**Computers and Concrete**, *Vol. 24, No. 1 (2019) 7-17*
DOI: https://doi.org/10.12989/cac.2019.24.1.007


# Applied linear and nonlinear statistical models for evaluating strength of Geopolymer concrete

Prabhat Ranjan Prem*, A. Thirumalaiselvi[a] and Mohit Verma[a]

*CSIR-Structural Engineering Research Centre, Chennai, India*

*(Received November 5, 2018, Revised April 26, 2019, Accepted April 30, 2019)*

**Abstract.** The complex phenomenon of the bond formation in geopolymer is not well understood and therefore, difficult to model. This paper present applied statistical models for evaluating the compressive strength of geopolymer. The applied statistical models studied are divided into three different categories - linear regression [least absolute shrinkage and selection operator (LASSO) and elastic net], tree regression [decision and bagging tree] and kernel methods (support vector regression (SVR), kernel ridge regression (KRR), Gaussian process regression (GPR), relevance vector machine (RVM)]. The performance of the methods is compared in terms of error indices, computational effort, convergence and residuals. Based on the present study, kernel based methods (GPR and KRR) are recommended for evaluating compressive strength of Geopolymer concrete.

**Keywords:** geopolymer concrete; modelling; compressive strength; kernel

## 1. Introduction

In view of the ongoing enormous infrastructure growth all around the globe, the consumption of cement by concrete industry has grown exponentially. Almost one tonne of carbon dioxide is released for each tonne of production ordinary Portland cement (OPC) (Pacheco-Torgal *et al*. 2008). A notable achievement to battle the demand of cement and reduce carbon dioxide emissions, research has yielded development of a cement free concrete known as Geopolymer. Geopolymer is an inorganic polymer developed by Davidovits (1991), is one such material which can replace cement in future. In geopolymer, the binders are produced through the process of polymerization activated by the reaction of an alkaline liquid with the Silicon and Aluminum yielding Si-O-Al-O bonds. Some of the popular materials used for development of geopolymer are kalonitic clays, coarse and fine fly ash development of geopolymer are kalonitic clays, coarse and fine fly ash (Ambily *et al*. 2014, Vijai *et al*. 2015, Jindal *et al*. 2018, Senthamilselvi and Palanisamy 2018, Yadollahia and Benli 2017, Katpady *et al*. 2017), fine and coarse rice husk bark ash (He *et al*. 2013), metakaolin (Marin-Lopez *et al*.2009), granulated blast furnace slag (Mozumder *et al*. 2017, Kurklu 2016, Prem *et al*. 2018b) and red mud (Cundi *et al*. 2005, Zhang *et al*. 2014). The combination of NaOH or KOH and sodium silicate or potassium silicate is usually used as alkaline liquid. The mechanical properties of the geopolymer are also dependent on the optimum curing conditions (Bakharev 2005). In general, the concrete compressive strength is dependent on various factors such as physical and chemical properties, curing conditions, the particle size of aggregates, etc. Hence to overcome such randomness various researchers have applied numerical (Prem *et al*. 2017, Arani *et al*. 2019), theoretical (Verma *et al*. 2016a, Prem *et al*. 2018a, Al-Rousan *et al*. 2018) or machine learning approach to examine its behavior (Verma *et al*. 2015, Dutta *et al*. 2018, Erdal *et al*. 2018, Verma *et al*. 2019). But such studies are available for conventional concrete and very limited for geopolymer concrete. The present research is motivated from this gap and hence detailed investigations on geopolymer concrete are presented in this study.

*Corresponding author, Ph.D.
E-mail: prabhat@serc.res.in
[a]Scientist

## 2. Research significance

The complex phenomenon of the bond formation in geopolymer is not well understood. In the absence of any theoretical model, it becomes difficult to arrive at the optimum mix to achieve desired mechanical properties. In order to provide guidance towards optimum selection of raw materials and curing procedure, authors in this study propose to predict the geopolymer compressive strength using regression algorithms. The use of computational method will enable to cut down costly lab experiments and provide inputs for designing geopolymer mix. The methods reported earlier by most of the researchers for the evaluation of the properties of cementitious composites are based on artificial neural networks (ANN) and fuzzy logic. The disadvantage of fuzzy logic is that the rules are based on the human heuristic knowledge while the ANN tends to over fit the data. In the present study, methods based on linear regression (Least absolute Shrinkage and Selection Operator (LASSO) and Elastic Net), tree regression (decision and bagging tree) and kernel methods (Support vector regression (SVR), Kernel ridge regression (KRR),



ISSN: 1598-8198 (Print), 1598-818X (Online)

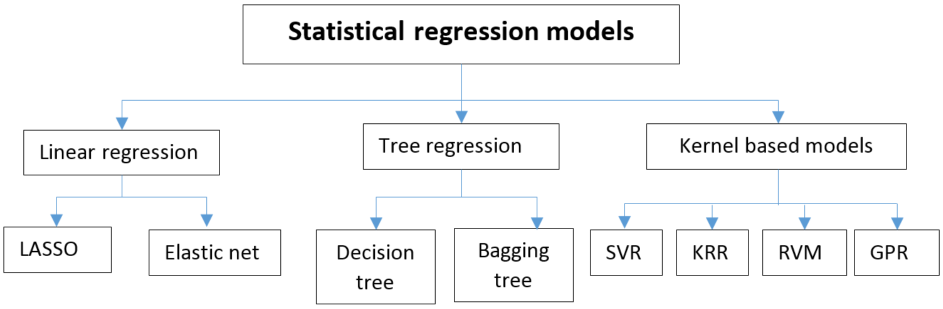



Fig. 1 Regression methods used in the present study

Table 1 Statistical properties of the dataset

| | FFA (%) | CFA (%) | FRHBA (%) | CRHBA (%) | T (o C) | C (days) | CS (MPa) |
|---|---|---|---|---|---|---|---|
| Maximum | 80.00 | 80.00 | 40.00 | 40.00 | 90.00 | 28.00 | 58.90 |
| Minimum | 0.00 | 0.00 | 0.00 | 0.00 | 25.00 | 7.00 | 14.80 |
| Mean | 35.00 | 35.00 | 15.00 | 15.00 | 59.00 | 17.50 | 33.89 |
| Standard deviation | 35.62 | 35.62 | 16.14 | 16.14 | 24.27 | 10.54 | 9.78 |
| Median | 30.00 | 30.00 | 10.00 | 10.00 | 60.00 | 17.50 | 32.90 |
| Variance | 1268.91 | 1268.91 | 260.50 | 260.50 | 588.91 | 111.18 | 95.58 |

Gaussian process regression (GPR) and Relevance vector machine (RVM)) are presented for evaluating compressive strength of geopolymer. From the review, it is found that the critical parameters which affects the mechanical properties of geopolymer are (i) type of raw material (ii) curing temperature and (iii) curing duration. The data published in Bohlooli *et al*. (2012) is adopted in the presented study. The performance of the methods is then compared in terms of error indices, convergence, residuals and computational cost. The methods were also validated with the trends reported in the literature. The present study focuses on the performance of the different applied statistical models for evaluating the compressive strength of geopolymer.

## 3. Data set

Geopolymer compressive strength largely depends upon the content of Si and Al. Fly ash (FA) is a rich source of silicon dioxide and aluminum oxide. It has been extensively used in geopolymer production. Rice husk bark ash (RHBA) is also high silica source (containing as high as 75% silicon dioxide). The ratio of Si to Al can be adjusted by blending fly ash with RHBA. The seeded distribution of fly ash and RHBA is utilized for production of geopolymer. Two different seeded distribution of RHBA and FA forms the input to the methods along with curing time and temperature. The data used in the current study has been adopted from Bohlooli *et al*. (2012). The method has been trained for 70 datasets while remaining 50 are used for testing. The statistics of the data used are given in Table 1. where FFA, CFA, FRHBA, CRHBA, T, C, and CS represents fine FA, coarse FA, fine RHBA, coarse RHBA, temperature, curing age and compressive strength of geopolymer concrete, respectively.

Table 2 Theoretical comparison of different regression methods

| | Core algorithm | Loss function | Regularization |
|---|---|---|---|
| LASSO | Penalized regression | Quadratic | $l_1$-norm |
| Elastic net | Penalized regression | Quadratic | $l_1$; $l_2$-norm |
| Decision tree | Sorting & grouping | Quadratic | $l_2$-norm |
| Bagging tree | Bootstrap aggregation | Quadratic | $l_2$-norm |
| SVR | Structural risk minimization | Quadratic, hinge, $\epsilon$-insensitive | $l_2$-norm |
| KRR | Matrix inversion | Quadratic | $l_2$-norm |
| RVM | Bayesian statistical inference | Marginal likelihood | Probabilistic |
| GPR | Bayesian statistical inference | Marginal likelihood | Probabilistic |

## 4. Methodology and method application

In this section, the different applied statistical models used for evaluating geopolymer compressive strength are described briefly. These methods are broadly classified into three categories linear, tree and kernel methods as shown in Fig 1. The methods are then applied to evaluate the compressive strength of geopolymer using simpleR toolbox (Camps-Valls *et al*. 2013) in Matlab. Theoretically, the objective function in each of the methods consists of two parts - loss function which measures how well method fit on training data and regularization which measures the complexity of method. The type of loss function and regularization used by the regression methods are summarized in Table 2.

### *4.1 Lasso*

LASSO is a shrinkage and selection method for performing linear regression. The size of the predicted coefficients is constrained using a penalty term. The coefficients are generated in a way that the bias is small. Consider the set of data represented by $\{(x_i, y_i), i = 1, .., N\}$, $X \in \mathbb{R}^n$, $Y \in \mathbb{R}$.

A formulation based on LASSO was developed by Tibshirani (1996). The following optimization problem is solved in LASSO for a given $\lambda$

$$min_{\beta_o \beta} \left( \frac{1}{2N} \sum_{i=1}^{N} (y_i - \beta_o - x_i^T \beta)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right) \quad (1)$$

where $x_i$ is input data vector ($p \times N$), $y_i$ is the output at $i^{th}$

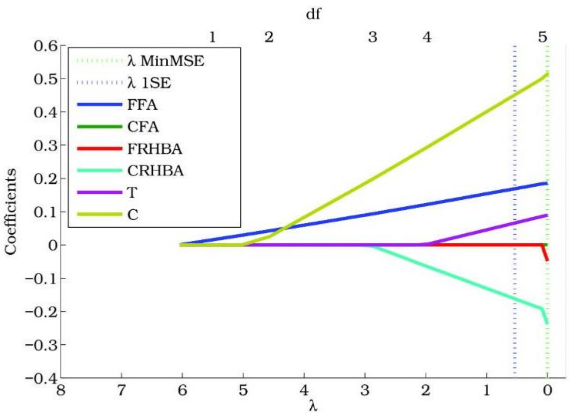



Fig. 2 Trace plot of coefficients for LASSO

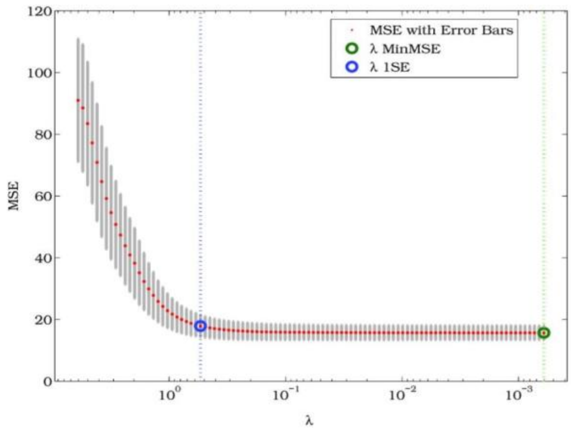



Fig. 3 Cross-validation for LASSO

observation, $N$ is the number of inputs, $\lambda$ is regularization parameter. The parameters $\beta_0$ and $\beta$ in the above equation are scalar and $p$-vector, respectively.

The training data is used for LASSO fit. The trace plot of the coefficients fit by LASSO in Fig. 2 shows the variation of the nonzero coefficients for the different values of $\lambda$. The larger value of implies more regularization and hence, results in lesser nonzero coefficients. The green dashed vertical lines represent the $\lambda$ value corresponding to the minimum mean squared error (MSE) and the blue one with MSE plus one standard deviation. The top portion of the plot shows the number of nonzero coefficients (degree of freedom, *df*) in the regression, as a function of $\lambda$. From the plot, it is found that important parameters affecting the geopolymer compressive strength are curing time, fine fly ash, coarse RHBA and the curing temperature. The cross-validation plot of the LASSO is shown in Fig. 3. Five-fold cross-validation is used in the present study to avoid overfitting and to generalize the method to an independent dataset. As $\lambda$ increases, MSE is also found to increase drastically and results in overfitting of the method. The feature relevance plot of the LASSO is shown in Fig. 4. The correlation between the experimental and predicted compressive strength of geopolymer is shown in Fig. 11(a).

### 4.2 Elastic net

The major drawback with the LASSO is that the selection of the number of variables is limited by number of observations. Overcoming this drawback, Zou and Hastie (2005) introduced elastic net which can be stated as

$$min_{\beta_o\beta}\left(\frac{1}{2N}\sum_{i=1}^{N}(y_i-\beta_o-x_i^T\beta)^2+\lambda P_\alpha(\beta)\right) \quad (2)$$

Where

$$P_\alpha(\beta)=\frac{(1-\alpha)}{2}\left|\left|\beta\right|\right|_2^2+\alpha\left|\left|\beta\right|\right|_1$$
$$=\sum_{j=1}^{p}(\frac{(1-\alpha)}{2}\beta_j^2+\alpha|\beta_j|)$$

Elastic net reduces to LASSO for $\alpha$=1. The value of α used in the present study is 0.5. Five-fold cross-validation is used for elastic net. The trace plot of the coefficients fit by elastic net is shown in Fig. 5. The parameters affecting the compressive strength of the geopolymer according to elastic net are curing time, fine fly ash, coarse RHBA, curing temperature and coarse fly ash. The feature relevance plot of the elastic net is shown in Fig. 6. The corresponding cross-validation plot is shown in Fig. 7. The correlation between the experimental and predicted compressive strength of geopolymer is shown in Fig. 11(b).

### 4.3 Decision tree

Decision tree typically represents a set of constraints which are hierarchically ordered, and are successively applied from root to terminal mode (Quinlan 1993). In

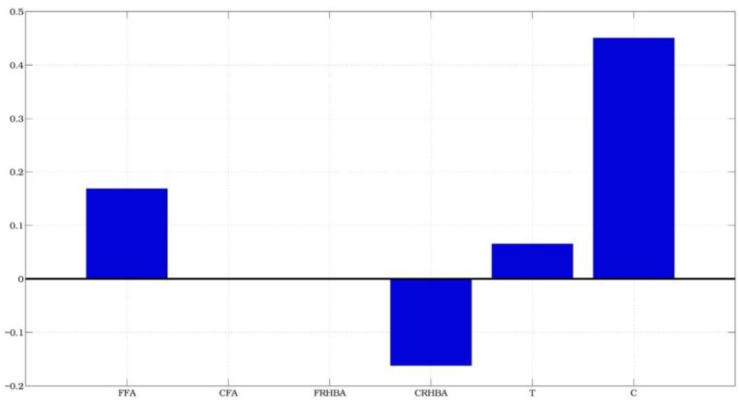



Fig. 4 Feature relevance for LASSO

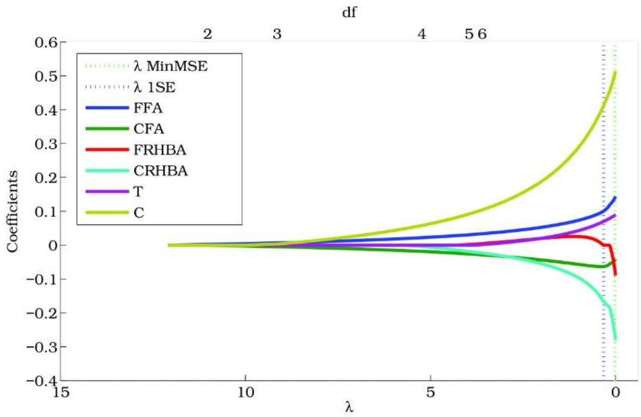



Fig. 5 Trace plot of coefficients for elastic net

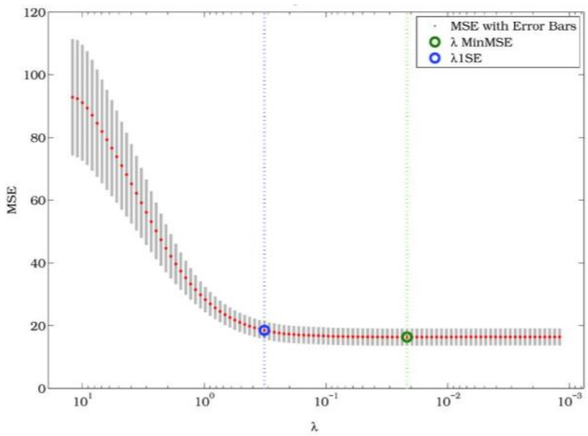



Fig. 7 Cross-validation for elastic net

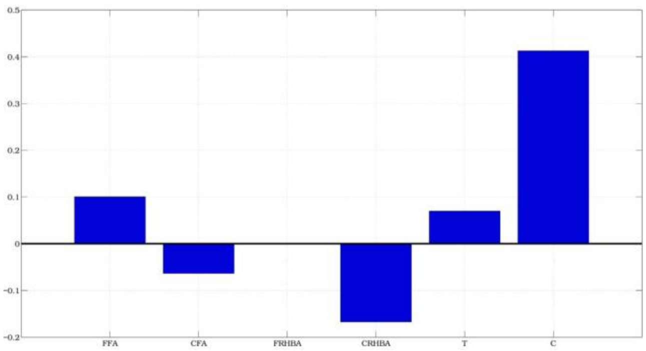



Fig. 6 Feature relevance for elastic net

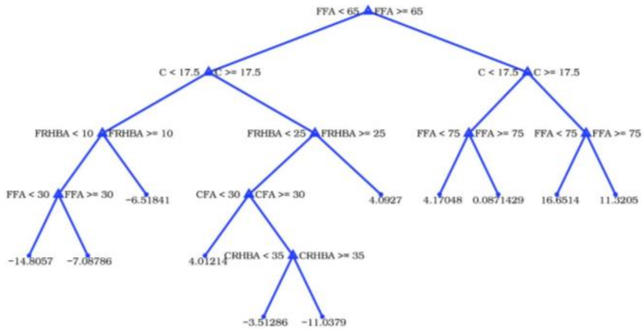



Fig. 8 Structure of the decision tree

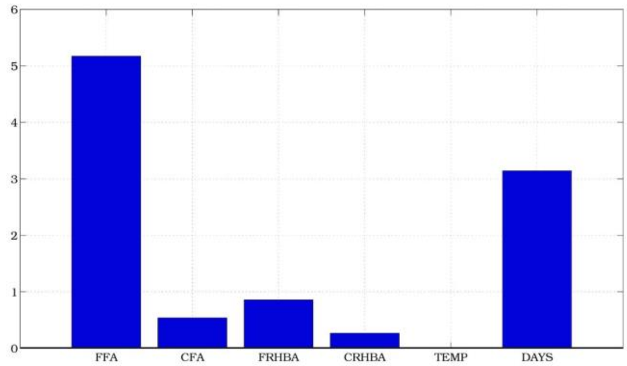



Fig. 9 Feature relevance for decision tree

decision tree, the parent node is split into binary pieces such that the following impurity function is minimized (Breiman *et al*. 1993)

$$min\, \Delta j(r,u) = j(u) - W_L\, j(u_L) \ - W_R j(u_R) \qquad (3)$$

In the equation the candidate split ($r$) at each node ($u$) is divided into left ($u_L$) and right node ($u_R$) having weights $W_L$ and $W_R$, respectively. The impurity in the system is denoted by $j(u)$, while $j(u_L)$ and $j(u_R)$ indicates impurity after splitting. $\Delta j(r,\, t)$ denotes impurity after split $r$.

From the forecasting tree, positive weight $w_n(x_n, \Omega)$ for each case $x \in \mathbb{R}$ *is* computed. If $l(x,\Omega,t)$ is assumed to be a node $t$ then each conditions of $x_n \in l(x,\Omega,t)$ are assigned equal weight $w_n\,(x,\Omega) = 1/N\,(t)$, where $N(t)$ denotes all cases in $l(x,\Omega,t)$. Further on considering $X = x$, the output $y$ is evaluated from Eq. (4)

$$y \ = \sum_{n=1}^{N} w_n\,(x,\Omega) Y_n$$
$$= \sum_{x,X_n \in l(x,\Omega,t)} w_n\,(x,\Omega) Y_n \qquad (4)$$

The decision tree obtained using the training data is shown in Fig. 8. The feature relevance for decision tree is shown in Fig. 9. Ten-fold cross- validation is used for elastic net. The corresponding cross-validation plot is shown in Fig. 10. The square marker shows the minimum cost and the dashed line is one standard deviation from the minimum. The correlation between experimental and elastic net fit is shown in Fig. 11(c).

### *4.4 Bagging tree*

Decision tress is a non-parametric regression method that generally gives satisfactory results for nonlinear cases. But the main drawback of this method is absence of stable prediction rule. To overcome this limitation all, the decision trees obtained from the bagged samples are considered. The output by bagging is the average prediction at $x$ from $L$ trees (Breiman 1996), defined by

$$y = \frac{1}{L}\sum_{i=1}^{K} y_k\,(x) \qquad (5)$$

The correlation between experimental and compressive strength of geopolymer predicted by bagging tree is shown in Fig. 11(d).

### *4.5 SVR*

In SVR, the aim is to approximate the set of data given

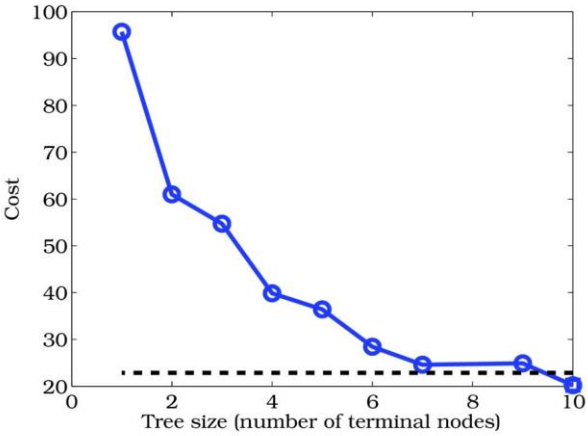



Fig. 10 Cross-validation for decision tree

by $\{(x_i, y_i), i = 1, .., N\}$, $X \in \mathbb{R}^n$, $Y \in \mathbb{R}$. Mathematically, it can be expressed as (CampsValls *et al*. 2006, Vapnik 2013)

$$y = f(x) = \langle w, \emptyset(x) \rangle + b \tag{6}$$

where $\langle . \rangle$ is the inner product operator, $w$ and $b$ are the function parameters and $\varphi(x)$ is the kernel function. SVR then runs the linear regression in the output space. The objective of the support vector approach is to minimize the following equation

$$\frac{1}{2}||w||_2^2 = C \sum_{n=1}^{N} L_1^{\epsilon} (y_n, \langle w, \emptyset(x_n) \rangle + b) \tag{7}$$

where $L_1^{\epsilon}$ is the loss function which determines how the SVR error is penalized and $C$ is the hyperparameter which is varied to arrive at the optimal solution.

Radial basis and $\epsilon$-insensitive loss function are used for SVR. The value of the hyper- parameters in evaluated using grid search which involves exhaustive searching through a manually specified range of the hyperparameter. The hyperparameter $C$, and radial basis function width used in the present study are 1000, 0.1 and 55.15, respectively. More details of the SVR method is given in Verma *et al*. (2016b). The correlation between the experimental and compressive strength evaluated using SVR is shown in Fig. 11(e).

### *4.6 KRR*

KRR is also known as least square SVR. KRR finds the solution by solving a set of linear equations while SVR needs to solve a quadratic optimization program. The optimization problem for KRR can be written as (Camps-Valls *et al*. 2012)

$$\min \frac{1}{2}||w||_2^2 + \frac{1}{2} C \sum_{n=1}^{N} e_n^2 \tag{8}$$

subject to: $y_n = \langle w, \emptyset(x) \rangle + b + e_n \, , n = 1 \ldots\ldots N$

where $e_n$ represents the error from the training set and $C$ is the penalty parameter. The kernel function used for KRR is radial basis. The hyperparameters for KKR were obtained using grid search. About one-third of the training data is kept aside for validation. The value of the penalty parameter $C$ and the width of the radial basis function obtained after grid search are $1.64 \times 10^{-5}$ and 145.42, respectively. The correlation between the experimental and predicted compressive strength of geopolymer is shown in Fig. 11(f).

### *4.7 RVM*

RVM, proposed by Tipping (2001), uses a sparse Bayesian learning whose functional form is similar to SVM. The output of the RVM is given by

$$y(x) = \sum_{n=1}^{N} w_n K(x, x_n) + w_0 \tag{9}$$

where $K(x, x_n)$ is a kernel function, $w_n$ are the model weights and $w_0$ is bias. For an input-target pair $(x, t)$ it is assumed that $p(t|x)$ is Gaussian $N(t|y(x), \sigma^2)$. The dataset's likelihood can be expressed as (Tipping 2001)

$$p(t|w, \sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} exp\left\{-\frac{1}{2\sigma^2}||t - \emptyset w||^2\right\} \tag{10}$$

where $\emptyset$ is $N \times (N + 1)$ with $\emptyset_{nm} = K(x, x_{m-1}$ and $\emptyset_{nm} = 1$. The posterior over the weights is written as (Tipping 2001)

$$p(w|t, \alpha, \sigma^2) = (2\pi)^{-\frac{N+1}{2}} |\Sigma|^{-\frac{1}{2}} exp \left\{-\frac{1}{2}(w - \mu)^T \Sigma^{-1}(w - \mu)\right\} \tag{11}$$

$$\Sigma = (\emptyset^T B \emptyset + A)^{-1} and\ \mu = \Sigma \emptyset^T B t$$

where $A = \mathrm{diag}(\alpha_0, \cdots, \alpha_n)$, $B = \sigma^{-2} I_n$ with ($\alpha_i$'s) as the hyperparmeters whose marginal likelihood is given by (Tipping 2001)

$$p(t|\alpha, \sigma^2) = (2\pi)^{-\frac{N}{2}}$$

$$|B^{-1} + \emptyset A^{-1} \emptyset^T|^{-\frac{1}{2}} exp\left\{-\frac{1}{2} t^T (B^{-1} + \emptyset A^{-1} \emptyset^T)^{-1} t\right\} \tag{12}$$

Finally, the marginal likelihood for the hyperparameters is maximized over $\sigma^2$

Radial basis function is used as kernel function. For this method also the hyperparameters are obtained using grid search. About one-third of the training data is kept aside for validation. The value of the parameter $\sigma^2$ and the width of the radial basis function obtained after grid search are 0:0694 and 145.42, respectively. The correlation between the experimental and predicted compressive strength of geopolymer is shown in Fig. 11(g).

### *4.8 GPR*

GPR is a non-parametric technique which uses Bayesian framework for solving nonlinear problems. More details on the GPR can be found in (Rasmussen and Williams 2006, Verrelst *et al*. 2012, Verma *et al*. 2016). The output of the GPR is given by

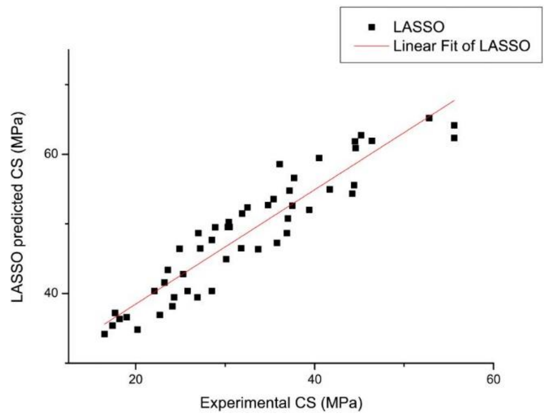
$$y = f(x) + \epsilon \tag{13}$$

where $(f(x))$ is a latent function and $(\epsilon)$ is the Gaussian noise. The output of GPR follows the following distribution

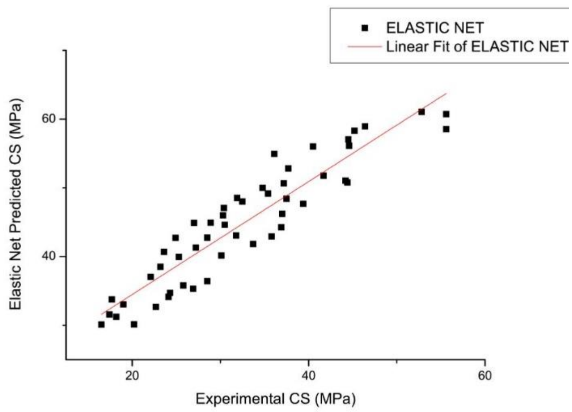
$$\begin{pmatrix} y \\ y_t \end{pmatrix} \sim N(0, M_t)$$

$$M_t = \begin{bmatrix} m(x) + \sigma^2 I & M(x, x_t) \\ M(x, x_t)^T & m(x_t) \end{bmatrix} \tag{14}$$

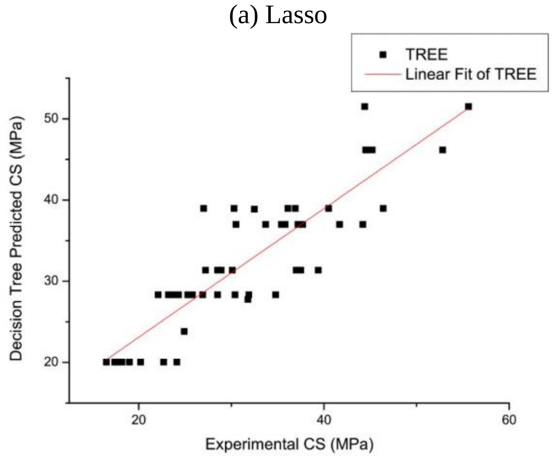
where $M(x, x_t)$ represents the training inputs and test input covariance, $M(x_t)$ is auto covariance of the testing data and $M(x)$ is the training data auto covariance. The mean and



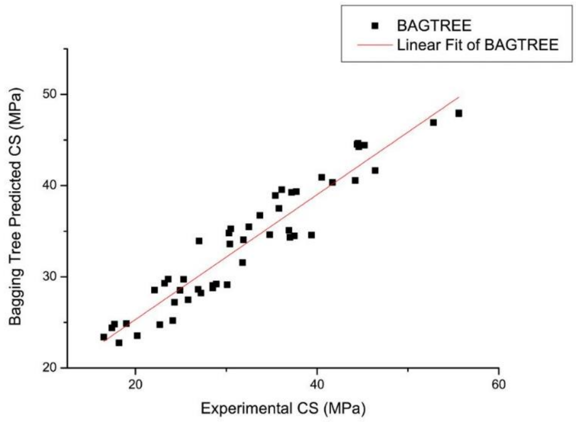
(a) Lasso



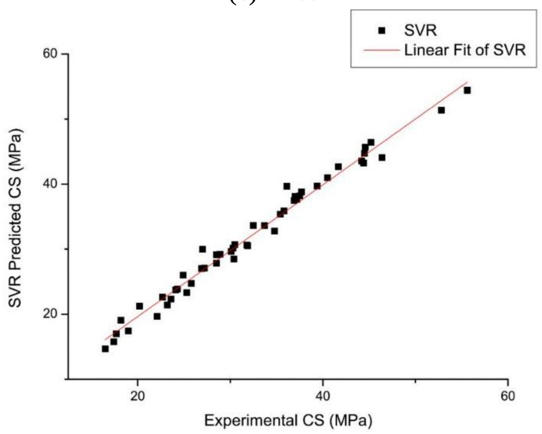
(b) Elastic Net



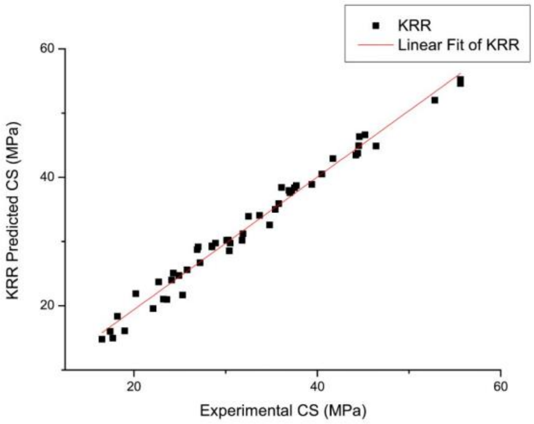
(c) Tree



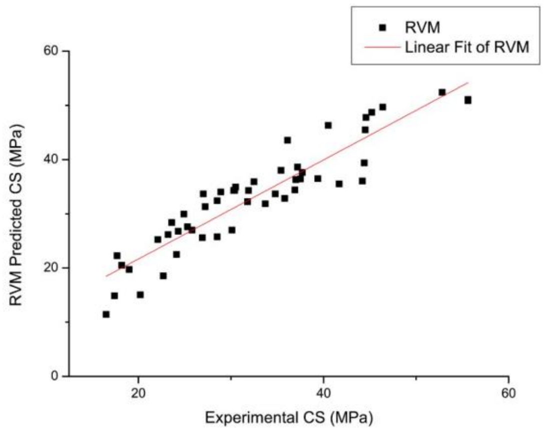
(d) Bagtree



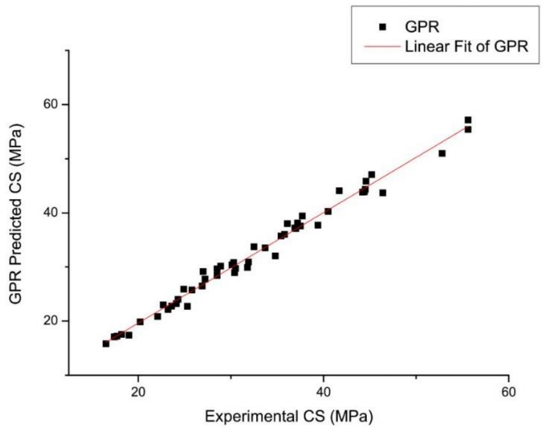
(e) SVR



(f) KRR



(g) RVM



(h) GPR

Fig. 11 Correlation between the experimental and predicted compressive strength

Table 3 Error indices for different method

| ` | ME | MAE | RMSE | R |
|---|---|---|---|---|
| LASSO | 16.73 | 17.09 | 16.73 | 0.93 |
| Elastic net | 11.73 | 12.31 | 11.73 | 0.93 |
| Decision tree | 0.49 | 4.50 | 3.72 | 0.89 |
| Bagging tree | 1.19 | 3.94 | 3.25 | 0.96 |
| SVR | 0.22 | 1.29 | 1.03 | 0.99 |
| KRR | 0.13 | 1.26 | 1.02 | 0.99 |
| RVM | 0.64 | 3.75 | 3.29 | 0.93 |
| GPR | 0.12 | 1.23 | 0.96 | 0.99 |

variance of the output is given by

$$\mu = K(x, x_t)^T (k(x) + \sigma^2 I)^{-1} y$$
$$\sigma = K(x_t) - K(x, x_t)^T (k(x) + \sigma^2 I)^{-1} K(x, x_t) \quad (15)$$

Composite kernel consisting of radial basis function with adaptive length scale and a diagonal noise covariance matrix is used in the current study. The parameters of the method are obtained by maximizing the marginal likelihood instead of grid search as the computational cost is high with the composite kernel. The signal scaling factor for the kernel and the standard deviation of the noise are evaluated as 3.54 and -0.08, respectively. The correlation between the experimental and predicted compressive strength of geopolymer is shown in Fig. 11(h).

## 5. Results and discussions

Based on the results obtained from the application of different methods in the previous section, the performance of the method is benchmarked in terms of (i) error indices (ii) residuals (iii) convergence (iv) computational time and (v) validation.

### *5.1 Error indices*

Four different types of error indices are used - mean error (ME), mean absolute error (MAE), root mean squared error (RMSE) and correlation coefficient (R). These error indices are evaluated for the testing dataset. The values of the indices obtained are given in Table 3. The error indices for the linear regression methods like LASSO and Elastic net are found to be worst. SVR, KRR and GPR are found to

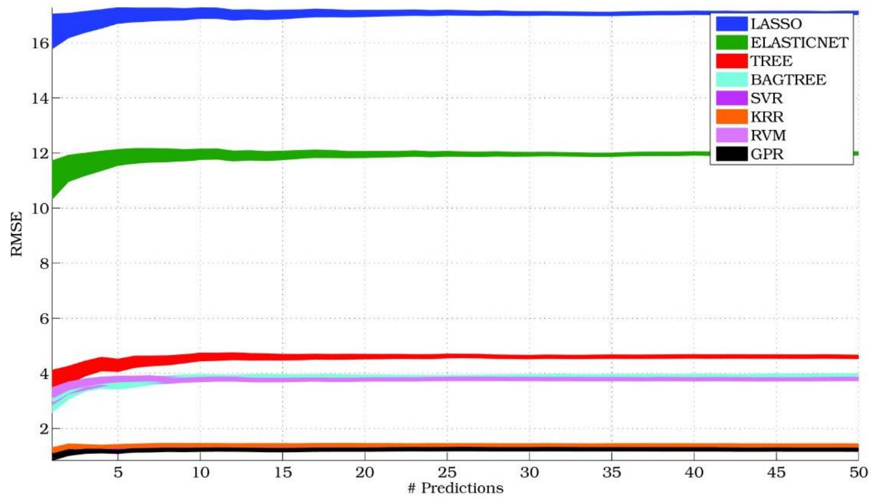



Fig. 12 Convergence of RMSE with number of predictions

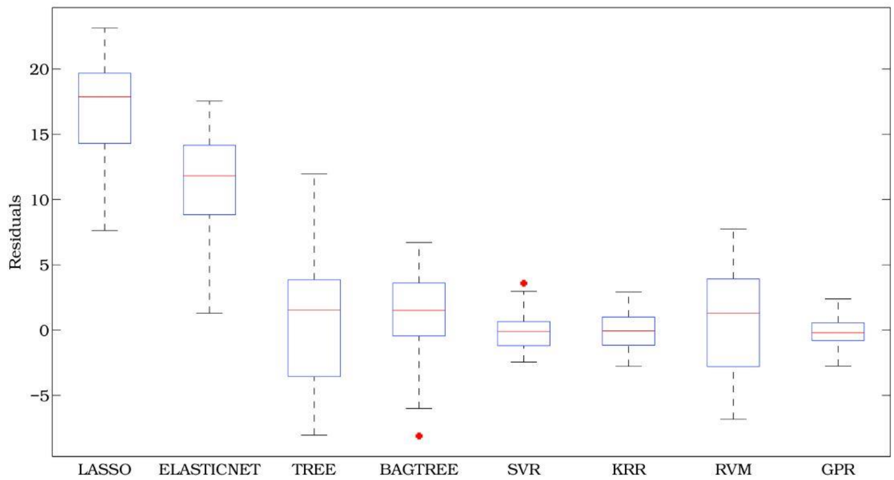



Fig. 13 Variation of residuals for different methods

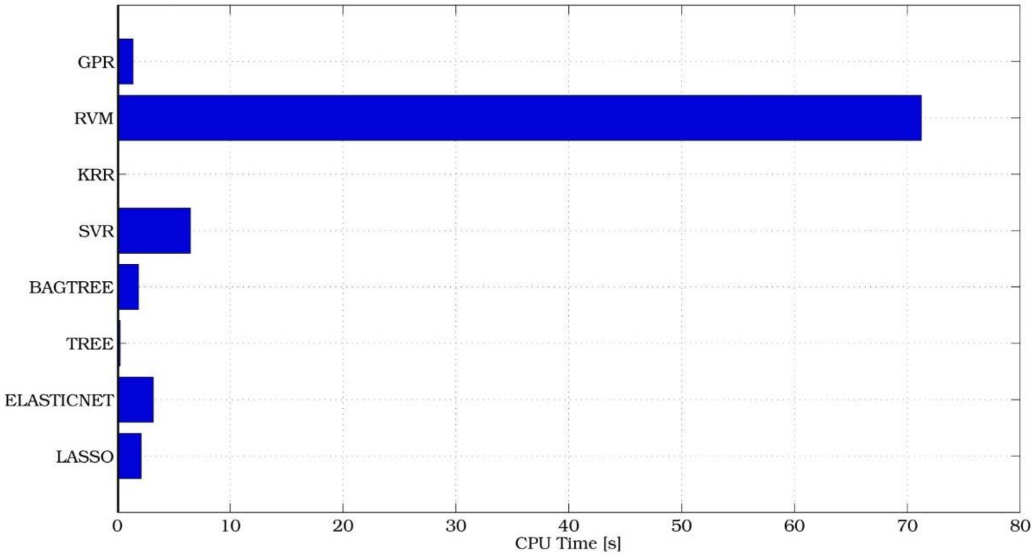



Fig. 14 Computational time for different methods

perform better than the RVM among kernel based methods. The best method to predict the compressive strength of geopolymer is found to be GPR based on the error indices.

### *5.2 Convergence*

The plot of RMSE versus the number of predictions is shown in Fig. 12. No abrupt jumps are observed indicating the consistency of the method in evaluating the compressive strength with the number of predictions. GPR and KRR are found to converge fast with the least RMSE value.

### *5.3 Residuals*

Fig. 13 Shows the box plot of the residuals for testing dataset. The mean residual is found to be least for GPR and maximum for LASSO. The maximum and the minimum residual or GPR are found to 2.38 and -2.75.

### *5.4 Computational time*

The variation of the computational time taken by the different algorithms for training and testing is shown in Fig. 14. The computational time is found to be maximum for RVM. Therefore, the use of RVM should be avoided in case of larger dataset. The computational time is found to be least for KRR.

### *5.5 Validation*

The purpose of this section is to see if the output from the methods follows the trend reported in the literature. For this purpose, three parameters are considered - (i) Curing temperature, (ii) Curing Duration and (iii) Ratio of FA/RHBA. The data was selected from the testing dataset in such a way that only the parameter studied was varying while other parameters were same. The trend followed by the predicted compressive strength from the methods are compared with that reported in the literature.

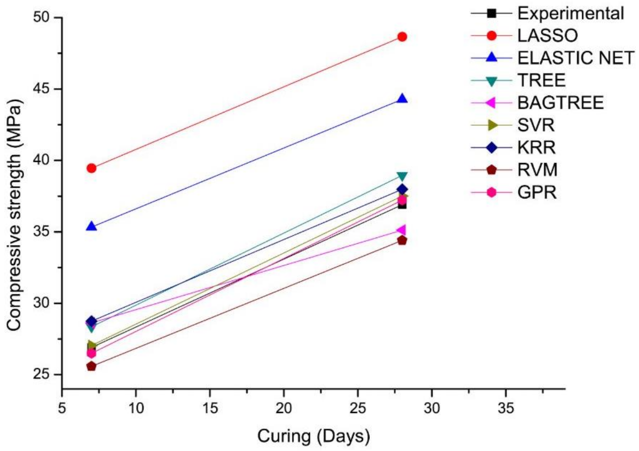



Fig. 15 Variation of compressive strength with curing duration

The gain in compressive strength of any normal concrete or geopolymer with the passage of curing duration is quite common and same was observed in Fig. 15**.** However, it is interesting to note that many researchers have focused to subject concrete to thermal curing (Prem *et al*. 2013, 2015) for accelerated strength development. The mechanical activation of geopolymer causes microstructure and structural changes (Kumar and Kumar 2011).

Bakharev (2005) reported that compressive strength gained after one month of curing can be immediately obtained only after 24 hour of heat curing. Temuujin *et al*. (2009) observed compressive strength variations from 16 MPa for room temperature cured samples to 45 MPa for activated fly ash based samples. The main cause for the sudden enhancement of mechanical properties is due to the speedy dissipation of silicate monomer and oligomer from RHBA surfaces leading to the formation of supersaturated aluminosilicate (Kusbiantoro *et al*. 2012).

The variation of the compressive strength with the curing temperature is shown in Fig. 16. In the experimental results, it is observed that there is an increase in the

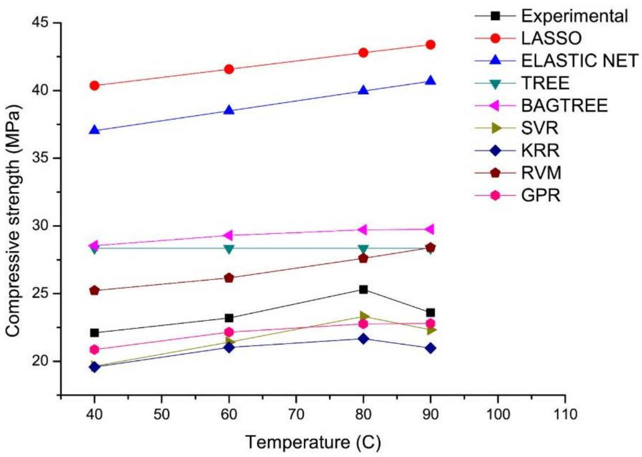



Fig. 16 Variation of compressive strength with curing temperature

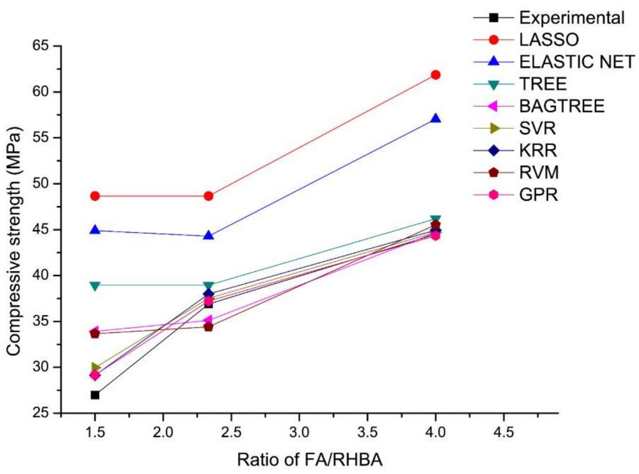



Fig. 17 Variation of compressive strength with FA/RHBA

compressive strength upto 80$^o$C after which it drops. The similar trend has been reported by Atiş *et al*. (2015) and Huseien *et al*. (2016). SVR and KRR are found to capture variation of compressive strength with curing temperature effectively.

Detphan and Chindaprasirt (2009) reported that the strength (ranging from 12.5-56.0 MPa) of the geopolymer depends upon the FA/RHBA mass proportion, RHBA fineness and the ratios of sodium silicate to sodium hydroxide. The variation of the compressive strength with the FA/RHBA is shown in Fig. 17. According the studies reported in the literature and the experimental results of Bohlooli *et al*. (2012), it is observed that compressive strength increases with the increase in FA/RHBA. SVR, KRR and GPR are able to capture the trend reflected by the experimental results.

## 6. Conclusions

The compressive strength of the geopolymer is predicted using different applied statistical models - LASSO, elastic net, decision tree, bagging tree, SVR, KRR, RVM and GPR. The inputs for the method are taken as the weight percentage of fine and coarse fly ash, fine and coarse rice husk bark ash, the temperature and time of water curing. The performance of the methods is compared in terms of error indices (ME, MAE, RMSE and R), convergence, residuals and computational time. Following conclusions are made from the present study:

1. Being the linear regression methods, the performance of LASSO and elastic net is not at par with other methods. Therefore, their application to predict compressive strength of geopolymer should be avoided.
2. The correlation between the actual and the compressive strength predicted by the methods for decision tree is found to be least.
3. The computational time for RVM is found to be more. A highly nonlinear optimization problem is solved in the training phase of RVM. The makes RVM unsuitable for larger datasets.
4. SVR, KRR and GPR are able to effectively capture the effect of the curing temperature, duration and FA/RHBA ratio on the geopolymer compressive strength
5. Overall, GPR and KRR are found to be have better compressive strength predicting capabilities compared to other methods in terms of all the performance criteria. There- fore, GPR and KRR are recommended for the evaluation of compressive strength of geopolymer.

There are several other factors which can affect the compressive strength of the geopolymer (like alkali content etc.). These parameters can be included as input in the present methods to improve their accuracy. These methods can be further extended to predict various other mechanical and fracture properties of geopolymer.

## Acknowledgements

The authors would like to thank the anonymous reviewers for their constructive comments

## References

Al-Rousan, R.Z., Alhassan, M.A. and Hejazi, M.A. (2018), “Novel nonlinear stiffness parameters and constitutive curves for concrete”, *Comput. Concrete*, **22**(6), 539-550. https://doi.org/10.12989/cac.2018.22.6.539.

Ambily, P.S., Ravisankar, K., Umarani, C., Dattatreya, J.K. and Iyer, N.R. (2014), “Development of ultra-high-performance geopolymer concrete”, *Mag. Concrete Res.*, **66**(2), 82-89. http://dx.doi.org/10.1680/macr.13.00057

Arani, K.S., Zandi, Y., Pham, B.T., Mu'azu, M., Katebi, J., Mohammadhassani, M., Khala, S., Mohamad, E.T., Wakil, K. and Khorami, M. (2019), “Computational optimized nite element modelling of mechanical interaction of concrete with fiber reinforced polymer”, *Comput. Concrete*, **23**(1), 61-68. http://dx.doi.org/10.12989/cac.2019.23.1.061.

Atis, C., Gorur, E., Karahan, O., Bilim, C., Ilkentapar, S. and Luga, E. (2015), “Very high strength (120mpa) class F fly ash geopolymer mortar activated at different NaOH amount, heat curing temperature and heat curing duration”, *Constr. Build. Mater.*, **96**, 673-678.

https://doi.org/10.1016/j.conbuildmat.2015.08.089.

Bakharev, T. (2005), "Geopolymeric materials prepared using class f fly ash and elevated temperature curing", *Cement Concrete Res.*, **35**(6), 1224-1232. https://doi.org/10.1016/j.cemconres.2004.06.031.

Breiman, L. (1996), "Bagging predictors", *Mach. Learn.*, **24**(2), 123-140. https://doi.org/10.1007/BF00058655.

Breiman, L., Friedman, J., Olshen, R. and Stone, C. (1993), "Classification and regression trees, wadsworth international group, belmont, ca, 1984" *Case Description Feature Subset Correct Missed FA Misclass*, **1**, 1-3.

Camps-Valls, G., Bruzzone, L., Rojo-Alvarez, J.L. and Melgani, F. (2006), "Robust support vector regression for biophysical variable estimation from remotely sensed images", *Geosci. Remote Sens. Lett.*, IEEE, **3**(3), 339-343. https://doi.org/10.1109/LGRS.2006.871748.

Camps-Valls, G., Gmez-Chova, L., Muoz-Mar, J., Lzaro-Gredilla, M. and Ver relst, J. (2013), "simpleR: A simple educational Matlab toolbox for statistical regression", V2.1, http://www.uv.es/gcamps/code/simpleR.html.

Camps-Valls, G., Munoz-Mari, J., Gomez-Chova, L., Guanter, L. and Calbet, X. (2012), "Nonlinear statistical retrieval of atmospheric profiles from metop-iasi and mtg-irs infrared sounding data", *Geosci. Remote Sens*., IEEE Tran. on, **50**(5), 1759-1769. https://doi.org/10.1109/TGRS.2011.2168963

Cundi, W., Hirano, Y., Terai, T., Vallepu, R., Mikuni, A. and Ikeda, K. (2005), "Preparation of geopolymeric monoliths from red mud-pfbc ash fillers at ambient temperature", *Proc. World Cong. Geopolymer, Saint-Quentin*, Franca, 85-87.

Davidovits, J. (1991), "Geopolymers", *J. Therm. Anal.*, **37**(8), 1633-1656.

Detphan, S. and Chindaprasirt, P. (2009), "Preparation of fly ash and rice husk ash geopolymer", *Int. J. Min. Metal. Mater.*, **16**(6), 720-726. https://doi.org/10.1016/S1674-4799(10)60019-2.

Dutta, S., Samui, P. and Kim, D. (2018), "Comparison of machine learning techniques to predict compressive strength of concrete", *Comput. Concrete*, **21**(4), 463-470. https://doi.org/10.12989/cac.2018.21.4.463.

Erdal, H., Erdal, M., Simsek, O. and Erdal, H. I. (2018), "Prediction of concrete compressive strength using non-destructive test results", *Comput. Concrete*, **21**(4), 407-417. https://doi.org/10.12989/cac.2018.21.4.407.

He, J., Jie, Y., Zhang, J., Yu, Y. and Zhang, G. (2013), "Synthesis and characterization of red mud and rice husk ash-based geopolymer composites", *Cement Concrete Compos.*, **37**, 108-118. https://doi.org/10.1016/j.cemconcomp.2012.11.010.

Huseien, G.F., Mirza, J., Ismail, M. and Hussin, M.W. (2016), "Influence of different curing temperatures and alkali activators on properties of ggbs geopolymer mortars containing fly ash and palm-oil fuel ash", *Constr. Build. Mater.*, **125**, 1229-1240. https://doi.org/10.1016/j.conbuildmat.2016.08.153.

Jindal, B.B., Singhal, D., Sharma, S. and Parveen, (2018), "Enhancing mechanical and durability properties of geopolymer concrete with mineral admixture", *Comput. Concrete*, **21**(3), 345-353. https://doi.org/10.12989/cac.2018.21.3.345.

Katpady, D.N., Takewaka, K. and Yamaguchi, T. (2017), "Development of slag based shirasu geopolymer", *Comput. Concrete*, **20**(1), 77-84. https://doi.org/10.12989/cac.2017.20.1.077.

Kumar, S. and Kumar, R. (2011), "Mechanical activation of fly ash: effect on reaction, structure and properties of resulting geopolymer", *Ceram. Int.*, **37**(2), 533-541. https://doi.org/10.1016/j.ceramint.2010.09.038.

Kurklu, G. (2016), "The effect of high temperature on the design of blast furnace slag and coarse fly ash-based geopolymer mortar", *Compos. Part B: Eng.*, **92**, 9-18. https://doi.org/10.1016/j.compositesb.2016.02.043.

Kusbiantoro, A., Nuruddin, M.F., Shafiq, N. and Qazi, S.A. (2012), "The effect of microwave incinerated rice husk ash on the compressive and bond strength of fly ash based geopolymer concrete", *Constr. Build. Mater.*, **36**, 695-703.

Marın-Lopez, C., Araiza, J. R., Manzano-Ramırez, A., Avalos, J. R., Perez-Bueno, J., Muniz-Villareal, M., Ventura-Ramos, E. and Vorobiev, Y. (2009), "Synthesis and characterization of a concrete based on metakaolin geopolymer", *Inorganic Mater.*, **45**(12), 1429-1432. https://doi.org/10.1134/S0020168509120231.

Mozumder, R.A., Laskar, A.I. and Hussain, M. (2017), "Empirical approach for strength prediction of geopolymer stabilized clayey soil using support vector machines", *Constr. Build. Mater.*, **132**, 412-424. https://doi.org/10.1016/j.conbuildmat.2016.12.012.

Pacheco-Torgal, F., Castro-Gomes, J. and Jalali, S. (2008), "Alkali-activated binders: A review: Part 1. historical background, terminology, reaction mechanisms and hydration products", *Constr. Build. Mater.*, **22**(7), 1305-1314. https://doi.org/10.1016/j.conbuildmat.2007.10.015.

Prem, P.R., Bharatkumar, B. and Iyer, N.R. (2013), "Influence of curing regimes on compressive strength of ultra high performance concrete", *Sadhana*, **38**(6), 1421-1431. https://doi.org/10.1007/s12046-013-0159-8.

Prem, P.R., Murthy, A.R. and Bharatkumar, B.H. (2015), "Influence of curing regime and steel fibres on the mechanical properties of UHPC", *Mag. Concrete Res.*, **67**(18), 988-1002. http://dx.doi.org/10.1680/macr.14.00333

Prem, P.R., Murthy, A.R. and Verma, M. (2018a), "Theoretical modelling and acoustic emission monitoring of RC beams strengthened with UHPC", *Constr. Build. Mater.*, **158**, 670-682. https://doi.org/10.1016/j.conbuildmat.2017.10.063.

Prem, P.R., Verma, M. and Ambily, P. (2018b), "Sustainable cleaner production of concrete with high volume copper slag", *J. Clean. Prod.*, **193**, 43-58. https://doi.org/10.1016/j.jclepro.2018.04.245.

Prem, P.R., Verma, M., Murthy, A.R., Rajasankar, J. and Bharatkumar, B. (2017), "Numerical and theoretical modelling of low velocity impact on UHPC panels", *Struct. Eng. Mech.*, **63**(2), 207-215. https://doi.org/10.12989/sem.2017.63.2.207.

Quinlan, J.R. (1993), "Combining instance-based and model-based learning", *Proceedings of the Tenth International Conference on Machine Learning*, 236-243.

Senthamilselvi, P. and Palanisamy, T. (2018), "Experimental and analytical study on flexural behaviour of fly ash and paper sludge ash based geopolymer concrete", *Comput. Concrete*, **21**(2), 157-166. https://doi.org/10.12989/cac.2018.21.2.157.

Temuujin, J., Williams, R. and Van Riessen, A. (2009), "Effect of mechanical activation of fly ash on the properties of geopolymer cured at ambient temperature", *J. Mater. Proc. Technol.*, **209**(12), 5276-5280. https://doi.org/10.1016/j.jmatprotec.2009.03.016.

Tibshirani, R. (1996), "Regression shrinkage and selection via the lasso", *J. R. Stat. Soc., Ser. B (Method.)*, **58**(1), 267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

Tipping, M.E. (2001), "Sparse bayesian learning and the relevance vector machine", *J. Mach. Learn. Res.*, **1**, 211-244.

Vapnik, V. (2013), *The Nature of Statistical Learning Theory*, Springer Science & Business Media.

Verma, M., Anandavalli, N., Prakash, A. and Rajasankar, J. (2019), "Laced steel- concrete composite sandwich system in flexure: Experiments, analysis and design", *J. Sandw. Struct. Mater.*, 1099636219842286. https://doi.org/10.1177/1099636219842286.

Verma, M., Prem, P.R., Rajasankar, J. and Bharatkumar, B. (2016a), "On low- energy impact response of ultra-high performance concrete (UHPC) panels", *Mater. Des.*, **92**, 853-

865. https://doi.org/10.1016/j.matdes.2015.12.065.

Verma, M., Rajasankar, J., Anandavalli, N., Prakash, A. and Iyer, N.R. (2015), "Fuzzy similarity approach for ranking and health assessment of towers based on visual inspection", *Adv. Struct. Eng.*, **18**(9), 1399-1414. https://doi.org/10.1260/1369-4332.18.9.1399.

Verma, M., Thirumalaiselvi, A. and Rajasankar, J. (2016b), "Kernel-based mod- els for prediction of cement compressive strength", *Neur. Comput. Appl.*, **28**(1), 1083-1100. https://doi.org/10.1007/s00521-016-2419-0.

Verrelst, J., Alonso, L., Camps-Valls, G., Delegido, J. and Moreno, J. (2012), "Re- trieval of vegetation biophysical parameters using gaussian process techniques", *Geosci. Remote Sens., IEEE Tran. on*, **50**(5), 1832-1843. https://doi.org/10.1109/TGRS.2011.2168962

Vijai, K., Kumutha, R. and Vishnuram, B. (2015), "Flexural behaviour of fibre reinforced geopolymer concrete composite beams", *Comput. Concrete*, **15**(3), 437-459. https://doi.org/10.12989/cac.2015.15.3.437.

Williams, C.K. (1998), "Prediction with gaussian processes: From linear regression to linear prediction and beyond", *Learning in Graphical Models*, Springer.

Williams, C.K. and Rasmussen, C.E. (2006), *Gaussian Processes for Machine Learning*, Vol. 2, No. 3, MIT Press, Cambridge, MA.

Yadollahia, M.M. and Benli, A. (2017), "Stress-strain behavior of geopolymer under uniaxial compression", *Comput. Concrete*, **20**(4), 381-389. https://doi.org/10.12989/cac.2017.20.4.381.

Zhang, M., El-Korchi, T., Zhang, G., Liang, J. and Tao, M. (2014), "Synthesis factors affecting mechanical properties, microstructure, and chemical com- position of red mud-fly ash based geopolymers", *Fuel*, **134**, 315-325. https://doi.org/10.1016/j.fuel.2014.05.058.

Zou, H. and Hastie, T. (2005), "Regularization and variable selection via the elastic net", *J. R. Stat. Soc.: Ser. B (Stat. Method.)*, **67**(2), 301-320. https://doi.org/10.1111/j.1467-9868.2005.00503.x.

*CC*