Journal of Science and Transport Technology

University of Transport Technology

# Prediction and sensitivity analysis of self compacting concrete slump flow by random forest algorithm

Raghvendra Kumar[1], Hai-Van Thi Mai[2,*]

[1]Department of Computer Science and Engineering, GIET University, Gunupur-765022, India

[2]University of Transport Technology, Ha Noi 100000, Vietnam

**Article info**
**Type of article:**
Original research paper

**Corresponding author:**
E-mail address:
vanmth@utt.edu.vn

**Received:**
December 09, 2021
**Accepted:**
March 11, 2022
**Published**:
March 20, 2022

**Abstract**: Self-compacting concrete (SCC) is a construction material with many advantages, including high performance and the capacity to self-compact without mechanical vibration. As a result, SCC is widely used in construction, especially at locations where concrete structures are difficult to construct. Filling ability is one of the three basic requirements that must be met when designing the SCC mix. The slump flow (SF) is used to determine the SCC mixture's filling capacity. As a result, it is critical to estimate this number fast and precisely. The purpose of this study is to propose the use of a random forest (RF) model to predict the SF of SCC and to assess the effect of input parameters on output parameters. The study constructed the RF model using a dataset of 507 experimental results collected, which is the biggest data collection compared to previous studies on this subject. Additionally, a 10-fold cross-validation approach is used to improve the model's prediction performance. As a result, the performance assessment criteria for the testing dataset have values of RMSE = 59.5664 mm, MAE = 32.4483 mm, and R = 0.8614, respectively. This result shows that the RF model is an effective tool in predicting the SF of SCC.

**Keywords**: Self-compacting concrete, Machine learning, Random forest, Slump flow, Filling ability.

## 1. Introduction

Currently, in the construction industry, it is usual to come across heavy reinforced concrete structures, as well as structures with complex construction locations, such as overhead bridge girders, cement concrete road pavements, and high-rise skyscrapers. One of the most significant when constructing these structures is the compaction of concrete. Inadequate compaction may result in substandard construction and low construction performance. Mechanical vibrating or hand compaction techniques are ineffective in these instances. Mechanical vibrating or hand compaction techniques are ineffective in these situations. Then self-compacting concrete (SCC) is the best solution. It is a material that compacts under its own weight, self-flows, and fills all positions, even those with complicated topography, forms, and reinforcing dense. Without any mechanical activity, SCC forms a homogenous, non-stratified, non-flowing material [1]. There are three essential conditions that the SCC mixture must fulfill in order to be designed at the uncured stage, namely the Filling ability,

Passing ability, and segregation resistance. In which the slump flow (SF) is used to determine the SCC mixture's filling ability [2]. Therefore, it is necessary to determine the SF of SCC when designing a mixture of SCC.

Typically, the SF of SCC is determined experimentally. If the acquired result is unsatisfactory, the tests must be redone from the mixed design phase, which is time-consuming. Another roadblock is that determining the optimal SCC blend is very time-consuming. To circumvent these constraints, an advanced technique for anticipating the SF of SCC is necessary.

In recent decades, artificial intelligence (AI) simulations or machine learning (ML) models have developed and are applied in many fields [3]–[5]. Many scientists have used AI models to predict the mechanical properties of SCC. For example, Saha et al. [6] used the support vector regression approach (SVR) to predict SCC properties. This study used 115 data samples with six input parameters, namely binder content, fly ash, water-powder ratio, fine aggregate, coarse aggregate, and superplasticizer. The results show that the SVR model is better than the ANN model in predicting the properties of SCC. In some other studies, Sathyan et al. [7] employed the Random Kitchen Sink Algorithm, whereas Sonebi et al. [8] used a support vector machine (SVM) technique to forecast SF. These studies analyzed relatively small datasets, including just 40 and 20 experimental outcomes, respectively. Additionally, Kaloop et al. [9] conducted a study in which they employed an emotional neural network to predict the SF of SCC based on just 90 experimental data collected. On the other hand, among machine learning models, random forest (RF) is a powerful machine learning method that has been effectively employed in a variety of domains, including civil engineering [10], earth science [11], and building materials [12], [13]. However, there have been no studies using this model to predict the SF of SCC.

The random forest model will be used in this research to estimate the SF of SCC using a pretty large data set of 507 experimental outcomes. Three measures are used to assess the model's prediction ability: the correlation coefficient (R), root mean square error (RMSE) and mean absolute error (MAE). In addition, the cross-validation approach is used to counteract the over-fitting phenomena and improve the model's forecast accuracy. Finally, the RF model assesses the significance of input factors that influence the SF of SCC.

## 2. Database

This section describes the database that was utilized in this research. SF prediction RF model was developed using a dataset of 507 experimental findings from seven studies published in the relevant literature [14]–[20]. There are twelve input parameters marked by the numbers (1) to (12): Cement content (kg/m$^3$), fly ash content (kg/m$^3$), water content (kg/m$^3$), sand or coarse aggregate content (kg/m$^3$), superplasticizer content (kg/m$^3$), limestone powder content (kg/m$^3$), fine blast furnace slag content (kg/m$^3$), silica fume content (kg/m$^3$), metakaolin content (kg/m$^3$), rice husk ash content (kg/m$^3$), viscosity reducing additive (kg/m$^3$). The cement content of the acquired data set varies from 83 to 670 kg/m$^3$, whereas the fly ash concentration ranges from 0 to 525 kg/m$^3$, showing that certain SCC samples do not include fly ash. The water content ranges between 126 ÷ 331.5 kg/m$^3$, whereas the fine and coarse aggregate contents vary between 240 ÷ 1180 kg/m$^3$ and 500 ÷ 1531 kg/m$^3$, respectively. The concentration of the superplasticizer is in the range of (0÷22.5) kg/m$^3$. The limestone powder, fine blast furnace slag, silica powder, metakaolin, and rice husk ash powder contents varied between (0÷376) kg/m$^3$, (0÷440) kg/m$^3$, (0÷82.5) kg/m$^3$, (0÷82.5) kg/m$^3$, (0÷200) kg/m$^3$. The viscosity-reducing additive content varies in a rather narrow range, with the smallest value being 0 kg/m$^3$, the highest value being 4.46 kg/m$^3$. The output parameter (SF) has a very broad data spectrum, ranging from (0÷920) mm. It should be noted that the lowest value (min) of SF is 0 mm

since the research employed just one sample with SF = 0 mm out of 507 experimental findings. Additionally, to reduce mistakes associated with RF model simulation, this data set is standardized to the range of values [0-1]. This is a frequently used technique in artificial intelligence challenges to restrict the mistakes created by numerical simulations, such as huge differences in fine or coarse aggregate content (hundreds) and viscosity-reducing additive content (less than 1). Table 1 contains information on the symbols, units, and statistical analysis of both input and output parameters.

**Table 1**. Statistical analysis of database

| No | Parameters | Symbol | Unit | Min | Mean | Max | Std |
|---|---|---|---|---|---|---|---|
| 1 | Cement | (1) | ($kg/m^3$) | 83 | 357.81 | 670 | 109.78 |
| 2 | Fly ash | (2) | ($kg/m^3$) | 0 | 88.39 | 525 | 98.61 |
| 3 | Water | (3) | ($kg/m^3$) | 126 | 184.68 | 331.50 | 27.23 |
| 4 | Sand or fine aggregate | (4) | ($kg/m^3$) | 240 | 816.34 | 1180 | 128.05 |
| 5 | Coarse aggregate | (5) | ($kg/m^3$) | 500 | 804.07 | 1531 | 130.33 |
| 6 | Superplasticizes additive | (6) | ($kg/m^3$) | 0 | 6.12 | 22.50 | 4.51 |
| 7 | Limestone powder | (7) | ($kg/m^3$) | 0 | 30.97 | 376 | 70.11 |
| 8 | Blast furnace slag | (8) | ($kg/m^3$) | 0 | 20.75 | 440 | 58.97 |
| 9 | Silica fume | (9) | ($kg/m^3$) | 0 | 6.93 | 82.50 | 17.80 |
| 10 | Metakaolin | (10) | ($kg/m^3$) | 0 | 1.79 | 82.50 | 9.81 |
| 11 | Rice husk ash | (11) | ($kg/m^3$) | 0 | 1.38 | 200 | 12.69 |
| 12 | Viscosity reducing additive | (12) | ($kg/m^3$) | 0 | 0.11 | 4.46 | 0.43 |
| | Slump flow | (0) | mm | 0 | 667.05 | 920 | 112.99 |

*Std=Standard deviation*

Next, Fig. 1 shows the distribution graph of 12 input parameters, one output parameter, and the correlation between those parameters. For each pair of parameters, the Pearson correlation coefficient ($r_s$) was computed and shown. Correlation may be classified into the following levels based on the value of the correlation index ($r_s$): $r_s$ values between 0 and 0.19 indicate very weak correlation, $r_s$ values between 0.2 and 0.39 indicate weak correlation, $r_s$ values between 0.4 and 0.59 indicate moderate correlation, $r_s$ values between 0.6 and 0.79 indicate a strong correlation and rs values between 0.8 and 1 indicate extremely strong correlation. As seen in Fig. 1, the correlation between the input variables and the SF is quite poor; the majority of the $r_s$ values are less than 0.4 (indicating a weak correlation).

Only one measurement, $r_s$ = 0.6, demonstrates a weak association between the fly ash content (2) and the cement content (1). The research reveals that 12 of the data set's input parameters may be deemed independent variables. As a result, all twelve input parameters will be addressed in this research in order to improve the accuracy and generality of the output parameter prediction model (SF).

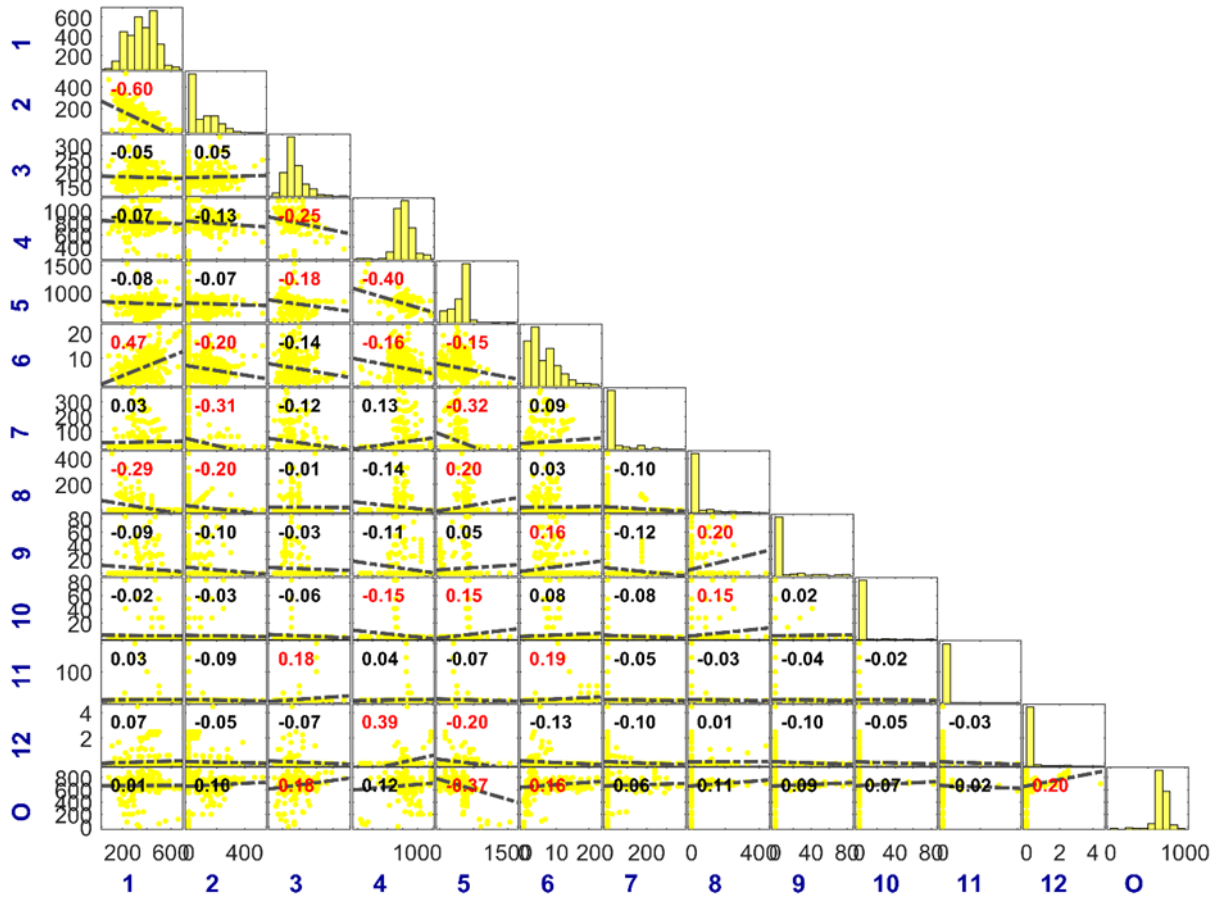


**Fig. 1.** Correlation analysis and distribution between input and output parameters

## 3. Machine learning

### 3.1. Random forest (RF)

Breiman introduced and developed RF as a supervised machine learning method [21]. The benefit of RF is that it may be used to solve issues involving classification and regression. RF model uses trees as a foundation. Whereas a random forest is a collection of decision trees, each of which is chosen randomly. RF operates by evaluating a large number of random decision trees and selecting the best-evaluated result from the pool of returning outcomes. The learning processes involve the construction of a collection of decision trees, each one driven by a Bootstrap subset. Each tree in the forest is trained using a randomly distributed subset of data, using the bagging and random feature principles. For regression problems, the final result is presented as the mean of each decision tree; for classification problems, the final result is decided by the majority result. The number of trees in the forest should be sufficient to guarantee that each characteristic is utilized many times. Typically, 500 trees are used to solve classification issues, while 1000 trees are used to solve regression problems. In this work, RF is used to predict SCC using 500 trees with an ideal leaf count of 20.

### 3.2. Cross validation (CV)

In machine learning, overfitting is the phenomenon where the found model over-fits the training data. This overfitting may result in incorrect predictions, noise, and a model with poor predictive performance on validation data. Cross-validation is often used to circumvent this. To train the network using cross-validation, the whole database is randomly partitioned into three sections: training data set, validation data set, and testing data set. If the dataset is separated into just two parts, training and testing, the testing data set will be kept separate for the model verification stage. During the model training process, this data set is not known. The training

data collection will include information about the model's training and validation processes. This will be accomplished by randomly dividing the training data set into K equal halves. The model will be trained K times, with each training session selecting one portion as validation data and the remaining (K-1) as training data. The final assessment result will be the average of the evaluation results obtained for each of the K training sessions. Generally, K should not be selected too big since a high K results in a significantly larger training data set than the control data set. At that point, the assessment findings will no longer accurately represent the underlying nature of the machine learning technology, particularly when dealing with big data sets. Cross-validation with K = 10 was used in this investigation, which is compatible with the work of other scientists worldwide [22]. Fig. 2 illustrates the cross-validation approach using ten CVs and three datasets for training, validation, and verification.

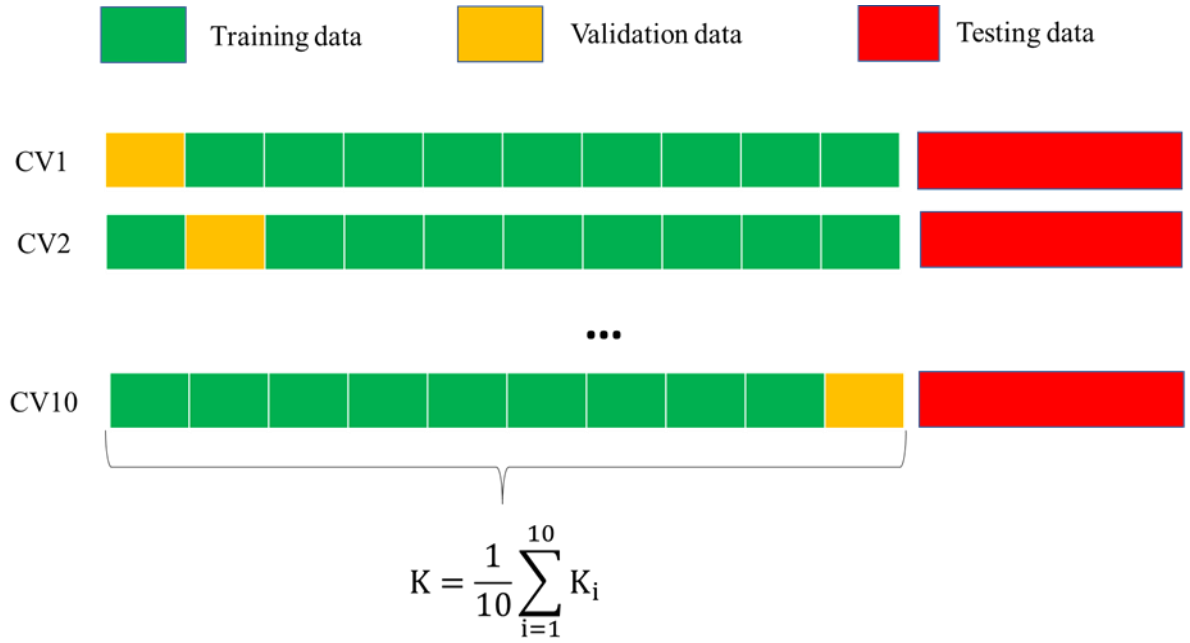


**Fig. 2.** Illustrate the cross-validation technique used in the paper

### 3.3. Evaluation of the RF model's predictive ability

This research employs three criteria to assess the SF prediction outcomes of SCC produced from the RF model: correlation coefficient (R), root mean square error (RMSE) and mean absolute error (MAE). The R-index indicates the correlation between the value predicted by the RF model and the observed value. A greater absolute value of R implies a stronger correlation between the anticipated and actual values, implying improved model performance (R values range from -1 to 1). The RMSE index is another way to calculate the error since it is based on the mean squared difference between the expected and actual output values, while the MAE index shows the average error of the actual value and the anticipated value. The low RMSE and MAE values indicate that the error between the actual and predicted values is minor, indicating that the RF model's prediction ability is excellent. The following formula is used to calculate these three indices.

$$R = \frac{\sum_{i=1}^{m}\left(s_{o,i} - \bar{s}_o\right)\left(s_{t,i} - \bar{s}_t\right)}{\sqrt{\sum_{i=1}^{m}\left(s_{o,i} - \bar{s}_o\right)^2 \sum_{i=1}^{m}\left(s_{t,i} - \bar{s}_t\right)^2}} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(s_{o,i} - \bar{s}_{t,i}\right)^2} \quad (2)$$

$$MAE = \frac{1}{m}\sum_{i=1}^{m}\left|s_{o,i} - s_{t,i}\right| \quad (3)$$

Where $s_o$ and $\bar{s}_0$ are the actual compressive strength and average actual

compressive strength (as established by experiment), $s_t$ and $\bar{s}_t$ are the predicted compressive strength and average predicted compressive strength (as determined by the model's prediction), and m is the number of samples in the database.

## 4. Results and discussion

This section details the process of developing an RF model for predicting the SF of SCC. The process of developing an RF model is divided into two phases: training (using a training dataset that contains 70% of the total data) and testing (using the remaining 30 percent of data). The first step (training phase) consists of the model's training, in combination with ten cross-validations (10-CV). To do this, the training dataset is partitioned into ten sections. Each simulation will consist of nine parts used to train the model, and one part used to verify it. The predicted performance assessment criteria for the model (R, RMSE, MAE) will be computed using the average value of the ten preceding runs. Overall, the 10-fold CV score of the models after 10 runs are reliable (R > 0.8), showing that the selected hyper-parameters are appropriate to conduct further investigation. When the RF model has attained the best-predicted performance on the training data set, the next step (test phase) is performed. The test data set is used to evaluate the model's predicting abilities for unknown data. In this work, the results of assessing the predictive performance of the RF model using the RMSE, MAE, and R criteria for both the training and testing data sets are displayed in Fig.s 3a, 3b, and 3c.

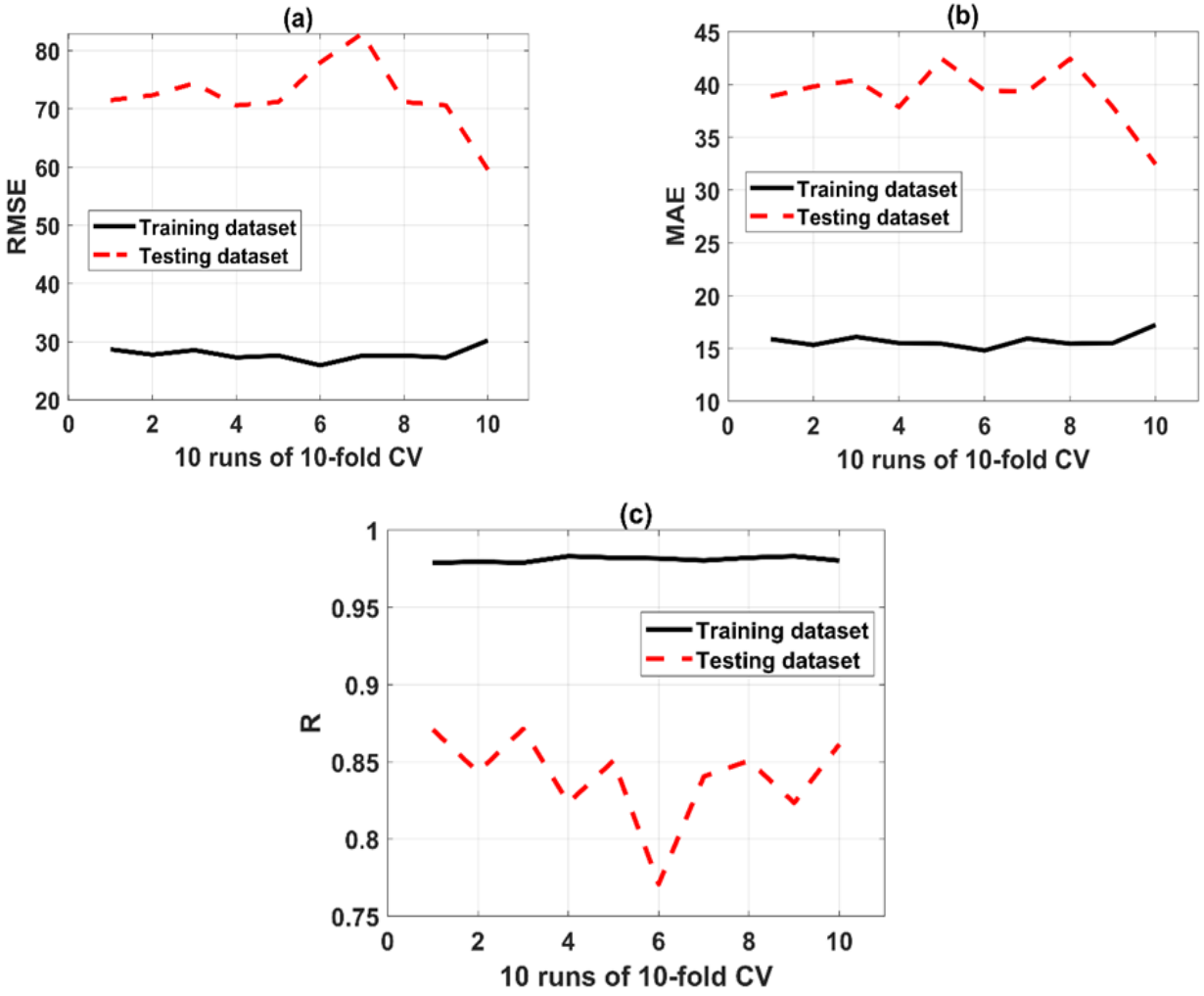


**Fig. 3.** Training and testing results of the RF model after ten cross-validations (10 CVs) based on the following evaluation criteria: (a) RMSE, (b) MAE, and (c) R

According to the Fig., with ten simulations, all three performance assessment criteria have a highly steady fluctuation amplitude: the RMSE fluctuates around 30 mm, the best run has an RMSE of 28.2, and the worst run has an RMSE of 30.3. Similarly, MAE fluctuates around 15.0 mm, with the best simulation occurring at the sixth random splitting of the training and testing dataset, denoted as CV6 (MAE=14.5), and the worst occurring at CV10 (MAE=17.1). Additionally, the evaluation criteria R has a pretty high and consistent value; after ten simulations, the R-value is almost the same and equivalent to 0.98. This signifies that the trained RF model has a high prediction power when used with the training data set and may be chosen for the testing dataset.

Following that, it can be shown that when the test data changes, the predictive capability of the RF model changes as well. However, the predictive performance of the RF model is pretty strong, as shown by R = 0.875 for the best runs (CV1, CV3). However, there are discrepancies in the testing data based on the model performance assessment criteria. Specifically, RMSE and MAE indicate that CV10 is the optimal run, but R indicates that CV1 and CV3 are the optimal runs. However, the difference is negligible. As a result, the suggested RF model may be utilized to accurately forecast the SF of SCC.

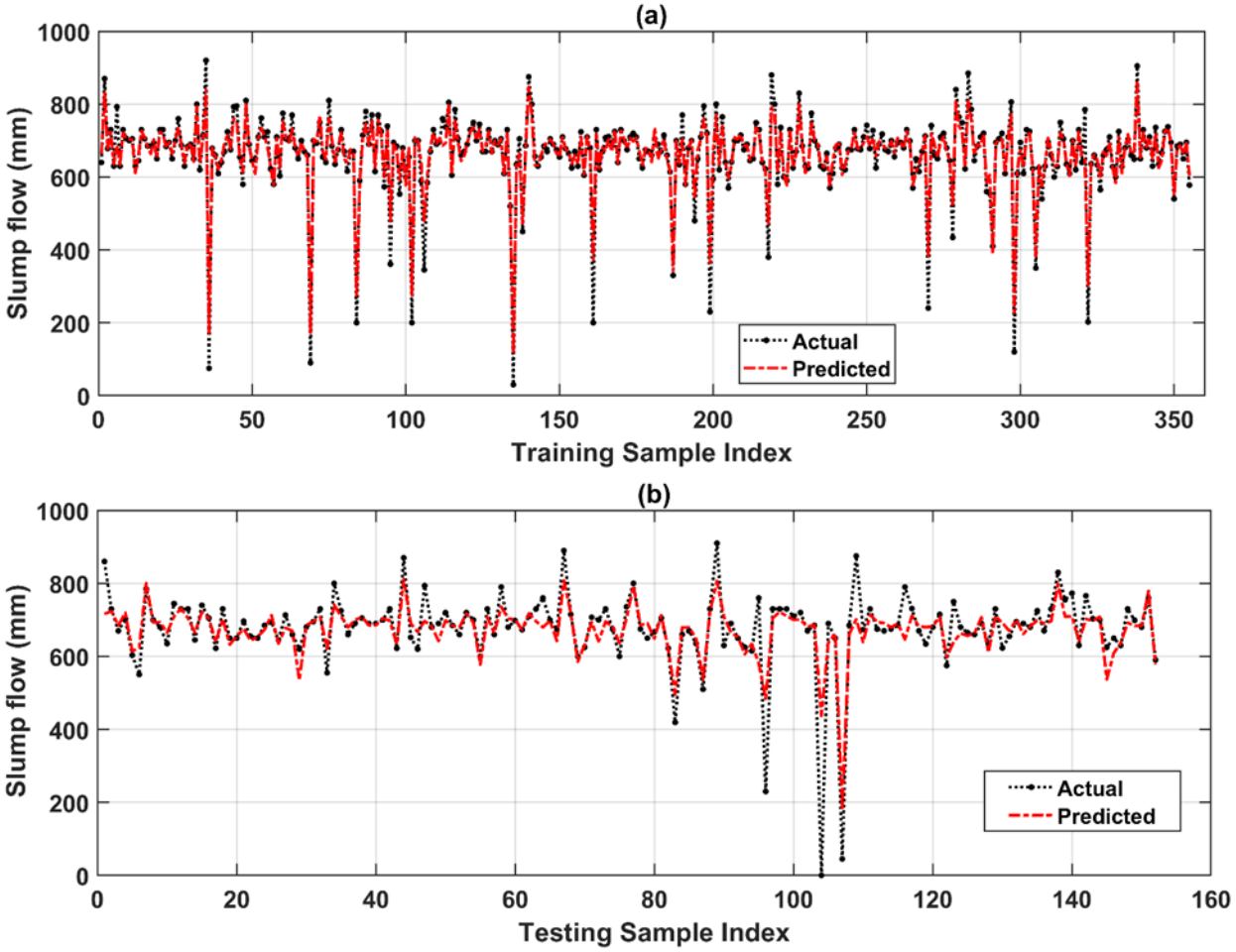


**Fig. 4.** Compare the predicted SF values from the RF model to the experimental values for the training (Fig. a) and testing data sets (Fig. b)

The best result (10th simulation - CV10) is shown in the next section. This result was chosen since the RMSE and MAE are the lowest, and the R is comparatively high among the ten simulations reported before. Fig. 4 illustrates the correlation between the real SF value achieved during the experiment (dotted line) and the simulated value produced using the RF model (dashed line) during

training and testing. Observations indicate that 355 samples in the training data set exhibit experimental outcomes that are reasonably compatible with the model's predicted values. Except for a few samples with variations, the experimental findings for the test data set are similarly reasonably close to the anticipated values (Sample No. 96 and Sample No. 104). However, the number of samples exhibiting this variation is negligible in comparison to the test's total of 152 samples. As a result, the RF model may be used to forecast the SF of SCC.

The regression model graph in Fig. 5 shows the correlation between the expected value predicted by the RF model and the actual value received from the experiment for the training and test data sets. Each Fig. shows the regression line as a solid blue line and the 95 percent confidence line as a dashed line. The closer the data is to the regression line, the more accurate the model's ability to forecast. It is seen that the predicted model's values for the training data set (Fig. 5a) and the test data set (Fig. 5b) are very similar to the experimental findings. This demonstrates the RF model's exceptional predictive capability.

Additionally, the performance of the RF model is assessed using the statistical criteria's value. The values for these criteria for the training and testing data sets are shown in Table 2. This is the best-predicted result (CV10). For the training set, R equals 0.9803, and for the test set, R equals 0.8614. The root mean square error (RMSE) values for the training and test data sets are 30.2382mm and 59.5664mm, respectively. The training set's MAE value is 17.2325 mm, whereas the test set's MAE value is 32.4483 mm. The reason for the significantly high RMSE and MAE results is because, out of 507 study data utilized, there is one with SF = 0 mm. As a consequence, with a data spectrum of the SF parameter ranging from 0 to 920 mm, the prediction results of this RF model are perfectly acceptable. These Fig.s demonstrate that it is feasible to utilize an RF model to forecast SF using SCC, so saving materials engineers time and money by eliminating the requirement for experimentation.

On the other hand, when a multivariable regression model (EQ4) is applied to the issue of forecasting SCC's SF, the following calculation formula is discovered:

$$Y = 0.5727X_1 - 0.0699X_2 + 1.1989X_3 + 0.5243X_4 + 0.8301X_5 + 1.3313X_6 + 1.5833X_7 - 0.1183X_8 + 34.9257X_9 + 0.7858X_{10} + 1.2823X_{11} + 0.275X_{12} \quad (4)$$

The assessment requirements for this technique are much lower than the prediction results from the RF model, notably for the testing set of the RF model, which has R = 0.8614, RMSE = 59.5664, MAE = 32.4483, compared to model EQ4, which has R = 0.5406, RMSE = 90.0212, MAE = 58.5860. Table 2 contains comprehensive results for R, RMSE, and MAE for both RF and EQ4 models on training and testing data sets.

Additionally, a regression Fig. illustrates the association between the predicted value according to the EQ4 model and the actual value for the training and test data sets (Fig. 6). The correlation between the predicted value predicted by the EQ4 model and the actual value for the training data set is shown in Fig. 6a; the correlation for the test data set is shown in Fig. 6b. Each image depicts the regression line as a solid blue line and the 95% confidence line as a dashed line. The model's prediction capacity increases in accuracy as the data gets closer to the regression line. However, as seen in Fig. 6, these data are significantly different from the regression line, particularly for the test set (Fig. 6b), where the 95 percent confidence line deviates significantly from the regression line. This demonstrates that the EQ4 model's predicting ability is much less accurate than that of the RF model.

The results proposed herein are compared with several existing models available in the literature. The RF model achieved higher

accuracy than the ANN ($R^2$=0.615) in Saha et al. [6], but lower than SVR ($R^2$=0.931). However, only 115 samples were considered in that work, compared with 507 samples in the present database. Additionaly, a wide range of material constituents are considered in this database, including limestone powder, blast furnace slag, silica fume, metakaolin, or rice husk ash.

Finally, feature important using permutation important is conducted, thanks to the built-in function of the RF model. Fig. 7 summarizes the findings from an analysis of the effect of 12 input parameters on output parameter (SF). The results indicate that the parameter with the greatest effect on SF is the content of superplasticizers additive. This parameter has a greater effect than the other eleven factors. Following that, in decreasing order of influence, the following factors have an effect on SF: fine aggregate content (4), coarse aggregate content (5), water (3), cement (1), fly ash (2), limestone powder (7), viscosity reducing additive (12), rice husk ash powder (11), blast furnace slag (8), silica powder (9), metakaolin (10). Thus, using RF to analyze the parameters controlling the SF of SCC, materials engineers may pre-orient and quantify the compositional components used in the SCC mixture. This enables them to create SCC that has SF fulfill the standards or are suitable for the intended usage.

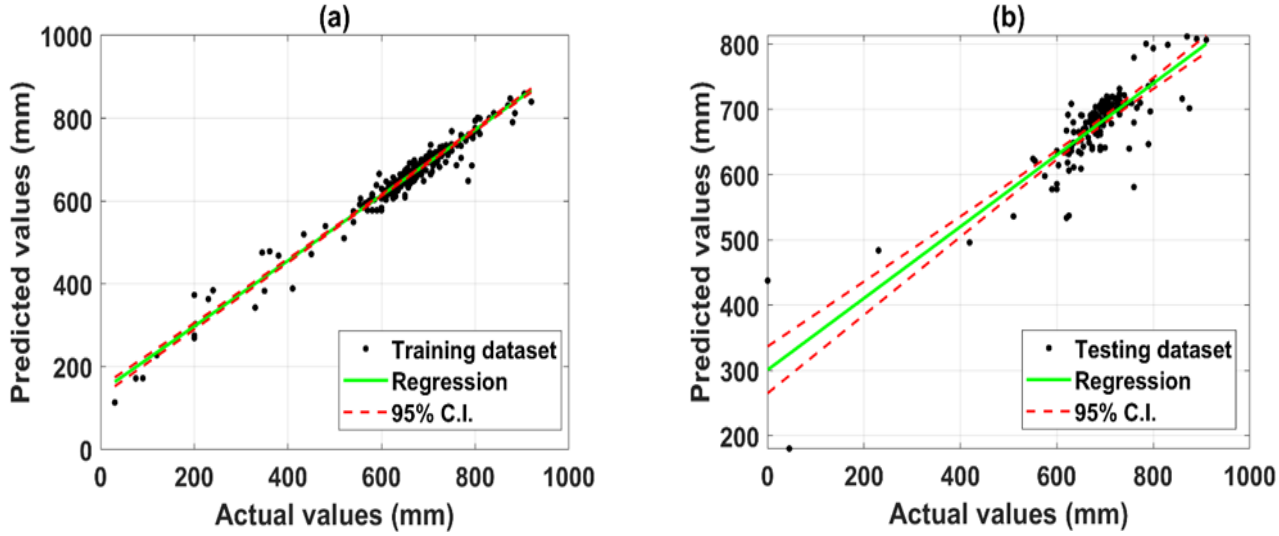


**Fig. 5.** Regression plot between experimental values and values simulated by RF model for (a) training set; and (b) testing set

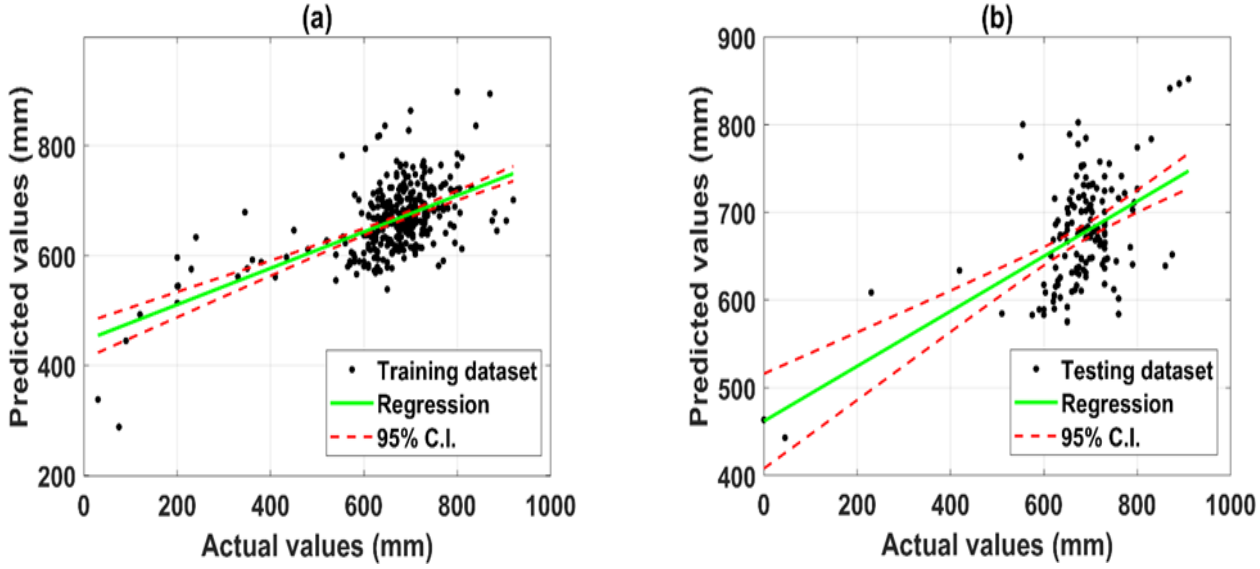


**Fig. 6.** Regression plot between experimental values and values simulated by EQ4 model for (a) training set and (b) testing set

**Table 2**. Summary of the criteria used to evaluate the prediction performance of the RF model for the best simulation and multivariate regression model

| | RMSE (mm) | MAE (mm) | R |
|---|---|---|---|
| RF training set | 30.2382 | 17.2325 | 0.9803 |
| RF testing set | 59.5664 | 32.4483 | 0.8614 |
| EQ4 training set | 93.5840 | 62.8994 | 0.5816 |
| EQ4 testing set | 90.0212 | 58.5860 | 0.5406 |

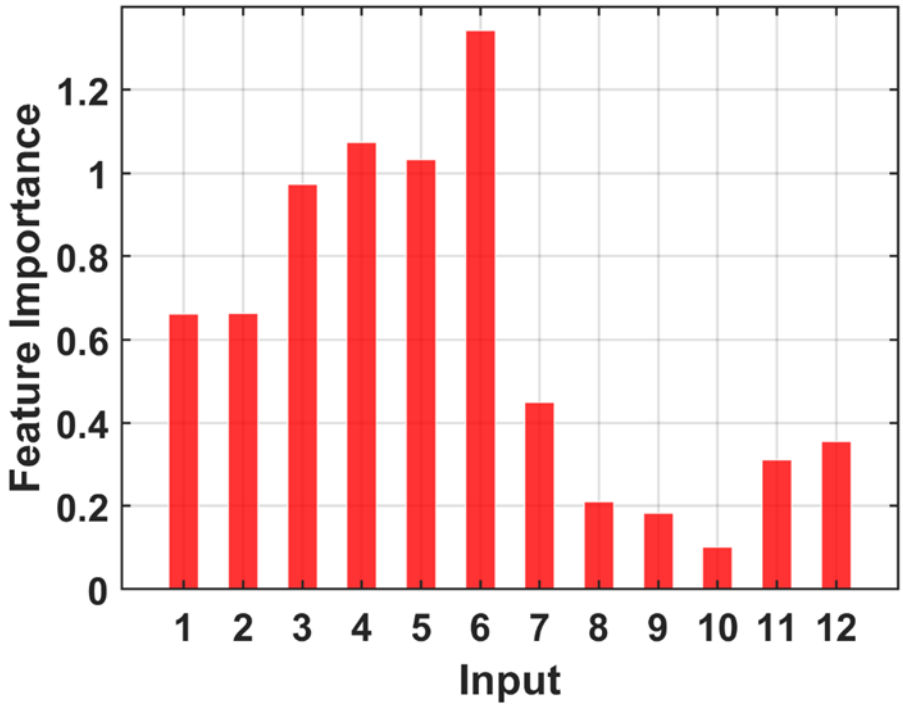


**Fig. 7.** The impact of the input parameters on the SF of the SCC

## 5. Conclusion

The purpose of this research was to use a RF model to forecast the SF of SCC while also evaluating the effect of input elements on the SF. The model was created using 507 experimental data points, including 12 input parameters and one output parameter. To increase the reliability of the simulation findings and to choose a model with more general predictive potential, this research performs 10-fold cross-validation on the training data set, coupling with random sampling technique. The study's findings indicate that the RF tool is capable of accurately predicting the SF of SCC, as shown by the RMSE = 59.5664 mm, MAE = 32.4483 mm, and R = 0.8614 for with the test data set. Finally, the RF model assesses the effect of input parameters on the SF of SCC. The findings indicated that the superplasticizer content had the most impact on SF, whereas the metakaolin content had the least. This forecasting tool will assist materials engineers in minimizing tests and optimizing the component material composition in the design of the SCC mix, hence saving money and time.

## References

[1] M. S. Abo Dhaheer, “Design and properties of self-compacting concrete mixes and their simulation in the J-ring test,” phd, Cardiff University, 2016. Accessed: Aug. 23, 2021. [Online]. Available: https://orca.cardiff.ac.uk/99770/

[2] “European Guidelines for Self Compacting Concrete (SCC) | EFCA.”

https://www.efca.info/download/european-guidelines-for-self-compacting-concrete-scc/ (accessed Feb. 18, 2022).

[3] T. A. Nguyen, H. B. Ly, H. V. T. Mai, and V. Q. Tran, “On the Training Algorithms for Artificial Neural Network in Predicting the Shear Strength of Deep Beams,” *Complexity*, vol. 2021, p. e5548988, May 2021, doi: 10.1155/2021/5548988.

[4] T. A. Nguyen, H. B. Ly, H. V. T. Mai, and V. Q. Tran, “Prediction of Later-Age Concrete Compressive Strength Using Feedforward Neural Network,” *Adv. Mater. Sci. Eng.*, vol. 2020, p. e9682740, Sep. 2020, doi: 10.1155/2020/9682740.

[5] H. V. T. Mai, V. Q. Tran, and T. A. Nguyen, “On the Training Algorithms for Artificial Neural Network in Predicting Compressive Strength of Recycled Aggregate Concrete,” in *CIGOS 2021, Emerging Technologies and Applications for Green Infrastructure*, Singapore, 2022, pp. 1867–1874. doi: 10.1007/978-981-16-7160-9_189.

[6] P. Saha, P. Debnath, and P. Thomas, “Prediction of fresh and hardened properties of self-compacting concrete using support vector regression approach,” *Neural Comput. Appl.*, vol. 32, no. 12, pp. 7995–8010, Jun. 2020, doi: 10.1007/s00521-019-04267-w.

[7] D. Sathyan, K. B. Anand, A. J. Prakash, and B. Premjith, “Modeling the Fresh and Hardened Stage Properties of Self-Compacting Concrete using Random Kitchen Sink Algorithm,” *Int. J. Concr. Struct. Mater.*, vol. 12, no. 1, p. 24, Mar. 2018, doi: 10.1186/s40069-018-0246-7.

[8] M. Sonebi, A. Cevik, S. Grünewald, and J. Walraven, “Modelling the fresh properties of self-compacting concrete using support vector machine approach,” *Constr. Build. Mater.*, vol. 106, pp. 55–64, Mar. 2016, doi: 10.1016/j.conbuildmat.2015.12.035.

[9] M. R. Kaloop, P. Samui, M. Shafeek, and J. W. Hu, “Estimating Slump Flow and Compressive Strength of Self-Compacting Concrete Using Emotional Neural Networks,” *Appl. Sci.*, vol. 10, no. 23, Art. no. 23, Jan. 2020, doi: 10.3390/app10238543.

[10] T. A. Nguyen, H. B. Ly, H. V. T. Mai, and V. Q. Tran, “Using ANN to Estimate the Critical Buckling Load of Y Shaped Cross-Section Steel Columns,” *Sci. Program.*, vol. 2021, p. e5530702, Apr. 2021, doi: 10.1155/2021/5530702.

[11] W. Ma, K. Tan, and P. Du, “Predicting soil heavy metal based on Random Forest model,” Jul. 2016, pp. 4331–4334. doi: 10.1109/IGARSS.2016.7730129.

[12] H. V. T. Mai, T. A. Nguyen, H. B. Ly, and V. Q. Tran, “Prediction Compressive Strength of Concrete Containing GGBFS using Random Forest Model,” *Adv. Civ. Eng.*, vol. 2021, p. e6671448, May 2021, doi: 10.1155/2021/6671448.

[13] H. V. T. Mai, V. Q. Tran, and T. A. Nguyen, “Using Random Forest for Predicting Compressive Strength of Self-compacting Concrete,” in *CIGOS 2021, Emerging Technologies and Applications for Green Infrastructure*, Singapore, 2022, pp. 1937–1944. doi: 10.1007/978-981-16-7160-9_196.

[14] M. Abu Yaman, M. Abd Elaty, and M. Taman, “Predicting the ingredients of self compacting concrete using artificial neural network,” *Alex. Eng. J.*, vol. 56, no. 4, pp. 523–532, Dec. 2017, doi: 10.1016/j.aej.2017.04.007.

[15] R. Siddique, P. Aggarwal, and Y. Aggarwal, “Prediction of compressive strength of self-compacting concrete containing bottom ash using artificial neural networks,” *Adv. Eng. Softw.*, vol. 42, no. 10, pp. 780–786, Oct. 2011, doi: 10.1016/j.advengsoft.2011.05.016.

[16] P. G. Asteris and K. G. Kolovos, “Self-compacting concrete strength prediction using surrogate models,” *Neural Comput. Appl.*, vol. 31, no. 1, pp. 409–424, Jan. 2019, doi: 10.1007/s00521-017-3007-7.

[17] P. Saha, P. Debnath, and P. Thomas, “Prediction of fresh and hardened properties of self-compacting concrete using support vector regression approach,” *Neural Comput. Appl.*, vol. 32, no. 12, pp. 7995–8010, Jun. 2020, doi: 10.1007/s00521-019-04267-w.

[18] S. Mahesh, “Self Compacting Concrete And Its Properties,” vol. 4, no. 8, p. 9, 2014.

[19] M. M. Al-Rubaye, “Self-compacting concrete: design, properties and simulation of the flow characteristics in the L-box,” phd, Cardiff University, 2016. Accessed: Nov. 13, 2021. [Online]. Available: https://orca.cardiff.ac.uk/98869/

[20] J. Seo, E. Torres, W. Schaffer, and South Dakota State University. Dept. of Civil and Environmental Engineering, “Self-Consolidating Concrete for Prestressed Bridge Girders,” 0092-15–03, Jul. 2017. Accessed: Nov. 13, 2021. [Online]. Available: https://rosap.ntl.bts.gov/view/dot/34197

[21] L. Breiman, “Random forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[22] R. Kohavi, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” in *Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2*, San Francisco, CA, USA, Aug. 1995, pp. 1137–1143.