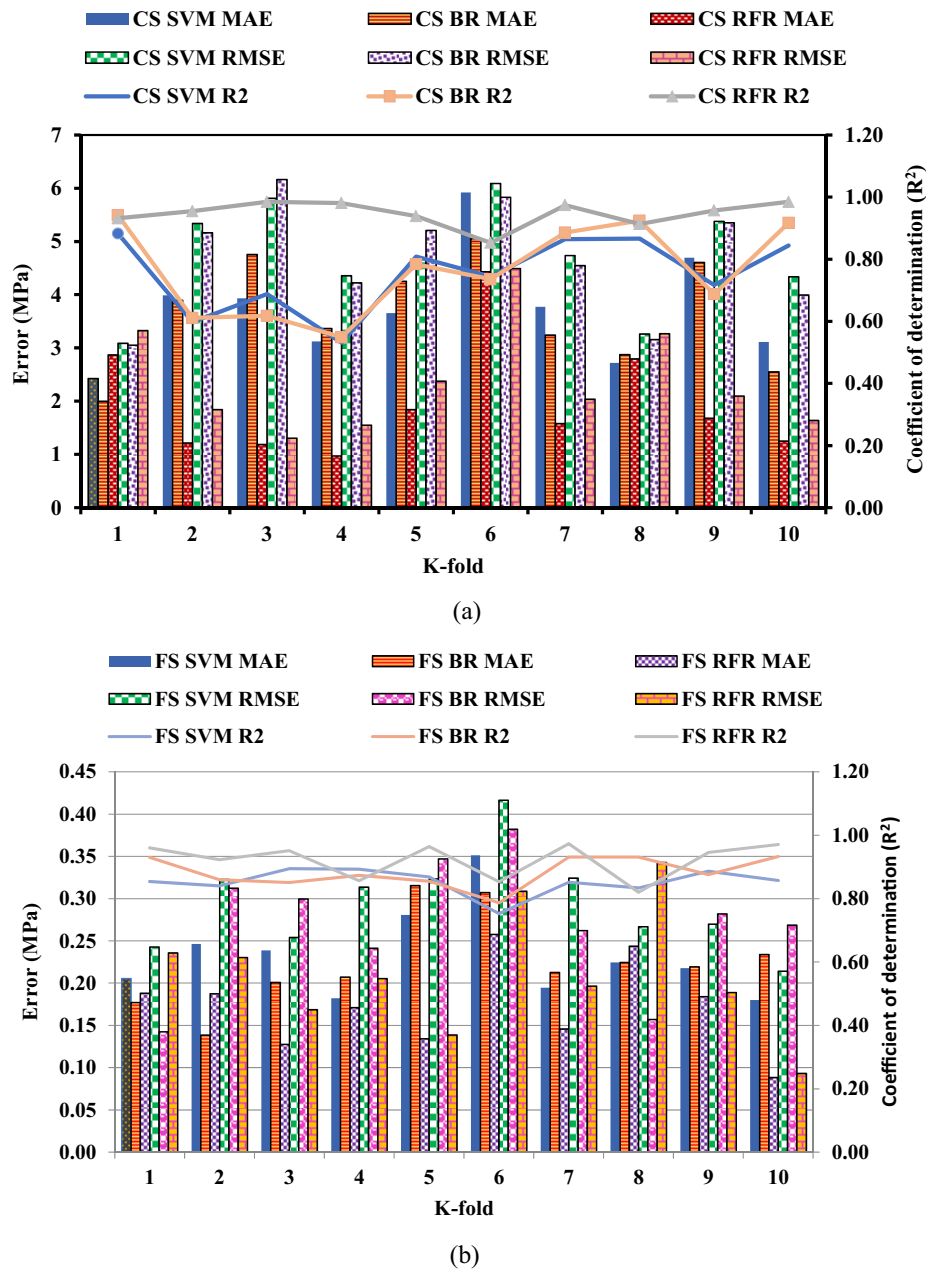


Figure 15: The K-fold analysis findings: (a) CS and (b) FS.

4 Discussions

In this study, the ML models are used to ensure that the predictions are specifically suited to GPC. This situation arises due to the limitation of these models to accept values from a constrained set of eight input variables. Given that all models utilize the same unit measurements and testing technique, it is feasible to depend on the CS and FS predictions generated by any of the models. If there are more than eight parameters in the composite analysis, it is possible that the projected models will not function properly.

If the data used to train these models differ significantly dissimilar to what they are intended to achieve, it is possible that they will not perform as predicted. It is dependent on the degree of consistency or variation in the units of the input parameters as to how accurately the models anticipate the results. For the models to correctly function, it is essential to keep the unit sizes consistent. ML models offer diverse applications within the construction industry, encompassing tasks like material strength forecasting, quality assurance, risk assessment, predictive maintenance, and improving energy efficiency. Nonetheless, these models

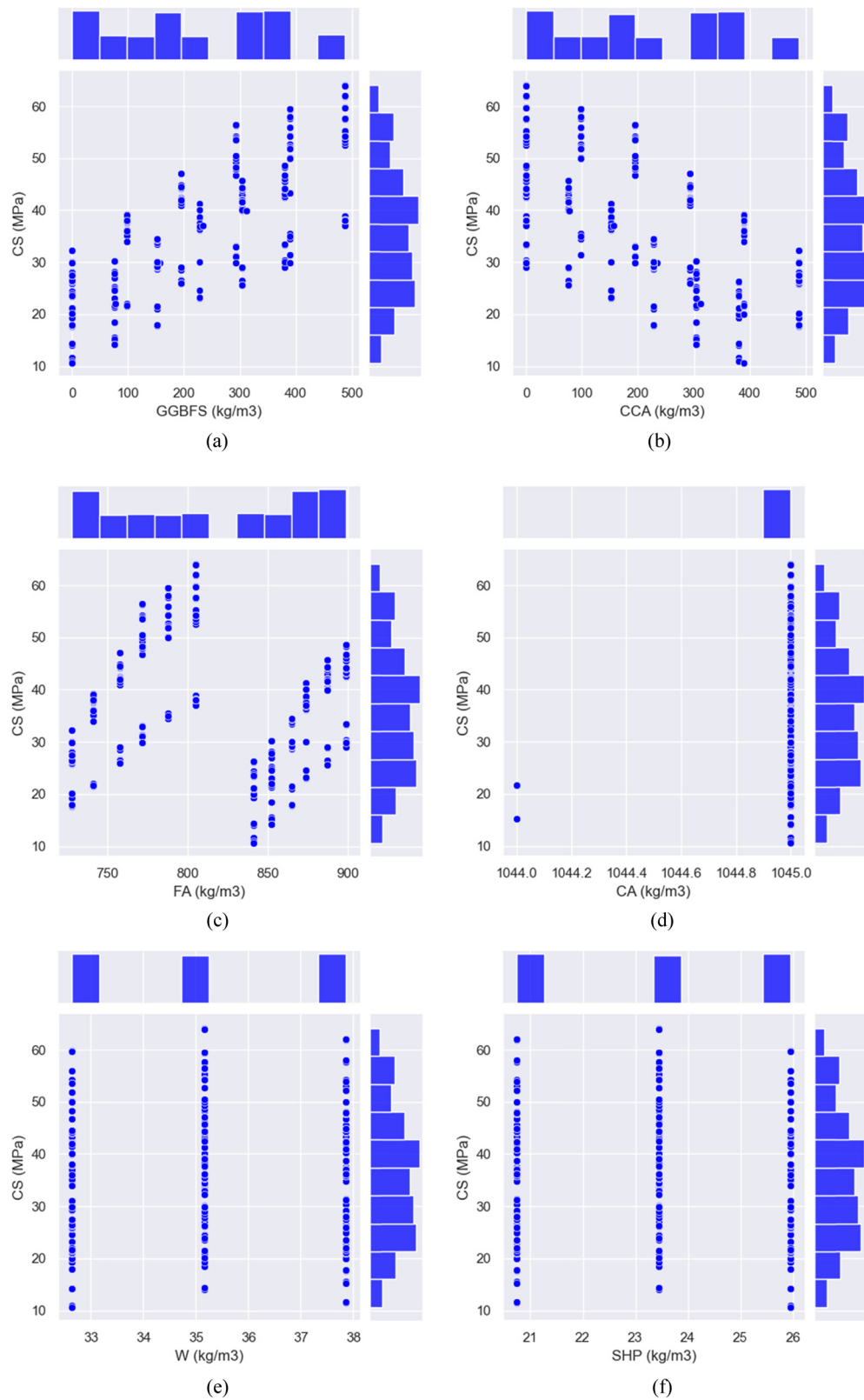


Figure 16: Interaction of input parameters with CS: (a) GGBFS, (b) CCA, (c) FA, (d) CA, (e) W, (f) SHP, (g) SHP, (h) CD, (i) MC, and (j) CG.

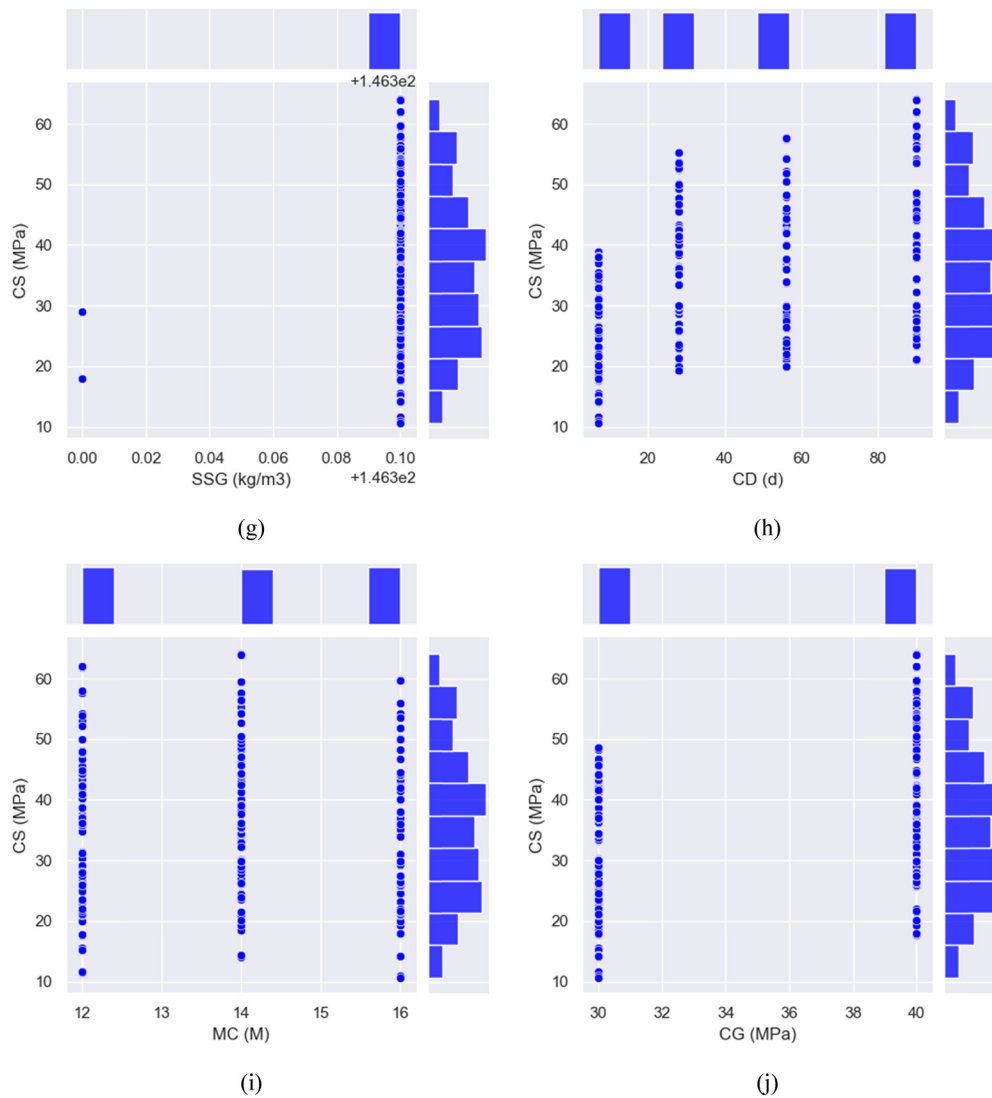


Figure 16: (Continued)

encounter challenges such as reliance on human input, utilization of potentially inaccurate data, and occasional errors in predictions. To overcome these hurdles and optimize ML-driven outcomes, future research avenues could include integrating Internet of Things (IoT) devices, developing hybrid models, adopting explainable AI methodologies, incorporating sustainability considerations, and tailoring data generation and dissemination processes for specific industrial sectors. These advancements in technology have the potential to yield significant advantages for the construction field, facilitating higher levels of efficiency, comprehension, accountability, and well-informed decision-making alongside enhanced safety and project efficacy. The findings of this study could also promote more environmentally responsible building practices in the construction industry, potentially increasing the adoption of GPC.

5 Conclusions

Using three different ML models, including SVM, BR, and RFR, the purpose of this work was to make a prediction about the mechanical properties of GPC (SCA-GPC) that was made up of slag and CCA. For the purpose of training and verifying the models that were produced, 260 different sets of data pertaining to mechanical characteristics were utilized. These sets included CS and FS. The following are some of the most significant findings that emerged from the research:

- The study's conclusion indicates that RFR models exhibited the highest accuracy in predicting the CS and FS of SCA-GPC among the models assessed. The R^2 values for the three ML models (SVM, BR, and RFR) that were created for SCA-GPC's CS and FS prediction were all greater than 0.85.

- Models were assessed for efficacy using statistical measures (MAE, RMSE, and MAPE). A more accurate ML model was represented by a lower error value. The lower error rates supported statements that RFR models accurately predicted SCA-GPC's CS and FS.
- K -fold analysis (MAE, RMSE, and R^2) also validated the RFR model's exceptional precision as paralleled to the commendable precision of SVM and BR models.
- The input/output interaction analysis revealed that the most important input parameters that had a stronger correlation with the CS and FS of SCA-GPC were FA, CCA, and GGBFS.

The methodology detailed in this article allows scientists and engineers to effectively assess, enhance, and validate GPC mixture proportioning. Nevertheless, additional research is needed to assemble a broader dataset encompassing a diverse range of strength grades to facilitate the development of prediction models.

Acknowledgments: The authors would like to express their deepest gratitude to the Scientific Research Deanship and to the College of Engineering at the University of Ha'il – Saudi Arabia for providing the necessary support for conducting this research through project number RG-23 062.

Funding information: This research has been funded by the Scientific Research Deanship at the University of Ha'il – Saudi Arabia, through project number RG-23 062.

Author contribution: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: The authors state no conflict of interest.

Data availability statement: The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

References

- [1] Ghosh, A. and G. D. Ransinchung. Application of machine learning algorithm to assess the efficacy of varying industrial wastes and curing methods on strength development of geopolymer concrete. *Construction and Building Materials*, Vol. 341, 2022, id. 127828.
- [2] Belaïd, F. How does concrete and cement industry transformation contribute to mitigating climate change challenges? *Resources, Conservation & Recycling Advances*, Vol. 15, 2022, id. 200084.
- [3] Andrew, R. M. Global CO₂ emissions from cement production, 1928–2018. *Earth System Science Data*, Vol. 11, 2019, pp. 1675–1710.
- [4] Ahmad, W., A. Ahmad, K. A. Ostrowski, F. Aslam, and P. Joyklad. A scientometric review of waste material utilization in concrete for sustainable construction. *Case Studies in Construction Materials*, Vol. 15, 2021, id. e00683.
- [5] Elmagarhe, A., Q. Lu, M. Alharthai, M. Alamri, and A. Elnihum. Performance of porous asphalt mixtures containing recycled concrete aggregate and fly ash. *Materials*, Vol. 15, 2022, id. 6363.
- [6] Schaubroeck, T., T. Gibon, E. Igos, and E. Benetto. Sustainability assessment of circular economy over time: Modelling of finite and variable loops & impact distribution among related products. *Resources, Conservation and Recycling*, Vol. 168, 2021, id. 105319.
- [7] Shaaban, I. G., J. P. Rizzuto, A. El-Nemr, L. Bohan, H. Ahmed, and H. Tindyebe. Mechanical properties and air permeability of concrete containing waste tires extracts. *Journal of materials in civil engineering*, Vol. 33, 2021, id. 04020472.
- [8] Nurrudin, M. F., H. Sani, B. S. Mohammed, and I. Shaaban. Methods of curing geopolymer concrete: A review. *International Journal of Advanced and Applied Sciences*, Vol. 5, 2018, pp. 31–36.
- [9] Saif, M. S., A. S. Shanour, G. E. Abdelaziz, H. I. Elsayad, I. G. Shaaban, B. A. Tayeh, et al. Influence of blended powders on properties of ultra-high strength fibre reinforced self compacting concrete subjected to elevated temperatures. *Case Studies in Construction Materials*, Vol. 18, 2023, id. e01793.
- [10] Oyeibisi, S., A. Ede, F. Olutoge, and D. Omole. Geopolymer concrete incorporating agro-industrial wastes: Effects on mechanical properties, microstructural behaviour and mineralogical phases. *Construction and Building Materials*, Vol. 256, 2020, id. 119390.
- [11] Oyeibisi, S., A. Ede, F. Olutoge, and B. Ngene. Assessment of activity indexes on the splitting tensile strengthening of geopolymer concrete incorporating supplementary cementitious materials. *Materials Today Communications*, Vol. 24, 2020, id. 101356.
- [12] Davidovits, J. Geopolymers: inorganic polymeric new materials. *Journal of Thermal Analysis and Calorimetry*, Vol. 37, 1991, pp. 1633–1656.
- [13] Pazouki, G. Fly ash-based geopolymer concrete's compressive strength estimation by applying artificial intelligence methods. *Measurement*, Vol. 203, 2022, id. 111916.
- [14] He, J., Y. Jie, J. Zhang, Y. Yu, and G. Zhang. Synthesis and characterization of red mud and rice husk ash-based geopolymer composites. *Cement and Concrete Composites*, Vol. 37, 2013, pp. 108–118.
- [15] Peng, Y. and C. Unluer. Analyzing the mechanical performance of fly ash-based geopolymer concrete with different machine learning techniques. *Construction and Building Materials*, Vol. 316, 2022, id. 125785.
- [16] Shahmansouri, A. A., M. Yazdani, S. Ghanbari, H. A. Bengar, A. Jafari, and H. F. Ghatte. Artificial neural network model to predict the compressive strength of eco-friendly geopolymer concrete incorporating silica fume and natural zeolite. *Journal of Cleaner Production*, Vol. 279, 2021, id. 123697.
- [17] Zhang, C., Z. Zhu, F. Liu, Y. Yang, Y. Wan, W. Huo, et al. Efficient machine learning method for evaluating compressive strength of cement stabilized soft soil. *Construction and Building Materials*, Vol. 392, 2023, id. 131887.
- [18] Singh, B., G. Ishwarya, M. Gupta, and S. K. Bhattacharyya. Geopolymer concrete: A review of some recent developments. *Construction and Building Materials*, Vol. 85, 2015, pp. 78–90.
- [19] Lenka, B. P., R. K. Majhi, S. Singh, and A. N. Nayak. Eco-friendly and cost-effective concrete utilizing high-volume blast furnace slag and

- demolition waste with lime. *European Journal of Environmental and Civil Engineering*, Vol. 26, 2022, pp. 5351–5373.
- [20] Majhi, R. K. and A. N. Nayak. Production of sustainable concrete utilising high-volume blast furnace slag and recycled aggregate with lime activator. *Journal of Cleaner Production*, Vol. 255, 2020, id. 120188.
- [21] Revilla-Cuesta, V., V. Ortega-López, M. Skaf, and J. M. Manso. Deformational behavior of self-compacting concrete containing recycled aggregate, slag cement and green powders under compression and bending: Description and prediction adjustment. *Journal of Building Engineering*, Vol. 54, 2022, id. 104611.
- [22] Ortega-López, V., F. Faleschini, C. Pellegrino, V. Revilla-Cuesta, and J. M. Manso. Validation of slag-binder fiber-reinforced self-compacting concrete with slag aggregate under field conditions: Durability and real strength development. *Construction and Building Materials*, Vol. 320, 2022, id. 126280.
- [23] Majhi, R. K., A. N. Nayak, and B. B. Mukharjee. Characterization of lime activated recycled aggregate concrete with high-volume ground granulated blast furnace slag. *Construction and Building Materials*, Vol. 259, 2020, id. 119882.
- [24] Zakka, W. P., N. H. A. S. Lim, and M. C. Khun. A scientometric review of geopolymer concrete. *Journal of Cleaner Production*, Vol. 280, 2021, id. 124353.
- [25] Farooq, F., X. Jin, M. F. Javed, A. Akbar, M. I. Shah, F. Aslam, et al. Geopolymer concrete as sustainable material: A state of the art review. *Construction and Building Materials*, Vol. 306, 2021, id. 124762.
- [26] Ahmed, H. U., A. A. Mohammed, and A. S. Mohammed. Effectiveness of silicon dioxide nanoparticles (Nano SiO₂) on the internal structures, electrical conductivity, and elevated temperature behaviors of geopolymer concrete composites. *Journal of Inorganic and Organometallic Polymers and Materials*, Vol. 33, 2023, pp. 1–21.
- [27] Ahmed, H. U., A. S. Mohammed, and A. A. Mohammed. Engineering properties of geopolymer concrete composites incorporated recycled plastic aggregates modified with nano-silica. *Journal of Building Engineering*, Vol. 75, 2023, id. 106942.
- [28] Ahmed, H. U., A. A. Mohammed, and A. S. Mohammed. Effectiveness of nano-SiO₂ on the mechanical, durability, and microstructural behavior of geopolymer concrete at different curing ages. *Archives of Civil and Mechanical Engineering*, Vol. 23, 2023, pp. 1–28.
- [29] Yang, H., L. Liu, W. Yang, H. Liu, W. Ahmad, A. Ahmad, et al. A comprehensive overview of geopolymer composites: A bibliometric analysis and literature review. *Case Studies in Construction Materials*, Vol. 16, 2022, id. e00830.
- [30] Avci, O., O. Abdeljaber, S. Kiranyaz, M. Hussein, M. Gabbouj, and D. J. Inman. A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications. *Mechanical systems and signal processing*, Vol. 147, 2021, id. 107077.
- [31] Ghahramani, Z. Probabilistic machine learning and artificial intelligence. *Nature*, Vol. 521, 2015, pp. 452–459.
- [32] Stocker, T. *Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change*, Cambridge university Press, Geneva, Switzerland, 2014.
- [33] Dietterich, T. G. Ensemble methods in machine learning, In: International workshop on multiple classifier systems, Springer Berlin Heidelberg, Berlin, Heidelberg, 2000, pp. 1–15.
- [34] Haenlein, M. and A. Kaplan. A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, Vol. 61, 2019, pp. 5–14.
- [35] Marani, A. and M. L. Nehdi. Machine learning prediction of compressive strength for phase change materials integrated cementitious composites. *Construction and Building Materials*, Vol. 265, 2020, id. 120286.
- [36] Ahmad, A., W. Ahmad, F. Aslam, and P. Joyklad. Compressive strength prediction of fly ash-based geopolymer concrete via advanced machine learning techniques. *Case Studies in Construction Materials*, Vol. 16, 2022, id. e00840.
- [37] Timur Cihan, M. Prediction of concrete compressive strength and slump by machine learning methods. *Advances in Civil Engineering*, Vol. 2019, 2019, pp. 1–11.
- [38] Raza, A., Khan, Q. U. Z., and Ahmad. A. Prediction of axial compressive strength for FRP-confined concrete compression members. *KSCE Journal of Civil Engineering*, Vol. 24, 2020, pp. 2099–2109.
- [39] Mansour, M. Y., M. Dicleli, J.-Y. Lee, and J. Zhang. Predicting the shear strength of reinforced concrete beams using artificial neural networks. *Engineering Structures*, Vol. 26, 2004, pp. 781–799.
- [40] Tamimi, A. K., J. A. Abdalla, and Z. I. Sakka. Prediction of long term chloride diffusion of concrete in harsh environment. *Construction and Building Materials*, Vol. 22, 2008, pp. 829–836.
- [41] Nazar, S., J. Yang, A. Ahmad, and S. F. A. Shah. Comparative study of evolutionary artificial intelligence approaches to predict the rheological properties of fresh concrete. *Materials Today Communications*, Vol. 32, 2022, id. 103964.
- [42] Song, H., A. Ahmad, F. Farooq, K. A. Ostrowski, M. Maślak, S. Czarnecki, et al. Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms. *Construction and Building Materials*, Vol. 308, 2021, id. 125021.
- [43] Moein, M. M., A. Saradar, K. Rahmati, S. H. G. Mousavinejad, J. Bristow, V. Aramali, et al. Predictive models for concrete properties using machine learning and deep learning approaches: A review. *Journal of Building Engineering*, Vol. 63, 2022, id. 105444.
- [44] Oyeibisi, S. and T. Alomayri. Artificial intelligence-based prediction of strengths of slag-ash-based geopolymer concrete using deep neural networks. *Construction and Building Materials*, Vol. 400, 2023, id. 132606.
- [45] Puth, M.-T., M. Neuhäuser, and G. D. Ruxton. Effective use of Pearson's product-moment correlation coefficient. *Animal behaviour*, Vol. 93, 2014, pp. 183–189.
- [46] Pearson, K. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, Vol. 50, 1900, pp. 157–175.
- [47] Gravier, J., V. Vignal, S. Bissey-Breton, and J. Farre. The use of linear regression methods and Pearson's correlation matrix to identify mechanical-physical-chemical parameters controlling the micro-electrochemical behaviour of machined copper. *Corrosion Science*, Vol. 50, 2008, pp. 2885–2894.
- [48] Zhou, J., Q. Tian, S. Nazar, and J. Huang. Hyper-tuning gene expression programming to develop interpretable prediction models for the strength of corncob ash-modified geopolymer concrete. *Materials Today Communications*, Vol. 38, 2024, id. 107885.
- [49] Wang, J., Q. Qu, S. A. Khan, B. S. Alotaibi, F. Althoey, Y. Gamil, et al. Experimenting the influence of corncob ash on the mechanical strength of slag-based geopolymer concrete. *Reviews on Advanced Materials Science*, Vol. 63, 2024, id. 20230187.

- [50] Zhou, J., Q. Tian, A. Ahmad, and J. Huang. Compressive and tensile strength estimation of sustainable geopolymers using contemporary boosting ensemble techniques. *Reviews on Advanced Materials Science*, Vol. 63, 2024, id. 20240014.
- [51] Lee, B. C. and D. M. Brooks. Accurate and efficient regression modeling for microarchitectural performance and power prediction. *ACM SIGOPS Operating Systems Review*, Vol. 40, 2006, pp. 185–194.
- [52] Ling, H., C. Qian, W. Kang, C. Liang, and H. Chen. Combination of support vector machine and *K*-Fold cross validation to predict compressive strength of concrete in marine environment. *Construction and Building Materials*, Vol. 206, 2019, pp. 355–363.
- [53] Suthaharan, S. Support vector machine. *Machine learning models and algorithms for big data classification: thinking with examples for effective learning*, Springer US, Boston, MA, 2016, pp. 207–235.
- [54] Huang, J., Y. Sun, and J. Zhang. Reduction of computational error by optimizing SVR kernel coefficients to simulate concrete compressive strength through the use of a human learning optimization algorithm. *Engineering with Computers*, Vol. 38, 2021, pp. 1–18.
- [55] Khan, K., W. Ahmad, M. N. Amin, A. Ahmad, S. Nazar, A. A. Alabdullah, et al. Exploring the use of waste marble powder in concrete and predicting its strength with different advanced algorithms. *Materials*, Vol. 15, 2022, id. 4108.
- [56] Han, Q., C. Gui, J. Xu, and G. Lacidogna. A generalized method to predict the compressive strength of high-performance concrete by improved random forest algorithm. *Construction and Building Materials*, Vol. 226, 2019, pp. 734–742.
- [57] Grömping, U. Variable importance assessment in regression: linear regression versus random forest. *The American Statistician*, Vol. 63, 2009, pp. 308–319.
- [58] Guo, K., X. Wan, L. Liu, Z. Gao, and M. Yang. Fault diagnosis of intelligent production line based on digital twin and improved random forest. *Applied Sciences*, Vol. 11, 2021, id. 7733.
- [59] Ahmad, A., K. Chaayasarn, F. Farooq, W. Ahmad, S. Suparp, and F. Aslam. Compressive strength prediction via gene expression programming (GEP) and artificial neural network (ANN) for concrete containing RCA. *Buildings*, Vol. 11, 2021, id. 324.
- [60] Farooq, F., W. Ahmed, A. Akbar, F. Aslam, and R. Alyousef. Predictive modeling for sustainable high-performance concrete from industrial wastes: A comparison and optimization of models using ensemble learners. *Journal of Cleaner Production*, Vol. 292, 2021, id. 126032.
- [61] Aslam, F., F. Farooq, M. N. Amin, K. Khan, A. Waheed, A. Akbar, et al. Applications of gene expression programming for estimating compressive strength of high-strength concrete. *Advances in Civil Engineering*, Vol. 2020, 2020, pp. 1–23.
- [62] Perkel, J. M. Why Jupyter is data scientists' computational notebook of choice. *Nature*, Vol. 563, 2018, pp. 145–147.
- [63] Ono, J. P., J. Freire, and C. T. Silva. Interactive data visualization in jupyter notebooks. *Computing in Science & Engineering*, Vol. 23, 2021, pp. 99–106.
- [64] Nongthombam, K. and D. Sharma. Data analysis using python. *International Journal of Engineering Research & Technology (IJERT)*, Vol. 10, 2021, pp. 463–468.
- [65] Kiusalaas, J. *Numerical methods in engineering with Python 3*, Cambridge university Press, Cambridge, United Kingdom, 2013.
- [66] Wang, N., M. Samavatian, V. Samavatian, and H. Sun. Bayesian machine learning-aided approach bridges between dynamic elasticity and compressive strength in the cement-based mortars. *Materials Today Communications*, Vol. 35, 2023, id. 106283.
- [67] Husnain, A., M. Iqbal, M. Ashraf, M. F. Javed, H. Alabduljabbar, and D. S. Abd Elminaam. Machine learning approaches for predicting shielding effectiveness of carbon fiber-reinforced mortars. *Case Studies in Construction Materials*, Vol. 20, 2024, id. e03189.
- [68] Hussain, F., I. Kaur, A. Hussain. Reviewing the influence of GGBFS on concrete properties. *Materials Today: Proceedings*, Vol. 32, 2020, pp. 997–1004.