Implementing Sobol's Global Sensitivity Analysis to SFRC's Flexural Strength Predictive Equation

*Tran T. H. Le ,Anh-Thang Le**

Faculty of Civil Engineering, Hochiminh City University of Technology and Education 70000, Vietnam; Email: huyenle3089@gmail.com

*Corresponding Author: Anh-Thang Le (Email:thangla@hcmute.edu.vn)
(Received: 27-May-2024; accepted: 29-August-2024; published: 30-September-2024)
http://dx.doi.org/10.55579/jaec.202483.464

Abstract. *Steel fibers are essential for SFRC since they strengthen the material's resistance to bending and cracking stresses and guarantee its endurance. However, the uncertainty of input parameters such as concrete compositions and steel fibers causes the stochasticity of flexural strength. The study uses big data to analyze the influence of fibers and other concrete compositions on the flexural strength properties of SFRC. For this purpose, the study focuses on developing predictive models for SFRC flexural properties based on a comprehensive database comprising two hundred and seven experimental results recorded by seventeen researchers. Bayesian Model Averaging is employed to identify significant components that influence the overall flexural strength and to develop a predictive flexural strength model. Monte Carlo simulation generates big data by utilizing the probability distribution of input variables and the predictive flexural strength model. The study used Sobol's global sensitivity analysis method to assess various input parameters' sensitivity to SFRC flexural strength based on the generated database. The impact order of input variables on the flexural strength is identified, as determined by the Sobol' Indice.*

Keywords: *Sobol' Method, Global Sensitivity Analysis, Flexural Strength, SFRC, Linear Regression, and Bayesian Model Averaging.*

1. Introduction

A significant quantity of concrete is produced annually, particularly in developing nations. Over fourteen billion cubic meters of concrete were used in the 2020 years, which is equivalent to approximately 4.4 tons for every person [1]. However, numerous barriers limit the performance of concrete in the field of construction engineering. A significant drawback of concrete is its relatively poor tensile strength, which is just 1/10–1/8 of its compressive strength. In order to address this problem, numerous studies are advocating the incorporation of steel fibers into the mixture for making Steel Fiber Reinforced Concrete (SFRC). This straightforward method is significantly efficient in improving the fracture resistance of concrete structural elements, which was first conducted by Bernard [2] as early as 1874. At that time, a wide range of researchers concentrated on identifying the mechanical properties of SFRC by conducting experiments at laboratories worldwide, especially flexural strength [3]].

Integrating steel fibers into plain concrete has been confirmed as an effective solution to enhance the capacity to bear the weight of structural components [4]. Flexural strength is a vital mechanical property of steel fiber reinforced concrete (SFRC), which is crucial for its performance in structural applications. Including steel fibers significantly enhances SFRC's ability to resist cracking and withstand bending loads, contributing to its toughness and durability in concrete structures. The impacts of fibers on the fundamental mechanical characteristics of SFRC and associated uncertainties are a topic of considerable study, and there are no good models to forecast them. Particularly remarkable is that these input random variables significantly affect the flexural strength of SFRC, which has a stochastic output value. This study is an effort to analyze the impact of input variables and flexural strength based on big data. The following steps were done to develop a predictive model. It includes building test databases, validating data distribution, analyzing and assessing correlations of the data's parameters, and developing a predictive model for estimating the flexural properties of SFRC.

A total of 207 experimental recordings were obtained from the publications of 17 researchers and inputted into a database. Bayesian Model Averaging (BMA) is employed to identify the most significant components that affect flexural strength and define the most optimal predictive model as well as a good simulation model.

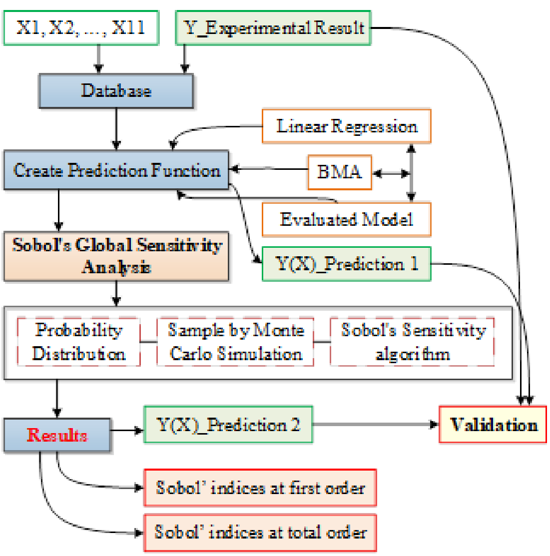

Fig. 1: Framework of the study.

Compared to methods like FAST, Morris, or others, the Sobol method is distinguished by its ability to provide precise global sensitivity analysis, particularly in capturing both main effects and interactions among input variables, although it requires more computational resources [5]. Thus, this study utilizes Sobol's method to perform global sensitivity analyses on the predictive model. Notably, most simulation models are complicated and nonlinear. Global sensitivity analysis is a common way to anticipate model performance and behavior. Sobol's method is a variance-based global sensitivity analysis approach that depends heavily on sampling and input parameter distribution, which is used in a nonlinear function to demonstrate its implementation and calculate input parameter sensitivities on model output. This study demonstrates the incorporation of Sobol's method into the predictive model to determine the sensitivity of a variety of input values, including water, cement, sand, coarse aggregate, fiber length, fiber diameter, fiber volume, and fiber tensile yield strength, to the flexural strength of SFRC. Fig. 1 shows the methodology within the context of this investigation.

2. Global sensitivity analysis

Sensitivity analysis, as defined by Saltelli et al. (2000) [6], is the investigation of how variations in the output of a model can be apportioned, qualitatively or quantitatively, to different sources of variations and how the input information influences the model's behavior. It can be another way known that sensitivity analysis examines how alterations in the input parameters of a mathematical model affect the variability of its output [7]. This study provides an overview of Sobol's global sensitivity analysis method applied to the model used for estimating a mechanical property of SFRC, which is the flexural strength via a predictive model related to 11 stochastic input variables.

Fig. 2 illustrates the comprehensive process of the global sensitivity analysis methods. For

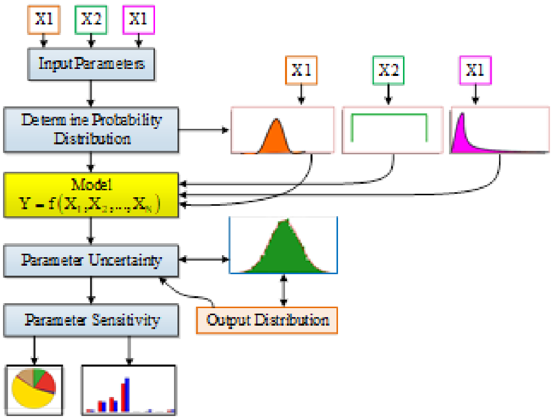

Fig. 2: A diagram for a global sensitivity analysis.

global sensitivity analysis to be conducted, the subsequent procedures are executed:

1. Probability and distributions are present in all input parameters;
2. The probability distribution of each parameter generates random input vectors;
3. Each random input vector has been provided with output values by the prediction model;
4. The probability distribution of the output is determined;
5. Sensitivity analysis is used to prioritize the input parameters based on their impact on the output.

If the input variables are not specified at the interval, only one variable change is required to determine the output's sensitivity to the variation of an input parameter. The following assumption must be made to implement global sensitivity analysis procedures: that defined $\phi \sim [0,1]^p$ on the interval. Within the given framework, it is thus feasible to demonstrate that ϕ may be broken down into basic functions.

$$\begin{aligned} Y = \phi(X) = \phi_0 + \textstyle\sum_{i=1}^{p} \phi_i\left(X^{(i)}\right) \quad + \ldots \\ \textstyle\sum_{1\leq i<j\leq p} \phi_{ij}\left(X^{(i)}, X^{(j)}\right) + \ldots + \ \phi_{1\ldots p} \end{aligned} \quad (1)$$

where ϕ_0 a constant, $\phi_{1\ldots p}\left(X^{(1)}, ..., X^{(p)}\right)$ verifies for

$$\begin{aligned} \forall k \in \{1, ..., s\} \text{ v } \{i_1, ..., i_s\} \subseteq \{1, ..., n\} : \\ \textstyle\int_0^1 \phi_{i_1,\ldots i_s}\left(X^{(i_1)}, \ldots, X^{(i_s)}\right) d \end{aligned} \quad (2)$$

Since the random and independent form of the data inputs $X^{(i)}$, by using the variance operator on the equation, one may derive the ANOVA decomposition.

$$Var(Y) = V = \textstyle\sum_{i=1}^{p} V_i + \sum_{1\leq i<j\leq p} V_{ij} \quad (3)$$

Through the variance operator VAR:

$$\begin{aligned} V_i &= Var\left(E\left(Y|\, X^{(i)}\right)\right) \\ V_{ij} &= Var\left(E\left(Y|\, X^{(i)}, X^{(j)}\right)\right) - V_i - V_j \\ V_{ijk} &= Var\left(E\left(Y|\, X^{(i)}, X^{(j)}, X^{(k)}\right)\right) - \ldots \\ & V_i - V_j - V_k - V_{ij} - V_{ik} - V_{jk} \end{aligned} \quad (4)$$

where E indicates the expectation in mathematics. Sobol sensitivity indices at first order S_i for $X^{(i)}$ as:

$$S_i = \frac{V_i}{V} = \frac{Var\left(E\left(Y|\, X^{(i)}\right)\right)}{Var(Y)} \quad (5)$$

Sobol sensitivity indices at second-order S_{ij} for $X^{(i)}$ and $X^{(j)}$ as:

$$S_{ij} = \frac{V_{ij}}{V} = \frac{Var\left(E\left(Y|\, X^{(i)}, X^{(j)}\right)\right) - V_i - V_j}{Var(Y)} \quad (6)$$

Sensitivity indices of higher order may be defined similarly. The interpretation of the sensitivity indices is straightforward since they range from 0 to 1, and their total sum is 1. If the value of S_i is near to 1, it indicates that the variable $X^{(i)}$ has a significant impact on Y. As the input dimension p rises, the number of Sobol indices grows exponentially, making it difficult to estimate all of these indices. To achieve that objective, total Sobol sensitivity indices S_{T_i} for each variable $X^{(i)}$ as:

$$S_{T_i} = \textstyle\sum_{k\#i} S_k = S_i + S_{ij} + S_{ik} + \ldots + S_{i\ldots n} \quad (7)$$

where $\#i$ signifies the collection of index sets that include i.

This study aims to analyze the Monte Carlo method for estimating Sobol sensitivity indices. The Monte Carlo method is a scientific approach that enables the acquisition of numerical results without physical experimentation. We can determine the probability distributions of our major parameters using experimental findings or other data. Subsequently, we use distribution data to construct a statistical sampling set with vast amounts of data. The estimation of a parameter using Monte Carlo methods is represented by the notation B in the subsequent expression.

Let's examine two realizations of N sample size.

$$Xi_k = \left(x^i_{k1}, x^i_{k2}, \ldots, x^i_{kp}\right), \quad k = 1\ldots N, \quad i = 1,2 \tag{8}$$

$$\widehat{S}_i = \frac{\widehat{V}_i}{\widehat{V}} = \frac{\widehat{U}_i - \widehat{\phi}_0^{\,2}}{\widehat{V}} \tag{9}$$

The mean is calculated using:

$$\widehat{\phi}_o = \frac{1}{N}\sum_{k=1}^{N} \phi(x^1_{k1}, ..., x^1_{kp}) \tag{10}$$

The variance is estimated by definition with

$$\widehat{V} = \tfrac{1}{N}\sum_{k=1}^{N} \phi^2(x^1_{k1}, ..., x^1_{kp}) - \widehat{\phi}_0^{\,2} \tag{11}$$

And finally, the term $\widehat{U}_i$ is derived by combining the two sample size realizations of X with N:

$$\widehat{U}_i = \tfrac{1}{N}\sum_{k=1}^{N} \phi\left(x^1_{k1}, \ldots, x^1_{k(i-1)}, x^1_{ki}, x^1_{k(i+1)}, \ldots, x^1_{kp}\right) \phi\left(x^2_{k1}, \ldots, x^2_{k(i-1)}, x^1_{ki}, x^2_{k(i+1)}, \ldots, x^2_{kp}\right) \tag{12}$$

By estimating $\widehat{U}_i$, the influence of the samples on dimension i can be assessed. Second-order sensitivity indices $\widehat{S}_{ij}$ can be approximated using the following equations:

$$\widehat{S}_{ij} = \frac{\widehat{U}_{ij} - \widehat{\phi}_0^{\,2} - \widehat{V}_i - \widehat{V}_j}{\widehat{V}} \tag{13}$$

with

$$\widehat{U}_{ij} = \tfrac{1}{N}\sum_{k=1}^{N} \phi\left(x^1_{k1}, ..., x^1_{k(i-1)}, x^1_{ki}, x^1_{k(i+1)}, ..., x^1_{k(j-1)}, x^1_{kj}, x^1_{k(j+1)}, ..., x^1_{kp}\right) \phi\left(x^2_{k1}, ..., x^2_{k(i-1)}, x^1_{ki}, x^2_{k(i+1)}, ..., x^2_{k(j-1)}, x^1_{kj}, x^2_{k(j+1)}, \ldots, x^2_{kp}\right) \tag{14}$$

Sensitivity indices with greater varieties are subsequently generated from these interactions. Total sensitivity indices can be approximated using the following equations:

$$\widehat{S}_{T_i} = 1 - \frac{\widehat{U}_{\widehat{i}} - \widehat{\phi}_0^{\,2}}{\widehat{V}} \tag{15}$$

where:

$$\widehat{U}_{\widehat{i}} = \tfrac{1}{N}\sum_{k=1}^{N} \phi\left(x^1_{k1}, \ldots, x^1_{k(i-1)}, x^1_{ki}, x^1_{k(i+1)}, \ldots, x^1_{kp}\right) \phi\left(x^1_{k1}, ..., x^1_{k(i-1)}, x^2_{ki}, x^1_{k(i+1)}, ..., x^1_{kp}\right) \tag{16}$$

Thus, in this section, one has described how Monte Carlo approaches are used to estimate sensitivity Sobol indices for various orders. These indices allow for the analysis of the impact of various inputs on the variance of Y.

3. Flexural strength predictive model

3.1. Selecting experimental database

Building models that estimate the 28-day flexural strength of SFRC using test data from 17 independent studies were combined into a database with 207 test results, shown in Table 2. In the database, variables are named and described in Table 1. Fiber shape constant includes the straight fiber (S), hooked-end fiber (H), crimped fiber (CR), mill-cut fiber (M), and flat-end fiber (F). The fiber shape constant values [8] comprise fiber with two hooked ends (=1.0), two crimped ends, or one hooked end and straight end (=0.75), in addition to straight, mill-cut fiber (=0.5).

3.2. Flexural strength equation

Regression analysis [9] is a statistical method utilized to estimate the correlation between variables that exhibit the relationship of reason and result. The primary objective of univariate regression is to establish a linear relationship between a dependent and an independent variable by analyzing their relationship and formulating an equation for that relationship. Linear regression refers to models of regression that consist of a single dependent variable and multiple independent variables. The data used for linear regression analysis consists of 11 independent input variables and one dependent variable of flexural strength of SFRC, and the relationship is utilized to make a predictive model. The formulation of a multivariate regression analysis model is as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \varepsilon \quad (17)$$

where, Y is the dependent variable.$X_1, X_2, ..., , X_p$ are dependent variables. $\beta_0, \beta_1, \beta_2, ..., \beta_p$ are parameters. Moreover, ε is an error.

Conventional linear regression presumes that a singular model with fixed coefficients accurately represents the output data. Nevertheless, there may be some degree of uncertainty regarding the coefficients of the variables and their significance. BMA [7] provides a more comprehensive understanding of the relationship between variables by considering a range of models. The study uses a statistical technique, which is BMA, to solve this problem.

BMA [10] addresses uncertainty in determining the appropriate regression model by simultaneously considering multiple plausible models about the flexural strength of SFRC. BMA weights these models based on their posterior probabilities and combines them to provide a comprehensive and flexible view of the relationships between variables. This helps improve predictions, identify important variables, and incorporate prior knowledge about variable relationships.

Tab. 1: Input variables of the Flexural Strength Predictive Equation.

Symbol	Description
X_1	Mass of water
X_2	Mass of cement
X_3	Mass of sand
X_4	Mass of coarse aggregate
X_5	Maximum aggregate size
X_6	Mass of admixture
X_7	Fiber shape constant
X_8	Fiber tensile yield strength
X_9	Fiber volume fraction
X_{10}	Fiber length
X_{11}	Fiber diameter

The distribution of λ given the data $X = [X_1, X_2, \ldots, X_{11}]$, which is illustrated in Table 1. Let $M = [M_1, M_2, \ldots, M_k]$ indicates the collection of all possible models.

$$p(\lambda|X) = \sum_{k=1}^{N} p(\lambda|M_k\ ,\ X)\, p(M_k\ |X) \quad (18)$$

In Eq. 18, the probability of each predictive model, $p(\lambda|M_k\ ,\ X)$ are weighted by the posterior model, $p(M_k\ |X)$ is the probability of the model. The term is produced by reapplying the Bayes rule, except this time, it is done at the level of models instead of parameters.

$$p(M_k\ |X) = \frac{p(X|M_k)\, p(M_k)}{\sum_{l=1}^{K} p(X|M_l\ )\, p(M_l)} \quad (19)$$

where $p(X|M_k) = \int p(X|\beta_k, M_k)\, p(\beta_k|M_k)\, d\beta_k$. The model's marginal likelihood is $p(\lambda|X)$. β_k, M_k is the vector of parameters of M_k model; $p(\beta_k|M_k)$ is the prior density of β_k under model M_k ; $p(X|\beta_k, M_k)$ is the probability.

The $p(M_k)$ is the prior likelihood, that M_k is the accurate model. The set M concludes all models, which depend on all probabilities. For this investigation, M includes all valid predictor combinations.

Just performing the calculation of the ratio between the posterior model probabilities of the alternative model (M_k) and the null model (M_l) is sufficient. This results in

$$\frac{p(M_k\ |X)}{p(M_l\ |X)} = \frac{p(X|M_k)}{p(X|M_l)} \frac{p(M_k)}{p(M_l)} = BF_{10} \frac{p(M_k)}{p(M_l)} \quad (20)$$

where BF_{10} is Bayes's factor. The previous model odds are represented by the ratio on the right-hand side, while the posterior model odds are represented by the ratio on the left. All of the models that are being considered here follow the form:

$$\lambda = \beta_0 + \sum_{i=1}^{m} \beta_i X_i + \varepsilon \quad (21)$$

The dependent variable's observed data is represented by the number of λ vectors. A normal distribution denoted as ε is provided, consisting of a variance of σ^2 and a mean of zero. It is assumed that the ε' values under various conditions are independent.

The Bayesian Model Averaging (BMA) [28] algorithm utilizes the Bayes Information Criterion

Tab. 2: Summary of the flexural strength database of SFRC.

Sources	Water (m^3)	Cement (kg)	Sand (kg)	Coarse aggregate (kg)	Admixture (kg)	s_{max} (mm)	Fiber shape	f_f (MPa)	V_f (%)	l_f (mm)	d_f (mm)	f_r (MPa)
Nili, Afroughsabet [11]	162-177	354-450	893-920	858-887	2.3-6.5	19	H	1050	0-1.00	60	0.75	4.24-9.43
Ibrahim, Chebaker [12]	230	338	1049	760	0.7	10	H	-	0-1.25	60	0.75	6.10-7.52
Sahin, Koksal [13]	170-179	325-487	785-891	850-965	0.9-1.8	19	H	1050-2000	0-1.00	60	0.71-0.75	3.30-17.30
Zhang [14]	164-185	308-529	633-764	929-1233	0-5.8	20	H	600-1000	0-2.00	35	0.55	3.10-8.32
Chen [15]	172-195	336-521	488-725	1080-1145	0-5.2	Oct-40	H	1100	0-1.00	30-60	0.75	4.68-9.31
Bai et al. [16]	165	367	702-765	1053-1146	2.2	16	CR	380	0-2.00	30	0.5	4.79-9.30
Abbass et al. [17]	137-157	350-550	682-798	1050-1078	3.5-3.7	10	H	1250	0-1.50	35	0.6	4.35-13.60
Yang et al. [18]	185	268-524	652-750	1056-1186	0.4-1.6	31.5	H	1000	0-0.63	40-60	0.62-0.75	2.50-4.80
Fan [19]	264	480	717-769	895-989	0	20	S, H	380-500	0-2.00	50	1	4.32-9.11
Raja, Sivakumar [20]	160	400	750	1140	0	20	H	1700	0-0.50	30.2-32.3	0.93-1.22	4.96-5.74
Gao [21]	215	500	556	1129	8	16	H	1345	0-1.0	30	0.5	7.61-19.50
Ma [22]	161	460	1150	1048	0	20	M	809	1-1.60	30	0.55	6.05-8.54
Zhu [23]	160-200	258-513	540-868	1012-1083	18.4	20	H, CR, M	500-1250	0-0.70	32	0.8	4.37-6.94
Niu et al. [24]	172	400	730	1046-110	0	15	H	-	0-2.00	32-50	0.55-1.15	4.59-10.42
Peng [25]	161-167	453-460	699	1594	0	20	CR	809	0-1.60	32	0.8	6.24-8.43
Soutsos et al. [26]	198	267	805	1190	18.1-18.4	20	H, CR, F	-	0-0.64	32-40	0.8	4.21-4.76
Pajak,Ponilieski [27]	205	490	808	808	0	8	S, H	1100-1250	0-1.00	50-60	0.9-1.0	2.45-8.31
Statistical indexes for the database of flexural strength (f_r)												
	160-264	267-500	556-1150	760-1594	0-18.4	8.0-40.0		380-2200	0-3.00	12.5-60	0-1.22	2.45-19.50

(BIC) to estimate the posterior probabilities of various models, which are then used to weigh these models in the averaging process. The influence of each input parameter on the final outcome is determined by its prevalence across the models and the weighted contributions of those models within the ensemble.

This paper employs the flexural strength of SFRC testing values to investigate a numerical model with 11 parameters corresponding to 207 times the variables for each test (i.e., the data contribution of 207 sets of X input and Y output). The R programming language is used to identify models that can forecast flexural strength based on input parameter variables.

The BMA discovers 26 possible models. Five of these are classified to be the best based on assessment using the determination factor (R^2), BIC criterion, and the number of variables required to predict outcomes. Fig. 3 shows the likelihood of predictors in models that predict SFRC's flexural strength. The results illustrate that the impact of uncertainty in input fiber shape constant values has a negligible effect on this strength. The mass of cement (X_2), the yield strength (X_8), and the volume (X_9) of fiber are very important since they appear in the models at 100% of the total. Their linear regression coefficients of the predicted models are 1.80E-02 with a standard deviation of 4.97E-03, 3.60E-03 with a standard deviation of 6.78E-04, and 2.86 with a standard deviation of 3.65E-01, as demonstrated in Model 1 column of Table 3. The analysis results of X_1 in Fig. 3 show that

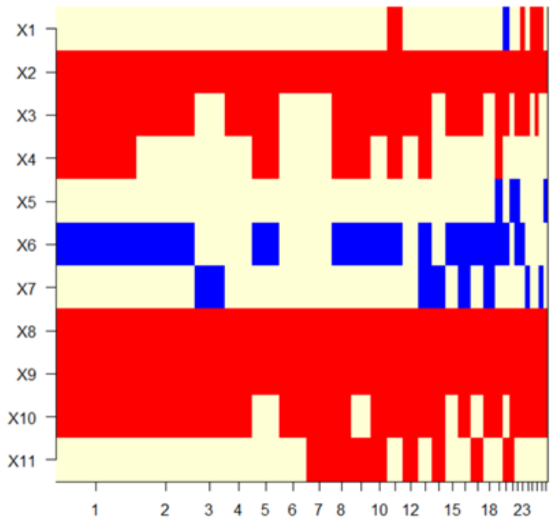

Fig. 3: Models of prediction selected via the BMA method.

adding the water mass to the linear regression model will reduce accuracy. The primary variable that must be incorporated into the regression function is the cement mass X_3. The actual water-to-cement (w/c) ratio is crucial. However, the cement mass X_3 could somehow represent it.

The results via the BMA method analysis of multivariable linear regression models for flexural strength, a mechanical property of SFRC, show that steel fiber's geometrical and material properties slightly affect the predictive models. In particular, the parameters such as the mass of cement, sand, aggregate, and the variables re-

Tab. 3: A summary of BMA's five best predictive models.

Variable	*Prob.*	*EV*	*SD*	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	100	-1.28E+01	7.31E+00	2.00E+01	1.12E+01	2.98E+00	8.49E+00	2.14E+01
X_1	8.3	3.11E-04	3.07E-03	-	-	-	-	-
X_2	100	1.34E-02	4.97E-03	1.80E-02	1.34E-02	7.95E-03	1.00E-02	1.99E-02
X_3	74.5	6.33E-03	4.91E-03	1.08E-02	6.86E-03	-	3.92E-03	1.29E-02
X_4	36.7	1.46E-03	2.13E-03	3.75E-03	-	-	-	4.11E-03
X_5	4.7	9.92E-04	8.61E-02	-	-	-	-	-
X_6	66.1	9.87E-02	1.17E+00	1.69E-01	1.08E-01	-	-	2.23E-01
X_7	18.3	4.65E-01	2.03E-01	-	-	2.84E+00	-	-
X_8	100	3.73E-03	6.78E-04	3.60E-03	3.40E-03	3.49E-03	3.21E-03	4.22E-03
X_9	100	2.86E+00	3.65E-01	2.89E+00	2.82E+00	2.82E+00	2.90E+00	2.80E+00
X_{10}	84.1	4.49E-02	2.72E-02	4.21E-02	4.53E-02	8.00E-02	6.27E-02	-
X_{11}	27.1	6.62E-01	1.29E+00	-	-	-	-	-
Number of variables (nVar)				7	6	5	5	6
Determination factor (R^2)				0.591	0.562	0.543	0.542	0.557
The Bayes Information Criterion (BIC)				-92.034	-91.39	-90.0334	-89.879	-89.875
Posterior probability				0.164	0.119	0.08	0.037	0.035

*Note: Probability (**Prob.**), Expected value (**EV**), Standard deviation (**SD**).

lated to the fiber's size play a critical role in building the SFRC's flexural strength predictive equation. Thus, when the step-by-step process is over, this study suggests a predictive equation as follows:

$$F_{pre} = \beta_0 + \beta_1 X_1 + \beta_2 X_2, \ldots, +\beta_{11} X_{11} \quad (22)$$

3.3. Evaluation

Numerous indicators and techniques are used to evaluate a linear regression equation [29]. This study used R-squared (R^2) and residual mean squared error (RMSE) for evaluating the proposed prediction equation. Fig. 4 evidence that the proposed model of this study outperforms the model of Wang [1] across both R^2 and RMSE metrics. The proposed model exhibits a significantly higher R^2 value of 0.591, indicating its superior ability to explain the variance in the dependent variable compared to Wang's model, which only achieves an R^2 of 0.431. This disparity suggests that the proposed model captures a larger proportion of the underlying patterns and relationships within the data, making it a more robust predictive model.

Furthermore, the proposed model boasts a lower RMSE of 1.865, implying that its predictions are, on average, closer to the observed values compared to Wang's model, which has an RMSE of 2.137. The smaller RMSE of the proposed model indicates that it produces more accurate predictions, enhancing its reliability for practical applications such as forecasting or decision-making.

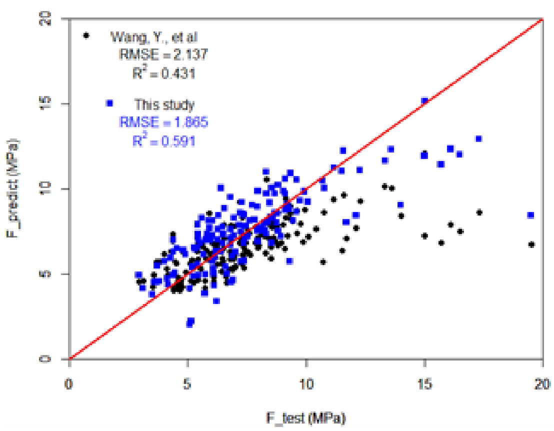

Fig. 4: Models of prediction selected via the BMA method.

4. Implementation of sobol's global sensitivity analysis method to the flexural strength of SFRC

4.1. Application of Sobol's methods of sensitivity analysis on a simple case

The study utilized a validation case study that employs the Ishigami function model to determine the reliability of the Sobol sensitivity analysis code. The Ishigami function, developed by Ishigami and Homma (1990) [30], is often used as an example of uncertainty and sensitivity analysis methods. This function is particularly useful due to its prominent nonlinearity and nonmonotonicity, making it an ideal case for studying these methods

$$Y(X) = \sin\left(X^{(1)}\right) + 7\sin^2\left(X^{(2)}\right) + ... \quad 0.1\left(X^{(3)}\right)^4 \sin\left(X^{(1)}\right) \tag{23}$$

where $X^{(1)}$, $X^{(2)}$ and $X^{(3)}$ are random variables that follow a uniform distribution between −ß and ß. By applying the global sensitivity analysis method and, more specifically, the estimation of the Sobol indices, the result can be generated using Monte Carlo methods to determine the mean through 20 retrials of 10.000 samples. The results of this sensitivity analysis were compared with previously published results [7,31], as shown in Table 5. The parameters and $X^{(2)}$ are the most influential when examining the variance of Y at first-order. Additionally, the Y value is least affected by $X^{(3)}$, as evidenced by a Sobol index of 0.001.

Table 3 shows that the theoretical analytical sensitivity results in the section are consistent with both prior experiments. The difference of random variables at the relative level is the highest at 5,49%, and there are 0% errors. The single in the $X^{(3)}$ variable has an error of up to 80%, but it can be seen that the first sensitivity indicator of the variable $X^{(3)}$ is approximately equal to 0 false candles. The conclusion is that the survey demonstrates the reliability of Sobol's methods of sensitivity analysis model in this study.

4.2. Application of Sobol's methods of sensitivity analysis on SFRC's flexural strength predictive equation

We use Sobol's sensitivity analysis method to look at how 11 different input variables affect the variation of the flexural strength of SFRC. We look at how these variables affect each other and how they affect the flexural strength of the material. Determining the probability distribution of all inputs plays a significant role in the analysis process. The probability distributions of 11 considered variables are shown in Table 4. The data generation is conducted using the Monte Carlo method based on input variable distributions, which randomly generates each variable with many data points.

For a sample size of 10000, Sobol's indices can be computed 10000 times by re-sampling the Monte Carlo draws, thereby providing various estimates. Fig. 5 depicts the comparison of the frequency distribution between the experimental data and the estimated value by Sobol's method. It is evident that there is a slight difference between them when examining their mean, skewness, and kurtosis. The difference between the mean of the experimental data (9.376) and that estimated by Sobol's method (7.606) is approximately 1.7 times. Similarly, the skewness and kurtosis values of the experimental data and estimation by Sobol's method are -0.027, 0.730, and 0.318, 3.327, respectively. Thus, the objective assessment reveals slight differences between the two models. This difference is mainly due to the different number of samples in each data set, and the larger the data, the more accurate it is, such as the 10000 samples estimated by Sobol's method. Fig. 6 shows the Sobol's sensitivity indices results. The figure displays the first and total order of all random variables. The significance of either the first-order or total-order indices when analyzing data depends on the specific context and analytical objectives. First-order indices are considered more

Tab. 4: Summary of comparisons of applied studies with the Sobol method.

X^i	The Sobol's indices						Errors of result (%)			
	Jérôme Morio [7]		DX Hung [31]		This study		[7]		[31]	
	S_i	S_{Ti}	S_i	S_{Ti}	S_i	S_{Ti}	S_i	S_{Ti}	S_i	S_{Ti}
$X^{(1)}$	0.313	0.576	0.313	0.56	0.313	0.565	0	1.91	0	1.802
$X^{(2)}$	0.434	0.438	0.44	0.434	0.426	0.446	1.843	1.826	3.182	2.765
$X^{(3)}$	0.001	0.254	0.004	0.241	0.005	0.255	80	0.394	25	5.49

Tab. 5: Summary of the variable distribution.

Variable	Distribution
X_1	Uniform
X_2	Uniform
X_3	Uniform
X_4	Uniform
X_5	Uniform
X_6	Uniform
X_7	Random with three determined values of 0.5, 0.75, and 1.0 [8]
X_8	Normal [32]
X_9	Uniform
X_{10}	Uniform
X_{11}	Uniform
SFRC's Flexural Strength (Testing)	Normal [33]

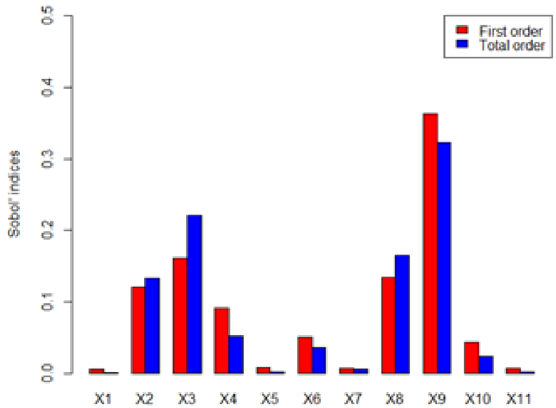

Fig. 6: Sobol's sensitivity indices from global sensitivity analysis.

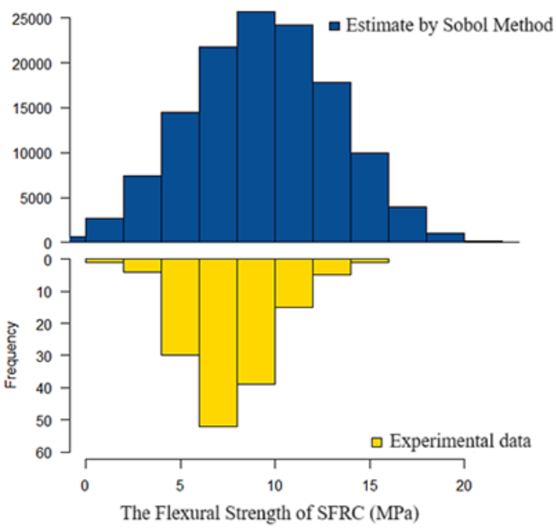

Fig. 5: Comparision of the frequency distribution.

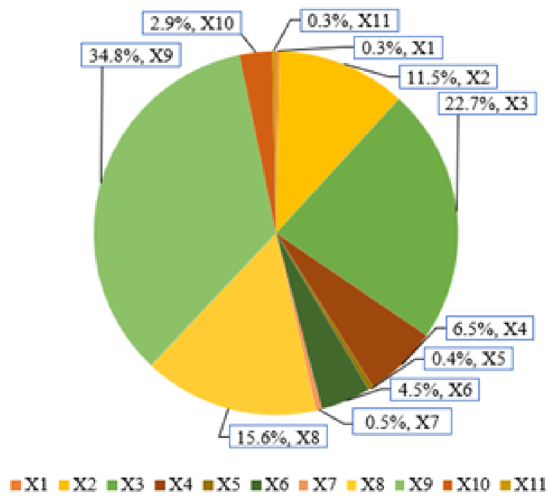

Fig. 7: First-order Sobol' Indice.

important since they demonstrate how certain input variables directly influence the variability of a model's output and identify the components with the most significant impact. On the other hand, total-order indices assess the overall influence of each input variable, considering both direct and interaction effects.

Fig. 7 demonstrates pie charts of first-order Sobol indices. The fiber volume (X_9) is the most sensitive, at approximately 34.8%. Next is the mass of sand (X_3) with a percentage ratio of about 22.7%, the tensile yield strength of the fibers (X_8), and the mass of cement (X_2) with sensitivity percentage ratios of approximately 15.6% and 11.5%, respectively.

The cement content significantly influences the flexural strength of concrete, as demonstrated by empirical studies, while the quantity of sand has a significantly smaller impact. The Sobol index is employed to assess the impact of a single input variable on the variability of the output. The Sobol index denotes the percentage of the output variation that the input variable can independently account for. Sobol's theory demonstrates that the sand mass (X_3) has a greater impact on the change in flexural strength than the cement mass (X_2) and yield strength (X_8) variables. The flexural strength variation is 22.7%, attributable to the sand mass variation. It does not imply that the sand mass influences the flexural strength values of concrete more than the yield strength (X_8), cement mass (X_2), and fiber volume (X_9).

The total-order Sobol index in Fig. 8 indicates that the fiber volume fraction (X_9) contributes most to the variability, with a value of approximately 37.2%. Following this are the cement mass (X_2) and fiber tensile yield strength (X_8), with variability contributions of roughly 15.1% and 14.8%, respectively. It could be seen that the order of influence remains consistent between the first-order indice and the total-order indice.

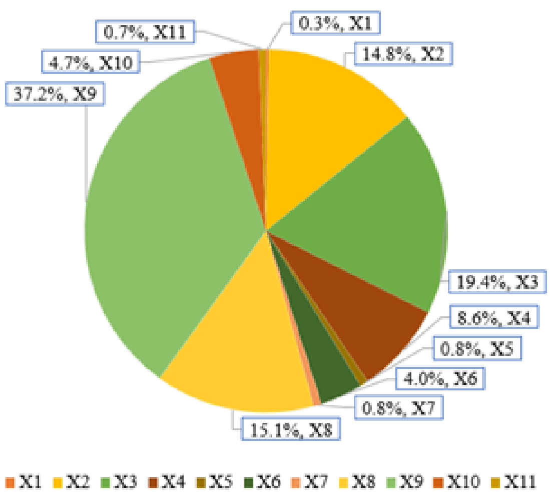

Fig. 8: First-order Sobol' Indice.

5. Conclusions

This study plays a critical role in constructing the SFRC Flexural Strength Predictive model in order to comprehend interrelationships among variables, generate comprehensible insights, and generate predictions predicated on linear assumptions. This model is expected to be a foundational tool in statistical analysis and various research and decision-making processes, especially sensitivity analysis. By capturing the linear relationships between input features and output responses, linear models offer simplicity, transparency, and often robust performance in implementing Sobol's Global Sensitivity Analysis to SFRC's Flexural Strength Predictive Equation.

This study uses the Sobol sensitivity index and Monte Carlo simulation technique to study the algorithm and create a global sensitivity analysis. The utilization of sensitivity data is anticipated to substantially influence the advancement of technical procedures and the production of building materials and structures. In reliability analysis, the input parameters showing a significant or minor global sensitivity to the output parameter are considered deterministic.

The proposed predictive equation for SFRC's flexural strength with high verification R^2=0.591. According to Sobol's Global Sensitivity Analysis, the fiber volume fraction exhibits the highest level of sensitivity, with a value of around 34.8%. The mass of sand accounts for approximately 22.7% of the total, while the tensile yield strength of fiber and the mass of cement have sensitivity percentages of around 15.6% and 11.5%, respectively. The order of influence is similar between the first-order and total-order indexes.

6. Acknowledgment

This research is supported by the Ho Chi Minh City University of Technology and Education.

References

[1] Y. Wang, H. Jin, C. Demartino, W. Chen, and Yong Yu. Mechanical properties of sfrc: Database construction and model prediction. *Case Stud. Constr. Mater.*, 17:e01484, 2022.

[2] Y. Yu, X.Y. Zhao, J. Xu, S.C. Wang, and T.Y. Xie. Evaluation of shear capacity of steel fiber reinforced concrete beams without stirrups using artificial intelligence models. *Materials*, 15(7), 2022.

[3] S.T. Kang and J.K. Kim. The relation between fiber orientation and tensile behavior in an ultra high performance fiber reinforced cementitious composites (uhpfrcc). *Cem. Concr. Res.*, 41(10):1001–1014, 2011.

[4] T. Job and R. Ananth. Mechanical properties of steel fiber-reinforced concrete. *J. materials civil engineering*, 19(5):385–392, 2007.

[5] A. Saltelli, S. Tarantola, F. Campolongo, and M. Ratto. Sensitivity analysis in practice: a guide to assessing scientific models. *Wiley Online Libr.*, 2019.

[6] A. Saltelli, S. Tarantola, and F. Campolongo. Sensitivity analysis as an ingredient of modeling. *Stat. science*, 15(4):377–395, 2000.

[7] J.E Morio. Global and local sensitivity analysis methods for a physical system. *Eur. journal physics*, 32(6):1577, 2011.

[8] E.O. Lantsoght. Database of shear experiments on steel fiber reinforced concrete beams without stirrups. *Materials*, 12(6):917, 2019.

[9] G.K. Uyanık and N. Güler. A study on multiple linear regression analysis. *Procedia-Social Behav. Sci.*, 106:234–240, 2013.

[10] J.A. Hoeting, M. David, A.E. Raftery, and C.T. Volinsky. Bayesian model averaging: a tutorial (with comments by m. clyde, david draper and ei george, and a rejoinder by the authors. *Stat. science*, 14(4):382–417, 1999.

[11] M. Nili and V. Afroughsabet. Combined effect of silica fume and steel fibers on the impact resistance and mechanical properties of concrete. *Int. journal impact engineering*, 37(8):879–886, 2010.

[12] I. Ibrahim and M.C. Bakar. Effects on mechanical properties of industrialised steel fibres addition to normal weight concrete. *Procedia engineering*, 14:2616–2626, 2011.

[13] Y. Şahin and F. Köksal. The influences of matrix and steel fibre tensile strengths on the fracture energy of high-strength concrete. *Constr. Build. Mater.*, 25(4):1801–1806, 2011.

[14] Z.C Zhang. *Experimental study on constitutive relation of steel fiber reinforced concrete under uniaxial compression.* Academic paper, Henan Province, China: Zhengzhou University, 2017.

[15] J. Chen. *The Study of the Effect of Aggregate Size and the Steel Fiber's Length on the Fracture Property of SFRC.* Master's thesis, Zhenzhou University, Zhenzhou, Republic of China, 2016.

[16] M. Bai, D. Niu, L. Jiang, and Y. Miao. Research on improving the mechanical properties and microstructure of concrete with steel fiber. *Bull. chinese ceramic society*, 32(10):2084–2089, 2013.

[17] W. Abbass, M.I. Khan, and S. Mourad. Evaluation of mechanical properties of steel fiber reinforced concrete with different strengths of concrete. *Constr. building materials*, 168:556–569, 2018.

[18] Y. Yang. *Effect of Fiber Dosage and Bond Length on Pullout Mechanical Property Study.* Master thesis, Guangzhou University, 2015.

[19] Fan Cy. *Study on Mechanical Properties of Hybrid Steel Fiber Reinforced Concrete.* PhD thesis, Zhengzhou University: Zhengzhou, China, 2017.

[20] R.B Rajeshwari and M. Sivakumar. Influence of coarse aggregate properties on

specific fracture energy of steel fiber reinforced self compacting concrete. *Adv. concrete construction*, 9(2):173–181, 2020.

[21] C. Gao. *Experimental Research on Mechanical Properties of Concrete and Reinfroced Oncrete after High Temperature*. PhD thesis, Yangzhou University: Yangzhou, China, 2013.

[22] Z. Ma. *An experimental study on the properties of steel fiber reinforced concrete at early ages*. PhD thesis, Beijing University of Technology, 2003.

[23] T. Zhu. *The mechanical properties of steel fiber-reinforced concrete at low fiber content*. PhD thesis, Zhengzhou: Zhengzhou University, 2011.

[24] L. Niu, S. Zhang, and Y. Wei. Effect of fiber dosage on the mechanical property of sfrc. *China Concr. Cem. Prod*, 2019.

[25] H. Peng. *Experimental Research of Steel Fiber Reinforced Concrete on Speed-Repairing the Airport Pavement*. Master thesis, Beijing University of Technology, 2022.

[26] M. Soutsos, T. Le, and A. Lampropoulos. Flexural performance of fibre reinforced concrete made with steel and synthetic fibres. *Constr. building materials*, 36:704–710, 2012.

[27] M. Pająk and T. Ponikiewski. Flexural behavior of self-compacting concrete reinforced with different types of steel fibers. *Constr. Build. materials*, 47:397–408, 2013.

[28] H.-K. Dang, D.-K. Thai, and S.-E. Kim. Stochastic analysis of steel frames considering the material, geometrical and loading uncertainties. *Adv. Eng. Softw.*, 179:103434, 2023.

[29] J. Gareth, W. Daniela, H. Trevor, and T. Robert. An introduction to statistical learning: with applications. *Spinger*, 2013.

[30] T. Ishigami and T. Homma. An importance quantification technique in uncertainty analysis for computer models. *Proceedings. First international symposium on uncertainty modeling analysis. IEEE*, pages 398–403, 1990.

[31] X.H. Dang. Anh huong cac tham so vat lieu, hinh hoc den on dinh cua ket cau vom hai khop. *Tap chi khoa hoc cong nghe xay dung*, 17(2V):134–143, 2023.

[32] A. Alshahrani and S. Kulasegaram. Effect of fibre diameter and tensile strength on the mechanical, fracture, and fibre distribution properties of eco-friendly high-strength self-compacting concrete. *Constr. Build. Mater.*, 403:133161, 2023.

[33] S. He and E.H. Yang. Non-normal distribution of residual flexural strengths of steel fiber reinforced concretes and its impacts on design and conformity assessment. *Cem. Concr. Compos.*, 123:104207, 2021.

About Authors

Tran T. H. LE has received Bachelor of Engineering and Master of Science in Civil Engineering from University of Technology and Education, Ho Chi Minh City in 2021 and 2024, respectively.

Anh-Thang LE received his bachelor's degree in civil engineering from the University of Technology in 1999 and in software engineering from the University of Science Ho Chi Minh City in 2000. In 2003, he obtained a Master of Science in Structural Engineering from the University of Technology, Vietnam, and in 2009, he received a PhD in Highway Engineering from Sejong University, Korea. He is now employed as an associate professor in the Faculty of Civil Engineering at the University of Technology and Education in Ho Chi Minh City, Vietnam. His interesting work includes recycled and waste materials, pavement engineering, and the constitutive modeling of asphalt and cement concrete.

