Arabian Journal of Geosciences (2021) 14: 2612
https://doi.org/10.1007/s12517-021-08869-4

ORIGINAL PAPER

Experimental investigation and prediction of strength development of GGBFS-, LFS- and SCBA-based green concrete using soft computing techniques

Kavita Rani[1] · Manju Suthar[2] · Parveen Sihag[2] · Amardeep Boora[3]

Received: 24 June 2021 / Accepted: 8 November 2021 / Published online: 23 November 2021

Abstract

The present research article is focusing on the utilization of ground granulated blast furnace slag (GGBFS), ladle furnace slag (LFS) and sugarcane bagasse ash (SCBA) as the partial substitution of ordinary Portland cement (OPC) in concrete mix in order to create a sustainable environment and enhance the engineering performance. The primary purpose of the study is to predict the optimum percentage of different additives in varying proportions, i.e. individually and in a combined manner also to construct a sustainable rigid pavement. Therefore, accurate estimation of strength evaluation is required to minimize the effort. In the current investigation, different modeling analysis techniques have been attempted with different soft computing tools, namely random forest (RF), random tree (RT), M5P, reduced-error pruning (REP) tree and linear regression (LR) to estimate the compressive strength of the concrete using the experimental data. In the present study, all the mentioned additives were added in the concrete mix as the replacement of OPC up to 35%. On the basis of the findings, it was observed that 20% of all the additives in individual form might be used as the partial substitute of OPC. While, in a combined form, concrete mix having 5% GGBFS, 10% LFS and 15% of SCBA was showing the optimum strength value. However, it was also observed that the greater percentage of all the additives can be utilized with an increment in the curing time period. RF approach was found most permissible approach to predict the strength value for green concrete as it was exhibiting higher value of coefficient of correlation, low value of mean absolute error and the root mean square error as revealed by outcomes of the models and statistical assessments methods. Sensitivity analysis is carried out and found that the curing time in days is the utmost effective input variable for estimating the compressive strength of concrete using this data set.

Keywords Soft computing technique · Ground granulated blast furnace slag · Ladle furnace slag · Sugarcane bagasse ash · Concrete

Introduction

Concrete has become the most often used construction materials, i.e. in rigid pavements, bridges and buildings which helps in building up of good infrastructure, though it also causes major environmental and health-related problems due to the production of cement on a large scale (Mishra et al., 2014; Gizaw et al., 2016; Mehraj et al., 2013). About 5 to 8% of global CO_2 emissions are produced by the cement industries in the manufacturing of cement (Mehta, 1997; Gartner, 2011; Ali et al., 2011; Flatt et al., 2012). Therefore, researchers across the world started focusing on the usage

Responsible Editor: Amjad Kallel

✉ Manju Suthar
manju.e10702@cumail.in; manju.manni4@gmail.com

Kavita Rani
kavita12aug@gmail.com

Parveen Sihag
parveen12sihag@gmail.com

Amardeep Boora
amardeepboora9@gmail.com

[1] Department of Civil Engineering, Dr B R Ambedkar National Institute of Technology, Jalandhar, Punjab 144011, India

[2] Department of Civil Engineering, Chandigarh University, Mohali, Punjab 140413, India

[3] Department of Civil Engineering, Jaypee University of Information Technology, Waknaghat 173234, Himachal Pradesh, India

of different left-over-things like slag, demolished waste, fly ash, rice husk ash and many more as an additive by substituting the cement partially or completely. In previous studies, Amin (2011) and Deepika et al. (2017) concluded that 20% of sugar cane bagasse ash (SCBA) can be used as the partial replacement of cement without affecting its required properties. Mangi et al. (2017) concluded that the utilization of 5% SCBA will result in the increment of compressive strength up to 12%, though Anupam et al. (2017) specify that the ideal percentage of the SCBA as the substitute of cement will depend on the properties of bagasse ash itself as it may vary from location to location. Besides this, several other studies were also conducted in order to utilized different other wastes like granulated blast furnace slag (GGBFS) (Das et al., 2015; Mehta and Siddique, 2018) and ladle furnace slag (LFS) (Manso et al., 2005; Rađenović et al., 2013; Borges et al., 2017; Maghool et al., 2017). On the basis of the outcomes, it was concluded that these different waste materials can be utilized in varying proportion as the replacement of cement without compromising its requisite characteristics. Besides these, there are many other waste materials which can be utilized as the same. But on the basis of the literature review, GGBFS, LFS and SCBA were identified as the more promising waste materials w.r.t the strength value of concrete. Therefore, these three different waste materials were selected for investigating compressive strength characteristics experimentally. Although various studies have been conducted by focusing these waste materials, yet the most of the studies used the activating compounds which may have adverse effect on the ground water also, i.e. leaching problem, etc. Secondly, no study was found by focusing on the utilization of these different waste materials in combined manner (i.e. increase the rate of utilization without using any activator). Therefore, a precise strength evaluation technique is required to minimize the effort, i.e. to calculate an optimum percentage. There are various mathematical tools accompanied by a software application that develops the different models based on an algorithm like support vector machine (SVM), M5P, RF, artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) for the solution of complex problems (Suthar and Aggarwal 2018, 2019; Sihag et al. 2019; Suthar 2020a, b; Behnood and Golafshani 2020; Mohammed et al. 2020a, b; Esmaeilbeiki et al. 2020; Kandiri et al. 2020; Sihag et al. 2021; Golafshani and Behnood 2021). Mohammed et al. (2020c) used ANN, M5P-tree and regression analyses for the estimation of early age compression strength of concrete modified and found that developed models were most reliable models for estimating compacting strength of concrete. Farooq et al. (2020) used RF and genetic engineering programming (GEP) to estimate the compressive strength of high strength concrete (HSC) and found that RF was performing better than GEP-based model for the estimation of compressive strength. Keeping in view the improved performance of M5P, RF and other models, in the current investigation, different modeling analysis techniques (M5P, RF, RT and REP tree and LR) have been used for the prediction of the concrete compressive strength by using the experimental data of this investigation for training and testing for development and validation of the developed models. In general, destructive tests are performed to study the different variables of hardened concrete. The replication of experimental tests to assess the different mechanical properties and features of concrete found a very time- and resource-consuming method. To economize the phenomenon, various artificial intelligence-based modeling approaches have offered another easy and economical way for predicting the concrete compressive strength in a hardened state. Sobhani et al. (2019) used the modeling analysis in association with various soft computing tools to predict the results of various structural properties and other problems related to concrete. The objective of the study was to investigate the versatility of three different waste materials GGBFS, LFS and SCBA to be used in cement concrete as a replacement of cement partially followed by development of soft computing model suitable for precise prediction of compressive strength of the cement concrete to reduce the experimental effort, save time and energy.

Background studies

Review on M5P tree technique

Basically, model trees depend on the idea of regression trees having linear functions on their leaves (Witten et al., 2005). In piece-wise type, they are similar to linear equations. According to Quinlan (1992), M5P model tree is a two-fold decision tree at the terminal (leaf) nodes with linear equations that can approximate continuous mathematical characteristics. In order to reduce the possibility of overfitting, trimming is implemented in this design. A separation technique is introduced for each node for procuring enhanced knowledge with less aberration within the subset class values in each branch. There are three main steps in preparing the M5P model: tree growth, pruning and smoothing. The generation of the model tree requires an estimation of standard deviations class values extending to nodes which is carried out utilizing separation criteria. Linear functions in each node are generated by this approach. It utilizes the standard deviation approach by measuring the predicted error at the terminal node.

The following is standard reduction:

$$\mathrm{SDR} = sd(Z) - \sum_{\mathrm{i}}^{\mathrm{n}} \left(\frac{|Z_{\mathrm{i}|}}{|Z|} sd(Zi \right) \tag{1}$$

where Z denotes the list of examples at the Zi node showing the outcome of the subset of possible set examples, and

Table 1 Chemical composition of GGBS, LFS and SCBA

Material and compound	SiO_2	Al_2O_3	Fe_2O_3	CaO	MgO	K_2O	Na_2O	SO_3	MnO	LOI
OPC-43	20.27	5.32	3.56	60.41	2.46	0.41	0.33	2.48	_	4.76
GGBFS	35.00	13.00	1.13	40.00	8.00	0.47	0.33	2.07	_	_
SCBA	78.34	8.55	3.61	2.15	1.65	3.46	0.12	0.18	0.10	1.84
LFS	15.00	14.30	1.54	48.37	15.25	0.36	0.43	0.93	3.82	_

sd is the standard equation. This technique yields a remarkable tree structure with high prediction precision.

Review on random forest

Breiman (1996) pioneered the use of random forest algorithm. This technique is versatile therefore is often chosen to solve various nonlinear or complex engineering issues. In this technique, numerous trees are created, where the root node achieves a dissimilar bootstrap (bagging) sample of the original data set. The division is performed at each node using a randomly chosen subcategory of the parameters of the estimator. The random forest algorithm is simple to use and is relatively indifferent to the training set characteristics yet is accomplished in attaining high prediction accuracy (Breiman, 1999 & 2001). It merely requires two user-defined parameters: the quantity of trees cultivated (k) and the quantity of input parameters (m). For model development, a hit or miss way is used. In the current investigation, this model was formed using the WEKA 3.9 software.

Reduces error pruning (REP) tree

The REP tree classifier is based on the concept of measuring entropy data acquisition and minimizing the variance error as a fast decision tree method (Witten et al., 2005). The REP tree produces several iterations updated by regression tree iterations. Then, it selects the best of the trees developed. Within the context of variance and the information gain method, this algorithm generates a regression/decision tree. This algorithm decreases the pruning error rate by using the form of linking. The calculation used in pruning the tree is the error expected by the tree in the average picture. At the beginning of the modelling process, the values of numerical attributes are sorted. This algorithm breaks the corresponding samples into pieces and processes the missing values (Quinlan, 1987).

Random tree

Without pruning, the random tree algorithm uses an assessment based on a particular number of random characteristics at every node. RT mainly uses arbitrary knowledge and have minute to do with machine learning (Hamoud et al, 2018). RT uses a bagging idea. Every node in a random forest is finest divided between the arbitrarily chosen forerunner subsets of that node. The algorithm deals with both problems with classification and regression. Random trees are a set of forests called tree estimators. The classification operates as follows: the classifier of random trees takes the vector of the input property, classifies it for each tree in the forest and extracts the category mark which receives maximum votes. In the event of a denial, the response of the classifier is the average response of all the trees in the forest (Cutler et al., 2012). RTs are basically a synthesis of two algorithms that exist in machine learning: RF principles and single model trees. Model trees are decision trees in which the linear pattern of each leaf is designed for the local subdomain that this leaf represents. The performance of single stable trees has been shown to significantly improve RFs; tree diversity is created by two random methods. First by removing each tree, the training data is sampled, as in Bagging. Second, instead of always calculating the best possible division for each node when growing a tree, only one random subset of all attributes is considered for each node, and the best part of that subset is determined. For the first time, random model trees merge random forests and model trees. RTs employ this result for dividing criteria and thus encourage

Table 2 Physical properties of OPC-43, GGBFS, SCBA and LFS

Physical characteristics	OPC-43	GGBFS	SCBA	LFS	Standard code
Consistency (%)	29.34%	50.5%	31.2%	34%	IS 4031(part 4)-2005
Specific gravity (g/cc)	3.12	2.83	1.90	3.04	IS 4031(part11)-1988b
Initial setting time (min)	103 min	311 min	190 min	282 min	IS 4031(part 5)-1988a
Final setting time (min)	241 min	371 min	328 min	375 min	IS 4031(part 5)-1988b

Table 3 Physical properties of aggregates

Physical characteristics	Coarse aggregate	Fine aggregate	Standard code
Specific gravity (g/cc)	2.74	2.65	IS 2386(part3)-1997
Water absorption (%)	1	2	IS 2386(part3)-1997
Bulk density (kg/m^3)	1510	1630	IS 2386(part3)-1997

considerately balanced trees where a spherical ridge environment runs on all leaves, thus simplifying the optimization method (Barddal et al., 2019).

Linear regression (LR)

LR is a process based on linear regression used for developing the relationship between variables (dependent and independent). The output variable is Z and the input variables are as follows: x_1, x_2, $x_{3,}$,…., x_n (Kashi et al., 2014). The commonly used equation for LR is as follows:

$$Z = co + c_1X_1 + c_2X_2 + c_3X_3 \ldots \ldots \cdots + c_nX_n \tag{2}$$

Multiple linear regression model entails upon how a single response variable Y relies linearly on a number of predictor variables.

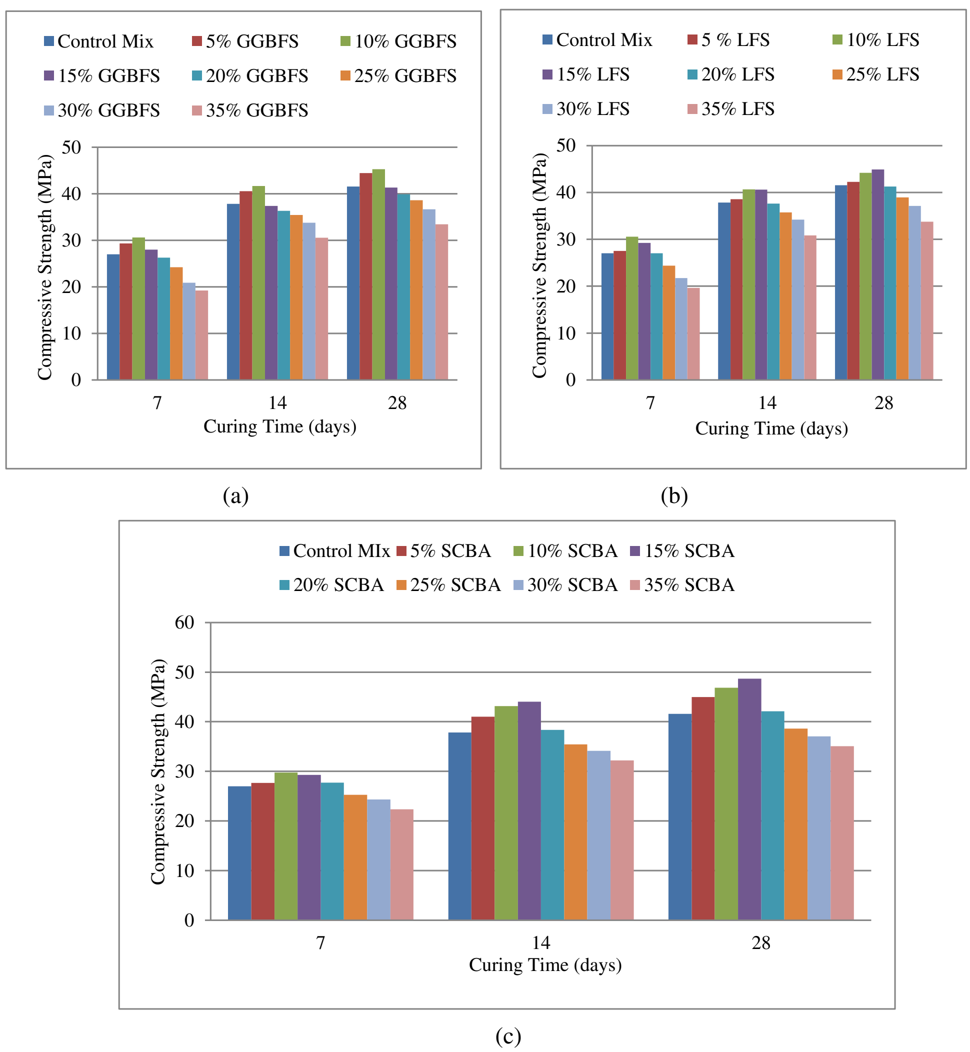


Fig. 1 Variation in strength values with curing time in days (d)

Materials, experiments and results

Materials

The 43-grade regular Portland cement was used to prepare the specimens. The cement's consistency was found to be 29.34%, while specific gravity of the cement was note be 3.12. Its compressive strength was 27.01, 37.82 and 41.56 MPa, respectively, for a period of 7, 14 and 28 days. GGBS and LFS used in this study were procured from a local vendor in New Delhi, India, while SCBA was collected from sugar mill located in the Saharanpur District, India. The chemical characteristics of the cement, GGBS, LFS and SCBA are presented in Table 1. Two types of aggregates were used in the concrete mixture, i.e. coarse aggregates and fine aggregates. The maximum and nominal size of the abrasive aggregate was found to be 20 and 10 mm, respectively, while the fine aggregates retained on 150 μ sieve were used in the current study. The guidelines of IS 383:1970 was followed to ascertain the portion of abrasive and fine aggregates. Physical properties of OPC along with different additives and aggregates are given in Tables 2 and 3, respectively. In the specimens, super-plasticizer naphtha-based superplasticizer confirming IS 9103:1999 which increases the workability of the mix was used as an additive material.

Methods

Initially, a conventional mix was designed as per Indian standards without mixing any waste materials, i.e. GGBFS, LFS and SCBA for M40 concrete grade followed by IS 10262:2019 guidelines for mixing the waste. The desired workability of the mix is attained by adjusting the amount of the super plasticizer. Specimens were casted in the cubes of dimension 150 × 150 × 150 (all mm) to administer the compressive strength test. Later on, different cubes were prepared from the mix of waste materials used in proportion (by weight) 0%, 5%, 10%, 15%, 20%, 25%, 30% and 35% as a replacement of cement (individually as well as in a combined form) by GGBFS, LFS and SCBA. The constant water-cement ratio was kept for all the mixes as 0.40 was set as per the specification of IS 10262:2019. All the specimens were made and cured an ambient temperature for 7, 14 and 28 days. Compressive strength tests were carried out in the Highway Engineering Lab. As per the specifications of the Bureau of Indian Standards (BIS: 516–1959), cubic specimens were tested after 7, 14 and 28 days in compressive testing machine having a capacity of 2000 KN. All the waste materials used in the current study, i.e. GGBFS, LFS and SCBA, can be utilized individually for the partial replacement of cement up to 15% without compromising recommended characteristics of the mix without utilizing any

Table 4 Design of samples (percentage by weight of mix)

Sample References	Formulation (%)	Cement (%)	Ground granulated blast furnace slag (GGBFS) (%)	Ladle furnace slag (LFS) (%)	Sugar cane bagasse ash (SCBA) (%)
A0	100C0G0L0S	100	0	0	0
A1	85C5G5L5S	85	5	5	5
A2	80C5G5L10S	80	5	5	10
A3	75C5G5L15S	75	5	5	15
A4	70C5G5L20S	70	5	5	20
A5	80C5G10L5S	80	5	10	5
A6	75C5G10L10S	75	5	10	5
A7	70C5G10L15S	70	5	10	15
A8	75C5G15L5S	75	5	15	5
A9	70C5G15L10S	70	5	15	10
A10	70C5G20L5S	70	5	20	5
A11	80C10G5L5S	80	10	5	5
A12	75C10G5L10S	75	10	5	10
A13	70C10G5L15S	70	10	5	15
A14	75C10G10L5S	75	10	10	5
A15	70C10G10L10S	70	10	10	10
A16	70C10G15L5S	70	10	15	5
A17	75C15G5L5S	75	15	5	5
A18	70C15G5L10S	70	15	5	10
A19	70C15G10L5S	70	15	10	5
A20	70C20G5L5S	70	20	5	5

Table 5 Combined effect of waste materials on compressive strength

Sample references	Formulation (%)	Average compressive strength (MPa)		
		7 days curing	14 days curing	28 days curing
A0	100C0G0L0S	27.01	37.82	41.56
A1	85C5G5L5S	29.3	40.75	44.71
A2	80C5G5L10S	30.0	41.24	45.76
A3	75C5G5L15S	27.2	37.83	41.55
A4	70C5G5L20S	26.5	36.45	40.45
A5	80C5G10L5S	28.3	39.33	43.15
A6	75C5G10L10S	29.4	40.42	44.75
A7	70C5G10L15S	30.5	42.37	46.50
A8	75C5G15L5S	29.2	40.31	44.58
A9	70C5G15L10S	26.5	36.42	40.40
A10	70C5G20L5S	28.8	40.06	43.96
A11	80C10G5L5S	28.7	40.02	43.98
A12	75C10G5L10S	26.5	36.42	40.40
A13	70C10G5L15S	26.6	36.94	40.53
A14	75C10G10L5S	27.3	38.32	41.96
A15	70C10G10L10S	28.4	39.59	43.40
A16	70C10G15L5S	28.0	39.20	42.92
A17	75C15G5L5S	27.4	38.05	41.75
A18	70C15G5L10S	27.8	38.74	42.50
A19	70C15G10L5S	25.9	36.05	39.55
A20	70C20G5L5S	25.6	35.64	39.11

Note: *C*, cement; *G*, GGBFS; *L*, LFS; and *S*, SCBA

*Values in brackets are the percent increment or decrement in compressive strength values corresponding to the controlled specimen

activating solution as shown in Fig. 1. Besides this, these different additives were successfully utilized up to 30% as the partial replacement of cement in combined form as shown in Table 4 for design of specimens, and Table 5 for results. It is also to be noted that the further utilization of these different additive can be done with the increment in the curing time period as concluded in a recent study (Raghavendra and Udayashankar, 2014). The aggregate quantities also influence the compressive strength to the larger extent. Since the study was concentrated on the impact of the three waste materials for their congruity in compressive strength, the impact of the aggregates on compressive strength was not included in it.

Models formation

In this study, five modeling techniques, namely linear regression (LR), M5P, random forest (RF), REP tree and random tree (RT) were used through WEKA 3.9 software using experimentally data.

Parameters for performance appraisal

For the assessment of the accuracy of the implemented models, three different types of statistical parameters were considered. The root mean squared error (RMSE), coefficient of

Table 6 Correlation matrix among all input and output variables

	Cement	*GGBS*	*LFS*	*SCBA*	Curing time in days (d)	*Compressive Strength MPa*
Cement	1					
GGBS	−0.3518	1				
LFS	−0.3518	−0.3143	1			
SCBA	−0.3518	−0.3143	−0.3143	1		
Curing time in days (d)	4.3E-17	9.08E-18	9.08E-18	−1.8E-17	1	
Compressive strength	0.2536	−0.1585	−0.0772	−0.0319	0.7976	1

Table 7 Features of the data set

Statistics	Data set	*Cement (%)*	*GGBS (%)*	*LFS (%)*	*SCBA (%)*	Curing time in days (d)	Compressive strength *(MPa)*
Minimum	Total	65	0	0	0	7	19.2
	Training	65	0	0	0	7	20.91
	Testing	65	0	0	0	7	19.2
Maximum	Total	100	35	35	35	28	48.67
	Training	100	30	35	35	28	48.67
	Testing	90	35	35	35	28	46.87
Mean	Total	77.50	7.50	7.50	7.50	16.33	35.50
	Training	78.49	6.51	7.44	7.56	16.20	35.75
	Testing	75.38	9.63	7.63	7.38	16.63	34.96
Standard deviation	Total	9.05	8.57	8.57	8.57	8.77	7.02
	Training	9.46	7.75	8.36	8.14	8.74	7.06
	Testing	7.79	9.90	9.13	9.54	8.92	7.01
Skewness	Total	0.78	1.42	1.42	1.42	0.39	−0.36
	Training	0.81	1.42	1.37	1.25	0.42	−0.33
	Testing	0.38	1.28	1.55	1.70	0.33	−0.44
Confidence level (95.0%)	Total	1.60	1.51	1.51	1.51	1.55	1.24
	Training	2.03	1.66	1.79	1.75	1.87	1.51
	Testing	2.49	3.16	2.92	3.05	2.85	2.24

correlation (CC) and mean absolute error (MAE) were used. The optimal model should have the minimum error value (RMSE and MAE) and CC should be close to 1. It is possible to quantify the three performance assessment parameters used in this research using Eqs. (3) to (5).

$$CC = \frac{N\sum_{i=1}^{N}(AP) - \left(\sum_{i=1}^{N}A\right)\left(\sum_{i=1}^{N}P\right)}{\sqrt{N\left(\sum_{i=1}^{N}A^2\right) - \left(\sum_{i=1}^{N}A\right)^2}\sqrt{N\left(\sum_{i=1}^{N}P^2\right) - \left(\sum_{i=1}^{N}P\right)^2}} \tag{3}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|P - A| \tag{4}$$

$$RMSE = \sqrt{\frac{1}{N}\left(\sum_{i=1}^{N}(P - A)^2\right)} \tag{5}$$

where N, A and P are the number of data set, actual values and predicted values, respectively.

In this investigation, compressive strength of the concrete specimens (CS) is dependent on a number of governing parameters like cement (%), ladle furnace slag (%), ground granulated blast furnace slag (%), sugar cane bagasse ash

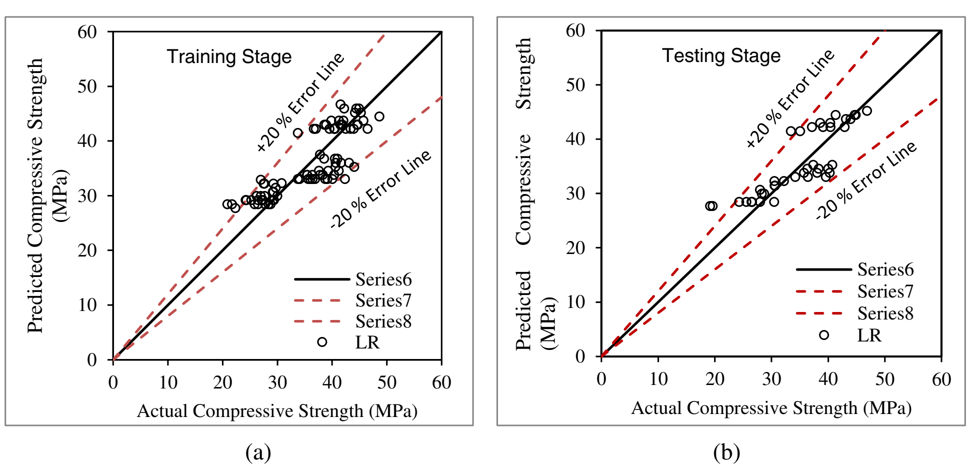


Fig. 2 Agreement plot of observed and predicted compressive strength of concrete by the LR-based model using training and testing data sets, respectively

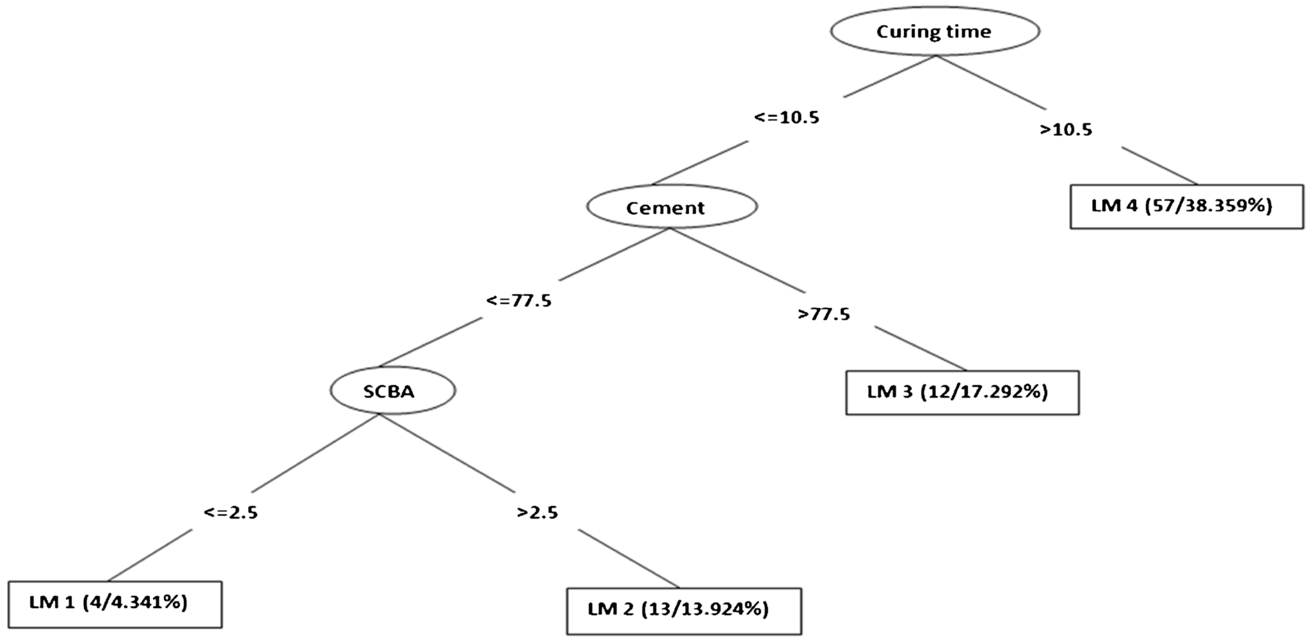


Fig. 3 Structure of M5P-based model

Table 8 Linear equations developed using M5P-based model

LM number	Linear equations
1	$CS = 0.2198C + 0.2235d + 9.3466$
2	$CS = 0.1479C - 0.0237GGBS - 0.0363SCBA + 0.2235d + 15.5616$
3	$CS = 0.0952C + 0.2235d + 19.2793$
4	$CS = 0.1562C + 0.3441d + 20.4735$

(%) and curing time in days (d). A total of 126 experimental data were used in the development of LR, M5P, RF, REP tree and RT models, while 86 randomly selected data were utilized for training and the rests of 40 data were employed for testing the model. To evolve a model to predict compressive strength, cement, ground granulated blast furnace slag (*GGBS: %*), ladle furnace slag (*LFSS: %)*, sugar cane bagasse ash (*SCBA:* %) and curing period (d*: day*) were used as input parameters in above said soft computing techniques, while compressive strength of concrete (CS) was a target for construction of model and validation goal. Table 6 shows the correlation matrix among the input and output parameters. Table 7 shows the characteristics of the data used for model formation and validation.

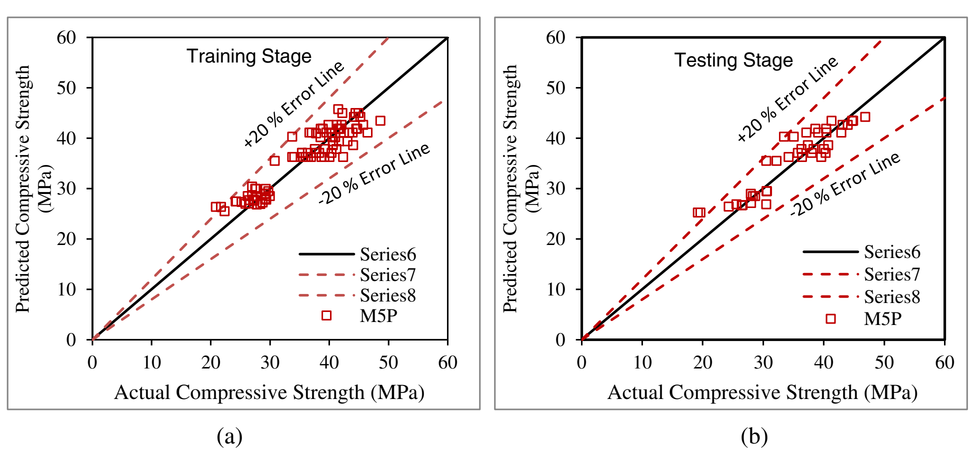


Fig. 4 Agreement plot of observed and predicted compressive strength of concrete by the M5P-based model using training and testing data sets, respectively

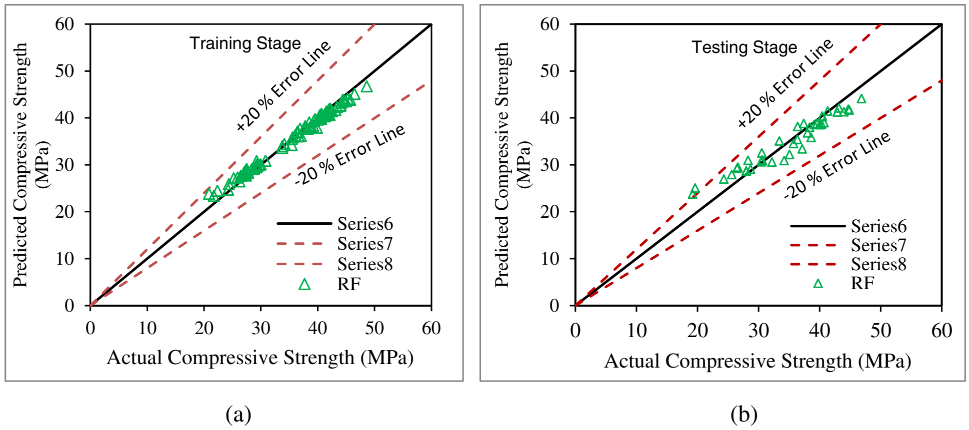


Fig. 5 Agreement plot of observed and predicted compressive strength of concrete by the RF-based model using training and testing data sets, respectively

Result and discussion

Assessment of LR-based model

In the current investigation, linear regression-based approaches for estimating compressive strength of concrete have also been developed. XLSTAT 2020 software employing the least square method has been used to develop this equation. Linear equation based on regression-based models is as follows:

LR equation:

$$CS = 0.1489C + 0.6555d + 13.4442 \tag{6}$$

Three statistical performance measurement statistics were employed to evaluate the working of the LR-based designs in terms of CC, RMSE and MAE. Overall performance of LR-based model is satisfactory with CC = 0.8323, 0.8410;

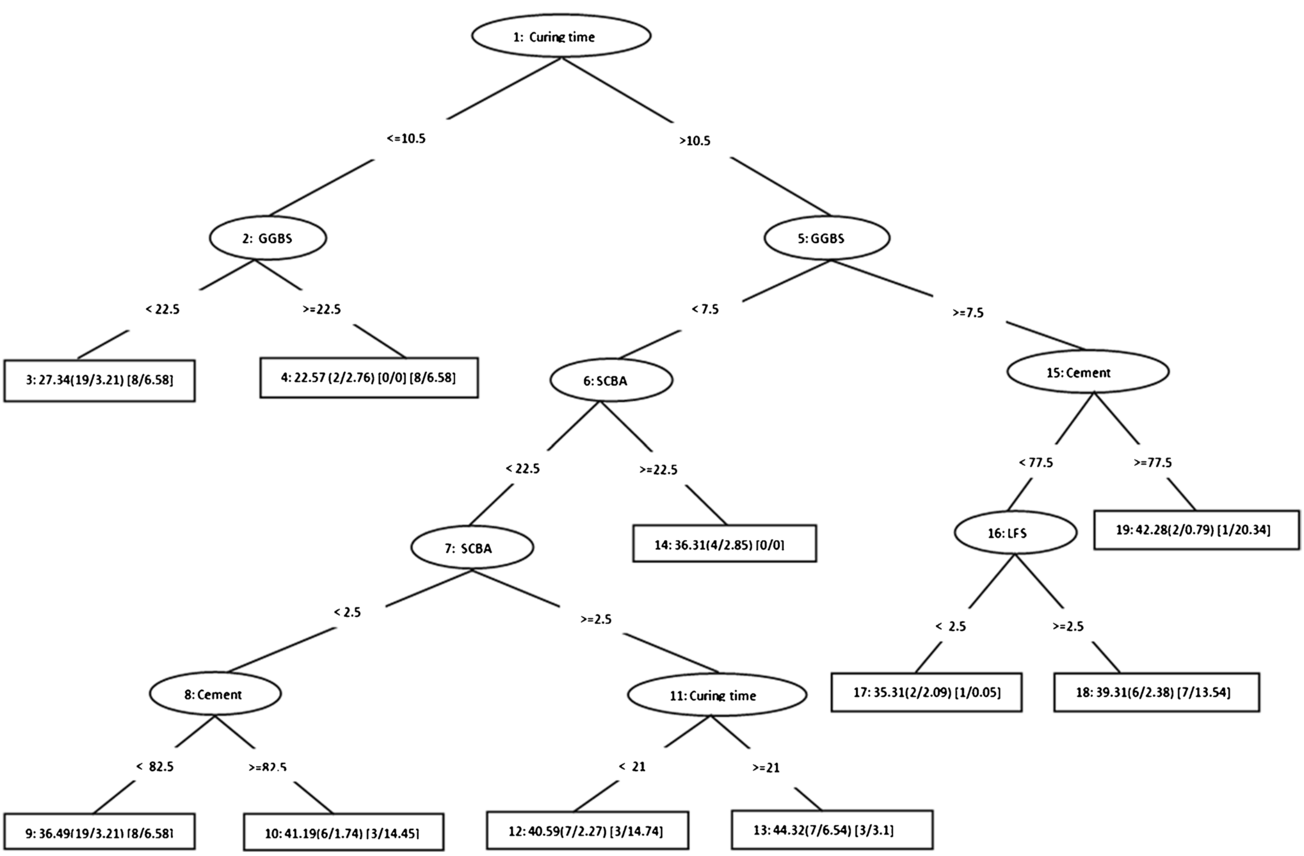
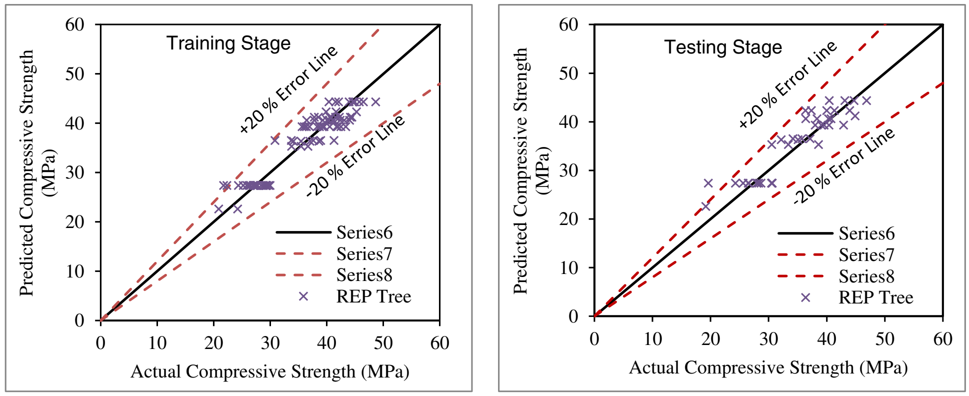


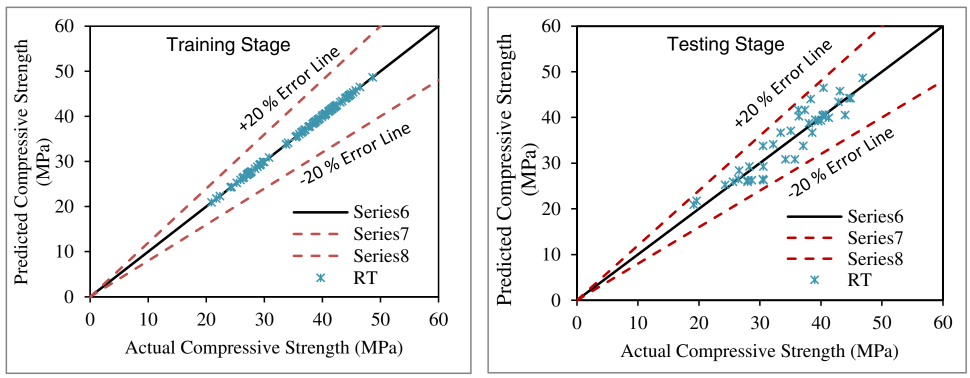
Fig. 6 Structure of REP tree-based model

Fig. 7 Agreement plot of observed and predicted compressive strength of concrete by the REP tree-based model using training and testing data sets, respectively



(a) (b)

Fig. 8 Agreement plot of observed and predicted compressive strength of concrete by the RT-based model using training and testing data sets, respectively



(a) (b)

RMSE = 3.2102 MPa, 2.9982 MPa; and MAE = 3.8891 MPa, 3.7981 MPa for training and testing stages of LR-based model, respectively. Figure 2 displays agreement plots considering training and testing data sets between observed and estimated concrete compressive strength through LR-based technique. As depicted in the graph, values predicted using LR-based model mostly lies within the ± 20% error band. Even in Fig. 2b testing stage, the error band is within the normal limits.

Assessment of M5P-based model:

M5P-based approach development is a hit and trial method. In this study, pruned type M5P-based model is developed and the structures are shown in Fig. 3.

Developed linear equations using pruned M5P-based model are enlisted in Table 8. These equations are formed according to the M5P structure shown in Fig. 3. Figure 4 provides plots of agreement among actual and predicted concrete compressive strength by M5P-based model for training and testing stages, respectively. Estimated values using M5P-based model mostly lie within the ± 20% error band. Overall M5P-based model outperforms LR-based model with CC = 0.9330, 0.9252; RMSE = 2.5635 MPa, 2.7149 MPa; and MAE = 2.0129 MPa, 2.1213 MPa for training and testing stages, respectively.

Assessment of RF-based technique

Formation of RF-based technique for predicting the values of the concrete's compressive strength is identical as M5P-based model development process. Figure 5 provides plots of agreement among real and predicted concrete compressive strength by RF-based model for training and testing stages, respectively. Estimated values from RF-based model closely follow the observed values.

Performance evaluation parameters depict that the work of the RF design is suitable for predicting concrete's compressive strength with CC, RMSE and MAE values which are 0.9967, 0.9791 MPa and 0.7918 MPa for training stage and 0.9687, 2.1887 MPa and 1.7808 MPa for testing stage, respectively.

Assessment of REP tree-based model

Development of REP tree-based design for estimating the values of the concrete compressive strength is identical as M5P- and RF-based model's development process. Structure

Table 9 Performance evaluation indices for all applied models

Models	Training stage			Testing stage		
	CC	MAE (MPa)	RMSE (MPa)	CC	MAE (MPa)	RMSE (MPa)
LR	0.8323	3.2102	3.8892	0.8410	2.9982	3.7981
M5P	0.9330	2.0129	2.5635	0.9252	2.1213	2.7149
RF	0.9967	0.7918	0.9791	0.9687	1.7808	2.1887
REP Tree	0.9461	1.8140	2.2722	0.9180	2.2613	2.8478
RT	1.0000	0.0115	0.0365	0.9323	2.1613	2.7331

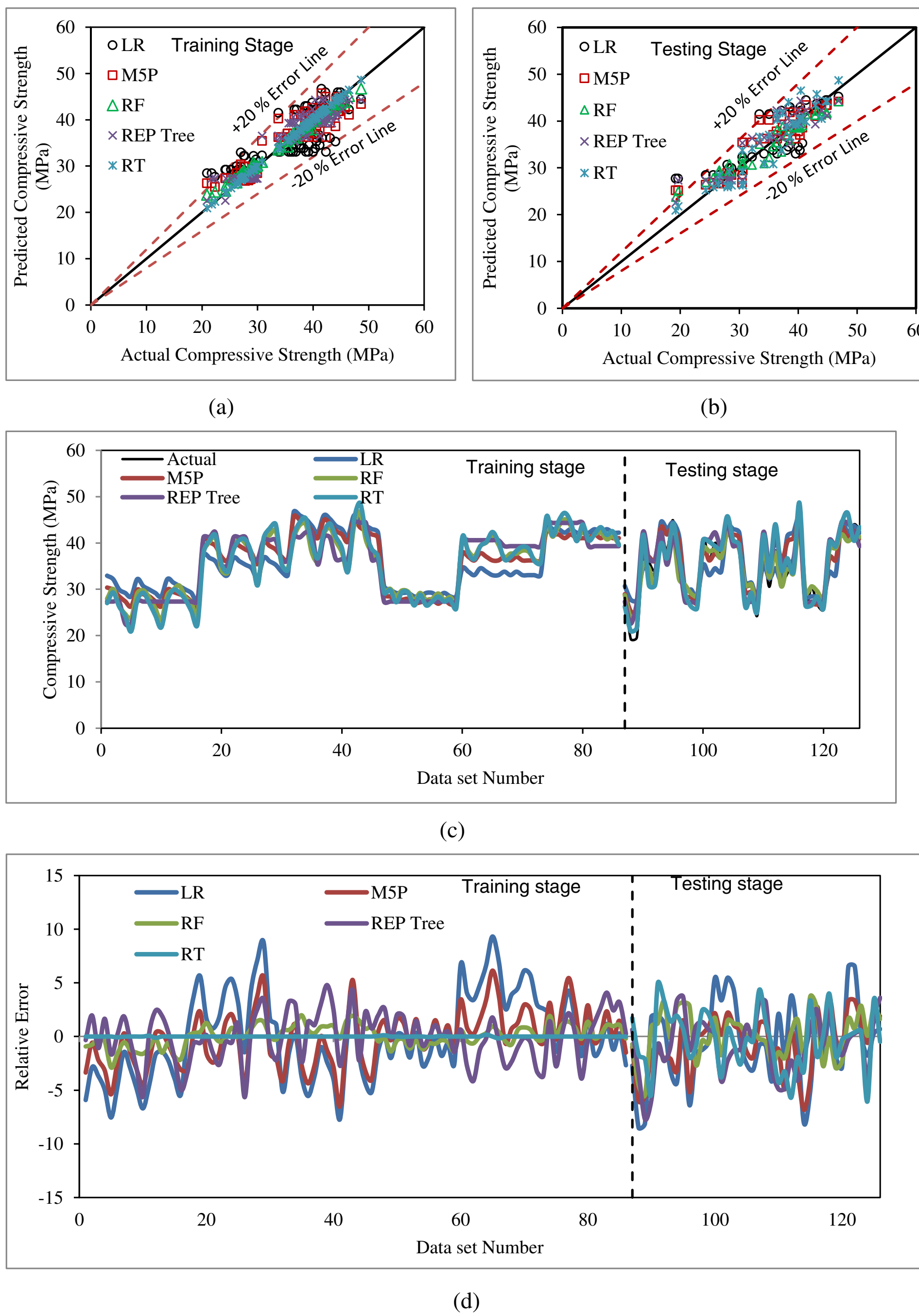


Fig. 9 Performance of various applied models for the prediction of compressive strength of concrete

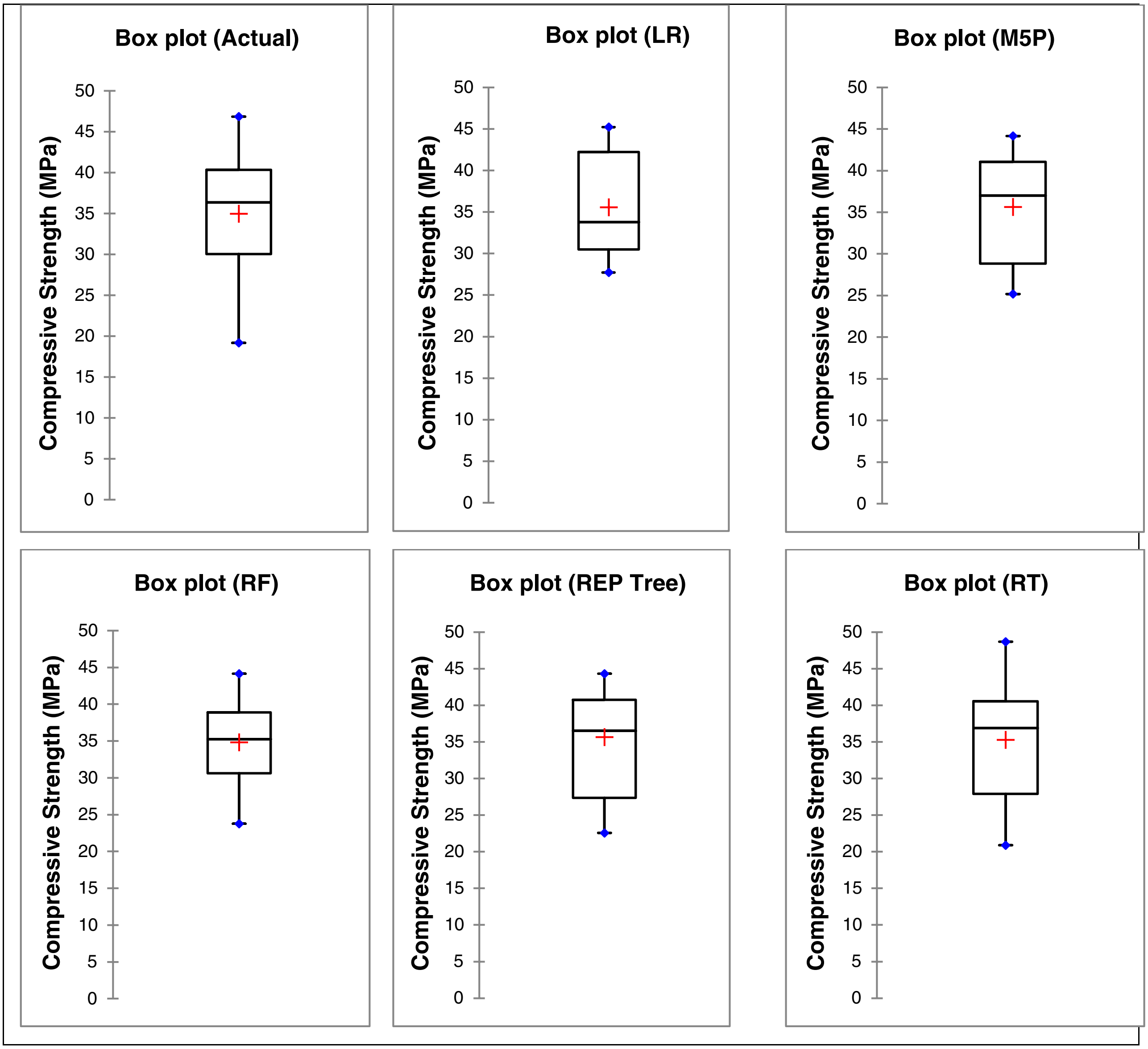


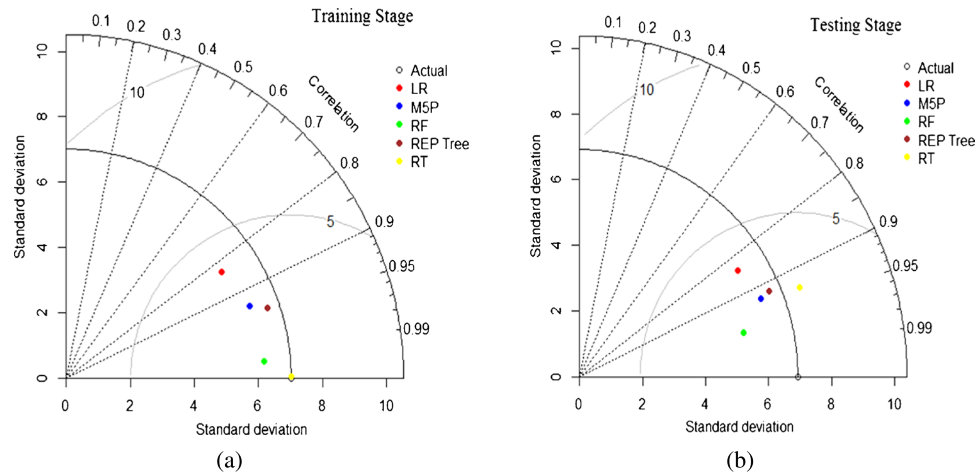
Fig. 10 Box plot of actual and applied models using testing stage

of REP tree-based model is shown in Fig. 6. Figure 7 provides plots of agreement among real and predicted concrete compressive strength by REP tree-based model for training and testing stages, respectively. Estimated values from REP tree-based model closely follow the observed values. Performance evaluation parameters depict that the work of the REP tree approach is suitable for predicting concrete's compressive strength with CC, RMSE and MAE values which are 0.9461,

Table 10 Descriptive statistics (actual vs. applied models: testing stage)

Statistic	Actual	LR	M5P	RF	REP Tree	RT
Minimum	19.200	27.714	25.199	23.755	22.570	20.910
Maximum	46.870	45.203	44.170	44.149	44.324	48.670
1st quartile	30.050	30.507	28.821	30.617	27.341	27.925
Mean	34.963	35.569	35.613	34.823	35.688	35.280
3rd quartile	40.333	42.225	41.045	38.884	40.741	40.560
IRQ	10.283	11.718	12.224	8.268	13.400	12.635

Fig. 11 Taylor's Diagram for actual and predicted values using all applied models for training and testing stages



2.2722 MPa and 1.8140 MPa for training stage and 0.9180, 2.8478 MPa and 2.2613 MPa for testing stage, respectively.

Assessment of RT-based model

RT-based model is formed using trial and error technique similar to other applied models development. Figure 8 provides plots of agreement among actual and predicted concrete compressive strength by RT-based model for training and testing stages, respectively. Estimated values using RT-based model mostly lies within the ± 20% error band. Overall, the performance of RT-based technique is suitable for forecasting the concrete compressive strength with CC = 1.00, 0.9323; RMSE = 0.0365 MPa, 2.7331 MPa; and MAE = 0.0115 MPa, 2.1613 MPa for training and testing stages, respectively.

Inter comparison among applied models

On comparing soft computing-based models (Table 9 and Fig. 9), it was found that RF-based model works better than other regression and soft computing-based applied designs. To evaluate the potential of regression and soft computing-based designs for estimating the concrete compressive strength, agreement, performance, and error were plotted in Fig. 9 for training and testing stages.

It is incidental from the plots that the estimated values produced by RF-based technique were in extremely proximity to the actual concrete compressive strength and predicted compressive strength of concrete values follow analogous design as that of actual compressive strength of concrete. M5P-based model outperforms RT-, REP- and LR-based models. Box plot is also plotted in Fig. 10 for the comparison of actual and predicted values considering various applied models for testing stage. Descriptive statistics of actual and applied models using testing stage are enlisted in Table 10.

Figure 10 and Table 10 suggest that RF model is outperforming in comparison to other applied models. Minimum and maximum values of actual and predicted values using RF model are very close. Widths of the lower and upper quartile are almost the same for actual and RF model in Fig. 10. Taylor diagram is a graphical illustration of the performance of developed models in terms of correlation, RMSE and standard deviation which is shown in Fig. 11.

Table 11 Sensitivity analysis using RF-based model with testing stage

Input combination					Target	RF based model (Testing Stage)		
Cement (%)	*GGBS* (%)	*LFS* (%)	*SCBA* (%)	*Curing time* (days)	CS (MPa)	CC	MAE (MPa)	RMSE (MPa)
						0.9687	1.7808	2.1887
**						0.9607	1.7342	2.1493
	**					0.9480	1.9304	2.3692
		**				0.9743	1.5246	1.8884
			**			0.9632	1.6549	2.0563
				**		**0.0085**	**7.3225**	**8.0989**

*Bold values shows most significant input variable

**White box in brown background shows the relevant column element is missing from the data

Figure 11 indicates that RF is best performing model and the performance of LR model is least for the estimation of concrete compressive strength using the current data set.

Sensitivity analysis

Sensitivity test was conducted to ascertain significant input variables in compressive strength using RF model. Different set of testing data was designed by eliminating one input parameter at a time as shown in Table10. The white box shows the elimination of that element from the dataset in analysis so as to observe its effect on stat diagnostic parameter. The effect of each input variable on the compressive strength was reported in terms of CC, MAE and RMSE. Results from Table 11 suggests that curing time in days (d) plays a significant role in predicting the compressive strength of concrete as compared to other input parameter using this data set. Even Table 6 shows that curing time in days (d) has the highest CC of 0.7976 in all other parameters validating the sensitivity analysis results.

Conclusion

In the present study, all the mentioned additives were added in the concrete mix as the replacement of OPC up to 35%. On the basis of the findings, it was observed that the 20% of all the additives in individual form maybe used as the partial substitution of OPC, while in a combined form, concrete mix having 5% GGBFS, 10% LFS and 15% of SCBA was showing the optimum strength value. However, it was also observed that the greater percentage of all the additives can be utilized with an increment in the curing time period. In this investigation, M5P, random forest (RF), random tree and REP tree have been developed to predict compressive strength of concrete and compared with linear regression (LR). For this purpose, the experiments were performed with the variation in proportion of waste materials, i.e. 0% 5%, 10%, 15%, 20%, 25%, 30% and 35% as replacement of cement (individually as well as in a combined form) by GGBFS, LFS and SCBA. The comparison analysis using performance evaluation indices concludes that developed RF approach outperformed the rest of the models (M5P, RT, REP tree and LR) using the given data set with CC = 0.9967, 0.9687; RMSE = 0.9791 MPa, 2.1887 MPa; and MAE = 0.7918 MPa, 1.7808 MPa for model development and validation period, correspondingly. Other major outcomes from this investigation are that M5P-based model performs better than RT-, REP tree- and LR-based models. To evaluate each input variable's effect on the proposed best performing model, sensitivity investigation is also carried out using best performing RF model. Outcomes of the sensitivity analysis indicates that curing time in days (d) is the most effective input variable for predicting the compressive strength of concrete using this data set.

Declarations

Conflict of interest The authors declare no competing interests.

References

Ali MB, Saidur R, Hossain MS (2011) A review on emission analysis in cement industries. Renew Sustain Energy Rev 15(5):2252–2261. https://doi.org/10.1016/j.rser.2011.02.014

Amin N (2011) Use of bagasse ash in concrete and its impact on the strength and chloride resistivity. J Mater Civ Eng 23(5):717–720. https://doi.org/10.1061/(ASCE)MT.1943-5533.0000227

Anupam AK, Kumar P, Ransinchung GD, Shah Yu (2017) Study on performance and efficacy of industrial waste materials in road construction: fly ash and bagasse ash. Airfield and Highway Pavements, DOI 10(1061/9780784480946):005

Barddal JP, Enembreck F, Gomes HM, Bifet A, Pfahringer B (2019) Merit-guided dynamic feature selection filter for data streams. Expert Syst Appl 116:227–242. https://doi.org/10.1016/j.eswa.2018.09.031

Behnood A, Golafshani EM (2020) Machine learning study of the mechanical properties of concretes containing waste foundry sand. Constr Build Mater 243:118152. https://doi.org/10.1016/j.conbuildmat.2020.118152

BIS (Bureau of Indian Standards). (1959). Methods of test for strength of concrete, IS: 516, BIS, New Delhi, India.

BIS (Bureau of Indian Standards). (1988a). Methods of physical tests for hydraulic cement. Determination of initial and final setting times, IS 4031 Part 5, New Delhi, India.

BIS (Bureau of Indian Standards) (1988) Methods of physical tests for hydraulic cement: Determination of density, IS 4031 Part 11. New Delhi, India

BIS (Bureau of Indian Standards). (1997). Methods of test for aggregates for concrete, IS: 2386, BIS, New Delhi, India.

BIS (Bureau of Indian Standards). (1999). Concrete Admixtures Specification, IS: 9103, BIS, New Delhi, India.

BIS (Bureau of Indian Standards). (2005). Methods of physical tests for hydraulic cement. Determination of consistency of standard cement paste, IS 4031 Part 4, New Delhi, India.

BIS (Bureau of Indian Standards). (2007) Specification for coarse and fine aggregates from natural sources for concrete, IS 383–07. New Delhi, India

BIS (Bureau of Indian Standards). (2019) Concrete Mix Proportioning – Guidelines, IS 10262. New Delhi, India

BIS (Bureau of Indian Standards). (1959). Method of tests for strength of Concrete, IS 516, Bureau of Indian Standards, New Delhi.

BIS (Bureau of Indian Standards) (1970) Specification for coarse and fine aggregates from natural sources for Concrete, IS 383, New Delhi.

Marinho BAL, Mol Santos CM, Carvalho, J.M.F.de., Mendes, J.C., Brigolini, G.J., & André Fiorotti Peixoto, R. (2017) Ladle furnace slag as binder for cement-based composites. J Mater Civ Eng 29(11):04017207. https://doi.org/10.1061/(asce)mt.1943-5533.0002061

Breiman L. (1996). Bagging predictors. *Machine learning*.24(2):123–40.
Breiman, L. (1999). Random forests. UC Berkeley TR567.
Breiman L (2001) Random forests. Mach Learn 45(1):5–32
Cutler A, Cutler DR, Stevens JR (2012) Random forests. Ensemble machine learning. Springer, Boston, MA, pp 157–175
Das S, Singh G, Ahmed AA, Saha S, Karmakar S (2015) Ground Granulated Blast Furnace Slag (GGBFS) based Concrete Exposed to Artificial Marine Environment (AME) and Sustainable Retrofitting using Glass Fiber Reinforced Polymer (GFRP) sheets. Procedia Soc Behav Sci 195:2804–2812. https://doi.org/10.1016/j.sbspro.2015.06.397
Deepika S, Anand G, Bahurudeen A, Santhanam M (2017) Construction products with sugarcane bagasse ash binder. J Mater Civ Eng 29(10):04017189. https://doi.org/10.1061/(asce)mt.1943-5533.0001999
Esmaeilbeiki, F., Nikpour, M. R., Singh, V. K., Kisi, O., Sihag, P., &Sanikhani, H. (2020). Exploring the application of soft computing techniques for spatial evaluation of groundwater quality variables. *Journal of Cleaner Production*, 276, 124206.
Farooq F, Nasir Amin M, Khan K, Rehan Sadiq M, Faisal Javed M, Aslam F, Alyousef R (2020) Comparative study of random forest and genetic engineering programming for the prediction of compressive strength of high strength concrete (HSC). Appl Sci 10(20):7330
Flatt RJ, Roussel N, Cheeseman CR (2012) Concrete: An eco material that needs to be improved. J Eur Ceram Soc 32(11):2787–2798
Gartner, E.M. (2011). Potential improvements in cement sustainability. *In Proceedings of the 31st Cement and Concrete Science Conference Novel Developments and Innovation in Cementitious Materials*, London, UK, 12 September; pp. 12–13.
Golafshani EM, Behnood A (2021) Predicting the mechanical properties of sustainable concrete containing waste foundry sand using multi-objective ANN approach. Constr Build Mater 291:123314. https://doi.org/10.1016/j.conbuildmat.2021.123314
Hamoud A, Hashim AS, Awadh WA (2018) Predicting student performance in higher education institutions using decision tree analysis. International Journal of Interactive Multimedia and Artificial Intelligence 5:26–31
Indian Road Congress (IRC). (2008). Tentative guidelines for cement concrete mix design for pavements. IRC: 44, New Delhi, India.
Kandiri A, Golafshani EM, Behnood A (2020) Estimation of the compressive strength of concretes containing ground granulated blast furnace slag using hybridized multi-objective ANN and salp swarm algorithm. Constr Build Mater 248:118676. https://doi.org/10.1016/j.conbuildmat.2020.118676
Kashi H, Emamgholizadeh S, Ghorbani H (2014) Estimation of soil infiltration and cation exchange capacity based on multiple regression, ANN (RBF, MLP), and ANFIS models. Commun Soil Sci Plant Anal 45:1195–1213
Maghool F, Arulrajah A, Haghighi H, Horpibulsuk S, Du Y-J (2017) The influence of a curing regime on the geotechnical properties of ladle furnace slag as used in pavement applications. Geotechnical Frontiers DOI 10(1061/9780784480441):034
Mangi, S.A., Jamaluddin, N., Wan Ibrahim, M.H., Abdullah, A.H., Abdul Awal, A.S.M.,Sohu, S., &Ali, N. (2017). Utilization of sugarcane bagasse ash in concrete as partial replacement of cement. *IOP Conf. Ser. Mater. Sci. Eng*, 271.
Manso JM, Losañez M, Polanco JA, Gonzalez JJ (2005) Ladle furnace slag in construction. J Mater Civ Eng 17(5):513–518. https://doi.org/10.1061/(asce)0899-1561
Mehta A, Siddique R (2018) Sustainable geopolymer concrete using ground granulated blast furnace slag and rice husk ash: strength and permeability properties. J Clean Prod. https://doi.org/10.1016/j.jclepro.2018.08.313
Mehta, P.K. (1997). Bringing the concrete industry into a new era of sustainable development. *In Proceedings of the Symposium on Advances in Concrete Science and Technology, Fifth CANMET/ACI International Conference Proceedings*, Rome, Italy, pp. 49–67.
Mohammed A, Burhan L, Ghafor K, Sarwar W, Mahmood W (2020c) Artificial neural network (ANN), M5P-tree, and regression analyses to predict the early age compression strength of concrete modified with DBC-21 and VK-98 polymers. Neural Comput Appl. https://doi.org/10.1007/s00521-020-05525-y
Mohammed, A., Rafiq, S., Sihag, P., Kurda, R., & Mahmood, W. (2020a). Soft computing techniques: systematic multiscale models to predict the compressive strength of HVFA concrete based on mix proportions and curing times. *Journal of Building Engineering*, 101851.
Mohammed A, Rafiq S, Sihag P, Mahmood W, Ghafor K, Sarwar W (2020b) ANN, M5P-tree model, and nonlinear regression approaches to predict the compression strength of cement-based mortar modified by quicklime at various water/cement ratios and curing times. Arab J Geosci 13(22):1–16
Quinlan, J.R. (1987). Generating production rules from decision trees. *Inijcai*, (Vol. 87, pp. 304–307). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.98.9054&rep=rep1&type=pdf
Quinlan, J.R. (1992). Learning with continuous classes. *In5th Australian joint conference on artificial intelligence.* (Vol. 92, pp. 343–348). https://doi.org/10.1142/9789814536271
Rađenović A, Malina J, Sofilić T (2013) Characterization of ladle furnace slag from carbon steel production as a potential adsorbent. Adv Mater Sci Eng 1–6. https://doi.org/10.1155/2013/198240
Raghavendra T, Udayashankar BC (2014) Flow and strength characteristics of CLSM using ground granulated blast furnace slag. J Mater Civ Eng 26(9):04014050. https://doi.org/10.1061/(asce)mt.1943-5533.0000927
Sihag P, Fatemeh E, Balraj S, Isha E, Hossein B (2019) Modelling unsaturated hydraulic conductivity by hybrid soft computing techniques. Springer-Verlag GmbH Germany, Part of Springer Nature. https://doi.org/10.1007/s00500/019-03847-1
Sihag P, Suthar M, Mohanty S (2021) Estimation of UCS-FT of dispersive soil stabilized with fly ash, cement clinker and GGBS by artificial intelligence. Iran J Sci Technol 45(2):901–912. https://doi.org/10.1007/s40996-019-00329-0
Sobhani, J., Ejtemaei, M., Sadrmomtazi, A., &Mirgozar, M.A. (2019). Modeling flexural strength EPS light weight concrete using regression, neural network and ANFIS. *International Journal of optimization in Civil Engineering*, 9(2): 313–329.URL: http://ijoce.just.ac.ir/article-1-392-en.html.
Suthar M, Aggarwal P (2018) Predicting CBR value of stabilized pond ash with lime and lime sludge using ANN and MR models. International Journal of Geosynthetics and Ground Engineering 4(1):6.https://doi.org/10.1007/s40891-017-0125-3
Suthar, M., & Aggarwal, P. (2019). Modeling CBR value using RF and M5P techniques. *MENDEL*, 25(1), 73–78. https://doi.org/10.13164/mendel.2019.1.073
Suthar M (2020a) Applying several machine learning approaches for prediction of unconfined compressive strength of stabilized pond ashes. Neural Computing and Application 32:9019–9028. https://doi.org/10.1007/s00521-019-04411-6
Suthar M (2020b) Modeling of UCS value of stabilized pond ashes using adaptive neuro-fuzzy inference system and artificial neural network. Soft Comput 24:14561–14575. https://doi.org/10.1007/s00500-020-04806-x
Witten, I.H., Frank, E., Hall, M.A., & Pal, C.J. (2005). Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann.