Research Article

Compressive Strength of Fly-Ash-Based Geopolymer Concrete by Gene Expression Programming and Random Forest

Mohsin Ali Khan [ID],[1] **Shazim Ali Memon** [ID],[2] **Furqan Farooq** [ID],[3] **Muhammad Faisal Javed** [ID],[3] **Fahid Aslam** [ID],[4] **and Rayed Alyousef** [ID][4]

[1]*Department of Structural Engineering, Military College of Engineering (MCE),*
 National University of Science and Technology (NUST), Islamabad 44000, Pakistan
[2]*Department of Civil and Environmental Engineering, School of Engineering and Digital Sciences, Nazarbayev University,*
 Nur-Sultan 010000, Kazakhstan
[3]*Department of Civil Engineering, COMSATS University Islamabad, Abbottabad Campus, Abbottabad 22060, Pakistan*
[4]*Department of Civil Engineering, College of Engineering in Al-Kharj, Prince Sattam bin Abdulaziz University,*
 Al-Kharj 11942, Saudi Arabia

Correspondence should be addressed to Mohsin Ali Khan; moak.pg18mce@student.nust.edu.pk and Shazim Ali Memon; shazim.memon@nu.edu.kz

Fly ash (FA) is a residual from thermal industries that has been effectively utilized in the production of FA-based geopolymer concrete (FGPC). To avoid time-consuming and costly experimental procedures, soft computing techniques, namely, random forest regression (RFR) and gene expression programming (GEP), are used in this study to develop an empirical model for the prediction of compressive strength of FGPC. A widespread, reliable, and consistent database of compressive strength of FGPC is set up via a comprehensive literature review. The database consists of 298 compressive strength data points. The influential parameters that are considered as input variables for modelling are curing temperature (T), curing time (t), age of the specimen (A), the molarity of NaOH solution (M), percent SiO$_2$ solids to water ratio (% S/W) in sodium silicate (Na$_2$SiO$_3$) solution, percent volume of total aggregate (% A_G), fine aggregate to the total aggregate ratio (F/A_G), sodium oxide (Na$_2$O) to water ratio (N/W) in Na$_2$SiO$_3$ solution, alkali or activator to the FA ratio (A_L/F_A), Na$_2$SiO$_3$ to NaOH ratio (N_s/N_o), percent plasticizer (% P), and extra water added as percent FA (E_W%). RFR is an ensemble algorithm and gives outburst performance as compared to GEP. However, GEP proposed an empirical expression that can be used to estimate the compressive strength of FGPC. The accuracy and performance of both models are evaluated via statistical error checks, and external validation is considered. The proposed GEP equation is used for sensitivity analysis and parametric study and then compared with nonlinear and linear regression expressions.

1. Introduction

Fly ash (FA) is considered as waste material resulted from thermal coal production [1]. It is carried by the gases released from the boiler and collected via electrostatic or mechanical separator [2]. The annual production of FA is 375 million tons, and its disposal cost per ton is $20 to $40 [3]. Dumping into landfills without prior treatment causes a malicious effect on the environment [4]. To sustain safe environment, effective management of waste is needed [5]. Fine particles of FA act as poisons when entering the respiratory system. Furthermore, it pollutes underground water, which is harmful to aquatic life and causes diarrhea and skin cancer [6].

Concrete is the second most usable material after water, as three tons of concrete is produced per person [7, 8]. In the world, every year 25 billion tons of concrete is produced that acquires 2.6 billion tons of cement, which will be increased

by 25% in the next ten years [9, 10]. Cement production causes a nasty impact on the atmosphere as one ton of cement emits one ton of CO_2 in the air, which alarms the ecology [11]. Also, limestone is the main source of cement, and its severe shortage may occur after 25–30 years [12, 13]. Therefore, green concrete production is needed to decline its malignant impact on the natural environment [14]. FA is used as supplementary cementitious material to produce green concrete [15]. It is worthy as it reduces the cement utilization and also its harmful effects on the ecology when dumped into landfills.

Since last two decades, the use of FA-based geopolymer concrete (GPC) is rising constantly as it reduces the consumption of cement [16–19]. FA-based GPC has been widely used in construction, but still no empirical model is developed to predict its compressive strength (f_c') on the basis of mix proportion with maximum input parameters. f_c' of FA-dependent GPC varies with different parameters like specimen age (A), curing time (t), initial curing temperature (T), molarity of NaOH solution (M), percent SiO_2 solids to water ratio (% S/W) in the formation of sodium silicate (Na_2SiO_3) solution, ratio of alkali to FA (A_L/F_A), ratio of Na_2SiO_3 to NaOH (N_s/N_o), addition of extra water as percent FA (% E_W), percentage of total aggregate by volume (% A_G), ratio of fine to total aggregate (F/A_G), and percentage of plasticizer (%P) [10, 20–27]. This generates lack of clarity in the prediction of f_c' of FA-dependent GPC. Furthermore, rapid growth of soft computing techniques for the development of empirical equation by using experimental data has been just noticed [28, 29].

Artificial intelligence (AI) techniques have been used widely in the civil engineering field for the prediction of different mechanical properties of concrete. Methods like random forest (RF) [30, 31], support vector machine (SVM) [32], artificial neural networks (ANNs) [33], adaptive neuro fuzzy interface (ANFIS) [34], decision tree (DT) [35], multivariate adaptive regression spline (MARS) [36], genetic programming (GP) [37], and gene programming (GEP) [38] were used vastly by different researchers [39–41]. Recently, ANN was used to accurately predict the mechanical properties of rice husk ash concrete [33] and workability of self-compacting concrete [42]. No empirical equation was provided in these models, which can be used practically, although these models show a strong correlation. This is due to the complex model of ANN structure which limits the wide scale adoption of ANN techniques [43]. The multicollinearity is the main issue in these models [44]. Likewise, an updated ANN technique was used to predict the compressive strength of silica fume concrete and elastic moduli of recycled aggregate concrete. Due to the complications in the proposed relationship, just a graphical interface was developed to make the model functional [45]. The comparative study of ANN and ANFIS was carried out for the prediction of compressive strength of concrete which revealed that ANFIS provides better and strong correlation than ANN [46]. RF is an ensemble machine learning technique which has been effectively used in the prediction of uniaxial compressive strength of rubberized concrete [30]. The RF gives outburst performance in modelling strength of

coal grout material [31]. The adamant results were obtained in the prediction of compressive strength of self-compacting concrete with antenna search-based RF algorithm [47].

Genetic programming (GP), one of the strong soft computing techniques, is worthy as it develops a model without considering the previously established relationships [48, 49]. Recently, GP is extended to gene expression programming (GEP), which uses linear chromosomes of fixed length and encodes a small program [50]. GEP is advantageous as it gives a simple and reliable mathematical equation that can be used practically. In civil engineering, it is used as a substitute for ordinary prediction techniques [39, 51–57]. GEP is employed to predict the influence of the strength class of cement on the compressive strength of mortar, the split tensile strength of concrete, and the fresh and hardened properties of the self-compacting mix [39, 51–57]. Farooq et al. [58] predicted the compressive strength of high-strength concrete using RF, ANN, DT, and GEP, providing coefficient of determination equal to 0.96, 0.89, 0.90, and 0.90, respectively. In RF, weak learners are used as base learners. This bagging mechanism of RF provides obstinate results. GEP leads RF as it is an individual model that provides an empirical relation between input and output parameters that can be used in field calculation.

Compressive strength is the major factor to be considered in the design and analysis of concrete [59]. Intensive research is carried out to find the compressive strength of FA-dependent GPC [60, 61]. For the sustainability of FA and to save cost and time, it is needed to develop a reliable and accurate equation that would relate mix proportion and compressive strength of FA-dependent GPC. The comprehensive revision of literature reveals few empirical equations for the prediction of compressive strength of FA-dependent GPC [39, 54, 57]. However, the prediction from these empirical equations are limited to a specified experimental study and is not practicable and reliable beyond the specified dataset. Alkaroosh et al. [62] established a model to predict the compressive strength of FA-dependent GPC using 56 datasets extracted from a particular research paper [63]. The suggested equation uses no variable to counter the formation of sodium silicate solution. Also, the model illustrates a linear relation in the molarity of NaOH and compressive strength. However, other studies reported an inverse relationship between the compressive strength and molarity of NaOH solution [64]. To cover this lack of correspondence, RF and GEP techniques are used to develop a more accurate, effective, and generalized model that predicts the compressive strength of FA-dependent GPC with acceptable error. A comprehensive and detailed dataset file is established from the literature that includes cube samples of size 150 mm and 100 mm and cylindrical samples having size (200×100) mm (height × diameter). The ample number of data points guarantees the reliability of the model for data points outside the dataset file. The performance of the RF and GEP model is tested through statistical checks, parametric study, and sensitivity analysis and compared with nonlinear and linear regression models.

2. Research Methodology

This section covers the methodology to develop GEP and RF models for the compressive strength (f'_c) of FA-dependent GPC.

2.1. Brief Overview of Gene Expression Programming.
Koza recommended the GP technique as an alternative to genetic algorithm (GA) which uses fixed length binary strings [65]. The use of nonlinear parse tree structure marks the GP as an acceptable technique. It considers the initial nonlinearity of the data. The same nonlinearity has been exercised before [62, 65]. GP is inadequate as it ignores the independent genome. The nonlinear structure of GP works as both the phenotype and genotype. It fails in the development of basic and simple model. To overcome inconsistencies in the GP algorithm, Ferreira suggested its modified version known as GEP technique [65]. It is based on the evolutionary theorem of population. The major change in GEP is the transmission of the genome towards successive generations. Another notable feature is the creation of entities using chromosome which is comprised of different genes [66]. In GEP, each gene originates from terminal set of constants, fixed length parameters, and arithmetic operations used as a function. There is a stabilized and smooth interface between chromosome level and allied functions. Chromosomes record the essential information needed for the establishment of model, and for the deduction of this information, a new language, i.e., Karva, is developed.

The flow diagram of the GEP algorithm is shown in Figure 1. The algorithm begins with the random creation of fixed length chromosomes for each individual. Then, these are similar to the expression trees (ETs). Afterward, the fitness of each individual is evaluated. For many generations, the reiteration begins with different individuals till the development of the finest outcome. For the reiteration of the population, genetic function as mutation, reproduction, and crossover are executed.

2.2. Brief Review of Random Forest Regression.
In 2001, Breiman proposed an improved regression technique known as random forest regression (RFR) [67]. The key characteristics of RFR are flexibility and speediness in the development of the relation between output and input parameters. Also, random forest handles large datasets more effectively than other machine learning algorithms. It has been used in different fields like in banking for the prediction of response of customer [68], prices direction in stock exchange [69], in pharmaceutical and medicine production [70], and so on. It has also been used in various engineering fields like potential mapping of ground water using geoinformatics system- (GIS-) based data [71], compressive strength prediction of high-performance concrete [35], self-compacting light-weight concrete [48], high-strength concrete [58], and so on.

The RF technique is comprised of three main steps that include the assembling of trained regression trees via training dataset, calculation of the mean value of single
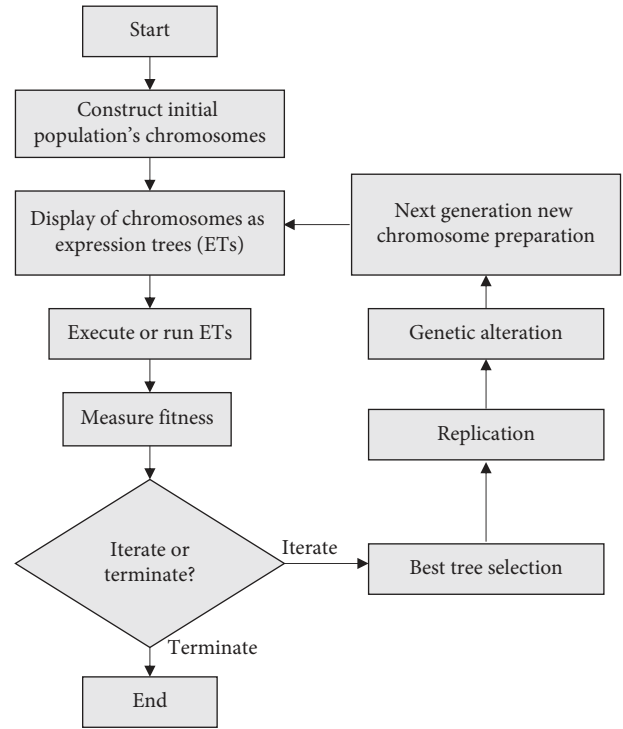

Figure 1: Flow diagram of gene expression programming [29].

regression tree outcome, and the validation of predicted results via validation dataset. The original trained set is used to calculate a new trained dataset comprising of boot-strap data. In this step, some of the data points are removed and swapped with the present data points. The removed data points assembled in other dataset are called out-of-bag data points. Then, the regression function is estimated using 2/3rd of the data points, and the out-of-bag data points are used in validating the model. This process is continued till the achievement of the required accuracy.

RFR is a built-in process that deletes the data points from out-of-bag data points and uses them for validation. This is the distinctive characteristic of RFR. Finally, for each expression tree, the total error is computed showing the efficiency and accuracy of each expression tree.

2.3. Data Collection.
Compressive strength (f'_c) is the key factor to design and analyze concrete. For the sustainability of FA and to save cost and time, it is needed to develop a reliable and accurate model that would relate mix proportion and f'_c of FA-based GPC.

Comprehensive dataset file was compiled from the literature [62, 63, 72–105]. The whole dataset is comprised of 298 experimental results of f'_c of FA-based GPC, which includes 31 and 166 cube samples having 100 mm and 150 mm size respectively and 101 cylindrical samples having size (200×100) mm (height × diameter). f'_c of cylindrical and cube samples is dependent on length to diameter (L/D) ratio [106, 107]. Also, f'_c of 150 mm cubes is 5% lesser than 100 mm cubes. The normalization of cube samples with cylindrical samples is shown in Table 1. The accomplishment

TABLE 1: Collection of data and normalization of compressive strength.

Type of sample	Number of data points	Normalization factor
Cylindrical (200×100) mm)	101	1
Cube $(150$ mm)	166	1×0.8
Cube $(100$ mm)	31	0.95×0.8

of detailed dataset file guarantees the accessibility and reliability of the GEP model to the data which are not utilized for the establishment of the model.

The dataset file contains data of f'_c as a response against input parameters, i.e., samples age (A) in days, initial curing temperature (T) of samples in degree Celsius, molarity of NaOH solution (M), percent SiO_2 solids to water ratio $(\% S/W)$ in the formation of sodium silicate (Na_2SiO_3) solution, ratio of alkali to FA (A_L/F_A), ratio of Na_2SiO_3 to NaOH (N_s/N_o), addition of extra water as percent FA $(\% E_W)$, percentage of total aggregate by volume $(\% A_G)$, ratio of fine to total aggregate (F/A_G), and percentage of plasticizer $(\% P)$. The collected samples are all heat cured for 24-hour duration at various curing temperatures as the increase in f'_c after 24-hour curing time is insignificant [63]. Due to the geo-polymerization, GPC shows higher early strength; therefore, less research is found in the literature for prolonged curing time. Also, Van Jaarsveld et al. [108] reported no increment in f'_c for prolonged curing duration after 24 hours. The distribution of explanatory variables on wide range guarantees the best performance of the model [109]. For all the selected explanatory parameters, the frequency distribution and cumulative percent are shown in Figure 2.

To develop a more generalized model, both cylindrical and cube samples are considered. The range, mean values, and standard deviation of response and explanatory parameter are presented in Table 2. To achieve a reliable prediction of f'_c, it is recommended to use the model within the specified range.

To evaluate the reliability and validity of the data points, several trials were performed. The divergence of data points greater than 20% was excluded in the development of the model and performance evaluation phase. A total of 298 data points were used to establish a reliable model for f'_c of FA-dependent GPC. The data points were randomly divided into two statistically consistent datasets, i.e., training set (30%–90% data points) and a validation set (70%–208% data points) [29]. Training data points are used to train the model, that is, genetic evolution and validation data points are utilized to verify and calibrate the generalization capability of the developed model as described in the literature [56].

2.4. *GEP Model Development.* In the first step, the most effective parameters for compressive strength (f'_c) of FA-dependent GPC were chosen to establish a simplified model. The performance evaluation via multiple initial runs indicates to calculate f'_c of FA based GPC as a function of the following equation.

$$f'_c = f\left(T, A, M, \% \frac{S}{W}, \frac{A_L}{F_A}, \frac{N_S}{N_O} A_G, \frac{F}{A_G}, \% P, \% E_W\right). \quad (1)$$

The number of genes, chromosomes, and expression trees (ETs) are distinguished to develop the GEP expression. The execution time of the program is controlled using population size (number of chromosomes). The combination of genes leads to chromosomes that are used in coding the sub-expression trees (sub-ETs). The complexity of the predictive GEP model reflects to use population size of 150. The configuration of the model in the program relies on the head size, the number of genes that decide the complexity of each term, and the sum of sub-ETs of the model. Hence, the genes and head size which are 3 and 10, respectively, are used for the establishment of the reliable model. The genetic operators are used for the genetic variation of chromosomes. During mutation, the random selection of tail or head of genes is executed and substituted with component of function or terminal sets randomly. The transposition performs the substitution of insertion sequence (IS) and the root insertion sequence (RIS) inside the chromosome. Then, in recombination, chromosomes are combined and divided into two to replace their components. To obtain good algorithm, the suggested setting in the previous study has been exercised [39]. GeneXproTool is used for the execution of the GEP algorithm. Table 3 presents the settings of the parameters used in the execution of the GEP algorithm, to develop a good model.

2.5. *Criteria for Evaluation of Model Performance.* To verify the performance of the developed models, the coefficient of correlation (R) is usually used. Because of its insensitivity to division and multiplication of output to a constant, it cannot be merely utilized for studying the performance of the model [110]. Therefore, root mean squared error (RMSE), relative square error (RSE), mean absolute error (MAE), and relative root mean square error (RRMSE) are also checked. The performance index (ρ) covers the function of both RRMSE and R, so the performance evaluation of the predictive models using ρ is highly recommended [109]. The error checks equations are given as equations (2)–(7).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (exp_i - pred_i)^2}{n}}, \quad (2)$$

$$MAE = \frac{\sum_{i=1}^{n} |exp_i - pred_i|}{n}, \quad (3)$$

$$RSE = \frac{\sum_{i=1}^{n} (exp_i - pred_i)^2}{\sum_{i=1}^{n} (\bar{e} - te_i)^2}, \quad (4)$$

$$RRMSE = \frac{1}{|\bar{e}|} \sqrt{\frac{\sum_{i=1}^{n} (exp_i - pred_i)^2}{n}}, \quad (5)$$
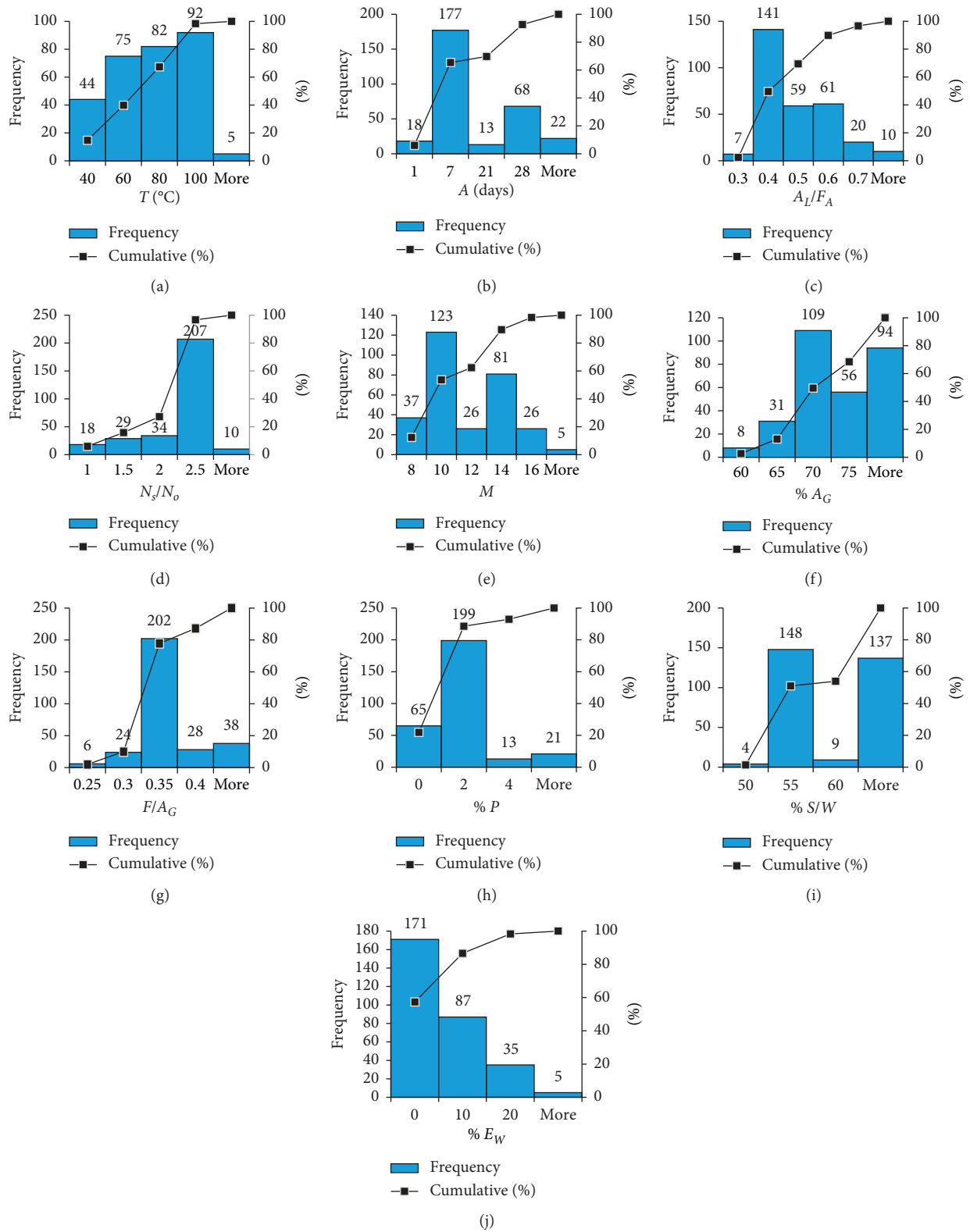
FIGURE 2: Frequency and cumulative percent of selected explanatory variables.

TABLE 2: Range, mean, and standard deviation of response and explanatory variable.

Parameters	Maximum value	Minimum value	Mean value	Standard deviation
Output variables				
T (°C)	120	23	71.57	24.61
A (days)	540	1	20.87	45.73
A/F	0.3	1	0.4545	0.1187
N_S/N_O	4	0.4	2.275	0.5168
M	20	8	11.68	2.6415
(% A_G)	80	60	72	4.753
F/A_G	0.5	0.2	0.3568	0.0493
% P	11.3	0	1.998	2.326
% S/W	81.4	43.4	61.68	10.167
% E_W	35	0	3.889	6.341
Response				
f'_c (MPa)	63	8.2	37	11.154

TABLE 3: The setting of parameters of the GEP algorithm.

Parameters	Adjusted GEP setting
General parameters	
Number of chromosomes	150
Number of genes	3
Head size	10
Linking function	Multiplication
Functions set	$+, -, /, \times, 3$
Arithmetical operators	
Constants/gene	10
Type of data	Floating data
Upper bound value	10
Lower bound value	−10
Gene operator	
Mutation rate	0.001380
Inversion rate	0.005460
IS transportation rate	0.005460
RIS transportation rate	0.005460
Gene recombination rate	0.007550
Gene transportation rate	0.002770

$$R = \frac{\sum_{i=1}^{n} (\exp_i - \overline{\exp_i})(\text{pred}_i - \overline{\text{pred}_i})}{\sqrt{\sum_{i=1}^{n} (\exp_i - \overline{\text{pred}_i})^2 \sum_{i=1}^{n} (\text{pred}_i - \overline{\text{pred}_i})^2}}, \quad (6)$$

$$\rho = \frac{\text{RRMSE}}{1 + R}, \quad (7)$$

where $\exp_i$, pred_i, $\overline{\exp_i}$, and $\overline{\text{pred}_i}$ are the ith experimental outcome, predicted model outcome, experimental average value, and average predicted model outcome, respectively while n indicates the total number of data samples. The higher R value and lower MAE, RMSE, RRMSE, and RSE values replicate the fineness of models. For a strong correlation, the R values should be higher than 0.8 (1 for the ideal model) [111]. Also, the ρ value would be nearly equal to zero.

3. Results and Discussion

3.1. GEP Expression for Compressive Strength of FA-Dependent Geopolymer Concrete.
The expression tree given by the GEP algorithm is shown in Figure 3, which is further

decoded to get an empirical equation for the compressive strength of FA-dependent GPC. The ETs are comprised of five arithmetic operators, i.e., $-, +, /, \times, \sqrt[3]{\cdot}$.

d_o: curing temperature (T) in degree Celsius, d_1: age of the specimen (A), d_2: alkali or activator to the FA ratio (A_L/F_A), d_3: Na$_2$SiO$_3$ to NaOH ratio (N_s/N_o), d_4: molarity of NaOH solution (M), d_5: percent volume of total aggregate (%A_G), d_6: fine aggregate to total aggregate ratio (F/A_G) d_7: percent plasticizer (% P), d_8: percent SiO$_2$ solids to water ratio (%S/W), and d_9: extra water added as percent FA (%E_W).

Equation (8) can be used for the prediction of compressive strength (f'_c) of FA-dependent GPC (MPa). It consists of four variables, i.e., A, B, C, and D, presented as equations (9)–(12), which are extracted from sub-ETs 1, 2, 3, and 4, respectively, as presented in Figure 3.

$$f'_c(\text{MPa}) = A \times B \times C \times D, \quad (8)$$

where

$$A = \sqrt[3]{\% \frac{S}{W}} - \% P + \left(M \times \frac{F}{A_G} \times \frac{A_L}{F_A} \times 6.6 \right) + \% E_W - \% A_G, \quad (9)$$
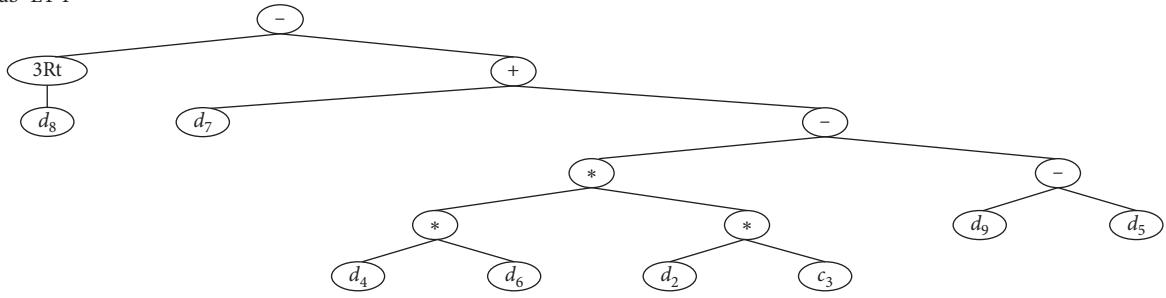
$$B = -\sqrt[3]{\frac{A + 80}{0.08(T - 18)} + \frac{N_S}{N_O} + M}, \quad (10)$$

$$C = \frac{F}{A_G} - \left((M \times \% E_W) - \frac{0.0003}{((N_S/N_O) - \% E_W)) - 0.0003} \right), \quad (11)$$
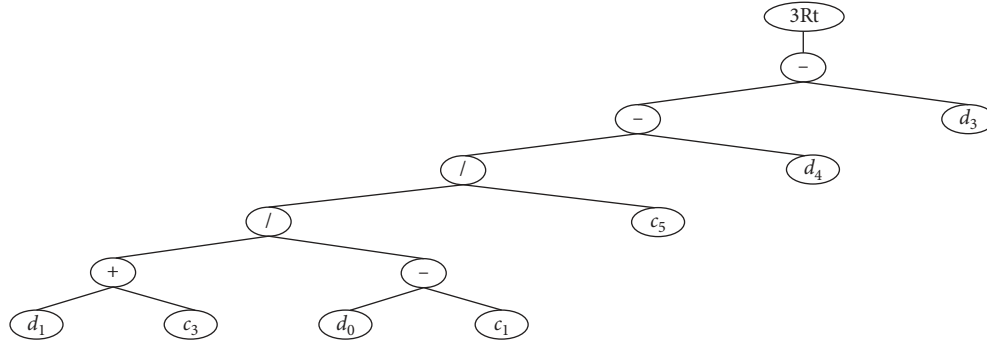
$$D = \sqrt[3]{\frac{1.2(\% P - \% (S/W))}{T}} + \sqrt[3]{\frac{0.2}{((F/A_G))}} + 0.8. \quad (12)$$

3.2. Evaluation of the Performance of the GEP Model.
Figure 4(a) shows the comparison of two regression lines, namely, the GEP model output values and experimental values for both the validation set data and training set data.
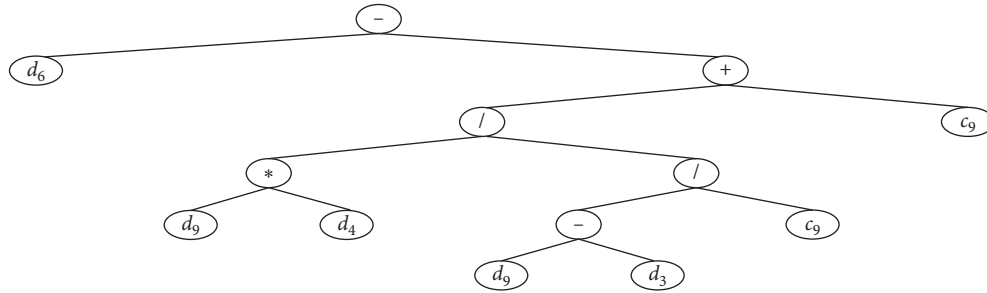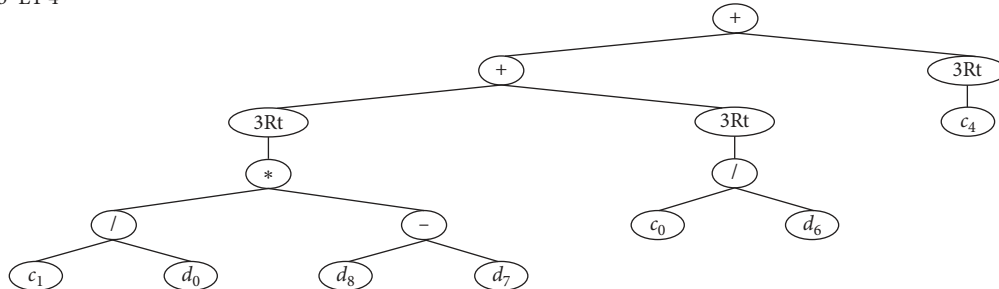
Figure 3: Expression trees (ETs) provided by the GEP algorithm.

The slope of the regression lines shows a strong correlation, i.e., 0.9892 and 1.000 for validation set data and training set data, respectively.

The absolute error between the output of GEP model and experimental values is shown in Figure 4(b). It provides an idea of maximum percent error in the GEP model. The maximum error percentage and mean error percentage are computed as 8.32% and 6.47%, respectively, which approves the similarity between GEP model outcomes and experimental values. Also, the frequency of the maximum error is less. Nearly 90% of GEP model outcomes of the validation dataset have an error of less than 10%, and the average percent error is below 5.56%. This confirms the reliability and generalization capability of the GEP model.
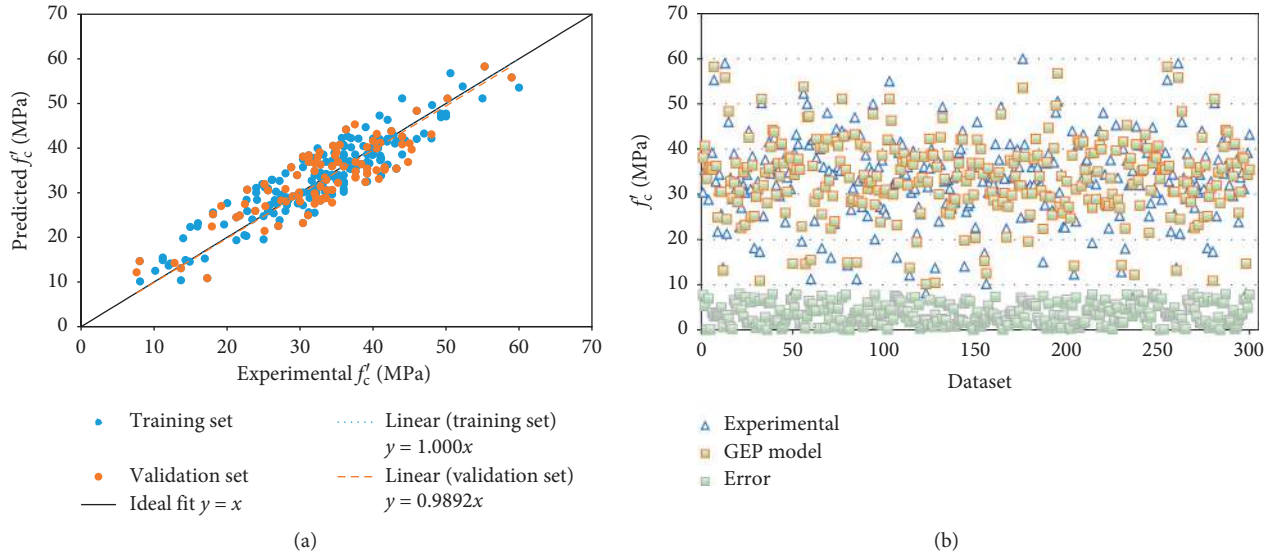
FIGURE 4: Performance evaluation of the GEP model. (a) Comparison between model and experimental outcomes for compressive strength from training and validation set data. (b) Absolute error plot of GEP predicted outcomes and experimental values.

For the reliable and accurate GEP model, the ratio of total data points to the total input variables should be minimum three [109]. This research uses a higher value equal to 30. The statistical checks for both validation data points and training data points are listed in Table 4. For the GEP model, MAE, RMSE, and RSE of training data points are calculated as 5.832, 5.971, and 0.325, respectively, and 2.057, 2.643, and 0.0675 for validation data points. The similarity in the statistical checks guarantees the generalization capability of the GEP model. Table 4 also shows that ρ for both sets reaches zero. So, the presented GEP model could be valid for new data points.

Different statistical measures are also considered for the external validation of the GEP model. The literature recommended that the inclination (slope) of one of the regression lines (k' or k) crossing the origin should be nearly equal to 1 [38]. Table 5 shows that the slope of regression lines is 0.995 and 1.001, which verifies the correlation and correctness. The literature also recommended that the square of correlation coefficient between the experimental and model predictive output ($R_o'^2$) or between model predictive output and experimental values (R_o^2) should come near 1 [112]. Table 5 confirms the validity of the GEP model. So, the proposed GEP model is not just a correlation.

3.3. Evaluation of Random Forest Regression Model.
Random forest regression technique is an ensemble algorithm that utilizes weak learner as a supervised learner and provides a best-performed model based on the coefficient of correlation (R) as shown in Figure 5. This algorithm divides the model into twenty submodels based on different n-estimator and gives model with maximum R. The mean ensemble R is equal to 0.9732 which depicts that all the twenty submodels strongly correlate with the predicted and experimental values. Amongst all these, the submodel with 40 estimators gives outburst

performance with maximum R equal to 0.9826. It is attributable to the use of weak learners as a decision which is used in ensemble algorithm [58].

The relation between the response and the predictor is shown via the slope of regression lines in Figure 6(a). The RF algorithm gives noticeable slope of the regression line as 1.000 and 0.9913 for training set data and validation set data, respectively, which proves the superiority of the RF algorithm.

The absolute error plot between the RF algorithm predicted values and experimental values is presented in Figure 6(b). In comparison with the GEP model, the RF model shows less error as the maximum percent error and average percent error are calculated as 4.89% and 2.14%. The RF algorithm yields outstanding results but does not provide an empirical equation like GEP.

Furthermore, the performance of the RF algorithm-based model is also verified through statistical error checks. Table 4 shows that statistical error checks for RF algorithm-predicted values are lesser than those of the GEP model predicted outputs, in both the training and validation stage. This confirms that the RF algorithm gives good performance than GEP model as it is an ensemble one that uses the decision trees as weak learners [58]. Also, $R_o'^2$ and R_o^2 are used for its external validation of RF model as tabulated in Table 5. Their values are calculated near to 1, which verifies that RF algorithm does not work as simple correlation.

3.4. GEP Model Comparison with Linear and Nonlinear Regression Models.
The past research reveals that for f_c' of FA-dependent GPC, no GEP model has been developed using the influential input parameters considered in this study. So, it is needed to develop nonlinear and linear regression expressions, for the same dataset, and compare it with the GEP model presented as equation (8). Equations (13) and (14) present the linear and nonlinear regression equations, respectively.

TABLE 4: Comparison of statistical measures amongst GEP, RF, nonlinear, and linear regression models.

Model	RMSE		RSE		MAE		RRMSE (%)		R		ρ	
	T_{RNG}[1]	V_{LDN}[2]	T_{RNG}	V_{LDN}	T_{RNG}	V_{LDN}	T_{RNG}	V_{LDN}	T_{RNG}	V_{LDN}	T_{RNG}	V_{LDN}
RF	3.034	1.986	0.193	0.0350	2.876	1.862	10.084	4.163	0.9826	0.9943	0.0546	0.02087
GEP	5.971	2.643	0.325	0.0675	5.823	2.057	16.949	4.949	0.8586	0.9643	0.0911	0.02519
Linear	6.986	5.546	0.589	0.3040	6.543	4.967	19.20	10.21	0.8074	0.8976	0.1062	0.05382
Nonlinear	6.593	5.054	0.497	0.2980	6.053	4.875	18.53	9.021	0.8357	0.9247	0.1009	0.04687

[1]TRNG shows training set data. [2]VLDN shows validation set data.

TABLE 5: External validity of the proposed GEP and RF models.

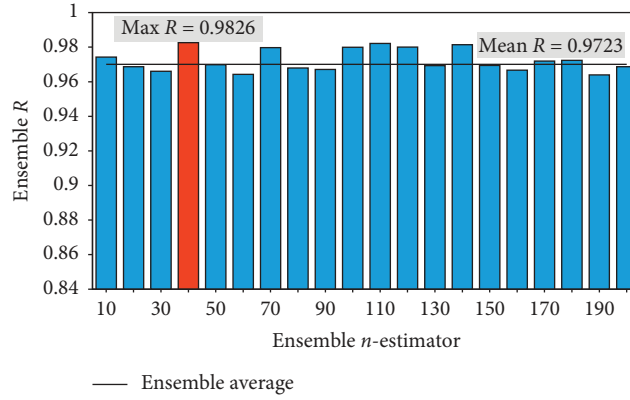| Expression	Constraint	GEP model	RF model
$k = \sum_{i=1}^{n} (\exp_i \times \text{pred}_i) / \sum_{i=1}^{n} (\exp_i^2)$	$0.85 < k < 1.15$	1.001	1.000
$k' = \sum_{i=1}^{n} (\exp_i \times \text{pred}_i) / \sum_{i=1}^{n} (\text{pred}_i^2)$	$0.85 < k' < 1.15$	0.995	0.9995
$R_o^2 = 1 - (\sum_{i=1}^{n} (\text{pred}_i - \exp_i^o)^2 / \sum_{i=1}^{n} (\text{pred}_i - \overline{\text{pred}_i^o})^2), \exp_i^o = k \times \text{pred}_i$	$R_o^2 \cong 1.0$	0.9998	0.9965
$R_o'^2 = 1 - (\sum_{i=1}^{n} (\exp_i - \text{pred}_i^o)^2 / \sum_{i=1}^{n} (\exp_i - \overline{\exp_i^o})^2), \text{pred}_i^o = k' \times \exp_i$	$R_o'^2 \cong 1.0$	0.9849	0.9994

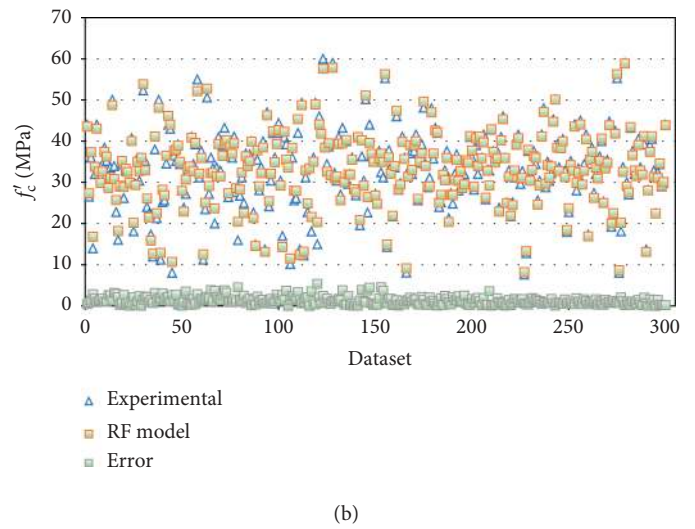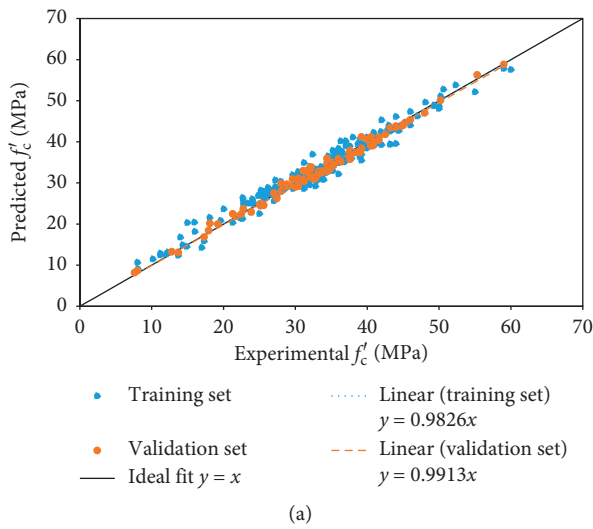FIGURE 5: A random forest regression model with twenty submodels.

(a)

(b)

FIGURE 6: Performance evaluation of RF model. (a) Comparison between model and experimental outcomes for compressive strength from training and validation set data. (b) Absolute error plot of RF predicted outcomes and experimental values.

$$f'_c = 12.8 + 0.23T + 0.04A - 27\frac{A_L}{F_A} + 1.13\frac{N_S}{N_o} - 0.4M + 0.64A_G\% - 0.4\frac{F}{A_G} + 1.3P\% - 0.45\frac{S}{W}\% - 0.7E_W\%,$$

$$\tag{13}$$

$$f'_c = -7.6 + 1.18T^{0.68} + 0.35A^{0.63} - 25.8\left(\frac{A_L}{F_A}\right)^{2.9} + 1.8\left(\frac{N_S}{N_o}\right)^{0.44}$$

$$- 0.009M^{2.24} + 0.76\left(A_G\%\right)^{0.93} - 0.37\left(\frac{F}{A_G}\right)^{1.06}$$

$$+ 2.25\left(P\%\right)^{0.72} - 0.08\left(\frac{S}{W}\%\right)^{1.34} - 0.27\left(E_W\%\right)^{1.32}.$$

$$\tag{14}$$

Figure 7 compares the results of the GEP model and nonlinear and linear regression models. For all three models, the statistical checks like RSE, MAE, RMSE, RMSE%, R, and ρ are mentioned in Table 4. ρ and RMSE of the GEP model for both validation set and training set are lesser than those of the linear and nonlinear regression models. $\rho_{training}$ and $RMSE_{training}$ for the GEP model are 14% and 14.5% lower than those of the linear expression, respectively. Also, in the validation phase, the GEP model performs better than nonlinear regression expression as $\rho_{validation}$ differs by 44%. Figure 7 illustrates that linear and nonlinear regression models fail to cover a large range of f'_c effectively. Hence, the application of regression expression is restricted.

Some limitation of regression analysis like the use of predefined equations either nonlinear or linear and pre-assumption of residuals normality restricts its application [111], while GEP modelling chooses the nonlinear relation between input and output parameters effectively and provides a higher generalized model, which significantly reduces the error as compared to regression analysis.

3.5. Sensitivity and Parametric Analysis. Sensitivity analysis (SA) checks the relative contribution of input parameters considered to predict the compressive strength (f'_c) of FA-dependent GPC, via equation (15) and (16). SA shows the reliance of output on input parameters.

$$N_j = f_{\max}\left(y_j\right) - f_{\min}\left(y_j\right),\tag{15}$$

$$SA = \frac{N_j}{\sum_n^{i=1} N_j},\tag{16}$$

where $f_{\min}(y_j)$ and $f_{\max}(y_j)$ are the jth minimum and maximum predictive model output, respectively while input values are kept constant at mean value. N_j gives the range of jth input variable by taking the difference between $f_{\max}(y_j)$ and $f_{\min}(y_j)$. Both training data points and validation points are consistent; therefore, SA and a parametric study were carried out for only training data points [39, 111]. The result of the sensitivity analysis is presented in Figure 8. It verifies that the relative contribution of input variables is similar in the perspective of material engineering.

The GEP empirical equation, i.e., equation (8), is used to evaluate the effectiveness of influential input parameters by conducting parametric study. The parametric analysis of the GEP model is presented in Figure 9. The changes in compressive strength were noted against the change in the value of only one input parameter from maximum to minimum, and the rest of all input variables are kept at mean value.

The curing temperature in the most important parameter to control the compressive strength (f'_c) of FA-dependent GPC, as shown in Figure 8 which reflects that curing temperature comparatively contributes 25.3%. Figure 9 illustrates an increase in f'_c at different rates with an increase in A, T, (N_s/N_o), %A_G, (F/A_G), and %P while it decreases with (A_L/F_A), % E_W, (% S/W), and M.

The alkali-activating solution being used in the GPC liberates silicates and hydroxides that form strong alumina silicate polymeric structure. As to speed up its reaction process with the source material, the GPC needs additional heat: to improve the mechanical properties of GPC. f'_c increases as curing temperature increases up to $100°C$ as shown in Figure 9. After $100°C$, the loss in moisture from concrete decreases its strength [64]. Wardhono et al. [77] showed through scanning electron microscopy (SEM) that after 240 days, the gel fills out the interior voids, which results in the formation of semihomogenous, but compacted, microstructure. Therefore, after 240 days, the decline in the incremental rate is noted. The change in total aggregate is related to the fine aggregate to total aggregate ratio. f'_c increases with increment in total aggregate amount as shown in Figure 9.

The ratio of alkali-to-FA and sodium silicate-to-sodium hydroxide and molarity of NaOH are all linked. Sodium silicate changes the microstructure and significantly increases the compressive strength. Therefore, preparation of sodium silicate solution with high ratio of percent silica-to-water is needed. Low alkali-to-FA ratio combines with high sodium silicate-to-sodium hydroxide ratio, and less molar solution of NaOH will result in greater f'_c. However, the NaOH solution should be sufficient to finish the dissolution process. Same results have also been reported in the literature [78].

Total water used in GPC is the combination of water needed for the preparation of sodium hydroxide solution and sodium silicate solution and the extra water added for adjusting the workability. For the workable GPC mix and to
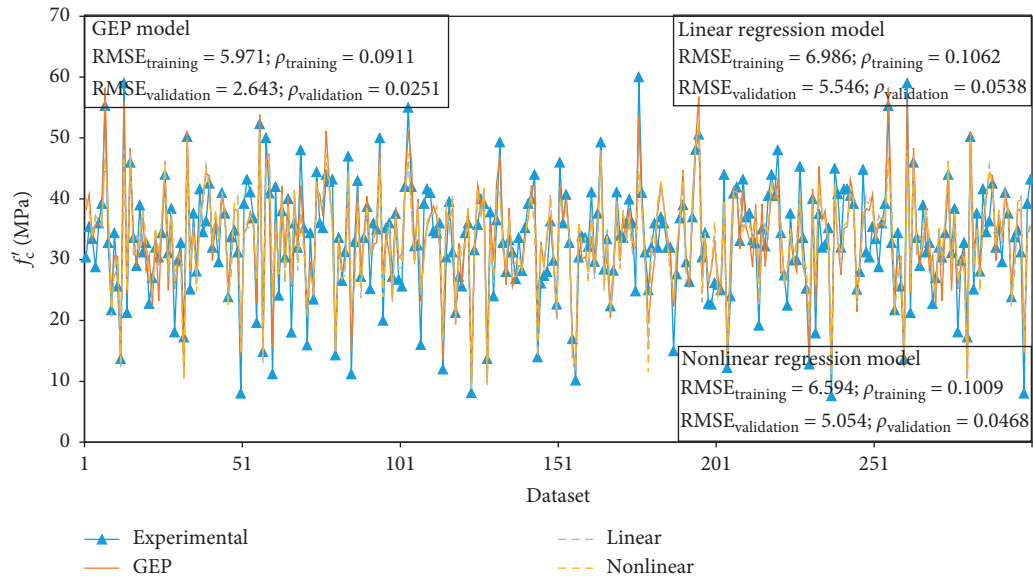
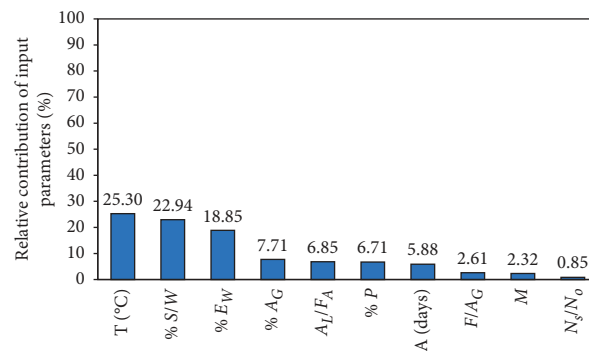Figure 7: The divergence between GEP and nonlinear and linear regression models.

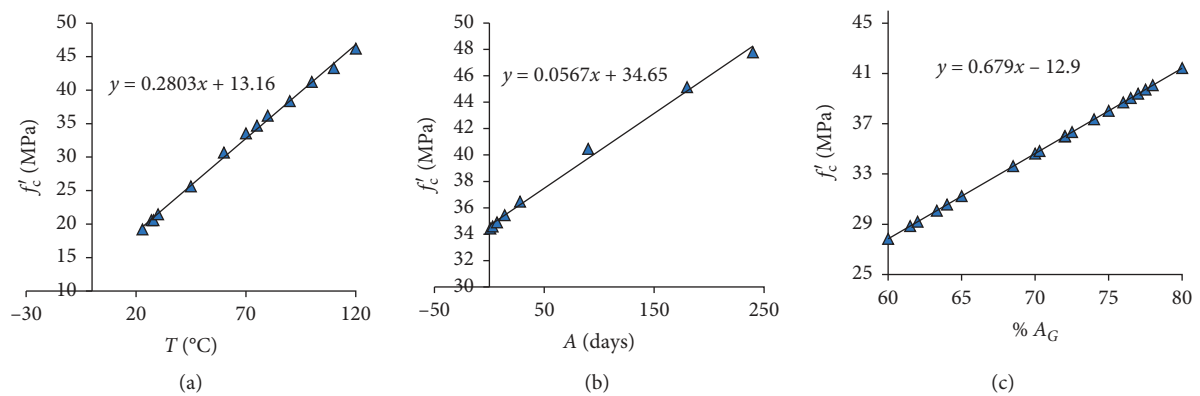Figure 8: Percent contribution of chosen input parameters.

(a)   (b)   (c)

Figure 9: Continued.
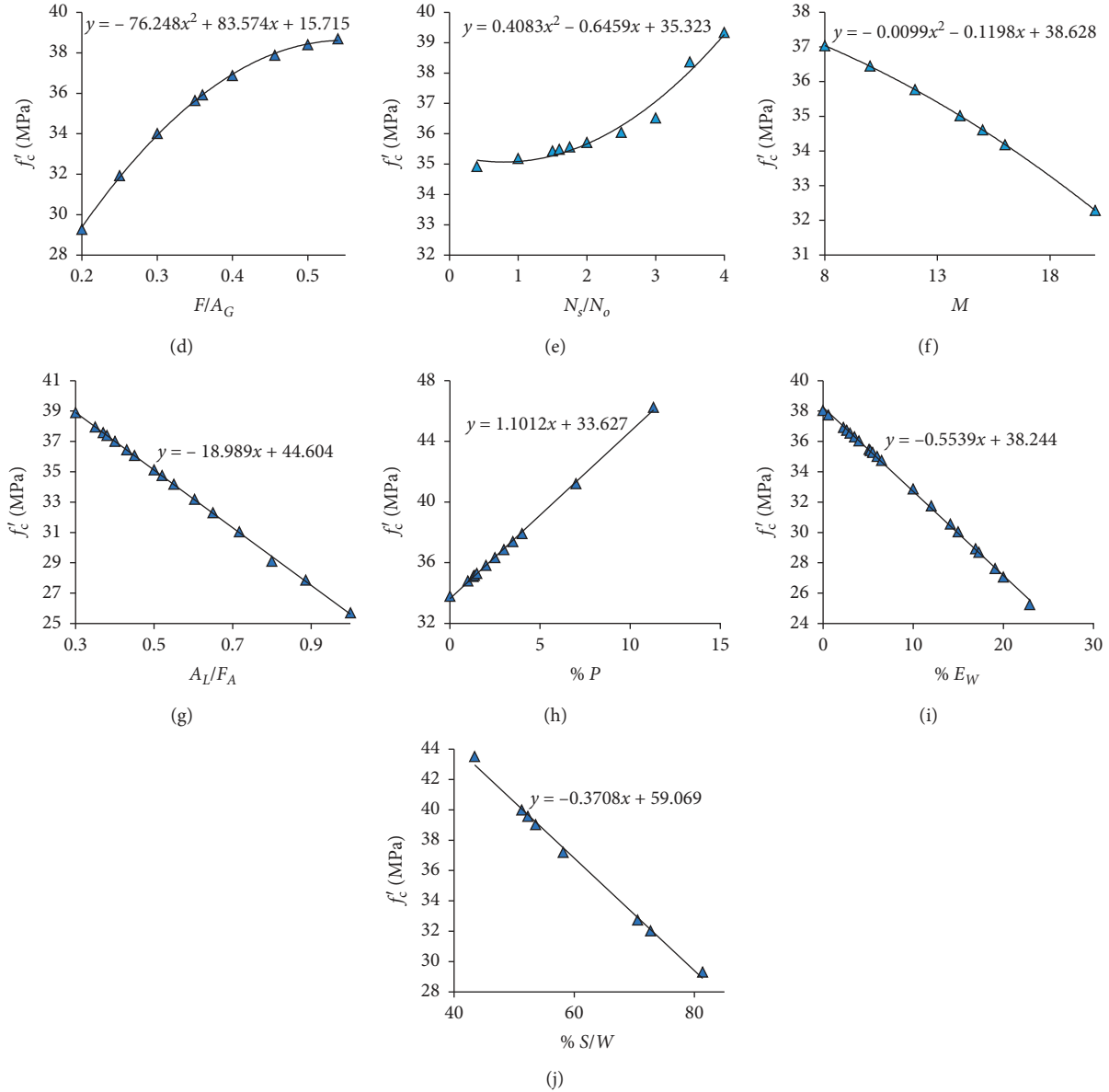
(d)

(e)

(f)

(g)

(h)

(i)

(j)

FIGURE 9: The trend of compressive strength with chosen input parameters.

avoid cracks, it is essential to add a plasticizer and extra water [95]. Figure 9 shows that the relative contribution of plasticizer or extra added water to f_c' is 6.71% and 18.85%, respectively. The extra added water beyond certain limit leads to segregation and bleeding of green concrete.

The results in Figure 9 are linked with previous literature [78, 95]. The parametric analysis accurately shows the effect of input parameters to predict f_c' of FA-dependent GPC.

4. Limitations and Recommendation for Future Work

The research work performed in this article does have certain drawbacks; however, it can be counted as data-mining-based research. The broadness and comprehensiveness of the data is essential for the reliability and proficiency of the predictive models. The range of the datasets used in this research was restricted to 298 experimental data points. This research did not consider the compressive strength of fly-ash-based geopolymer concrete at elevated temperature. Also, this study lacks in providing the empirical relation for other mechanical properties of FGPC like split tensile strength and flexural strength as limited research is available in the literature for both the mechanical properties. In fact, an appropriate testing dataset should be completed as it is essential part in engineering viewpoint. However, this research considered a wide range dataset with ten most influential parameters for modelling compressive strength of FA-dependent GPC.

Furthermore, it is also recommended that the new database developed should be investigated with various supervised machine learning techniques like artificial neural network (ANN), recurrent neural network (RNN), support

vector machine learning (SVM), adaptive neuro fuzzy interface (ANFIS), and multivariate adaptive regression spline (MARS).

5. Conclusions

In this study, random forest (RF) and gene expression programming (GEP) are used to develop a mathematical expression for the compressive strength f_c' of fly-ash- (FA-) dependent geopolymer concrete (GPC). The RF and GEP models are developed on the data collected from the past research, and the most effective variables are considered as input parameters. The proposed GEP empirical expression can be used for the utilization of toxic FA in place of dumping into landfills. This would eventually lead to sustainable green construction. Following are the conclusions deducted via a supervised machine learning algorithm.

(1) The highest R and lowest error checks are observed in the RF model as compared to GEP, nonlinear, and linear regression models. The RF as ensemble machine learning algorithm gives a remarkable performance with R, MAE, RMSE, RSE, and ρ equal to 0.9826, 2.896, 3.034, 0.193, and 0.0546 for training dataset, respectively, and 0.9943, 1.862, 1.986, 0.0350, and 0.02087 for validation dataset, respectively. Also, RF and GEP model accurately meets the specifications for external validation.

(2) RF model outburts performance but lacks in providing an empirical equation. In comparison with nonlinear and linear regression models, the GEP model gives outburst performance and provides an empirical expression, which is suitable for the preliminary design of FA-dependent GPC.

(3) The sensitivity analysis reveals that curing temperature is the most sensitive and dominant parameter in handling the production of FA-dependent GPC. The parametric study of the GEP model shows that the model correctly covers the effect of all explanatory variables.

(4) Furthermore, it is recommended to perform a leachate study before the addition of FA as geopolymer material.

Data Availability

The detailed excel data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] S. E. Aprianti, "A huge number of artificial waste material can be supplementary cementitious material (SCM) for concrete production-a review part II," *Journal of Cleaner Production*, vol. 142, pp. 4178–4194, 2017.

[2] A. Akbar, F. Farooq, M. Shafique, F. Aslam, R. Alyousef, and H. Alabduljabbar, "Sugarcane bagasse ash-based engineered geopolymer mortar incorporating propylene fibers," *Journal of Building Engineering*, vol. 33, Article ID 101492, 2021.

[3] A. Dwivedi and M. K. Jain, "Fly ash–waste management and overview: a review," *Recent Research in Science and Technology*, vol. 6, pp. 30–35, 2014.

[4] O. A. Abdulkareem, A. M. Mustafa Al Bakri, H. Kamarudin, I. Khairul Nizar, and A. e. A. Saif, "Effects of elevated temperatures on the thermal behavior and mechanical performance of fly ash geopolymer paste, mortar and lightweight concrete," *Construction and Building Materials*, vol. 50, pp. 377–387, 2014.

[5] N. Nordin, M. M. A. B. Abdullah, M. F. M. Tahir, A. V. Sandu, and K. Hussin, "Utilization of fly ash waste as construction material," *International Journal of Conservation Science*, vol. 7, pp. 161–166, 2016.

[6] N. Ghazali, K. Muthusamy, and S. Wan Ahmad, "Utilization of fly ash in construction,"in Proceedings of the IOP ConferenceSeries: Materials Science and Engineering, PostgraduateSymposium in Civil and Environmental Engineering 2019, Batu Pahat, Malaysia, Universiti Tun Hussein Onn Malaysia, IOP Publishing, pp. 012023, March 2019.

[7] F. Farooq, A. Akbar, R. A. Khushnood, W. L. B. Muhammad, S. K. U. Rehman, and M. F. Javed, "Experimental investigation of hybrid carbon nanotubes and graphite nanoplatelets on rheology, shrinkage, mechanical, and microstructure of SCCM," *Materials*, vol. 13, p. 230, 2020.

[8] K. M. Liew and A. Akbar, "The recent progress of recycled steel fiber reinforced concrete," *Construction and Building Materials*, vol. 232, Article ID 117232, 2020.

[9] K. A. Baumert, *Navigating the Numbers: Greenhouse Gas Data and International Climate Policy*, World Resources Institute, Washington, DC, USA, 2005, http://http//www.wri.org/publication/navigating-the-numbers.

[10] A. Wongsa, A. Siriwattanakarn, P. Nuaklong, V. Sata, P. Sukontasukkul, and P. Chindaprasirt, "Use of recycled aggregates in pressed fly ash geopolymer concrete," *Environmental Progress & Sustainable Energy*, vol. 39, no. 2, 2020.

[11] A. Akbar and K. M. Liew, "Influence of elevated temperature on the microstructure and mechanical performance of cement composites reinforced with recycled carbon fibers," *Composites Part B: Engineering*, vol. 198, Article ID 108245, 2020.

[12] B. Sumanth Kumar, A. Sen, and D. Rama Seshu, "Shear strength of fly ash and GGBS based geopolymer concrete," in *Lecture Notes in Civil Engineering*, vol. 68, pp. 105–117, Springer, Berlin, Germany, 2020.

[13] F. Farooq, S. K. U. Rahman, A. Akbar et al., "A comparative study on performance evaluation of hybrid GNPs/CNTs in conventional and self-compacting mortar," *Alexandria Engineering Journal*, vol. 59, no. 1, pp. 369–379, 2020.

[14] K. Snehal, B. B. Das, and M. Akanksha, "Early age, hydration, mechanical and microstructure properties of nano-silica blended cementitious composites," *Construction and Building Materials*, vol. 233, Article ID 117212, 2020.

[15] Q. Wang, D. Wang, and H. Chen, "The role of fly ash microsphere in the microstructure and macroscopic properties

of high-strength concrete," *Cement and Concrete Composites*, vol. 83, pp. 125–137, 2017.

[16] H. Y. Zhang, G. H. Qiu, V. Kodur, and Z. S. Yuan, "Spalling behavior of metakaolin-fly ash based geopolymer concrete under elevated temperature exposure," *Cement and Concrete Composites*, vol. 106, Article ID 103483, 2020.

[17] C. R. Meesala, N. K. Verma, and S. Kumar, "Critical review on fly-ash based geopolymer concrete," *Structural Concrete*, vol. 21, no. 3, pp. 1013–1028, 2020.

[18] M. E. Gülşan, R. Alzeebaree, A. A. Rasheed, A. Niş, and A. E. Kurtoğlu, "Development of fly ash/slag based self-compacting geopolymer concrete using nano-silica and steel fiber," *Construction and Building Materials*, vol. 211, pp. 271–283, 2019.

[19] K. Kondepudi and K. V. L. Subramaniam, "Rheological characterization of low-calcium fly ash suspensions in alkaline silicate colloidal solutions for geopolymer concrete production," *Journal of Cleaner Production*, vol. 234, pp. 690–701, 2019.

[20] N. Li, C. Shi, Z. Zhang, H. Wang, and Y. Liu, "A review on mixture design methods for geopolymer concrete," *Composites Part B: Engineering*, vol. 178, Article ID 107490, 2019.

[21] T. T. Tran, T. M. Pham, and H. Hao, "Rectangular stress-block parameters for fly-ash and slag based geopolymer concrete," *Structures*, vol. 19, pp. 143–155, 2019.

[22] P. Zhang, Z. Gao, J. Wang, J. Guo, S. Hu, and Y. Ling, "Properties of fresh and hardened fly ash/slag based geopolymer concrete: a review," *Journal of Cleaner Production*, vol. 270, 2020.

[23] W. Prachasaree, S. Limkatanyu, A. Hawa, P. Sukontasukkul, and P. Chindaprasirt, "Manuscript title: development of strength prediction models for fly ash based geopolymer concrete," *Journal of Building Engineering*, vol. 32, Article ID 101704, 2020.

[24] H. Zhang, L. Li, P. K. Sarker et al., "Investigating various factors affecting the long-term compressive strength of heat-cured fly ash geopolymer concrete and the use of orthogonal experimental design method," *International Journal of Concrete Structures and Materials*, vol. 13, no. 1, 2019.

[25] D. Dao, H.-B. Ly, S. Trinh, T.-T. Le, and B. Pham, "Artificial intelligence approaches for prediction of compressive strength of geopolymer concrete," *Materials*, vol. 12, no. 6, p. 983, 2019.

[26] S. Luhar, S. Chaudhary, and I. Luhar, "Development of rubberized geopolymer concrete: strength and durability studies," *Construction and Building Materials*, vol. 204, pp. 740–753, 2019.

[27] Y. Wang, S. Hu, and Z. He, "Mechanical and fracture properties of fly ash geopolymer concrete addictive with calcium aluminate cement," *Materials*, vol. 12, no. 18, p. 2982, 2019.

[28] M. F. Javed, M. N. Amin, M. I. Shah et al., "Applications of gene expression programming and regression techniques for estimating compressive strength of bagasse ash based concrete," *Crystals*, vol. 10, no. 9, p. 737, 2020.

[29] M. F. Javed, F. Farooq, S. A. Memon et al., "New prediction model for the ultimate axial capacity of concrete-filled steel tubes: an evolutionary approach," *Crystals*, vol. 10, no. 9, p. 741, 2020.

[30] Y. Sun, G. Li, J. Zhang, and D. Qian, "Prediction of the strength of rubberized concrete by an evolved random forest model," *Advances in Civil Engineering*, vol. 2019, p. 1, 2019.

[31] Y. Sun, G. Li, N. Zhang, Q. Chang, J. Xu, and J. Zhang, "Development of ensemble learning models to evaluate the strength of coal-grout materials," *International Journal of Mining Science and Technology*, 2020.

[32] A. M. Abd and S. M. Abd, "Modelling the strength of lightweight foamed concrete using support vector machine (SVM)," *Case Studies in Construction Materials*, vol. 6, pp. 8–15, 2017.

[33] M. A. Getahun, S. M. Shitote, and Z. C. Abiero Gariy, "Artificial neural network based modelling approach for strength prediction of concrete incorporating agricultural and construction wastes," *Construction and Building Materials*, vol. 190, pp. 517–525, 2018.

[34] B. Vakhshouri and S. Nejadi, "Predicition of compressive strength in light-weight self-compacting concrete by ANFIS analytical model," *Archives of Civil Engineering*, vol. 61, no. 2, pp. 53–72, 2015.

[35] Q. Han, 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, pp. 734–742, 2019.

[36] P. Samui, "Multivariate adaptive regression spline (Mars) for prediction of elastic modulus of jointed rock mass," *Geotechnical and Geological Engineering*, vol. 31, no. 1, pp. 249–253, 2013.

[37] M. Sonebi and A. Cevik, "Genetic programming based formulation for fresh and hardened properties of self-compacting concrete containing pulverised fuel ash," *Construction and Building Materials*, vol. 23, no. 7, pp. 2614–2622, 2009.

[38] F. Aslam, F. Farooq, M. N. Amin et al., "Applications of gene expression programming for estimating compressive strength of high-strength concrete," *Advances in Civil Engineering*, vol. 2020, Article ID 8850535, 23 pages, 2020.

[39] M. F. Iqbal, Q.-f. Liu, I. Azim et al., "Prediction of mechanical properties of green concrete incorporating waste foundry sand based on gene expression programming," *Journal of Hazardous Materials*, vol. 384, Article ID 121322, 2020.

[40] M. Jalal, Z. Grasley, C. Gurganus, and J. W. Bullard, "Experimental investigation and comparative machine-learning prediction of strength behavior of optimized recycled rubber concrete," *Construction and Building Materials*, vol. 256, Article ID 119478, 2020.

[41] J.-S. Chou and A.-D. Pham, "Enhanced artificial intelligence for ensemble approach to predicting high performance concrete compressive strength," *Construction and Building Materials*, vol. 49, pp. 554–563, 2013.

[42] H. Mashhadban, S. S. Kutanaei, and M. A. Sayarinejad, "Prediction and modeling of mechanical properties in fiber reinforced self-compacting concrete using particle swarm optimization algorithm and artificial neural network," *Construction and Building Materials*, vol. 119, pp. 277–287, 2016.

[43] H. Sebaaly, S. Varma, and J. W. Maina, "Optimizing asphalt mix design process using artificial neural network and genetic algorithm," *Construction and Building Materials*, vol. 168, pp. 660–670, 2018.

[44] R. Sudin and N. Swamy, "Bamboo and wood fibre cement composites for sustainable infrastructure regeneration," *Journal of Materials Science*, vol. 41, no. 21, pp. 6917–6924, 2006.

[45] A. Behnood and E. M. Golafshani, "Predicting the compressive strength of silica fume concrete using hybrid artificial neural network with multi-objective grey wolves," *Journal of Cleaner Production*, vol. 202, pp. 54–64, 2018.

[46] Z. Keshavarz and H. Torkian, "Application of ANN and ANFIS models in determining compressive strength of

concrete," *Journal of Soft Computing in Civil Engineering*, vol. 2, no. 1, pp. 62–70, 2018.

[47] J. Zhang, G. Ma, Y. Huang, J. Sun, F. Aslani, and B. Nener, "Modelling uniaxial compressive strength of lightweight self-compacting concrete using random forest regression," *Construction and Building Materials*, vol. 210, pp. 713–719, 2019.

[48] A. H. Gandomi, S. K. Babanajad, A. H. Alavi, and Y. Farnam, "Novel approach to strength modeling of concrete under triaxial compression," *Journal of Materials in Civil Engineering*, vol. 24, no. 9, pp. 1132–1143, 2012.

[49] A. H. Gandomi, G. J. Yun, and A. H. Alavi, "An evolutionary approach for modeling of shear strength of RC deep beams," *Materials and Structures*, vol. 46, no. 12, pp. 2109–2119, 2013.

[50] C. Ferreira, "Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence," *Springer*, Vol. 21, Berlin, Germany, 2006.

[51] L. Chen, C.-H. Kou, and S.-W. Ma, "Prediction of slump flow of high-performance concrete via parallel hyper-cubic gene-expression programming," *Engineering Applications of Artificial Intelligence*, vol. 34, pp. 66–74, 2014.

[52] S. B. Beheshti Aval, H. Ketabdari, and S. Asil Gharebaghi, "Estimating shear strength of short rectangular reinforced concrete columns using nonlinear regression and gene expression programming," *Structures*, vol. 12, pp. 13–23, 2017.

[53] I. F. Kara, "Empirical modeling of shear strength of steel fiber reinforced concrete beams by gene expression programming," *Neural Computing and Applications*, vol. 23, no. 3-4, pp. 823–834, 2013.

[54] E. Sadrossadat, B. Ghorbani, M. Hamooni, and M. H. Moradpoor Sheikhkanloo, "Numerical formulation of confined compressive strength and strain of circular reinforced concrete columns using gene expression programming approach," *Structural Concrete*, vol. 19, no. 3, pp. 783–794, 2018.

[55] F. Özcan, "Gene expression programming based formulations for splitting tensile strength of concrete," *Construction and Building Materials*, vol. 26, no. 1, pp. 404–410, 2012.

[56] A. Gholampour, A. H. Gandomi, and T. Ozbakkaloglu, "New formulations for mechanical properties of recycled aggregate concrete using gene expression programming," *Construction and Building Materials*, vol. 130, pp. 122–145, 2017.

[57] D. Behnia, K. Ahangari, A. Noorzad, and S. R. Moeinossadat, "Predicting crest settlement in concrete face rockfill dams using adaptive neuro-fuzzy inference system and gene expression programming intelligent methods," *Journal of Zhejiang University Science A*, vol. 14, no. 8, pp. 589–602, 2013.

[58] F. Farooq, M. Nasir Amin, K. Khan et al., "A comparative study of random forest and genetic engineering programming for the prediction of compressive strength of high strength concrete (HSC)," *Applied Sciences*, vol. 10, no. 20, pp. 7330–7418, 2020.

[59] A. Akbar, K. M. Liew, F. Farooq, and R. A. Khushnood, "Exploring mechanical performance of hybrid MWCNT and GNMP reinforced cementitious composites," *Construction and Building Materials*, Article ID 120721, 2020.

[60] S. Ishak, H.-S. Lee, J. K. Singh et al., "Materials performance of fly ash geopolymer concrete incorporating bamboo ash at elevated temperature," *Materials*, vol. 12, no. 20, p. 3404, 2019.

[61] M. Albitar, P. Visintin, M. S. Mohamed Ali, and M. Drechsler, "Assessing behaviour of fresh and hardened geopolymer concrete mixed with class-F fly ash," *KSCE Journal of Civil Engineering*, vol. 19, no. 5, pp. 1445–1455, 2015.

[62] I. S. Alkroosh and P. K. Sarker, "Prediction of the compressive strength of fly ash geopolymer concrete using gene expression programming," *Computers and Concrete*, vol. 24, pp. 295–302, 2019.

[63] D. Hardjito and B. V. Rangan, "Development and properties of low-calcium fly ash-based geopolymer concrete," *Research report GC1, Curtin Research Publications*, Curtin University of Technology, Perth, Australia, 2005.

[64] B. Joseph and G. Mathew, "Influence of aggregate content on the behavior of fly ash based geopolymer concrete," *Scientia Iranica*, vol. 19, no. 5, pp. 1188–1194, 2012.

[65] J. R. Koza and R. Poli, "Genetic programming," in *Search Methodologies Book*, pp. 127–164, Springer, Boston, MA, USA, 2005.

[66] M. Saridemir, "Genetic programming approach for prediction of compressive strength of concretes containing rice husk ash," *Construction and Building Materials*, vol. 24, pp. 1911–1919, 2010.

[67] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[68] V. Svetnik, A. Liaw, C. Tong, J. C. Culberson, R. P. Sheridan, and B. P. Feuston, "Random forest: a classification and regression tool for compound classification and QSAR modeling," *Journal of Chemical Information and Computer Sciences*, vol. 43, no. 6, pp. 1947–1958, 2003.

[69] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, "Predicting stock market index using fusion of machine learning techniques," *Expert Systems with Applications*, vol. 42, no. 4, pp. 2162–2172, 2015.

[70] H. Jiang, Y. Deng, H.-S. Chen et al., "Joint analysis of two microarray gene-expression data sets to select lung adenocarcinoma marker genes," *BMC Bioinformatics*, vol. 5, no. 1, p. 81, 2004.

[71] O. Rahmati, H. R. Pourghasemi, and A. M. Melesse, "Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: a case study at Mehran Region, Iran," *Catena*, vol. 137, pp. 360–372, 2016.

[72] B. Joseph and G. Mathew, "Behavior of geopolymer concrete exposed to elevated temperatures," *Ph.D. Dissertation.*, Cochin University of Science and Technology, Kochi, India, 2015.

[73] P. K. Sarker, R. Haque, and K. V. Ramgolam, "Fracture behaviour of heat cured fly ash based geopolymer concrete," *Materials & Design*, vol. 44, pp. 580–586, 2013.

[74] T. Long, X. S. Shi, Q. Y. Wang, and L. Li, "Mechanical properties and microstructure of fly ash based geopolymeric polymer recycled concrete," *Sichuan Daxue Xuebao (Gongcheng Kexue Ban)/Journal of Sichuan University (Engineering Science Edition)*, vol. 45, pp. 43–47, 2013.

[75] T. Sujatha, K. Kannapiran, and S. Nagan, "Strength assessment of heat cured geopolymer concrete slender column," *Asian Journal of Civil Engineering*, vol. 13, pp. 635–646, 2012.

[76] P. R. Vora and U. V. Dave, "Parametric studies on compressive strength of geopolymer concrete," *Procedia Engineering*, vol. 51, pp. 210–219, 2013.

[77] A. Wardhono, C. Gunasekara, D. W. Law, and S. Setunge, "Comparison of long term performance between alkali activated slag and fly ash geopolymer concretes," *Construction and Building Materials*, vol. 143, pp. 272–279, 2017.

[78] W. Lokuge, A. Wilson, C. Gunasekara, D. W. Law, and S. Setunge, "Design of fly ash geopolymer concrete mix proportions using Multivariate Adaptive Regression Spline

model," *Construction and Building Materials*, vol. 166, pp. 472–481, 2018.

[79] A. Mehta and R. Siddique, "Properties of low-calcium fly ash based geopolymer concrete incorporating OPC as partial replacement of fly ash," *Construction and Building Materials*, vol. 150, pp. 792–807, 2017.

[80] K. Ramujee and M. Potharaju, "Mechanical properties of geopolymer concrete composites," *Materials Today: Proceedings*, vol. 4, pp. 2937–2945, 2017.

[81] T. Sathanandam, P. O. Awoyera, V. Vijayan, and K. Sathishkumar, "Low carbon building: experimental insight on the use of fly ash and glass fibre for making geopolymer concrete," *Sustainable Environment Research*, vol. 27, no. 3, pp. 146–153, 2017.

[82] P. Nuaklong, V. Sata, and P. Chindaprasirt, "Influence of recycled aggregate on fly ash geopolymer concrete properties," *Journal of Cleaner Production*, vol. 112, pp. 2300–2307, 2016.

[83] M. Fareed Ahmed, M. Fadhil Nuruddin, and N. Shafiq, "Compressive strength and workability characteristics of low-calcium fly ash-based self-compacting geopolymer concrete," *World Academy of Science, Engineering and Technology*, vol. 74, pp. 8–14, 2011.

[84] A. Wongsa, Y. Zaetang, V. Sata, and P. Chindaprasirt, "Properties of lightweight fly ash geopolymer concrete containing bottom ash as aggregates," *Construction and Building Materials*, vol. 111, pp. 637–643, 2016.

[85] F. U. A. Shaikh, "Mechanical and durability properties of fly ash geopolymer concrete containing recycled coarse aggregates," *International Journal of Sustainable Built Environment*, vol. 5, no. 2, pp. 277–287, 2016.

[86] H. K. Shehab, A. S. Eisa, and A. M. Wahba, "Mechanical properties of fly ash based geopolymer concrete with full and partial cement replacement," *Construction and Building Materials*, vol. 126, pp. 560–565, 2016.

[87] A. A. Aliabdo, A. E. M. Abd Elmoaty, and H. A. Salem, "Effect of cement addition, solution resting time and curing characteristics on fly ash based geopolymer concrete performance," *Construction and Building Materials*, vol. 123, pp. 581–593, 2016.

[88] F. N. Okoye, J. Durgaprasad, and N. B. Singh, "Mechanical properties of alkali activated flyash/Kaolin based geopolymer concrete," *Construction and Building Materials*, vol. 98, pp. 685–691, 2015.

[89] N. Ganesan, R. Abraham, and S. Deepa Raj, "Durability characteristics of steel fibre reinforced geopolymer concrete," *Construction and Building Materials*, vol. 93, pp. 471–476, 2015.

[90] L. N. Assi, E. Deaver, M. K. Elbatanouny, and P. Ziehl, "Investigation of early compressive strength of fly ash-based geopolymer concrete," *Construction and Building Materials*, vol. 112, pp. 807–815, 2016.

[91] F. U. A. Shaikh and V. Vimonsatit, "Compressive strength of fly-ash-based geopolymer concrete at elevated temperatures," *Fire and Materials*, vol. 39, no. 2, pp. 174–188, 2015.

[92] S. V. Joshi and M. S. Kadu, "Role of alkaline activator in development of eco-friendly fly ash based geo polymer concrete," *International Journal of Environmental Science and Development*, vol. 3, no. 5, p. 417, 2012.

[93] G. Lavanya and J. Jegan, "Evaluation of relationship between split tensile strength and compressive strength for geopolymer concrete of varying grades and molarity," *International Journal of Applied Engineering Research*, vol. 10, pp. 35523–35529, 2015.

[94] P. S. Deb, P. Nath, and P. K. Sarker, "Properties of fly ash and slag blended geopolymer concrete cured at ambient temperature," in *Proceedings of the ISEC 2013-7th International Structural Engineering and Construction Conference*, pp. 571–576, New Developments in Structural Engineering and Construction, Honolulu, HI, USA, 2013.

[95] M. F. Nuruddin, S. Demie, and N. Shafiq, "Effect of mix composition on workability and compressive strength of self-compacting geopolymer concrete," *Canadian Journal of Civil Engineering*, vol. 38, no. 11, pp. 1196–1203, 2011.

[96] S. Ishak, H. S. Lee, J. K. Singh, M. A. M. Ariffin, N. H. A. S. Lim, and H. M. Yang, "Performance of fly ash geopolymer concrete incorporating bamboo ash at elevated temperature," *Materials*, vol. 12, no. 20, p. 3404, 2019.

[97] S. V. Patankar, Y. M. Ghugal, and S. S. Jamkar, "Mix design of fly ash based geopolymer concrete," in *Advances in Structural Engineering: Materials*, pp. 1619–1634, Springer India, New Delhi, India, 2015, ISBN 9788132221876.

[98] S. V. Patankar, S. S. Jamkar, and Y. M. Ghugal, "Effect of water-to-geopolymer Binder ratio on the production of fly ash based geopolymer concrete," *International Journal of Advanced Technology in Civil Engineering*, pp. 79–83, 2013.

[99] M. D. J. Sumajouw and B. V. Rangan, "Low-calcium fly ash-based geopolymer concrete: reinforced beams and columns," *Research report GC 3, Curtin Research Publications*, Curtin University of Technology, Perth, Australia, 2006.

[100] P. S. Deb, P. K. Sarker, and P. Nath, "Sulphate resistance of slag blended fly ash based geopolymer concrete," in *Proceedings of the 26th Biennial National Conference of the Concrete Institute of Australia, Gold Coast, Australia*, Concrete Institute of Australia, October 2013.

[101] B. Galvin and N. Lloyd, "Fly ash based geopolymer concrete with recycled concrete aggregate," in *Proceedings of the Concrete 2011 Conference. The Concrete Institute of Australia*, Perth, Australia, October 2011.

[102] A. Kusbiantoro, M. F. Nuruddin, N. Shafiq, and S. A. Qazi, "The effect of microwave incinerated rice husk ash on the compressive and bond strength of fly ash based geopolymer concrete," *Construction and Building Materials*, vol. 36, pp. 695–703, 2012.

[103] M. F. Nuruddin, S. A. Qazi, A. Kusbiantoro, and N. Shafiq, "Utilisation of waste material in geopolymeric concrete," *Proceedings of the Institution of Civil Engineers-Construction Materials*, vol. 164, no. 6, pp. 315–327, 2011.

[104] P. Nath and P. K. Sarker, "Flexural strength and elastic modulus of ambient-cured blended low-calcium fly ash geopolymer concrete," *Construction and Building Materials*, vol. 130, pp. 22–31, 2017.

[105] M. Olivia and H. Nikraz, "Properties of fly ash geopolymer concrete designed by Taguchi method," *Materials & Design*, vol. 36, pp. 191–198, 2012.

[106] A. J. Hamad, "Size and shape effect of specimen on the compressive strength of HPLWFC reinforced with glass fibres," *Journal of King Saud University - Engineering Sciences*, vol. 29, no. 4, pp. 373–380, 2017.

[107] J. R. del Viso, J. R. Carmona, and G. Ruiz, "Shape and size effects on the compressive strength of high-strength concrete," *Cement and Concrete Research*, vol. 38, no. 3, pp. 386–395, 2008.

[108] J. G. S. Van Jaarsveld, J. S. J. Van Deventer, and G. C. Lukey, "The effect of composition and temperature on the properties of fly ash- and kaolinite-based geopolymers," *Chemical Engineering Journal*, vol. 89, pp. 63–73, 2002.

[109] A. H. Gandomi and D. A. Roke, "Assessment of artificial neural network and genetic programming as predictive tools," *Advances in Engineering Software*, vol. 88, pp. 63–72, 2015.

[110] S. K. Babanajad, A. H. Gandomi, and A. H. Alavi, "New prediction models for concrete ultimate strength under true-triaxial stress states: an evolutionary approach," *Advances in Engineering Software*, vol. 110, pp. 55–68, 2017.

[111] A. H. Gandomi, A. H. Alavi, M. R. Mirzahosseini, and F. M. Nejad, "Nonlinear genetic-based models for prediction of flow number of asphalt mixtures," *Journal of Materials in Civil Engineering*, vol. 23, no. 3, pp. 248–263, 2011.

[112] A. H. Alavi, M. Ameri, A. H. Gandomi, and M. R. Mirzahosseini, "Formulation of flow number of asphalt mixes using a hybrid computational method," *Construction and Building Materials*, vol. 25, no. 3, pp. 1338–1355, 2011.